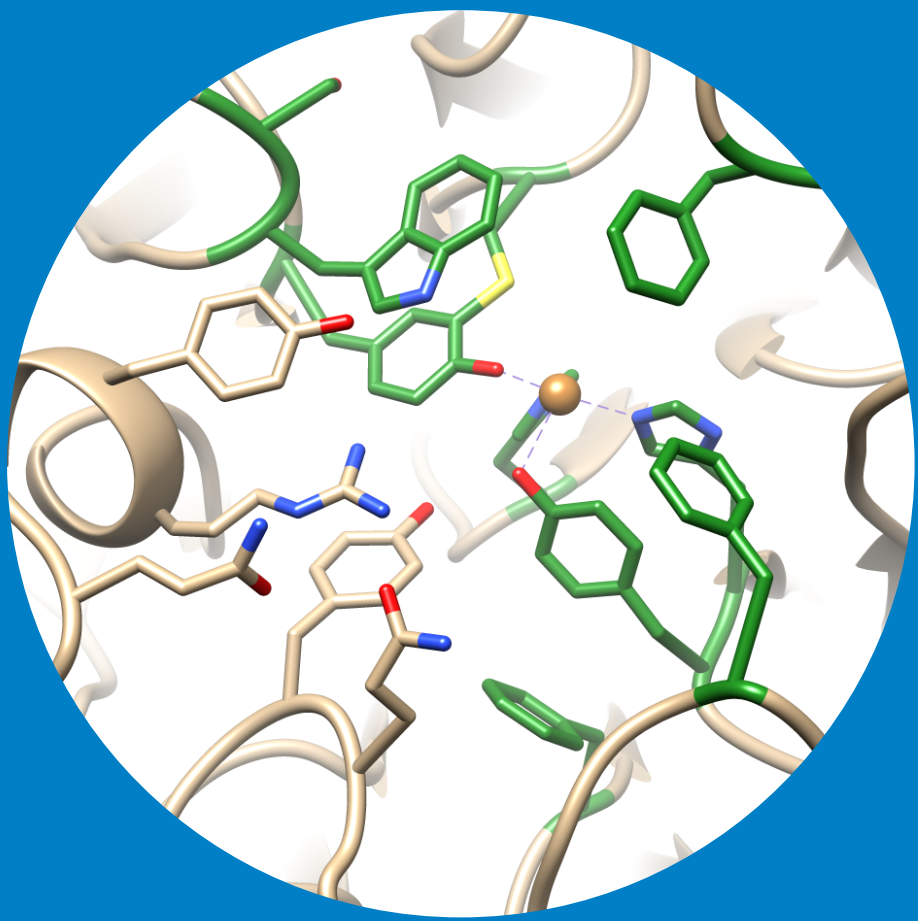
Department of Bioproducts and Biosystems

Engineering & Discovery of CAZyme AA5_2 Oxidases

Filip Mollerup



Aalto University

DOCTORAL DISSERTATIONS

Aalto University publication series
DOCTORAL DISSERTATIONS 232/2018

Engineering & Discovery of CAZyme AA5_2 Oxidases

Department of Bioproducts and Biosystems

Filip Mollerup

A doctoral dissertation completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Chemical Technology, at a public examination held at the lecture hall Ke2 of the school on 30th November 2018 at noon.

Aalto University
School of Chemical Engineering
Department of Bioproducts and Biosystems
Protein Technology research group

Supervising professor
Prof. Emma Master, Aalto University, Finland / University of Toronto, Canada

Preliminary examiners
Doc. Nina Hakulinen, University of Eastern Finland, Finland
Dr. Mikael Lafond, Aix-Marseille Université, France

Opponent
Prof. Dietmar Haltrich, University of Natural Resources and Life Sciences, Vienna, Austria

Aalto University publication series
DOCTORAL DISSERTATIONS 232/2018

ISBN 978-952-60-8309-4 (printed)
ISBN 978-952-60-8310-0 (pdf)
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)
http://urn.fi/URN:ISBN:978-952-60-8310-0

Unigrafia Oy
Helsinki 2018

Finland

Publication orders (printed book):
mollerup.filip@gmail.com

441 697
Printed matter

Aalto University

Abstract

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Author
Filip Mollerup

Name of the doctoral dissertation
Engineering & Discovery of CAZyme AA5_2 Oxidases

Publisher School of Chemical Engineering

Unit Department of Bioproducts and Biosystems

Series Aalto University publication series DOCTORAL DISSERTATIONS 232/2018

Field of research Biotechnology

Manuscript submitted 14 May 2018 **Date of the defence** 30 November 2018

Permission to publish granted (date) 21 August 2018 **Language** English

☐ **Monograph** ☒ **Article dissertation** ☐ **Essay dissertation**

Abstract

The carbohydrate active enzyme (CAZyme) database includes carbohydrate oxidases that have been classified into various auxiliary activity (AA) families. Among these, copper-radical oxidases that target primary hydroxyls are grouped into the AA5_2 subfamily. The most well-characterized to date is galactose oxidase from the aggressive plant pathogenic fungus *Fusarium graminearum* (*Fgr*GaOx). *Fgr*GaOx is one of the most utilized biocatalysts for detection or oxidation of primary hydroxyls in galactose-containing carbohydrates and aliphatic alcohols due to its robustness, high turnover, and use of atmospheric di-oxygen as the terminal electron acceptor, allowing generation of aldehyde products in pH-neutral aqueous solutions at room temperature. The enzyme has been studied and applied in a wide range of green processes, most notably for specific activation of oligo- and polysaccharides in carbohydrate engineering for hydro- and aerogels, carbohydrate-based films for food packaging, and coating of paper surfaces with oxidized hemicellulose. However, applications of *Fgr*GaOx are limited to galactose-containing carbohydrates as the enzyme strictly oxidizes galactosyl residues. This thesis aims to discover and develop family AA5_2 carbohydrate oxidases with distinct activity relative to *Fgr*GaOx. First, an engineering approach was applied to enhance *Fgr*GaOx action. This approach included site-directed mutagenesis and construction of fusion proteins comprising galactomannan or cellulose binding domains. Impacts of fusing the carbohydrate binding modules (CBMs) to *Fgr*GaOx were consistently higher when fusing the CBM to the C-terminus of the enzyme. While CBM fusion increased substrate binding and confirmed that *Fgr*GaOx remained active following adsorption to cellulosic materials, it did not increase the end-point oxidation of polysaccharides by *Fgr*GaOx. Moreover, site directed mutagenesis of *Fgr*GaOx-CBM fusions revealed an apparent trade-off between the catalytic efficiency and substrate range of the enzyme. Comprehensive consideration of characterized AA5_2 oxidases, including engineered variants of *Fgr*GaOx, identified amino acid positions important to catalysis, substrate selectivity, and catalytic efficiency. Resulting sequence-function relationships were used to identify consensus patterns among copper-radical oxidases in AA5_2. This analysis informed the selection of two new AA5_2 members, namely *Cgr*RaOx from *Colletotrichum graminicola* and *Pru*AA5_2A from *Penicillium rubens Wisconsin 54-1255*, for detailed biochemical characterization. In both cases, the substrate preference of the enzymes was clearly distinguished from that of known galactose oxidases. In particular, despite sharing only 23% identity, *Cgr*RaOx and *Pru*AA5_2A preferred raffinose and glycolaldehyde dimer substrates. This discovery uncovered a new connection across the AA5 family, and sheds new light on the biological roles of AA5_2 oxidases.

Keywords Galactose oxidase, protein engineering, CAZYme, enzyme discovery

ISBN (printed) 978-952-60-8309-4 **ISBN (pdf)** 978-952-60-8310-0

ISSN (printed) 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Helsinki **Location of printing** Helsinki **Year** 2018

Pages 152 **urn** http://urn.fi/URN:ISBN:978-952-60-8310-0

Preface

The work comprising this doctoral thesis took place in the Protein Technology research group at Department of Bioproducts and Biosystems in the School of Chemical Technology, Aalto University, Finland. This work was based on funding support from Academy of Finland (ENOX consortium - 1252183); the Finnish Cultural Foundation (individual research grant - 00140620) and the European Research Council (BHIVE project - 648925).

First and foremost, I am deeply grateful to my supervisor and mentor, Prof. Emma Master. You are a true source of inspiration and motivation. I am particularly grateful for all the chances you invested in me, allowing me to endow my own ideas and giving me liberty to shape my own research path. This have taught me so many important lessons. Besides scientific education, a doctoral degree is a character-building transition into professional adulthood, and I am happy I got to mature under your mentorship - thank you Emma.

I like to express my gratitude to my opponent Prof. Dietmar Haltrich (University of Natural Resources and Life Science, Vienna, Austria) for taking on the task to examine this thesis. Likewise, I like to thank Doc. Nina Hakulinen (University of Eastern Finland, Finland) and Dr. Mikael Lafond (Aix-Marseille Université, France) for acting as pre-examiners. I highly appreciate the good feedback which lead to further improvement of this thesis.

I like to thank Prof. Ossi Turunen and Prof. Emer. Matti Leisola for first welcoming me to the laboratory when I came here as a master student. Prof. Alexander Frey acted at my formal supervisor during the first half of my studies. I will like to express my gratefulness to your supporting role and accepting me into your own group during this period. Most of my thesis work was done in collaborations across research institutions. I thank my collaborators from Helsinki University, Dr. Kirsti Parikka and Prof. Maija Tenkanen. Maija, you have been like an instructor and role model to me. I like to thank Dr. Martina Andberg and Prof. Kristiina Kruus from VTT for good cooperation on the challenging work with the raffinose oxidase. It was a pleasure to work with such nice, and skilled people. I also thank Dr. Thu Voung (University of Toronto) for your kind collaboration and help with my work and my visit to Toronto.

I am happy to have so great colleagues and friends at Aalto University. It has been a pleasure to work with my current and former group members, Galina, Ville, Johanna, Mareike, Leamon, Antti, Vera, Martina, Edita and Kim. I am grateful to have worked with my first master's thesis student Laia. Thank you all, for the time we spend together, our interesting (and long) meetings and the support I got from you. Ville, my friend, I love our philosophical lunches and all the stuff we have done outside the lab, trying not to talk about our research but always fail! I would have gone crazy while writing this thesis if it wasn't for you Chris. I will never forget our walks and the sunsets over Otaniemi with a comforting bottle of Writer's Tears. I am also happy for all the colleagues and friends I've gotten to know at the lab: Bart, Esti, Salla, Jorg, Ulla, Janne, Markus, Siiri, Pehzmann, Anne, Mari, Laura, Sasikala, Essi, He Li, Heli, Katja, Suvi, Sesilja, Sami, Georg, Wenwen, Pihla, Katja, Siiri, Salem and Julie-Anne. I have so many memories from here and I am grateful to have met so many interesting people at one place. Jorg and Salem, thank you for all the time we spend in the office together. Julie-Anne, thank for you friendship and help with my stay in Toronto and with my grant application to Finnish Cultural Foundation. I also think of all the people that have been part of my life in Finland and feel lucky to have so many nice and kind people in my life here. My friends from Denmark are a huge part of me and there is nearly nothing better than coming home to visit all of you, it is never often enough.

I would never have made it this far, had it not been for my mom, Jane Mollerup. You always supported me in all my choices through my development and education. That I should become a Doctor of Science was not necessarily in the cards for me, but you have always encouraged me to follow my dreams. Support, hope and perseverance from you is always uncompromised. I learned from you, the courage to step into unknown territories with a belief that I would grow to overcome the challenges on my path. This is a major part of who I am today and I am forever grateful.

Last, I am deeply grateful for all the support and love from you Sini. Thank you for all your loving and unconditional support, you have been my biggest encouragement during the writing of this thesis. Your patience has been immense during the spring of 2018 and I cannot thank you enough for being there for me. I am so happy to have you in my life.

Sincerely,

Filip Mollerup
Helsinki, November 3st, 2018

Contents

List of Publications

This doctoral dissertation consists of a summary and of the following publications which are referred in the text by their numerals.

1. Mollerup, Filip; Master, Emma: (2016) Influence of a family 29 carbohydrate binding module on the recombinant production of galactose oxidase in *Pichia pastoris*. Elsevier, Data in Brief, volume 6, pages 176-183. ISSN 2352-3409. DOI: 10.1016/j.dib.2015.11.032

2. Mollerup, Filip; Parikka, Kirsti; Vuong, Thu V.; Tenkanen, Maija; Master, Emma: (2016) Influence of a family 29 carbohydrate binding module on the activity of galactose oxidase from *Fusarium graminearum*. Elsevier, Biochimica et Biophysica Acta – general subjects, 2, volume 1860, pages 354-362. ISSN 0304-4165, DOI: 10.1016/j.bbagen.2015.10.023

3. Andberg, Martina; Mollerup, Filip; Parikka, Kirsti; Koutaniemi, Sanna; Boer, Harri, Juvonnen, Minna, Master, Emma; Tenkanen, Maija; Kruus, Kristiina: (2017) Identification and characterization of a novel *Colletotrichum graminicola* raffinose oxidase in the AA5 family. ASM journals, Applied and Environmental Mibrobiology, Issue 20, volume 83, e01383-17. ISSN 0099-2240. DOI: 10.1128/AEM.01383-17

4. Mollerup, Filip; Aumalla, Ville; Parikka, Kirsti; Thiesen, Yann; Brumer, Harry; Tenkanen, Maija; Master, Emma: A family AA5_2 carbohydrate oxidase from *Penicilium rubens* displays activity towards glycolaldehyde and raffinose-like galacto-oligosaccharides. Manuscript submitted to PLOS ONE, eISSN: 1932-6203

In addition to above publications, this thesis includes sections (4.2.4, 4.2.5, 4.2.6) ewith unpublished data.

Author's Contribution

Publication 1: Influence of a family 29 carbohydrate binding module on the recombinant production of galactose oxidase in *Pichia pastoris*.
FM planned and executed all aspects of the research, interpreted and wrote the article under the supervision of EM.

Publication 2: Influence of a family 29 carbohydrate binding module on the activity of galactose oxidase from *Fusarium graminearum*.
FM planned the research with EM and with assistance from TM. FM performed the cloning and transformations of recombinant proteins, temperature stability, AGE, activity and kinetics. KP performed the experiments and analysed the data for the degree of oxidation under the supervision of TM. TV performed QCM-D and analysed data from QCM-D and AUC. FM interpreted all results and had preliminary responsibility of writing the article under the supervision of EM. TM assisted in interpretation of the results and writing of the article.

Publication 3: Identification and characterization of a novel *Colletotrichum graminicola* raffinose oxidase in the AA5 family.
Co-first authorship with MA. MA planned the research with KK and performed the sequence analysis, cloning and transformation of recombinant proteins, initial activity screen, shake flask productions, pH optimum and temperature stability under the supervision of HB and KK. FM performed the bioreactor productions, substrate range and kinetics under the supervision of EM. KP, SA and MJ performed NMR and MS and analysis interpretation of these results under the supervision of TM. MA and FM had the preliminary responsibility for writing the article. KK supervised all aspect of the work and writing of the article.

Publication 4: A family AA5_2 carbohydrate oxidase from *Penicilium rubens* displays functional overlap the AA5 family.
FM planned the research under supervision of EM. FM performed the sequence analysis, selection, cloning and transformation of *Pru*AA5_A2, protein models and structural interpretation, substrate range, lag-phase analysis and kinetics. FM and VA performed the expression and purification of *Pru*AA5_2A. VA performed the influence of buffer, pH and temperature optimum and stability. KP performed NMR and MS analysis under the supervision of TM. YT performed phylogenetic analysis under the supervision of HB. FM interpreted all results and wrote the article under supervision EM. KB, YT HB and TM contributed to the writing of the article.

FM planned the research with EM, performed all aspects of the unpublished sections and interpreted the results under the supervision of EM.

List of abbreviations and symbols

ABTS	2,2´-azino-bis(3-ethylbenzothiozoline-6-sulfonic acid)
AGE	Affinity gel electrophoresis
AUC	Analytical ultracentrifugation
CAZy	Carbohydrate Active Enzymes database (CAZy.org)
CAZyme	Carbohydrate active enzyme
CBM	Carbohydrate binding module
CBM29-GaOx	Fusion protein between family 29 carbohydrate binding domain from *Piromyces abies equi* and galactose oxidase from *Fusarium graminearum*
CRO	Copper-radical oxidase
*Cgr*AlcOx	Alcohol oxidase from *Colletotrichum graminicola*
*Cgr*RaOx	Raffinose oxidase from *Colletotrichum graminicola*
EPR	Electron paramagnetic resonance
*Fgr*GaOx	Galactose Oxidase from *Fusarium graminearum*
GaOx-CBM29	Fusion protein between galactose oxidase from *Fusarium graminearum* and a family 29 carbohydrate binding domain from *Piromyces abies equi*
HRP	Horseradish peroxidase
GaOx	Galactose oxidase
GG	Guar galactomannan
GGM	Galactoglucomannan
k_{cat}	Catalytic rate constant, s^{-1}
K_M	Michaelis-Menten constant, mM
LBG	Locust bean galactomannan
Native-PAGE	Polyacrylamide gel electrophoresis under non-denaturing conditions
NMR	Nuclear magnetic resonance

M3-CBM3	Fusion protein between a family 3 carbohydrate binding domain from *Clostridium thermocellum* and the M3 variant of galactose oxidase from *Fusarium graminearum*.
M6-CBM3	Fusion protein between a family 3 carbohydrate binding domain from *Clostridium thermocellum* and the M6 variant of galactose oxidase from *Fusarium graminearum*.
MS	Mass spectrometry
LC	Liquid chromatography
QCM-D	Quartz crystal microbalance with dissipation
*Pru*AA5_2A	AA5_2 oxidase from *Penicillium rubens*
U	Enzyme unit ($\mu mole*min^{-1}$)
XG	Galactoxyloglucan

List of Figures

List of tables

1. Introduction

Plant cell wall polysaccharides, including cellulose, hemicelluloses, and pectins, are the largest source of carbohydrates in the world. Existing applications of these biopolymers includes pulp for making paper and non-woven tissues, textile fabrication and as additives in food and chemical applications. Carbohydrate active enzymes are biocatalysts that play a key role in the essential carbohydrate metabolism in all living organisms, including digestion of carbohydrates for nutrition, modification into other metabolites, energy storage, regulation of cellular processes and structural components of cells. In microbial life, their main role is to facilitate bacterial and fungal extracellular decomposition and modification of lignocellulosic material (cellulose, hemicelluloses and lignin) and other plant carbohydrates during microbial decay of organic matter. Enzymes and associated proteins involved in this process are organized in the carbohydrate active enzyme database (CAZy.org) into families of glycoside hydrolases (GH families), glycosyl transferases (GT families), polysaccharide lyases (PL families), carbohydrate esterases (CE families), carbohydrate binding modules (CBM families) and auxiliary activity oxidoreductases (AA families).

Biotechnological developments have harnessed these biocatalysts mainly for enzymatic degradation of lignocellulosic biomass to corresponding mono- and oligosaccharides for fermentation to biofuels and biochemicals. However, in recent years biocatalysts have also been used in carbohydrate engineering whereby the backbone of intact polysaccharides is tailored for use in bio-based, biodegradable materials.

1.1 Carbohydrate oxidoreductases

The carbohydrate active enzyme (CAZy) database was expanded in 2013 to include auxiliary activities (AA families; Levasseur et al., 2013), as these activities are increasingly recognized as playing an important role in microbial degradation of biomass and modification of cell wall carbohydrates in plants and microorganisms. AA families found to contribute directly or indirectly to lignin or aromatics modification include AA1 (laccase and multicopper oxidases), AA2 (manganese and lignin peroxidases), AA4 (vanillyl-alcohol oxidase), AA5_subfamily 1 (glyoxal oxidases), AA6 (1,4-benzoquinone reductase), and AA8 (iron reductase domain). Lytic polysaccharide monooxygenases (LPMOs) are exclusively active on oligo- and polysaccharides and

present a unique mechanism of backbone cleavage through oxidation of the glycoside bond by oxidation of C1 or C4 positions, while some LPMO's potentially also oxidize C6-hydroxyls (Chen et al., 2018). Such LPMOs belong to families AA9, AA10, AA11, AA13, AA14, and AA15, and play an important role in lignocellulose degradation by facilitating the access of other CAZymes to recalcitrant lignocellulose (Katja Salomon Johansen, 2016). Notably, LPMOs have been shown to accept electrons and hydrogen peroxide generated from other CAZyme oxidoreductases (Kracher et al., 2016).

Families AA3, AA5, and AA7 include carbohydrate oxidoreductases that introduce a carbonyl at specific positions within carbohydrate substrates while using dioxygen or other electron acceptors. Enzymes from families AA3 and AA7 are FAD-dependent with reported activity on monosaccharides and oligosaccharides. Family AA5 members are copper dependent and further distinguished by the action of certain sub-family 2 members on polysaccharides, in addition to monosaccharide and oligosaccharide substrates (Baird and Smith, 1989; Delgrave et al., 2001; Parikka et al., 2010).

1.1.1 Overview of CAZy family AA5 enzymes

The CAZy family AA5 consists of copper-radical oxidase (CRO), containing a single copper-ion, and comprises two subfamilies. Subfamily 1 consists solely of the aldehyde-active glyoxal oxidase whereas subfamily 2 contains hydroxyl-targeting carbohydrate and alcohol oxidases. Both families adopt similar modular structure over their catalytic domains and follow the same catalytic mechanism. This thesis work investigates the galactose oxidase from subfamily 2 and new members of this subfamily. Below follows a short introduction to subfamily 1; the remainder of this thesis will then focus on subfamily 2 (AA5_2).

Glyoxal oxidases and their biological functions were recently reviewed in Isobe et al., 2012, and Kersten and Cullen, 2014. Glyoxal oxidase can be considered the physiological partner of lignin peroxidase and manganese peroxidase, given the substrate range and production of hydrogen peroxide (Kersten and Cullen, 2014; Whittaker, 2003). Glyoxal oxidase oxidizes aldehydes and α-hydroxycarbonyls, including glyoxal, glycolaldehyde, glycolaldehyde dimer and methylglyoxal, which are derived from the peroxidase-driven lignin degradation, free-radical mediated carbohydrate fragmentation, or fungal metabolites (Hammel et al., 1994; Kersten and Cullen, 2014; Kersten and Kirk, 1987). The glyoxal oxidases adopt a similar overall structure and redox chemistry as the core catalytic domain of the subfamily AA5_2 oxidases (Whittaker et al., 1996). They differ from the AA5_2 enzymes, however, in terms of the substrate binding site and by generally possessing higher redox potentials (640 versus 430 mV) and less stable Cys-Tyr radical cofactor compare to the archetypal AA5_2 oxidase: galactose

oxidase (Johnson et al., 1985; Whittaker et al., 1996). The reduced stability of the Cys-Tyr radical has been explained by the presence of a histidine in place of the stacking indole ring of W290 in *Fgr*GaOx (Rogers et al., 2007), which may also explain the inability of glyoxal oxidases to oxidize hydroxyl groups.

At the beginning of my doctoral studies, CAZyme family AA5_2 was mainly understood through detailed characterizations of a single representative enzyme: galactose oxidase *(Fgr*GaOx*)* from the aggressive plant pathogenic fungus *Fusarium graminearum*. Other known AA5_2 members from the *Fusarium* genus, such as, *F. fujikuro* (Aisaka and Terada, 1982), and *F. acuminatum* (Alberton et al. , 2007), as well as the ones added during the period of this thesis from *F. oxysporum* and *F. sambucinum* (*Fox*GaOx and *Fsa*GaOx) (Paukner et al., 2015, 2014), all show comparable substrate preference to *Fgr*GaOx. However, they have not been tested for oxidation of polysaccharides (non-lytic), which is so far suggested to be a unique feature galactose oxidase. Reported differences relative to *Fgr*GaOx include 3.8 and 3.0 times lower K_m values for *Fsa*GaOx on melibiose (16 mM) and raffinose (20 mM) compared galactose (61 mM) (Paukner, 2015), raising attention to these oligosaccharides. Recent characterization of AA5_2 enzymes from fungi besides *F. spp.* represent an important expansion of the known diversity within the family, including new enzyme functionality among AA5_2 copper-radical oxidases. These are alcohol oxidases from *Colletotrichum graminicola* (*Cgr*AlcOx) and *C. gleoeosporiodes* (*Cgl*AlcOx) (Yin et al. , 2015). The catalytic domain of *Cgr*AlcOx adopts a similar structure to that of *Fgr*GaOx, *Cgr*AlcOx lacks the N-terminal carbohydrate binding module present in *Fgr*GaOx (family CBM32) and does not oxidize carbohydrates but instead targets a broad range of aliphatic alcohols, such as 1-heptanol and cinnamyl alcohol, with greater catalytic efficiencies than *Fgr*GaOx on D-galactose.

1.2 Galactose oxidase

Galactose oxidase (GaOx, CAZy family AA5, E.C.:1.3.3.9) was considered the simplest member of the copper dependent oxidase family, which also include other famous members such as mono amine oxidase, laccase, ascorbate oxidase, tyrosinase, dopamine mono-hydroxylase and alcohol oxidases. An oxidase with the unique ability to oxidize the primary hydroxyl of D-galactose, thus named galactose oxidase, was first reported in 1959 from *Polyporus circinatus* cultures growing on galactose in mineral medium (Cooper et al., 1959). *P. circinatus* is from the Ascomycota division, and was later re-identified as the plant pathogen *F. graminearum,* the anamorph of *Gibberella zeae* (Ögel et al., 1994). Whereas Cooper et al already established in 1959 that galactose oxidase is a metalloenzyme, the copper requirement of the enzyme was confirmed by Amaral et al in 1963. (Amarel et al., 1963).

The range of substrates targeted by galactose oxidase has since then expanded to include galactose derivatives such as 1-methyl-β-galactopyranoside (Whittaker and Whittaker, 2001) and galactose-containing oligo-and polysaccharides including: lactose, melibiose, raffinose, galactoxyloglucan, galactomannan and galactoglucomannan (Delgrave et al., 2001; Parikka et al., 2010) (Figure 1), as well a broad range of aliphatic primary alcohols such as ethanol, 1-propanol, 1,2-propanol, butanol and polyols such as glycerol (Whittaker and Whittaker, 2001).

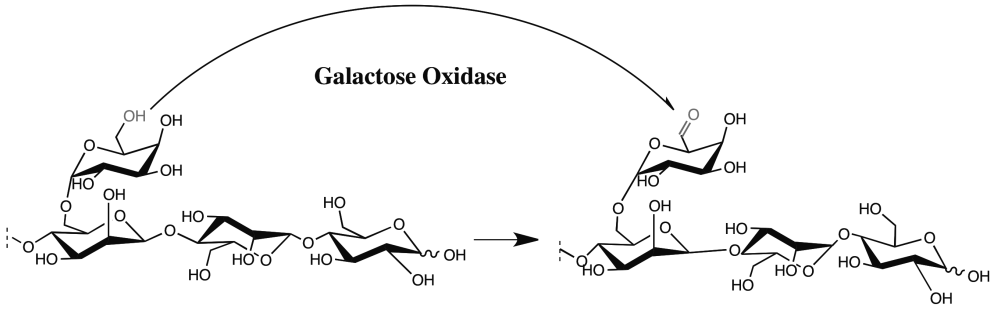


Figure 1. Reaction catalysed by galactose oxidase
*Fgr*GaOx oxidizes the C6-hydroxyl of D-galactose and terminal galactosyl residues in oligo- and polysaccharides into the corresponding aldehyde.

The broad range of galactose-containing oligo- and polysaccharides and small primary alcohols accepted by galactose oxidase is partly explained by the comparatively open substrate-binding and shallow active site (Firbank et al. , 2001, Ito et al. , 1991). Despite this, the selectivity among carbohydrates is strict for D-galactose where practically no activity on other sugars has been observed. Galactose oxidase from *F. graminearum* (*Fg*rGaOx) achieves its highest catalytic turnovers on mono-D-galactose, simple galactose derivatives and small oligosaccharides where turnover rates for galactose-containing polysaccharides are approximately 100-fold lower (Parikka et al., 2010).

Tracking galactose oxidase activity

Due to the instability and reactivity of carbohydrate-aldehydes, quantification of GaOx derived oxidation products is difficult. Nuclear Magnetic Resonance (NMR) is most commonly used for detection of aldehyde derivatives and can be used to estimate the degree of oxidation of monosaccharides when the sample is analysed fresh or carefully stored frozen (Parikka and Tenkanen, 2009). Quantification of the degree of oxidation of mono-, oligo- and polysaccharides can be analysed by gas-chromatographic mass-spectrometry (GC-MS) after deuterium labelling by reduction of the aldehyde with sodium borodeuteride (Hartmans et al., 2004; Parikka et al., 2012, 2010). This also stabilizes the oxidation products, eliminating formation of derivatives and crosslinking, and enables analysis by electron spray ionization mass-spectrometry (ESI-MS) (Parikka et al., 2012b; Xu et al., 2012). Detection of direct turnover rates

and *Fgr*GaOx activity is, however, most conveniently done by following oxygen consumption or using a peroxidase coupled hydrogen peroxide assay such as the HRP/ABTS and HRP/Amplex red assays used in thesis.

1.2.1 Structure and redox chemistry of galactose oxidase

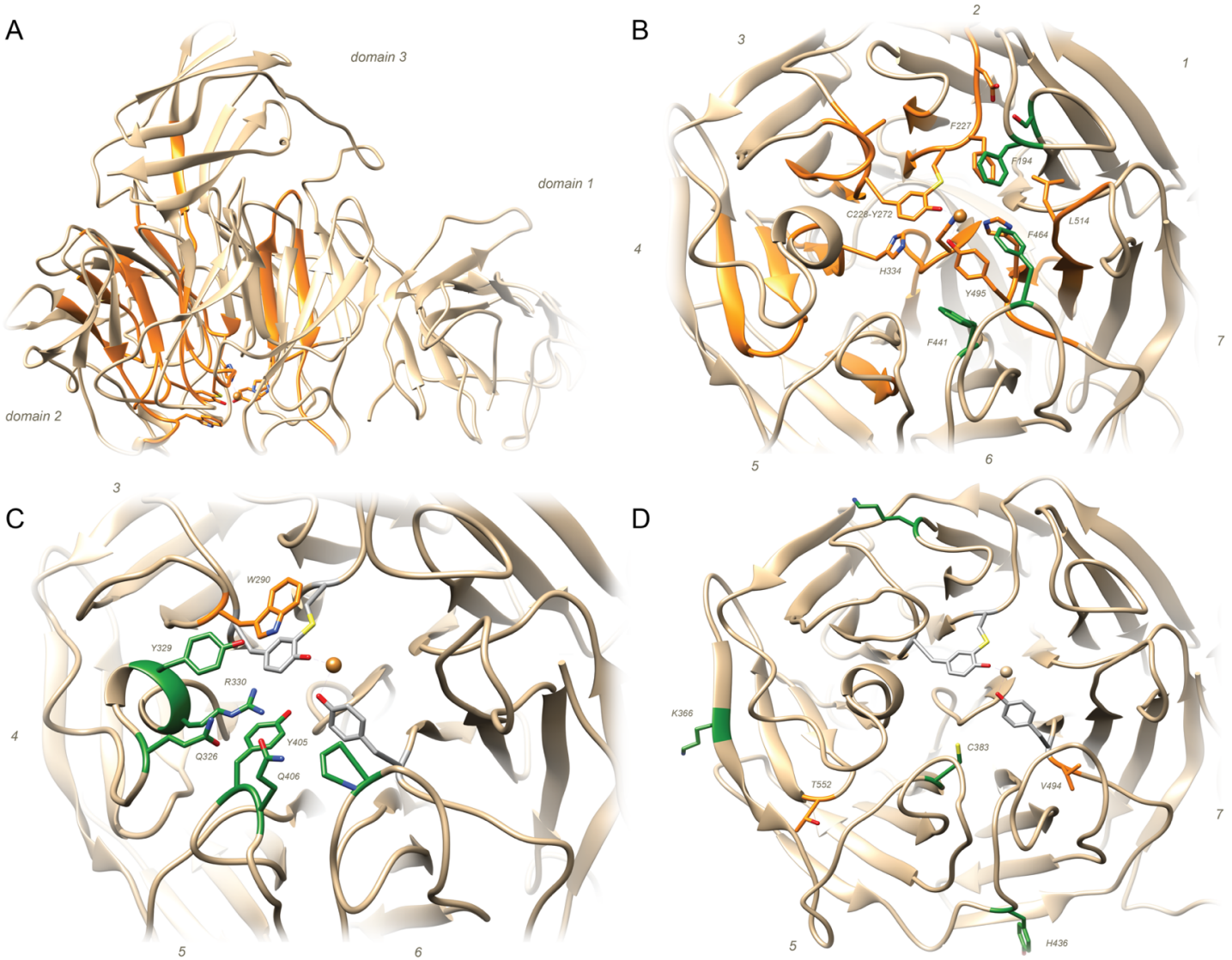
Overall structure

The *Fgr*GaOx structure is often used as an archetypal reference for CAZy family AA5_2 members and other CRO's exhibiting a similar structural fold. Accordingly, *Fgr*GaOx is a frequently used reference for studying structure-sequence relationships in AA5_2 oxidases. The first high resolution crystal structure of *Fgr*GaOx was published in 1991 (Ito et al., 1991). When first determined, the structure revealed a unique active site architecture including a novel internal tyro-cysteine co-factor, which made *Fgr*GaOx the subject of numerous crystallization, magnetic resonance and spectroscopy studies (Baron et al., 1994; Cowley et al., 2016; Firbank and Rogers, 2001; Ito et al., 1994; Lee et al., 2008; Reynolds et al., 1997). These studies show that *Fgr*GaOx (68 kDa) forms three closely embedded domains: An N-terminal carbohydrate binding module (Type F5/8), a central catalytic domain (β-propeller fold) and a C-terminal domain of unknown function (DUF1929) (β-barrel fold) (Figure 2A). The N-terminal domain (domain 1) consists of eight β-strands in a β-sandwich fold containing a cation ligand site and a subsidiary binding pocket, remote from the catalytic domain (Ito et al., 1994). Although the exact function of this domain remains unclear, crystals soaked in 400 mM D-galactose showed that galactose binds to the subsidiary binding site. The domain was later recognized as a carbohydrate binding module (CBM) from family 32 (CAZy database), whose other members bind galactose, lactose, N-acetyl-galactosamine, polygalacturonic acid, β-mannooligo- and polysaccharides as well as a wide range of galactose-containing oligosaccharides (Ficko-Bean and Boraston, 2018). Despite this, affinity studies to date have not detected galactose binding by *Fgr*GaOx. Likewise, analyses performed through this thesis did not reveal *Fgr*GaOx binding to galactose-containing polysaccharides (section 4.2.1). Alternatively, the N-terminal CBM32 domain of *Fgr*GaOx could potentially play a role in enzyme folding, due to its close association with the catalytic domain (Ito et al., 1994). Although the domain is a carbohydrate binding module, a chaperone like function is plausible considering that one of the N-terminal β-sheets continues into the β-fold of blade 7 in the catalytic domain (Figure 2A). Moreover, removal of the CBM32, or replacement with another carbohydrate binding domain, leads to loss of functional *Fgr*GaOx expression (McPherson et al., 1993; Mottiar, 2012).

The central domain (domain 2) is a 7-bladed β-propeller fold where each blade is made up of four antiparallel β-strands forming a kelch_1 motif (Figure 2A and B). Namely β-propeller folds are also known as kelch-repeat proteins. This fold is also adopted by eukaryotic ligand

binding proteins, neuramidase glycoside hydrolase, methylamine dehydrogenase, nitrile reductase (Whittaker, 2003), as well as a large range of CRO's, that often are classified as galactose oxidase-like enzymes; however, the majority of these are likely not galactose oxidases or even carbohydrate active oxidases despite the similarity. Instead, this fold may be characteristic for a range of CRO's with rather diverse substrate range of biological roles. Domain 2 is entirely made up of these β-structure repeats and loop regions, where only a single α-helix turn is present (Ito et al., 1994). Referral to this helix-turn is limited to a few remarks of its existence in the available literature, but it is positioned on blade-4 and forms part of the substrate binding region of the active pocket by providing two essential substrate ligands, namely Y329 and R330 in *Fgr*GaOx (Figure 2C). Accordingly, this secondary-structure motif may have importance for the substrate preference in galactose oxidases. Domain 2 contains the copper-radical site positioned at the bottom of the propeller fold, and three of four copper ligands (Y272, Y495 and H496). The fourth copper ligand (H581) is positioned on a long loop from the C-terminal DUF1929 domain (domain 3) (residues 576 – 588), which is positioned on top of the catalytic domain. The loop extends 20 ångstrom down through the centre of the catalytic domain to reach the active site and donate H581 to the copper coordination sphere. Besides the supply of a copper ligand, not much is known about the function of the DUF1929 domain in *Fgr*GaOx. However, its presence in sequences as a C-terminal addition to the β-propeller domain is used to identify genes encoding copper-radical oxidases similar to galactose oxidase, such as alcohol and glyoxal oxidase and other genes with low homology, like the raffinose oxidase discovered in this thesis or the GlxA from *Streptomyces lividans*.

Figure 2. Structure of galactose oxidase from *Fusarium graminearum*
The figure was produced using the crystal structure of FgrGaOx (pdb code 1gog). (A) Distribution of the conserved AA5_2 sequence segments (orange) in the structure of *Fgr*GaOx (for later analysis in 4.3.1, Figure 19). Domain 1 in *Fgr*GaOx is the N-terminal CBM32 (F5_F8_type_C domain). The Cys-Tyr cofactor is noted by C228-Y272; Domain 2, the central catalytic domain (β-propeller fold), contain the active site at the bottom; Domain 3 (the C-terminal DUF1929 domain) sits in top of the catalytic domain and contains one of the copper ligands (H581) (see also Figure 3). (B) Bottom-up view of the catalytic domain (domain 2 in Figure 1A) showing the tight clustering of amino acids important to catalytic function in AA5_2 (Group 1 in Table 1) towards one half of the active site, consistent of blade 7, 1, 2 and 3 in the β-propeller fold. Group 1 residues that are positioned in a conserved sequence segment are shown in orange, whereas conserved residues S188, F194, F441 and F464 (shown in green) are positioned in more variable regions of the *Fgr*GaOx sequence (see section 4.3.1). (C) Distribution of amino acids contributing to substrate preference in *Fgr*GaOx (Group 2, Table 1) are clustered towards the opposing half of the active site covering blade *4*, *5* and *6*. All residues (green), except W290 on blade *3* (orange), are positioned outside the conserved sequence segments. The redox active tyrosine's (grey) and the copper(II)-ion are shown for reference. (D) Amino acid positions where substitutions leading to enhanced catalytic activity of *Fgr*GaOx. Four of the six residues (shown in green) are outside a conserved sequence stretch whereas T352 and V494 (shown in orange) are within a conserved sequence stretch.


A
domain 3
domain 1
domain 2
B
2
3
1
F227
F194
C228-Y272
L514
F464
4
H334
Y495
F441
7
5
6
C
3
W290
Y329
R330
4
Y405
Q326
Q406
5
6
D
K366
C383
T552
V494
7
5
H436

Amino acids contributing to the redox chemistry of galactose oxidase

As described above, the active site of *Fgr*GaOx contains a copper-based redox centre at the bottom of a shallow pocket (Figure 2B), revealing that *Fgr*GaOx adopts an exposed redox-centre (Ito et al., 1994; McPherson et al., 1993). The copper-cofactor is coordinated by five ligands in a square-based pyramidal geometry, where the copper occupies the centre of a nearly square plane formed by the side-chains of Y272, H496, H581 and a solvent molecule expected to be water (Baron et al., 1994); however, the solvent molecule can be replaced with other compounds, such as acetate or azide (Rogers et al., 2007). The axial member in the pyramidal geometry is the phenol of a second tyrosine ligand (Y495), which is in a slightly distorted position from the ideal axis for the coordination, consequently reducing the bond strength to the copper-ion (Figure 3).

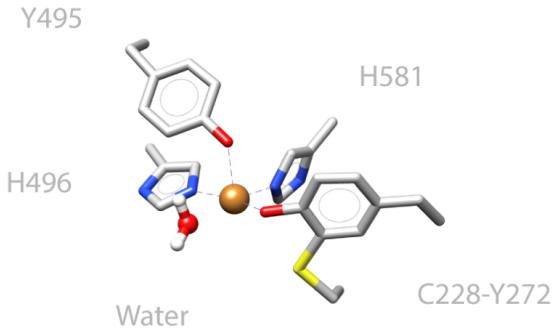


Figure 3. Copper ligands and water in the first coordination sphere of *Fgr*GaOx

*Fgr*GaOx contains two redox agents, namely the copper-ion at the centre of the coordination sphere and an internal Cys-Tyr radical cofactor which is formed by a linkage between the Y272 ligand and C228 generated during the maturation of the holoenzyme (Firbank and Rogers, 2001; Rogers et al., 2000). Formation of the Cys-Tyr cofactor follows occupation of a copper in the active site, where both maturation steps are processed by galactose oxidase alone, requiring only free copper and molecular oxygen. In addition to the copper, experiments and modelling have confirmed that the Cys-Tyr cofactor is also a redox agent in *Fgr*GaOx and is specifically adapted to accommodate the stable free radical electron present in the active form of the enzyme. In proteins where tyrosine plays an active role in electron transfer, covalent modification of the phenol sidechain particularly by redox active sulphur compounds such as the sidechain of C228, can significantly alter the redox potential of the corresponding tyrosine (Warren et al. , 2012). The C228-Y272 linkage also aids to fixate the orientation of the phenol ring of Y272 relative to the copper, permitting electron orbital overlap that facilitates

delocalization of the free radical over the copper ion (Baron et al., 1994; McPherson et al., 1993). Accordingly, the stability of the free radical in the active oxidation state is unusually robust in *Fgr*GaOx and presumably other family AA5_2 enzymes (Kempner et al. , 2010, Lee et al. , 2008).

The catalytic mechanism

During catalysis, *Fgr*GaOx switches between a fully oxidized (Cu(II)-Y•) and a fully reduced (Cu(I)-Y) state through a two-electron redox turnover (Figure 4). Owing to the presence of two redox agents, a third state is also possible which forms through a one-electron reduction of the cys-tyr free-radical site to generate the semi-reduced form of the enzyme (Cu(II)-Y). Due to the loss of the free radical, this state is stable but inactive, and so its spontaneous occurrence causes the activity of the enzyme to diminish during storage. However, the oxidized state (active) can be regenerated by oxidation of semi-reduced GaOx with a suitable oxidation agent, such as $K_3Fe(CN)_6$, H_2IrCl_6, EDTA-Mn^{2+} or some oxidoreductases such as peroxidase (Kwiatkowski and Kosman, 1973; Rogers et al., 2007; Whittaker and Whittaker, 2003). Accordingly, even though the free radical in the active form of GaOx is rather stable, the enzyme is often a mix between active (oxidized) and inactive (semi-reduced) states after storage, and it is recommended to incubate with an oxidation agent or peroxidase prior activity measurement to insure a solution of fully active *Fgr*GaOx.

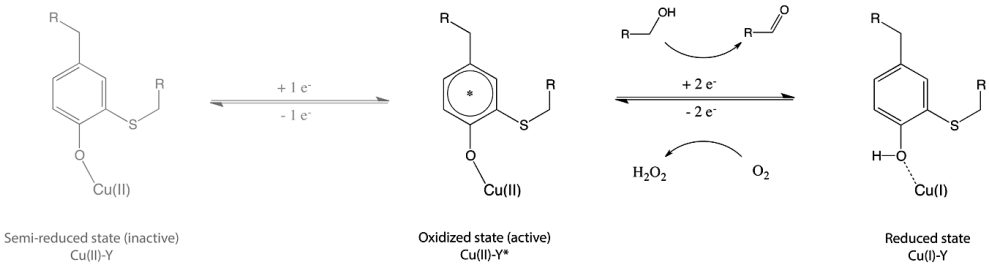


Figure 4. Structure of the cysteine-tyrosine radical cofactor and redox states of *Fgr*GaOx
The Cys-Tyr radical cofactor and copper-ion cofactors circulates between oxidation states during the catalytic $2e^-$ turnover cycle (right scheme). Spontaneous single electron reduction of the Cys-Tyr cofactor results on a semi-reduced state which is inactive (left). Activation of *Fgr*GaOx in the semi-reduced state is possible by treatment with a suitable oxidation agent. This figure was drawn on the basis of Rogers et al., 2007

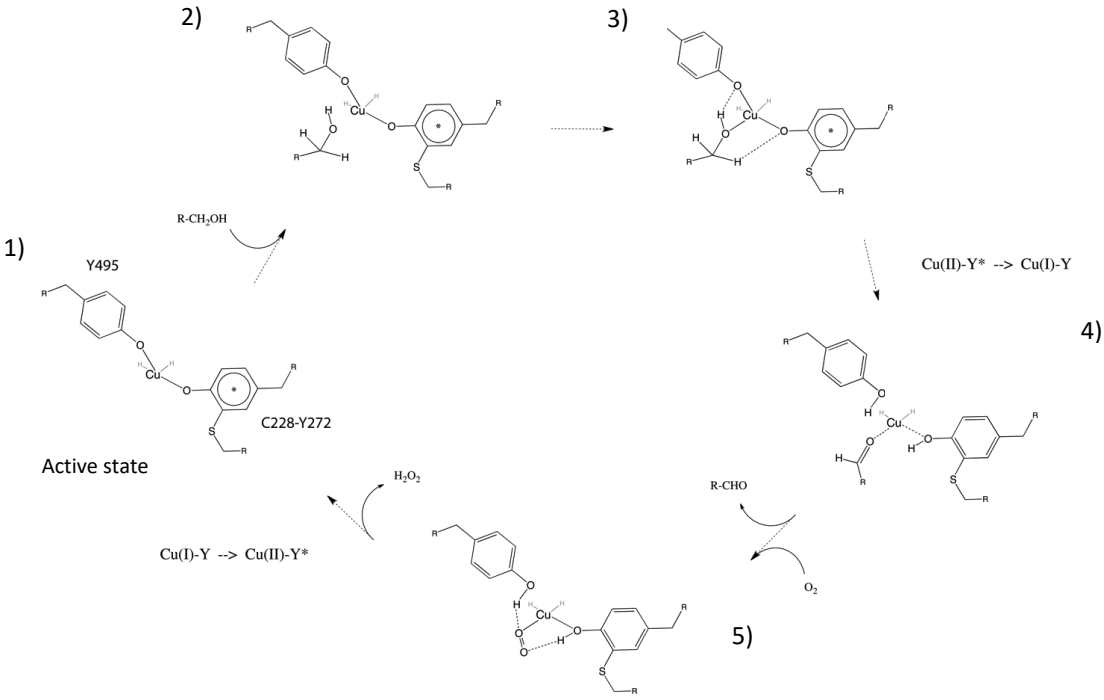


Figure 5. Putative reaction mechanism for galactose oxidase

Although the reaction mechanism of *Fgr*GaOx cannot be fully confirmed experimentally, there is consensus around a suggested model that which12 based on experimental data and molecular modelling, and accounts for the shift in oxidation states and the loss of two ligands to the reduced copper (Cu(I)). The reaction mechanism has been thoroughly discussed in reviews (Whittaker, 2005, 2003), and experiments designed to investigate the redox chemistry, particularly the role of the essential residues Y272, C228 and Y495, are in good agreement with the suggested mechanism (Baron et al., 1994; Firbank and Rogers, 2001; Kempner et al., 2010; Lee et al., 2008; McPherson et al., 1993; Reynolds et al., 1997; Rogers et al., 2000; Whittaker et al., 1989; Whittaker and Whittaker, 2001, 1993).

The catalytic turnover is based on two half-reactions, where the first half-reaction involves the two-electron oxidation of an alcohol substrate to the corresponding aldehyde (Figure 5 (1)). During substrate binding, the alcohol group replaces the solvent molecule in the copper coordination sphere forming a bond between the alcohol oxygen and Cu(II) (2). The phenol sidechain of Y495 functions as a base by abstracting the proton from the alcohol hydroxyl and thereby freeing electrons for the formation of the carbonyl bond of the aldehyde product (3). While accepting a proton, Y495 breaks its bond to copper, reducing the coordination number of the copper ion. Simultaneously, the phenoxyl radical of the Cys-Tyr cofactor catalyses the breakage of the substrate C_α-H bond by accepting the hydrogen atom (proton + electron) which effectively reduces the free radical and allows C_α to form the carbonyl. This generates a

surplus electron on the resulting aldehyde group which is then accepted by Cu(II) (4), to generate Cu(I). The reduced copper is only coordinated by Y272, H496 and H581 after product dissociation, which was confirmed by electron paramagnetic resonance (EPR) (Reynolds et al., 1997). In the second half-reaction, reduced GaOx (Cu(I)-Y) reduces dioxygen by transferring two protons and two electrons to generate hydrogen peroxide and the oxidized form of GaOx (Cu(II)-Y$^{\cdot}$) via an essentially reversed process (5).

The role of aromatic residues in the active pocket

One of the observations made during this thesis work suggest that the redox centre of galactose oxidase (i.e., Domain 2) is surrounded by a rich network of conserved aromatic residues important for CRO-functionality and substrate binding residues that together form the edge of the substrate-binding pocket (Figure 2B). The abundance of aromatic amino acids in *Fgr*GaOx was also noted in connection with the first crystal structure of *Fgr*GaOx (Ito et al., 1991) and also mentioned, although not thoroughly discussed in other publications. Though to be particularly important, the tryptophan residue (W290) plays a role in stabilizing the free radical in the Cys-Tyr cofactor (Figure 2C). The indole ring of the tryptophan sidechain is situated directly above the phenyl ring of Y272, providing π - π ring stacking interactions which shield the redox centre from solvent exposure and could further contribute to the stability of the oxidized form by delocalizing the free electron radical (Baron et al., 1994). However, EPR data indicate the radical delocalization of the tryptophan sidechain does not occur (Rogers and Dooley, 2001); nevertheless, the stacking tryptophan significantly impacts the Cys-Tyr cofactor by lowering the redox potential (Rogers et al., 2007). The residue is also suggested to take part in binding of galactose by supplying a hydrogen bond to the hemiacetal oxygen in the sugar ring (Ito et al., 1994; Lippow et al., 2010; Rannes et al., 2011).

Other residues in the active pocket of *Fgr*GaOx that are not considered to directly affect substrate selectivity, but appear to be important for CRO-functionality, include: F227 and H334, which surround the Cys-Tyr cofactor and could potentially play a role in the stability and redox potential of the cofactor in conjunction with W290 (Figure 2B). F194 is the near neighbor to W290 in the edge of the active pocket and may help shield the copper from solvent, however this residue may also have importance for substrate acceptance. The basic tyrosine, Y495 in *Fgr*GaOx, is surrounded by F463 and F441 which may shield it from solvent and coordinate the position of Y495. These amino acids are clustered in a half circle around the active site stretching clockwise from blade 3 to 6. Residues suggested via this thesis to play a fundamental role in the CRO-functionality of *Fgr*GaOx are collected in group 1 Table 1 and shown in the structure of *Fgr*GaOx in Figure 2B. Particular to this group is that these residues seem not to be significantly involved in coordination of D-galactose. This group also contains the copper-

ligands and other residues that are not mentioned in the text but likewise considered important elements in the active pocket, as will be further discussed in 4.3.1.

Substrate ligands

Although attempts have been made, there are no crystal structures of *Fgr*GaOx with galactose in the active site. Consequently, identification of the substrate interacting residues in *Fgr*GaOx is based on substrate binding models and empirical data from mutagenesis. The first model of the enzyme-substrate complex predicted that R330 forms bidentate hydrogen bonds to the C3- and C4-hydroxy groups of galactose, while Q406 forms an additional hydrogen bond to the C2-hydroxyl of galactose (Ito et al., 1991). Accordingly, R330 can be viewed as the main substrate ligand to depict the specificity of *Fgr*GaOx towards galactose (See also Figure 18B). This was later confirmed by the introduction of glucose activity in *Fgr*GaOx by W290F, R330K and Q406T substitutions in the M3 variant (Sun et al., 2002). Also, Qln326 forms a hydrogen bond to R330 and thereby influences the positioning of this substrate ligand (Lippow et al., 2010). Mutagenesis studies suggest that Y329 also plays an important role in substrate interaction, where its substitution to lysine can induce a new hydrogen bond to glucose at the C1-hydroxyl (Lippow et al., 2010). Substrate docking performed by me on *Fgr*GaOx with galactose, suggest that the phenol moiety of Y329 is 2.8 Å away from the C1-hydroxyl of galactose, which is considered the range of hydrogen bonds with moderate strength (Jeffrey, 1995) (results not showed). Although not thought to interact directly with galactose, Y405 and P463 may indirectly impact substrate binding as mutagenesis of these positions affect the substrate preference of *Fgr*GaOx variants by increasing the activity of the M3 variant on glucose, mannose and N-acetetylglucosamine (Rannes et al., 2011).

In addition to the residues highlighted above, W290 has drawn attention given its importance to the catalytic efficiency of *Fgr*GaOx on D-galactose. Based on mutagenesis studies, W290 was suggested to bind galactose through a hydrogen bond to the hemiacetal oxygen in the galactose ring (Rogers et al., 2007). A W290F substitution led to drastic decrease in affinity to galactose (K_M of 2950 mM versus 82 mM for *Fgr*GaOx), while the catalytic turnover of the mutant was only slightly reduced (k_{cat} of 371 s^{-1} versus 503 s^{-1}). As mentioned, W290F also leads to gain in glucose oxidation by the M3 variant of *Fgr*GaOx (Sun et al., 2002). Accordingly, W290 may play a double role in the active pocket as the aromatic stacking interaction to Y272 is clearly essential for the redox potential and stability of the Cys-Tyr cofactor. Yet, the W290F substitution does not impart a significant change in the redox potential or radical stability, suggesting that radical stabilization can be facilitated by other aromatic residues, while the indole ring of W290 also provides binding to D-galactose (Rogers et al., 2007).

Substrate ligands and residues expected to play an important role in substrate interaction are summarized in group 2 of Table 1. These residues cluster together on blade 4 and 5 primarily, forming a region of substrate interactive amino acids which is distinctively opposite the cluster of group 1 amino acids (Figure 2C).

Amino acids with impact on activity

Amino acids that are not essential for the CRO-functionality nor implicated in binding of D-galactose but seemingly impact the catalytic activity or stability of *Fgr*GaOx (Delgrave et al., 2001; Sun et al., 2001; Wilkinson et al., 2004) are listed in group 3 of Table 1 and displayed in Figure 2D. In contrast to the clustering of group 1 and 2, residues in group 3 seems broadly distributed throughout domain 2 of *Fgr*GaOx. Four of the six residues in group 3 are solvent exposed surface residues (K248, T352, K366 and Y436), whereas C383 is positioned in close proximity to Y405 and F441, and V494 to F464 and Y495. V494 is also a neighbour to Y495. Accordingly, deviations or mutations at C383 and V494 may affect core catalytic properties. For example, mutations V494A and C383S were shown to independently double V_{MAX}, and C383S alone decreased the K_M of *Fgr*GaOx on galactose (Delgrave et al., 2001; Wilkinson et al., 2004).

Table 1. Important amino acids in *Fgr*GaOx

Group 1			**Group 2**		
188	S	Hydrogen bonds to N225	290	W	Ring stacking interaction to Y272, substrate ligand, mutation entails great loss of catalytic activity
194	F	Shielding copper(II) from solvent, ring stacking interaction with F227	326	Q	Hydrogen bonding to R330
225	D	Hydrogen bonds to S188	329	Y	Substrate ligand and hydrogen bonding to R330
227	F	Ring stacking interaction with Y272	330	R	Substrate ligand, bidentate hydrogen bonds to C3-OH and C4-OH of galactose
228	C	Thioether linkage in C228-Y272 cofactor	405	Y	Positioned in the active site, hydrogen bonds to Y495 and R330
272	Y	The catalytic tyrosine, copper ligand	406	Q	Substrate ligand, hydrogen bond to C2-OH of gal
291	S	Positioned in internal cavity	463	P	Mutation considered detrimental to Gal or Glc activity
334	H	Hydrogen bond to Y405	**Group 3**		
441	F	Ring stacking interaction to Y495	248	K	Surface residue, substitution increases activity towards galactose
464	F	Potential substrate interaction	352	T	substitution increases activity of *Fgr*GaOx on gal
495	Y	Copper ligand	366	K	Surface residue, increases activity on gal
496	H	Copper ligand	383	C	Mutation increases V*max* on gal by 1.75x and lowers K_M by 3.6
514	L	Positioned in a hydrophobic cavity	436	Y	Mutation increases V*max* on gal by 2x
581	H	Copper ligand	494	V	Internal hydrophobic cavity, substitution increases V*max* on Gal by 1,75x

Amino acids considered important for the copper radical oxidase functionality in FgrGaOx (group 1), D-galactose (group 2), have impact on the activity of *Fgr*GaOx (group 3). These residues are used in further sequence analysis of AA5_2 oxidases (4.3.1)

1.3 Application and engineering of galactose oxidase

Oxidation of alcohols to carbonyls is one the most important reactions in synthetic chemistry and the application potential for galactose oxidase is therefore very high. Using *Fgr*GaOx for oxidation of alcohols and carbohydrates in general was recently reviewed (Goswami et al., 2013; Siebum et al., 2006), and specifically for oxidation of carbohydrates (Galante et al., 2018; Parikka et al., 2015). Due to the high catalytic turnover, use of dioxygen as electron acceptor, and oxidation of a broad range of aliphatic alcohols and galactose-containing oligo- and polysaccharides, *Fgr*GaOx is widely used in a range of different applications and, correspondingly, has been subject for various of engineering studies. Early engineering of *Fgr*GaOx aimed to allow study of the fundamental biochemical properties of the enzyme which was discussed above (1.2.1). When considering engineering of *Fgr*GaOx with a purpose of propagating the usability in relevant applications, there are generally three re-emerging engineering targets: (1) aiming to improve the catalytic efficiency on D-galactose and galactose-containing oligo- and polysaccharides (2) introducing oxidation of new carbohydrate substrates, such as D-

glucose, D-mannose or D-fructose, (3) gain of oxidation activity on chemical groups other than the primary hydroxyls, such as secondary hydroxyls.

1.3.1 Oxidation of galactose containing polysaccharides

The product from *Fgr*GaOx oxidation of galactose is the corresponding C6-galactoaldehyde, but due to the reactivity of the aldehyde, side products occur. Whereas unsaturated α / β -aldehydes have also been identified (Maradufu and Perlin, 1974; Parikka and Tenkanen, 2009), the aldehyde product is mainly present as its hydrate form (geminal diol) in aqueous systems. Depending on the reaction conditions and presence of proximate hydroxyl groups, intra- and intermolecular hemiacetals formation can occur, leading to molecular crosslinks in the latter case. Crosslinked dimer and multimer products are thought to form especially from the reactive unsaturated aldehyde during heating and alkaline pH in higher product concentrations (Kupper et al., 2012; Parikka and Tenkanen, 2009). However, carbohydrate dimer products are often observed after *Fgr*GaOx oxidation under standard conditions. In fact, as early as 2002 the idea of using *Fgr*GaOx-oxidized galactose as a chemical cross-linker was discussed (Schoevaart and Kieboom, 2002).

The high turnover on polysaccharides and ability to achieve high conversion yields makes *Fgr*GaOx, and particularly optimized variants of it, good catalyst for creating stable hydrogels and aerogels from galactose-containing polysaccharides, such as galactoglucomannan, galactoxyloglucan and galactomannan (Figure 6) (Campia et al., 2017; Ghafar et al., 2017, 2015, Mikkonen et al., 2014, 2013; Parikka et al., 2012b; Rossi et al., 2017, 2016; Silvetti et al., 2018). The strength of single-component hydrogels, and their ability to form stable aerogels, directly depends on the extent of crosslinks at a given polysaccharide concentration which in turn depends of the degree of oxidation.

*Fgr*GaOx oxidation can also facilitate regio-selective chemoenzymatic modification of galacto-polysaccharides to use as cellulose coatings, through subsequent chemical derivatization (Leppänen, 2013; Leppänen et al., 2014, 2010; Parikka et al., 2012b; Xu et al., 2012). However, internal experience from my co-authors and our own lab suggest that wild-type performance is insufficient to generate required degrees of oxidation at high substrate concentrations, which is considered to be partly caused by end-product inhibition. As will be discussed below, the performance of *Fgr*GaOx is highly dependent on the expression conditions, but also protein engineering is necessary to optimize the performance of galactose oxidase under high substrate concentrations, particularly to maximise end-point conversion yields (high degree of oxidation).

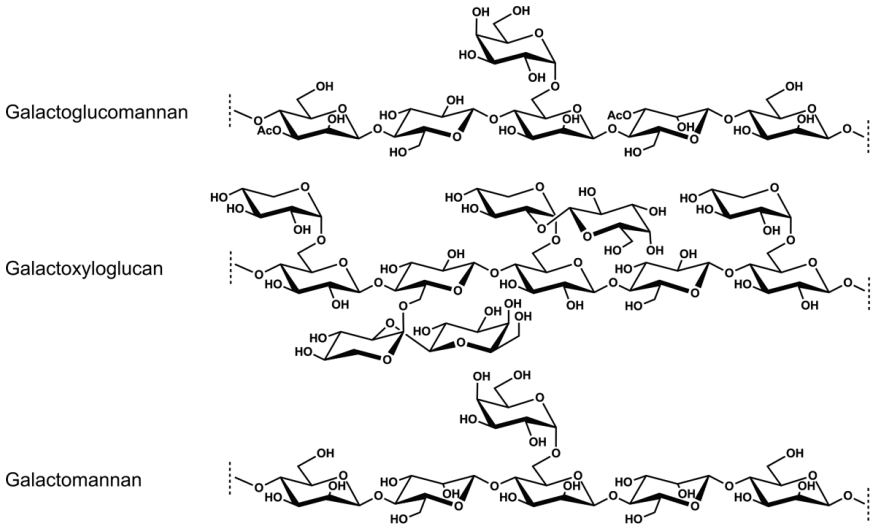


Figure 6. Examples of galactose containing hemicelluloses oxidized by *Fgr*GaOx
This figure was originally published in [2].

Reaction conditions also play a key role in conversion yields, which was also discussed by Parikka and colleagues (Parikka et al., 2015). That same year, Pedersen and colleagues published a collective and thorough investigation of the process condition requirements for *Fgr*GaOx (Pedersen et al., 2015). Like discussed in 1.2.1, spontaneous loss of the Cys-Tyr radical leads to a mix of active and inactive states of *Fgr*GaOx. Horse radish peroxidase (HRP) is added to reactions to regenerate and maintain the oxidised state of *Fgr*GaOx which increases conversion yields. Also, accumulation of hydrogen peroxide becomes detrimental for reactions and while HRP reduces hydrogen peroxide it requires a terminal reductant (such as 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid (ABTS), phenol or lignin-derivatives). Instead, a combination with HRP and an excess load of catalase (100-times excess) under highly aerobic conditions increases conversion yields and has proven important for achieving high conversion yields from *Fgr*GaOx.

2,2,6,6-Tetramethylpiperidinoxyl Radical (TEMPO) based oxidation systems have been efficient alternatives for chemo-enzymatic modification of cellulose and hemicellulose and the generation of hydro- and aerogels, particularly because of the high conversion yields and unspecific substrate acceptance that can facilitate direct oxidation of cellulose surfaces, as recently reviewed (Bragd et al., 2004; Díaz-Rodríguez et al., 2014; Pierre et al., 2017). However, TEMPO oxidization occasionally shows detrimental impacts to end-product integrity by unspecific degradation of polysaccharides, toxicity and potential environmental harmfulness of the end-product and residual catalyst, which limits the use of this method and also merits the continuous development of cleaner and unharmful enzymatic alternatives.

1.3.2 Use in analytical and diagnostic methods

Because galactose is an important component in carbohydrate metabolism there are numerous examples of applications of *Fgr*GaOx in diagnostic and biomedical applications. For example, *Fgr*GaOx has been applied in biosensor systems for diagnostic and analytical use (Charmantray et al., 2013; Evik et al., 2010; Sharma et al., 2006; Tkac et al., 2007, 2000; Xie et al., 2016). In fact, the use of galactose oxidase in biosensing methods was recently reviewed by Kanyong and colleagues (Kanyong et al., 2017). *Fgr*GaOx and variants described below have also been used to detect cell surface bound glycoproteins where glycosylation terminates on galactose or N-acetylglucosamine (Rannes et al., 2011).

1.3.3 Engineering to increase catalytic performance

Efforts to increase the catalytic performance of *Fgr*GaOx on D-galactose and galactose-contain polysaccharides has been motivated by the potential use of *Fgr*GaOx for bulk production of activated hemicellulose, used as an additive to increase paper strength and processability (Delgrave et al., 2001). Using high-throughput digital imaging and several rounds of directed evolution, the authors identified a *Fgr*GaOx variant (8-1; C383S, Y436H, V494A) with only 3 amino acid substitutions, but displaying 14-times improved catalytic efficiency on β-methylgalactose. Enhanced catalytic efficiency was driven by both increase of k_{cat} (5.2-times) and a significant reduction of K_M (3.5-times). The same study showed that each mutation also independently enhanced the catalytic turnover of *Fgr*GaOx, while C383S solely was assigned to increased affinity. Further mutagenesis of 8-1 generated the 7.3.2 variant (C383S, Y436H, V494A, K248E, T352S, K366R) which showed significantly increased activity on guar galactomannan but reduced catalytic efficiency on β-methylgalactose. Still, given its optimized performance on galactosyl-residues in polysaccharides, the 7.3.2 variant was identified as a novel candidate for industrial use. In another study, similar increase in catalytic efficiency was also found in the CS/VA variant (C383S, V494A) and the C383S substitution alone drawing attention to the importance of the C383 residue (Wilkinson et al., 2004). Wilkinson and colleagues solved the crystal structure of the C383S variant of *Fgr*GaOx in an attempt to explain the extensive reduction in K_M toward galactose-substrates but found no observable changes around the substrate ligands that could explain the effect of the serine sidechain.

1.3.4 Engineering for activity on other substrates

Although *Fgr*GaOx is the only carbohydrate oxidoreductase reported to efficiently act on polysaccharides, the strict selectivity for terminal galactosyl-residues limits the range of application. Efforts to expand the oxidation to a broader range of carbohydrates, with particular interesting on glucosyls and mannosyls, would therefore considerably broaden the applicable

potential, since there are no C-6 glucose or mannose oxidases found in nature so far. The team from nobel laurate Frances Anolds lab were the first to construct a *Fgr*GaOx variant (M3; W290F, R330K, Q406T) with activity on D-glucose and some aliphatic alcohols, through a combination of structure-guided combinatorial library construction, random mutagenesis and saturation mutagenesis (Sun et al. , 2002). While M3 displayed a new, but low activity on D-glucose (1.6 U/mg at 420 mM), the catalytic activity of M3 was reduced by 1000-fold on D-galactose relative to the compared wild-type, when compared at lower concentrations, rendering the M3 variant much less prominent for applicable use. The R330K variant, containing only a single substitution, was reported to have good activity on D-fructose (90.6 U/mg) while similarly resulting in decreased activity on D-galactose (Deacon et al. , 2004). Subsequent efforts to engineer *Fgr*GaOx for glucose activity, involving a structure-based computational defined library, resulted in Des3-2 (Q326E, Y329K, R330K) with 4 times higher catalytic efficiency on D-glucose compared to M3 (Lippow et al. , 2010). The enhanced activity of Des3-2 over M3 lead the authors to suggest that substitutions of the stacking tryptophan (like W290F in M3) reduces catalytic activity by compromising the stabilization of the Cys-Tyr radical in *Fgr*GaOx (Moon et al., 2012). However, as previously mentioned, the W290F substitution does not lead to accelerated radical decay but rather reduces affinity for D-galactose as the indole-ring of W290 is suggested to coordinate the hemiacetal oxygen in the sugar ring of D-galactose (Rogers et. al. 2007). Indeed, the M3 variant displays 8-times higher activity on 1-3-dihydroxyacetone than the wild-type, indicating that the CRO-functionality in M3 is intact. This is also in agreement with the division of the active site into two distinct regions dividing residues in their function (see 1.2.1). To broaden the applicability of M3, the variant was further engineered to optimize its activity on glucose, N-acetylglucosamine and mannose (Rannes et al., 2011). The corresponding F_2 variant (W290F, R330K, Y405F, Q406E) displayed 14-times higher activity on N-acetylglucosamine over M3, as well as increased activity on α- and β-methylglucose. Chemoenzymatic labelling of glycoproteins, namely asiolatransferrin (glycoprotein with terminal gal residues (Arndt, 2003)) by the M1 variant (discussed below), and carboxypeptidase Y (high mannose glycosylated) by the F2 mutant, demonstrated the potential of these oxidases in glycoprotein engineering and labelling. Despite their poor potential for oxidation of bulk polysaccharides, M3 and F2 variants *Fgr*GaOx have novelty in biomedical and diagnostic applications. Accordingly, there are several patents and patent applications on GaOx facilitated glycoprotein engineering (Arnold et al., 2000; Behrens et al., 2005; Hallahan and Gilbert, 1993; K. S. Johansen, 2016; Rouau et al., 1999; Stefano, 2006; Zhu, 2002).

*Fgr*GaOx have also been used in synthesis pathways involving small aliphatic and phenolic primary alcohols (Minasian et al., 2004; Whittaker and Whittaker, 2001). The inherent activity on aliphatic alcohols by *Fgr*GaOx makes it a potential catalyst in organic synthesis chemistry. Engineering of *Fgr*GaOx targeting carbohydrates, can also increase the performance on non-

carbohydrate substrates. For example, M3 displays two times higher activity on 2-pyridylcarbonol than *Fgr*GaOx (Sun et al., 2001), whereas the F2 variant was demonstrated as an efficient catalyst for a range of amino alcohols, such as 3-amino-1,2-propanediol, for subsequent aldolase reactions (Herter et al., 2015). Targeting secondary alcohols, directed evolution of the M3 variant yielded of enantioselective alcohol oxidases (M_{3-5} and $M_{3-5-215}$) with remarkably good oxidation yields on a broad range of optically pure aryl-based secondary alcohols (Escalettes and Turner, 2008). Like F2, the M_{3-5} variant also showed potential for oxidation of amino alcohols (Herter et al., 2015). By showing that M_{3-5} also oxidises the geminal diol-derivative of its own aldehyde product, the enzyme was used as a duo-functional catalyst in the cascade oxidation of aryl-based primary hydroxyls to generate carboxylic acids (Birmingham and Turner, 2018). Notably though, the ability and behind-lying mechanism of AA5_2 oxidases, like *Fgr*GaOx and *Cgr*RaOx, to oxidize the geminal diol derivatives of its own aldehyde products was first shown by my colleagues on the oxidation of raffinose in paper [3] of this thesis. Also, high performance of W290-containing variants on non-carbohydrates confirms that W290F does not significantly impair the CRO-functionality.

Table 2. Summary of FgrGaOx mutants described in sections 1.3.3 and 1.3.4.

Variant	Mutations	Notable impacts	Reference
8-1	C383S, Y436H, V494A	5-times higher catalytic efficiency on D-gal	(Delgrave et al., 2001)
7.5.2	C383S, Y436H, V494A, K248E, T352S, K366R	1.5-times higher activity on Guar galactomannan over 8-1	
M1	S10P, M70V, P136G, G195E, V494A, N535D	Enhanced stability and expression in *E. coli*	(Sun et al., 2001)
M3	W290F, R330K, Q406T	Activity on D-Glc	(Sun et al., 2002)
F2	W290F, R330K, Q406E, Y405F	Enhanced activity on D-Fru, D-Man and D-GlcNac over M3	(Rannes et al., 2011)
Des3-2	Q326E, Y329K, R330K	4-times higher activity on D-Glc over M3	(Lippow et al., 2010)
M3-5	W290F, R330M, Q406T	Oxidation of secondary alcohols	(Escalettes and Turner, 2008)
M3-5-215	W290F, R330M, Q406T, N413Y, Y436F		

1.3.5 CBM fusion proteins

In general, carbohydrate binding domains (CBM's) are thought to impart a targeting affect to particular polysaccharide structures, and/or a proximity affect where the CBM increases the effective concentration of the substrate presented to the enzyme (Boraston et al., 2004; Oliveira et al., 2015; Shoseyov et al., 2006; Várnai et al., 2014). Adherence of enzymes and other proteins to polysaccharides can be increased through addition of a CBM, and there are several reports that study the impact of fusing a CBM to various glycosidases. For example, the fusion of a CBM6 or CBM22 to family 10 xylanase from *B. halodurans* doubled the hydrolytic activity of the corresponding enzyme on insoluble and soluble oat spelt xylan (Mamo et al.,

2007; Selanere and Andersson, 2002). Likewise, fusion of a CBM3 increased the activity of Xyn10D from *Paenibacillus curdlanolyticus* by 1.5 times on insoluble wheat arabinoxylan (Sakka et al., 2011). Linking a CBM2 to Cel74 from *Thermotoga maritima* promoted hydrolysis of microcrystalline cellulose (Chhabra et al., 2003). Also relevant, the inherent CBM32 from *Clostridium thermocellum* Man5A has been shown to be important for substrate recognition and suppression of transglycosylation activity observed for Man5A (Mizutani et al., 2014).

At the initiation of this thesis work there were no published research articles showing successful attempts to fuse a CBM to galactose oxidase or any other carbohydrate oxidoreductase. CBMs are most commonly associated with glycoside hydrolases but known examples of carbohydrate oxidoreductases containing CBMs include but may not be limited to: cellobiose dehydrogenase from family AA3 (CBM1), galactose oxidase from family AA5_2 (CBM32), lytic polysaccharide monooxygenases from family AA9, AA10, AA11 and AA13 (CBM1, CBM18 and CBM20) (www.cazy.org). Particular to this thesis, it is important to note that galactose oxidase from *F. graminearum* contains an inherent N-terminal CBM32 which is embedded into the catalytic domain of the enzyme (see Figure 2). Deletion of the CBM32 domain or substitution with CBM29-1-2 from *Piromyces equi* results in inactive forms of galactose oxidase (McPherson et al., 1993; Mottiar, 2012). While the exact function of the CBM32 domain in *Fgr*GaOx is unknown, other family 32 CBMs are characterised by a type C carbohydrate binding domain that bind tightly to galactose.

A related study from my colleagues, published during the time of this thesis work, showed increased catalytic activity of the family AA7 glucooligosaccharide oxidase from *Sarocladium strictum*, GOOX-VN, through fusion of a family 22 carbohydrate binding module from *Clostridium thermocellum* (Vuong and Master, 2014). The resulting *Ct*CBM22_GOOX-VN fusion protein successfully showed induced binding to soluble oat spelt xylan, beechwood xylan and propoxylated wheat bran hemicellulose as well insoluble oat spelt xylan. Fusion of *Ct*CBM22 to the N-terminal of GOOX-VN also resulted in increased catalytic turnover rates, particularly at lower substrate concentrations, on cello- and xylo-oligosaccharides and soluble oat spelt xylan.

1.3.6 Recombinant production of galactose oxidase

Galactose oxidase has been recombinantly produced in *Escherichia coli, Pichia pastoris* and *Aspergillus nidulans*. Sun and colleagues optimized the *Fgr*GaOx encoding *gaoA* gene from *F. graminearum* for enhanced expression and stability in *E. coli* generating the M1 variant (Sun et al., 2001). The encoded protein that contains five amino acids substitutions (S10P, M70V, G115E, V494A and N535D) is reported to have wild-type substrate range and slightly enhanced

catalytic activity, which could be due to V494A. The M1 gene was then further optimized with a range of silent mutations to enhance expression in *E. coli* by 40-fold (Deacon and McPherson, 2011). A comparison between *E. coli* and *P. pastoris* expression systems for recombinant production of native *Fgr*GaOx and M1 was made by Spadiut and colleagues (Spadiut et al., 2010). The authors also investigated the impacts of intra- versus extracellular expression, codon optimization, and a C-terminal His6-tag to facilitate protein purification. Extracellular expression of the *Fgr*GaOx gene optimized for *P. pastoris* lead to highest yields and specific activity of *Fgr*GaOx, which also allows for a convenient purification strategy of the secreted protein. While the M1 gene increases expression yields in *E. coli*, this gene sequence lead to inferior production in *P. pastoris* compared to the *Pichia* optimized gene.

Due to the high transformation efficiency and ease of cultivation, *E. coli* has been the preferred host for high-throughput studies for mutagenesis studies of galactose oxidase (Delgrave et al., 2001; Lippow et al., 2010; Rannes et al., 2011; Sun et al., 2001; Wilkinson et al., 2004). On the other hand, when the aim is to produce *Fgr*GaOx at increased yields, *P. pastoris* is preferred given its scalable and optimized expression conditions (Deacon et al., 2004; Hartmans et al., 2004; Whittaker and Whittaker, 2000). Critically, the specific activity of galactose oxidase is highly dependent on the expression conditions, particularly when using the *P. pastoris* expression system. Using the same *P. pastoris* strain containing the *Pichia*-optimized *Fgr*GaOx gene (Spadiut et al., 2010), Anasontzis et al. (2014) found that the specific activity of *Fgr*GaOx could be increased 10-fold by optimizing the cultivation conditions in a suitable bioreactor system. Accordingly, and due to the obvious advantages, *P. pastoris* was chosen as the preferred expression host for production of AA5_2 oxidases in this thesis. At first, expression was attempted using shake flasks but due to low yields, bioreactor systems were also adopted. Optimization of bioreactor production of a new AA5_2 oxidase, *Pru*AA5_2A, was also performed by my master student during 2016, showing that control of methanol feed-rate in the fed-batch induction stage to ensure stable cultivation conditions is critical for expression yields (Pizarro, 2016).

2. Aims of thesis

The aims of this thesis were [1] to enhance the oxidative performance on polysaccharides of the galactose oxidase from *Fusarium graminearum* and M3 (the glucose active variant from Sun et al., 2002), through fusion with carbohydrate binding modules from family 29 and family 3, generating CBM29-GaOx, GaOx-CBM29 and M3-CBM3 respectively. [2] to increase the catalytic activity of M3-CBM3 on glucose through additional amino acid substitutions known to enhance the catalytic performance of wild-type galactose oxidase. [3] Investigate the diversity of CAZyme family AA5_2 by selection, expression and characterization of two uncharacterized AA5_2 members with low homology among themselves and relative to archetypal galactose oxidase from *F. graminearum*.

3. Materials and Methods

3.1 Commercial enzymes and substrates

Horse radish peroxidase, Bovine liver catalase and Phosphorylase b from rabbit muscle was purchased from Sigma Aldrich. All mono- and oligosaccharides, alcohols, aldehyde and acid substrates were likewise purchased from Sigma Aldrich if not mentioned otherwise. Locust bean galactomannan was purchased from Fluka, whereas galactoxyloglucan from tamarind seed was obtained from Megazyme, and guar galactomannan was obtained from Sigma Aldrich (batch 14K0168). Spruce galactoglucomannan was generously provided by Professor Stefan Wilför (Åbo Academy, Finland). Reported molecular weights of polysaccharides used in this analysis are as follows: 29 kDa for spruce galactoglucomannan (Willför et al., 2003) 470 kDa for tamarind galactoxyloglucan, 2600 kDa for guar galactomannan (Parikka et al., 2010), and 310 kDa for locust bean galactomannan (as stated by Sigma-Aldrich (G0753)).

3.2 Recombinant AA5_2 oxidases

A *Pichia pastoris* KM71H expressing wild-type *Fgr*GaOx was kindly supplied by Professor H. Brumer (University of British Columbia) (Spadiut et al., 2010). Prior gene synthesis, the native signal sequence of *Pru*AA5_2A from *Penicillium rubens* (strain ATCC 28089 / DSM 1075 / NRRL 1951 / Wisconsin 54-1255; Uniprot: B6HHT0; Genbank: 96757.1), was predicted using the SignalP server (Petersen et al., 2011) and removed from the amino acid sequence. The gene encoding the corresponding protein sequence, including prosequence, was optimized for expression in *P. pastoris* and then synthesized and cloned into pPICZalpha by Genscript (NJ, USA). The gene construction of *Cgr*RaOx from *Colletotrichum graminicola M1.001* is described in [3] as this was performed by my co-authors.

Construction of fusion proteins

The amino acid sequence of galactose oxidase from *Fusarium graminearum* (P0CS93) and the family 29 carbohydrate binding module (CBM29-1-2) from *Piromyces equi* (AAK20910) were used to design GaOx-CBM29 and CBM29-GaOx fusions, according to (Spadiut et al., 2010), both lacking the N-terminal pro-sequence of native *Fgr*GaOx, and containing a C-terminal His_6-tag as well as the following linker sequence: TPTKGATPTNTATPTKSATATPTRPSVPTNTPTNTPANTPM. M3-CBM3 and M6-CBM3 was designed similar to above

using the amino acid sequence of CBM3 from *Clostridium thermocellum* (A3DDF1|482-641) and by introducing amino acid substitutions in the sequence of galactose oxidase from *F. graminearum* (P0CS93) (M3-CBM3; W290F, R330K, Q406T and M6-CBM3; W290F, R330K, Q406T, C383S, Y436N, V494A). Genetic constructs were designed on basis of the provided amino acid sequence, optimized for expression in *P. pastoris* and cloned into the pJ912 expression plasmid by DNA 2.0 (USA), under promotion of the AOX1 methanol inducible promotor for expression and secretion of targeted proteins from *P. pastoris*. The pJ912 plasmids, obtained from DNA 2.0 were transformed into *E. coli* XL-1 (Agilent Technologies, USA) for storage and regeneration, and transformed into *P. pastoris* SMD1168H (Invitrogen) by electroporation for protein production. *P. pastoris* transformants were induced on buffered methanol-complex agar plates (BMMY agar (w/v): 1% yeast extract, 2 % peptone, 2 % agar, 100 mM potassium phosphate buffer (pH 6.0), 1.34 % yeast nitrogen base without amino acids (YNB), 4×10^{-5} % biotin, 0.5 % methanol), and then screened for protein expression by immuno-colony blot as previously described (Foumani, 2011). Integration of the plasmid into the *P. pastoris* genome was verified by colony PCR using gene specific primers.

3.3 Production and purification of AA5_2 oxidases

Recombinant expression using Bioreactors

All cultivations were performed using a Biostat B Plus bioreactor (Sartorius) equipped with pH and oxygen probes from Hamilton connected to a MFCSwin process control system (BBI B. Braun Biotech International). Cultivation conditions were based on *Pichia* Fermentation Process Guidelines provided by Invitrogen, with minor modifications. Briefly, 100 mL inoculum of *P. pastoris* KM71H (MUT^s) expressing *Fgr*GaOx and *P. pastoris* SMD1168H and X-33 (MUT^+) expressing the other protein constructs were cultivated in YPD at 30 °C to an OD_{600} between 2 to 6. Cells were harvested by centrifugation (3000 x g, 5 min) and then suspended in basal salts medium (26.6 mL 85 % phosphoric acid, 0.93 g calcium sulfate, 18.2 g potassium sulfate, 14.9 g magnesium sulfate-heptahydrate, 4.13 g potassium hydroxide, 40 g glycerol per liter). The initial fermentation medium comprised 1L basal salts medium containing 4% (w/v) glycerol and 0.435 % (v/v) PTM_1 trace salts (6.0 g cupric sulfate pentahydrate, 0.08 g sodium iodide, 3.0 g manganese sulfate monohydrate, 0.2 g sodium molybdate dihydrate, 0.02 g boric acid, 0.5 g cobalt chloride, 20.0 g zinc chloride, 65.0 g ferrous chloride heptahydrate, 0.2 g biotin, 5.0 ml sulfuric per liter). The pH was controlled to pH 6.0 by automatic addition of 15% (w/v) ammonium hydroxide, which also served as the sole nitrogen source during cultivation. The dissolved oxygen concentration was maintained above 40% by automatic cascade stirring and gas flow feedback-control, using the proportional-integral-derivative (PID) controller algorithm function in the bioreactor. Specifically, the stirrer speed was first increased from 300 to 1200 rpm and then the airflow was increased from 0.5 to 3.0 L min^{-1} until maximum oxygen transfer capacity of the reactor was reached. Antifoam (Struktol J 647) was added

automatically as required. Upon depletion of glycerol from the medium the glycerol fed-batch phase was initiated manually by feeding a 50% (w/v) glycerol mix, with 1.2 % (v/v) PTM_1 trace salts, to the reactor at initially 15 mL h^{-1}, and stepwise increased to 25 mL h^{-1} until the maximum oxygen transfer capacity was reached, while still maintaining a substrate-limited growth rate and dissolved oxygen concentration between 25-35%.

For induction, the temperature was lowered to 25 °C, the glycerol feed was stopped, and the dissolved oxygen concentration was allowed to increase as residual glycerol was consumed by the cultures. To express fusion constructs from *P. pastoris* MUT^+ strains, methanol with 1.2 % (v/v) PTM_1 trace salts was added stepwise from 0.5%, to 1.0% and then 2.0%, while allowing the dissolved oxygen to stabilize prior to the next methanol addition. In this way, SMD1168 cultures were acclimatized to methanol metabolism prior to feeding with methanol at a flow rate of 7.0 mL h^{-1}, while monitoring DO to ensure that oxygen consumption was limited by methanol metabolism. To induce *Fgr*GaOx production from KM71H, methanol addition was maintained at 2.7 mL h^{-1}.

Purification

All recombinant enzymes were purified using a sequence of hydrophobic interaction chromatography and following poly-His-tag metal-ion affinity chromatography as described in [1]. After purification, all protein samples were transferred to 50 mM sodium phosphate buffer (pH 7.5) with continuous dilution and concentration using a Vivaspin (MWCO 30.000 Da) before using the Bradford method (Bio-Rad Laboratories, USA) to measure protein concentration and SDS-PAGE to confirm enzyme purity. Purified samples were then flash frozen with liquid nitrogen and stored at -80°C. Enzyme identities were verified by mass spectrometric analysis of tryptic fragments (Mass Spectromic Fingerprinting.

3.3.1 Activation of AA5_2 oxidases

To ensure full activation of the AA5_2 oxidases, purified protein samples were treated with copper sulfate and potassium ferricyanide as previously described (Baron et al., 1994; Rogers et al., 2000). Specifically, for this study, 0.15 mg/mL of purified enzymes were incubated in 0.5 mM copper sulfate for up to 120 h at 4 °C. After 0.5, 1, 2, 4, 18, 96 and 120 h of incubation, samples were recovered for SDS-PAGE analyses and activity measurement using the standard activity assay. Purified enzymes were also separately treated with 0.46 mM potassium ferricyanide at room temperature for 10 min, and then analysed as above. In all cases, enzymes were diluted appropriately prior activity determination, and untreated enzymes were used as references.

3.4 Sequence and structure analysis

A sequence alignment was performed using the ClustalW Omega software (Sievers et al., 2011) and the nine *Fusarium* spp. listed in family AA5_2 as well as all 36 eukaryotic AA5_2 sequences (retrieved from the CAZy database on April 4, 2017). To identify amino acid positions likely to contribute to substrate range, an alignment was also performed for functionally characterized AA5_2 members, including the alcohol oxidase (AlcOx) (Genbank: EFQ30446.1) and raffinose oxidase (RaOx) (Genbank: EFQ36699) from *C. graminicola M1.001* (characterized in this thesis). Amino acids positions with importance to biochemical properties were also collected through literature review (summarized in Table 9). Amino acids were also mapped on the crystal structure of *Fgr*GaOx (pdb code 1gog) (Ito et al., 1991) and then grouped according to their postulated function, where Group 1 included catalytic residues and copper-ligands, Group 2 include amino acid positions implicated in substrate range, and Group 3 included amino acid positions identified through mutagenesis to increase catalytic activity or stability.

3.5 Characterization of copper-radical oxidases

3.5.1 Affinity gel electrophoresis

Binding of *Fgr*GaOx and *Ct*CBM29-1-2 fusion proteins to guar and locust bean galactomannan, tamarind galactoxyloglucan, and spruce galactoglucomannan was examined by native affinity gel electrophoresis as described by Freelove et al. (2001). Briefly, native polyacrylamide gels prepared for these analyses contained 7.5 % (w/v) bis-acrylamide, 25 mM Tris/250 mM glycine buffer (pH 8.8) and 0.01% and 0.005% (w/v) of each polysaccharide. Approximately 5 µg of GaOx and each fusion protein were loaded onto the gels and then run at 10 mA/gel for 3 to 5 h. Phosphorylase b from rabbit muscle (5 µg, Sigma Aldrich) was used as a reference for these analyses.

3.5.2 Immobilization on cellulose surfaces

The binding of M3-CBM3, M3-CBM3 and *Fgr*GaOx (results in Figure 16) was tested by incubation of 50 µg protein with 5 mg of cellulose filter paper (Whatmann #1) or microcrystalline cellulose (Avicel®) in 100 µL sodium phosphate buffer (pH 7.5) over night at +4 °C in PVDF 0,2 µm filter microtiter plates (Corning, cat. No. 3504). The liquid phase (containing unbound protein) was separated from the cellulose substrate (containing bound protein) by centrifugation at 4500 RCF at +4 °C for 10 min. The cellulose fractions were washed by adding with 100 µL buffer and immediately centrifuged. The bound protein was released from the cellulose substrate by heating (95 °C) in 30 µL SDS-PAGE loading buffer for 5 min prior gel loading for SDS-PAGE analysis (see also Figure 16). To detect activity of immobilized incubations with

cellulose materials were repeated with a reduced enzyme concentration (0.1 nmole) in order to compare on molar basis and due to fast saturation of the activity assay when higher concentrations were used. Same immobilization procedure was followed where galactose oxidase activity was tested in the cellulose containing bound protein, after centrifugation and washing, using a fluorescence based Amplex red/HRP hydrogen peroxide assay: 50 µL of assay mix (50 µM Amplex red and 2 U/mL HRP in 50 mM sodium phosphate buffer, pH 7.5) was added to the cellulose fractions and either 50 µL buffer (to follow oxidation of cellulose) or 50 µL of 300 mM D-galactose was added. Time-resolved fluorescence was detected using a Synergy H1 fluorescence spectrometer and the Gen5 software (version 2.09) (BioTek): emission 590 $\pm$ 17.5 nm and excitation at 530 $\pm$ 12.5 nm with a gain of 30 and the optics in the top position.

3.5.3 Analysis of oxidase activity

Activity was measured by following the formation of hydrogen peroxide using the previously described chromogenic ABTS (2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) and horseradish peroxidase (HRP) assay (Baron et al., 1994). The final reaction mixture (volume: 205 µL) contained 7 U/mL HRP, 2 mM ABTS, and between 50 and 300 mM substrate in 20 mM MOPS buffer (pH 7.5), as enzymes showed best performance at this pH. Prior to initiating the reaction, 5 µL of protein samples (5-60 ng depending on activity levels) were incubated for 30 min at 30 ^{o}C in 100 µL 2x assay mix (4 mM ABTS and 15 U/mL HRP in milliQ water) to ensure complete activation by HRP. The reaction was initiated by addition of a 2x substrate concentration (in 40 mM HEPES at pH 7.5) and continuously monitored for up to 3 h by reading the absorbance at 420 nm. Hydrogen peroxide production was calculated using the extinction coefficient of the ABTS radical as described in [4].

Substrate range

The substrate range of *Pru*AA5_2A and *Cgr*RaOx was determined using 300 mM D-glucose, L-arabinose, D-xylose, D-galactose, melibiose, sucrose, lactose, raffinose, stachyose, ethanol, 1-propanol, 2-propanol, 1-butanol, 1,2-butandiol, glycerol, D-sorbitol, benzyl alcohol; the exception was for the glycolaldehyde dimer, acetaldehyde, D-glyceraldehyde and glyoxalic acid where the activity was tested at 50, 25, and 15 mM of freshly prepared substrate solutions.

Kinetic analysis

Kinetics parameters of AA5_2 oxidases were determined using 10 mM to 800 mM substrate concentrations in the activity assay as described above. Kinetic parameters were calculated using the Michealis-Menten function in Origin Pro 2016 (OrginLab Corp., USA), with the exception of the glycolaldehyde dimer where the substrate inhibition function was used instead of the Michaelis-Menten function.

4. Results and discussion

4.1 Production of AA5_2 CAZymes in *Pichia pastoris*

All enzyme constructs used in the experiments were produced in *Pichia pastoris* using bioreactors to optimize yield and specific activities (summarized in Table 3). The biochemical characteristics of the engineered fusion proteins based on *Fgr*GaOx (GaOxCBM29-GaOx, GaOx-CBM29, M3-CBM3 and M6-CBM3) are addressed in Engineering of galactose oxidase (4.2.1), whereas the properties of new AA5_2 oxidases (*Cgr*RaOx and *Pru*AA5_2A) are addressed in New AA5_2 oxidases *Cgr*RaOx and *Pru*AA5_2A (4.3). My thesis work began with the expression of CBM29-GaOx and GaOx-CBM29 and *Cgr*RaOx in *P. pastoris* SMD11868H and X-33, first by using shake-flasks. However, due to low yields, bioreactor systems were later employed instead. Results from using bioreactor instead of shake-flask are published in [1] and [3]. Accordingly, this section briefly discusses the advantages of using bioreactors instead of shake-flasks for cultivation of *P. pastoris* on the basis of [1] and [3] as well as some unpublished data.

Table 3. Summary of proteins produced and characterized in this thesis work

Construct	Origin	Type	Size, KDa	Pichia strain	Yield, $mg*L^{-1}$	Main activities	Binding	$K_{cat,}$ s^{-1}	k_M
*Fgr*GaOx	*Fusarium graminearum*	Wild-type	68	KM71H	106.4	Galactose Xyloglucan	None	400 3,6	42 0,07
GaOx-CBM29	*Fusarium gr. + Piromyces eq.*	Fusion protein	107	SMD1168H	13.6	Galactose, Galactomannan	Glucans + mannans	600 6,7	60 0,03
CBM29-GaOx	*Fusarium gr. + Piromyces eq.*	Fusion protein	107	SMD1168H	3.4	Galactose, Galactomannan	Glucans + mannans	517 5,2	56 0,08
M3-CBM3	*Fusarium gr. + Clostridim th.*	Mutant (3) fusion prot.	93	SMD1168H	21.4	Glucose (low) Galactose	Crystalline cellulose	7,1	201
M6-CBM3	*Fusarium gr. + Clostridium th.*	Mutant (6) fusion prot.	93	SMD1168H	20.6	Galactose	Crystalline cellulose	230	56
*Cgr*RaOx	*Colletotrichum graminicola*	Wild-type	94	X-33	6.7	Raffinose Glycolaldehyde dimer	NA	13,9 8,8	481 185
*Pru*AA5_2A	*Penicilium ruben*	Wild-type	69	SMD1168H	34.6	Raffinose Glycolaldehyde dimer	NA	57,7 76,8	248 52,8

The yields reported post purification. Kinetic values, activities and binding properties will be discussed in following sections

Shake-flasks versus bioreactors

The SMD1168H strain was chosen for expression of fusion proteins (CBM29-GaOx, GaOx-CBM29, M3-CBM3 and M6-CBM3) because of its lack of extracellular protease activity. Standard shake flask cultivations for protein expression in *P. pastoris* used in our laboratory involves a 24-hour pre-cultivation in BMGY-medium at 30 °C prior induction in methanol containing BMMY medium at 15 - 25 °C. The lower induction temperature aids the aerobic protein secretion pathway in *P. pastoris* by slowing down the growth rate and securing enough oxygen. The cultivation conditions for shake-flasks expression of GaOx-CBM29 showed that lowering pre-incubation (BMGY stage) temperature equivalent to the induction temperature or adding a protease inhibitor (leupeptin) did not increase extracellular expression. However, addition of 0.5 mM copper sulphate to aid the autocatalyzed maturation process in galactose oxidase had significant impact on the expression of active GaOx-CBM29, which also corresponds to earlier studies (Spadiut et al., 2010).

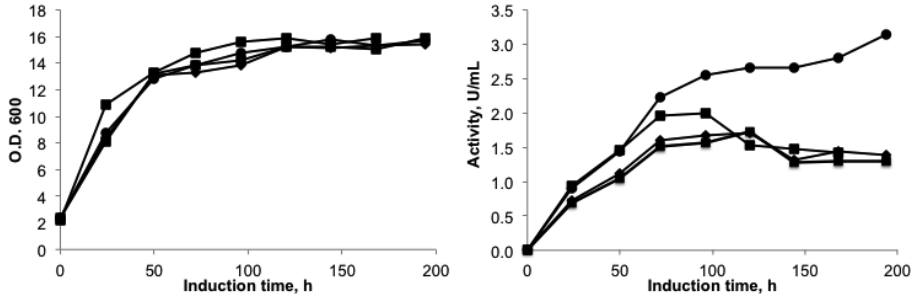


Figure 7. Expression of GaOx-CBM29 in shake flasks
Cell density (left) and extracellular oxidase activity on 300 mM D-Galactose (right) after induction. ■ Standard cultivation (pre-cultivation at 30 °C and induction at 15 °C); ◆ standard cultivation medium plus 10 µM leupeptin in the induction medium, ✕ pre-cultivation at 15 °C prior to induction; ● standard cultivation plus addition of 0.5 mM copper sulfate to the induction medium. This figure was originally published in [1].

The yields of purified enzyme increased by 5- to 19-fold when cultivations were performed in bioreactors rather than shake flasks (Table 4). Also, conditions provided by the bioreactor also impacted the specific activity of CBM29-GaOx and GaOx-CBM29 and particularly *Cgr*RaOx, where the yield increased 9.6 times and the specific activity 43 times leading to a significantly improved oxidase and, in fact, allowing for biochemical characterization of the *Cgr*RaOx. Similarly, the yield and specific activity of the fusion proteins CBM29-GaOx and GaOx-CBM29 also improved which facilitated more catalytic activity for characterization.

Table 4. Shake-flask versus bioreactor production

Enzyme	Pichia strain	Shake flask cultivation		Fermentation	
		Yield (mg/L) [a]	Specific Activity [b] (Units/ mg)	Yield (mg/L)	Specific Activity (Units/ mg)
GaOx-CBM29	SMD1168	1.5	355 ± 23	13.6	487 ± 10
CBM29-GaOx	SMD1168	0.64	364 ± 32	3.4	438 ± 29
*Cgr*RaOx	*X-33*	0.7	0.07 ± 0.01	6.7	3.0 ± 0.3

P. pastoris is capable of reaching high cell densities during cultivations under adequate oxygen supply. During shake-flask cultivations oxygen availability becomes the limiting factor once the cell mass becomes critical relative to the oxygen supply. In this case (Figure 7) oxygen availability became limiting near O.D. = 13 and the growth curve saturated at O.D. of 16 (approx. 30 g/L cell wet weight) and protein expression terminated after 100 hours. The main advantage of a bioreactor system over shake-flasks is the controlled capacity for oxygen-supply and substrate feed, which allows for much greater cell mass during cultivation while avoiding oxygen depletion. As exemplified with the fusion protein M6-CBM, expressed in the same *P. pastoris* strain, bioreactor cultivations reached a cell density near 250 g/L thus over 8 times more than shake flask cultivations. Since protein secretion in *P. pastoris* and the post-expression maturation in *Fgr*GaOx are oxygen dependent processes, the dissolved oxygen concentration was kept over 40 % and the growth rate was controlling by limiting the substrate feed, which resulted in higher specific activity of the expressed oxidases. The specific productivity of *P. pastoris* also increases when using bioreactors, where the productivity of GaOx-CBM29 was 77 µg/g of cells versus 50 µg/g cells in shake flasks. Correspondingly the specific productivity for M6-CBM3 in the bioreactor system was 82 µg/g cells which is similar to that of GaOx-CBM29.

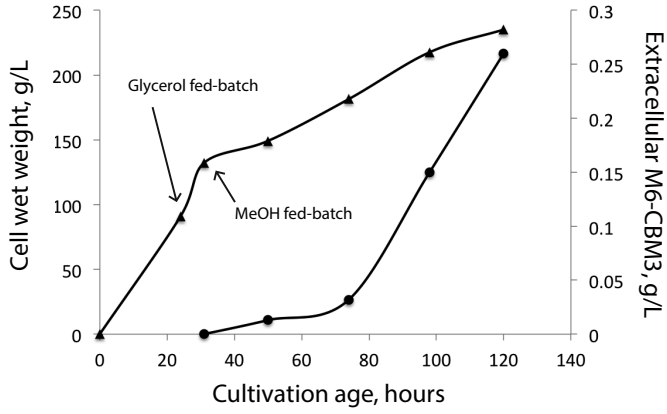


Figure 8. *P. pastoris* bioreactor fed-batch cultivation process of M6-CBM3

Post-expression activation

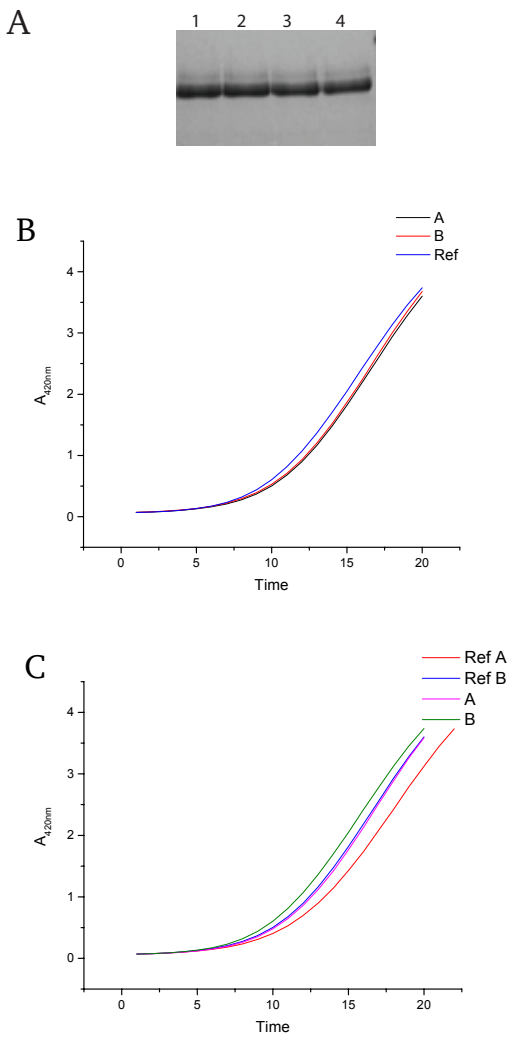


Figure 9. Post production activation of M6-CBM3
(A) SDS-PAGE analysis of purified M6-CBM3 (Lane 1), after incubation with 0.05 mM copper sulphate (lane 2), after incubation with 0.46 mM potassium ferricyanide (Lane 3) and after incubation with incubation with 0.05 mM copper sulphate and subsequently 0.46 mM potassium ferricyanide (Lane 4). (B) activity of purified M6-CBM3 (Ref) and after incubation with 0.05 mM copper sulphate (sample A and B). (C) activity of purified M6-CBM3 (Ref A) and copper sulphate treated sample (RefB) and after 0.46 mM potassium ferricyanide of refA (sample A) and refB (sample B).

Previous studies have shown that activation of *Fgr*GaOx, after expression in *E. coli* and *P. pastoris* shake-flask cultivations, is required to ensure annexation of the copper-ion, formation of the Cys-Tyr cofactor and the correct oxidized state (Cu(II)-Y*) (Firbank and Rogers, 2001; Rogers et al., 2000; Spadiut et al., 2010). The Y272-C228 linkage is stable under SDS-PAGE conditions and causes the electrophoretic mobility of mature *Fgr*GaOx (containing the Y272-C228 linkage) to correspond to 65 kDa despite the molecular weight of *Fgr*GaOx being 68 kDa (Rogers et al., 2007). Immature *Fgr*GaOx (lacking the Y272-C228 linkage) migrates according to its molecular weight, allowing the cofactor formation being monitored by SDS-PAGE. Accordingly, all the enzymes expressed in this thesis work were incubated with 0.05 mM copper sulphate and subsequently 0.46 mM potassium ferricyanide to ensure that there were fully matured and oxidized prior characterization.

The electrophoretic mobility and catalytic activity of M6-CBM3, before and after incubation with copper-sulphate and potassium ferricyanide, is shown in Figure 9, and is characteristic of other enzymes prepared in this thesis. In short, activation treatment did not affect the electrophoretic mobility or catalytic activity indicating that M6-CBM3 undergoes full maturation during the expression in the bioreactor.

4.2 Engineering of galactose oxidase

Engineering of the archetypal galactose oxidase from *Fusarium graminearum* aimed to improve the catalytic performance of galactose oxidase on galactose- and glucose-containing polysaccharides, and in case of the glucose-oxidizing fusion protein M3-CBM3, also glucose containing oligosaccharides. This research objective consisted of two main engineering approaches: (1) appending of carbohydrate binding modules in an effort to increase enzyme action on polymeric substrates, and (2) site-directed mutagenesis to enhance the catalysis rate of M3-CBM3 on D-glucose and glucose-containing oligosaccharides.

4.2.1 Fusion of CBM29-1-2 to the N- and C-terminals of galactose oxidase

Article [2]

Figure 10 shows structural models of the fusion proteins CBM29-GaOx and GaOx-CBM29 that were designed to enhance the performance of *Fgr*GaOx on galactose-containing polysaccharides in accordance to the background described in section 1.3.5. More specifically, given the recognized, limited extent of oxidation of galactomannan and other galactose-containing polysaccharides with *Fgr*GaOx oxidase (Parikka et al., 2015), this current study investigated the impact of a family 29 carbohydrate binding module (i.e., *Peq*CBM29-1-2 from *Piromyces abies equi*) on the activity of *Fgr*GaOx.

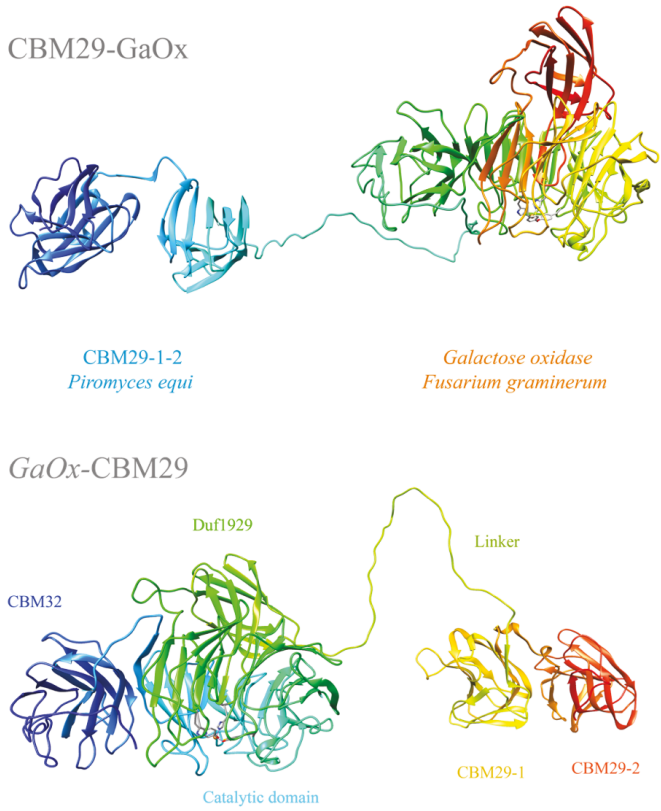


Figure 10. Models of CBM29-GaOx and GaOx-CBM29
Artistic models of CBM29-GaOx (top) and GaOx-CBM29 (bottom). Above structure is depicted with protein names whereas the individual domains are highlighted in the lower. The N-terminal CBM32 domain of *Fgr*GaOx and the tandem CBM's of CBM29-1-2 are highlighted in the bottom model. The models may not accurately depict actual structures. CBM29-GaOx: *Peq*CBM29-1-2 (1-305, 34.6 kDa), Tp-linker (306-345, 4.4 kDa), *Fgr*GaOx (346-984, 68 kDa) and GaOx-CBM29: *Fgr*GaOx (1-639), TP-linker (640-680), *Peq*CBM29-1-2 (681-984).

In *P. equi*, the CBM29-1-2 module is positioned at the C-terminus of the non-catalytic NCP1 protein, which forms part of a larger cellulase complex. Earlier studies, confirmed CBM29-1-2 binding to a broad range of β-1,4-mannan, xylan and glycans and the synergistic action of between the CBM29-1 and CBM29-2 domains when linked (Freelove et al., 2001). In the current study the tandem *Peq*CBM29-1-2 was fused to the N- and C-terminal of *Fgr*GaOx, generating CBM29-GaOx and GaOx-CBM29, to facilitate targeting to galactose containing hemicelluloses: galactomannan from guar and locust bean, galactoxyloglucan from tamarind, and galactoglucan from spruce (structures of these hemicelluloses are depicted in Figure 6).

Both CBM29-GaOx and GaOx-CBM29 showed binding to the galactose containing polysaccharides as analysed by affinity gel electrophoresis (AGE) (Figure 11), and thus confirming the function of the appended CBM29-1-2. The fusion proteins showed increased electrophoretic mobility over *Fgr*GaOx in the reference gels containing no polysaccharide (Native-PAGE gel), which can be explained by the lower pI of *Peq*CBM29-1-2. No binding was observed for *Fgr*GaOx on any of the polysaccharides, whereas GaOx-CBM29 showed strongest biding to guar galactomannan (GG) followed by galactoxyloglucan (XG), locust bean galactomannan (LBG) and galactoglucomannan (GGM). CBM29-GaOx showed a similar binding pattern but had slightly different electrophoretic mobility than GaOx-CBM29. Both fusion proteins showed increased binding strength when the polysaccharide concentration was increased from 0.005 % to 0.01 % (w/v).

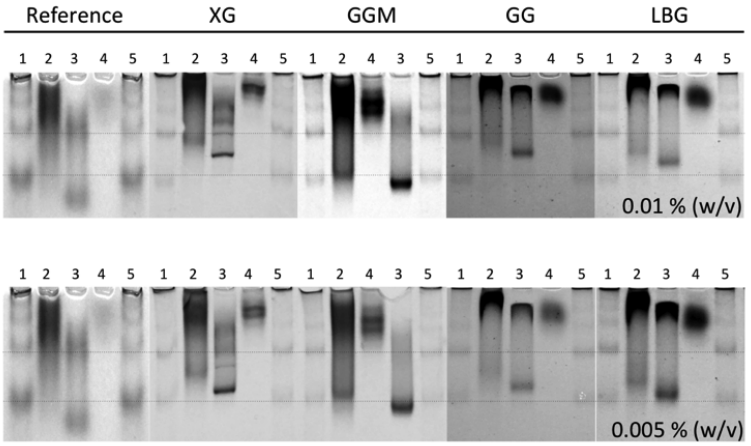


Figure 11. Affinity gel electrophoresis of GaOx, CBM29-GaO and CBM29-GaO
The binding of CBM29-GaOx and GaOx-CBM29 on galactoxyloglucan (XG), galactoglucomannan (GGM), Guar galactomannan (GG) and Locust bean galactomannan (LBG). Dotted lines indicate the migration of the Phosphorylase b (PhB). Lanes; 1: PhB, 2: CBM29-GaOx, 3: GaOx-CBM29, 4: FgrGaOx, 5: PhB. This figure was originally published in [2].

Binding of GaOx-CBM29 and CBM29-GaOx to solid surfaces coated with LBG was tested by quartz crystal microbalance with dissipation (QCM-D) (Figure 12). The greater frequency

decrease observed for GaOx-CBM29 relative to CBM29-GaOx indicated higher extents of GaOx-CBM29 binding to the LBG coated surface. Higher rates of GaOx-CBM29 binding also suggested higher affinity to LBG compared to CBM29-GaOx. This was further evidenced by passing CBM29-GaOx over the sensor after binding of CBM29-GaOx, which revealed the presence of unoccupied LBG. Similarly, the sensor first treated with GaOx-CBM29, did not allow binding of CBM29-GaOx.

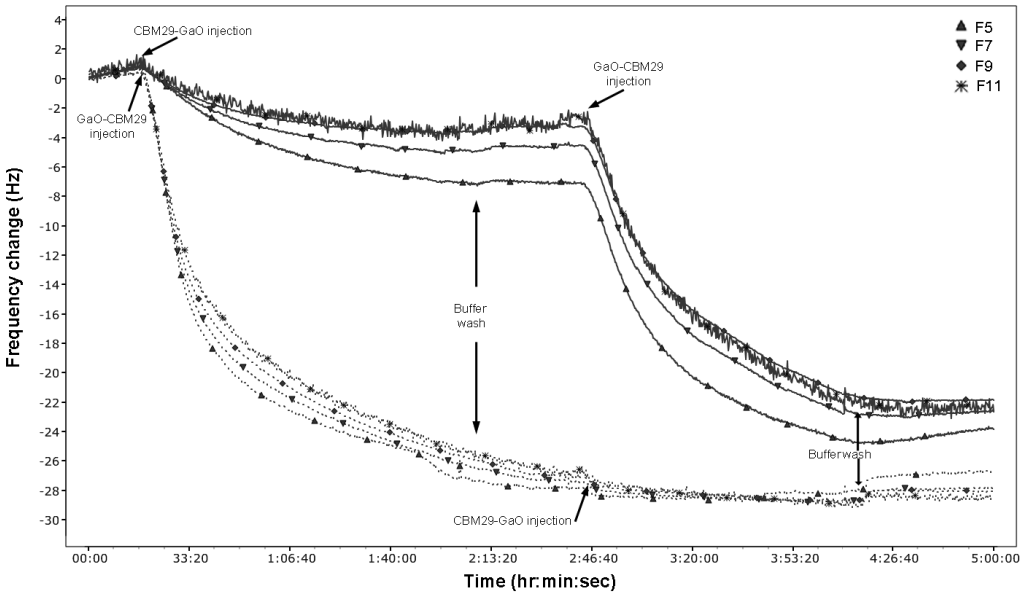


Figure 12. Binding of CBM29-GaOx and GaOx-CBM29 to LBG-coated QCM-D sensors
QCM-D frequency changes during the binding of GaO-CBM29 or CBM29-GaO with LBG-coated gold sensors. Protein samples (0.5 µg/mL) were passed over the coated sensors with a flow rate of 0.1 mL/min for 5 h. The solid lines represent frequency changes at the 3rd, 5th, 7th, 9th and 11th harmonic overtones of sensors first treated with CBM29-GaOx; dash lines show corresponding frequency changes of sensors first coated with GaOx-CBM29. Arrows indicate transition points within the experiment. This figure was originally published in [2].

4.2.2 Fusion of CBM29 causes oligomer formation in GaOx-CBM29

The appearance of two protein bands for CBM29-GaOx and GaOx-CBM29 in the AGE analysis (Figure 11) where only one band was observed in lanes with *Fgr*GaOx, suggests presence of multiple oligomeric states in the fusion proteins. This was especially apparent for GaOx-CBM29, with or without polysaccharide substrates. Supporting this, only single bands were detected in SDS-PAGE [1]. Whereas there are no reports of oligomerization of *Fgr*GaOx and wild-type CBM29-1-2, ligand-mediated dimerization has been detected for a E78R variant of CBM29-2 (Flint et al., 2004). The oligomer states of *Fgr*GaOx and GaOx-CBM29 were determined by analytical ultracentrifugation (AUC) in absence of polysaccharides: The ratios of the oligomer mass to the theoretical mass was 1.1:1 for *Fgr*GaOx and 2.5:1 for GaOx-CBM29,

confirming an induced oligomer state by CBM29-1-2 [2]. Due to insufficient protein stocks, AUC measurements were not performed for CBM29-GaOx.

The binding characteristics of *Peq*CBM29-1-2, which are abundant in previous studies (Charnock et al., 2002; Flint et al., 2005; Freelove et al., 2001), include no evidence supporting oligomerization in wild-type *Peq*CBM29-1-2 nor the CBM29-1 and CBM29-2 subunits isolated, albeit the exception of the E78R variant (Flint et al., 2004). Accordingly, unintended protein-protein integrations between *Peq*CBM29-1-2 domains and the *Fgr*GaOx domains of different fusion proteins presumably led to the oligomerization of GaOx-CBM29. GaOx-CBM29 exhibits binding affinity to polysaccharides despite the oligomer state, suggesting that at least one set of CBM29-1-2 domains must be functional in the oligomer. Accordingly, the following conclusions can be drawn:

(a) The presence of two bands in lanes "3" in Figure 11 suggests that GaOx-CBM29 is present in two oligomer (confirmed by AUC) states and both states appear to have binding affinity to the tested polysaccharides, particularly on Guar and Locust Bean galactomannans.
(b) The conclusions drawn in (a) appears to be true for CBM29-GaOx also (lanes "2" in Figure 11), but the smearing of protein bands makes interpretation difficult.
(c) The possibility of different oligomer masses between CBM29-GaOx and GaOx-CBM29 may impact the QCM-D data in Figure 12. Particularly, CBM29-GaOx and GaOx-CBM29 appear to have different electrophoretic mobility in Native-PAGE (reference gel in Figure 11)
(d) Protein-protein interactions between GaOx-CBM29 binding LBG on the QCM-D and GaOx-CBM29 in solution can add to the mass on the sensor and further reduce the readout frequency (Figure 12).

Despite this, both AGE and QCM-D show stronger binding by GaOx-CBM29 to polysaccharides compared to CBM29-GaOx, and thus the conclusion previously stated remains valid.

4.2.3 Oxidation of galactose containing polysaccharides

Michealis-Menten kinetics of *Fgr*GaOx, CBM29-GaOx and GaOx-CBM29 on galactose, XG, GGM, GG and LBG are presented in Table 5. Although the fusion of *Peq*CBM29-1-2 seemingly increased k_{cat} on galactose, up to 1.5 times for GaOx-CBM29. The increase could have been an effect expression variability. In particular, the expression of the fusion proteins in the SMD1168 strain versus KM71H for *Fgr*GaOx (See 4.1), where batch to batch variability may impart differences in specific activity (Anasontzis et al., 2014; Spadiut et al., 2010) and [1]. However, the similar catalytic efficiencies on D-galactose suggest that production variability

did not impact catalytic properties. Accordingly, significant gains in catalytic efficiency on polysaccharides was caused by fusion with CBM29-1-2, where the highest increases were observed for GaOx-CBM29 towards GG and LBG with 7.5 and 19.8 times higher catalytic efficiency relative to *Fgr*GaOx. Consistent with stronger binding of GaOx-CBM29 over CBM29-GaOx, the increase in catalytic efficiency was higher for the C-terminal fusion and was mainly driven by a significant decrease (up to 10-fold) of the K_M values GaOx-CBM29 compared to CBM29-GaOx. In fact, K_M values were only modestly lowered by N-terminal fusion of *Peq*CBM29-1-2.

Two particular trends arise from the binding and kinetics analyses; (1) outperformance of the C-terminal *Peq*CBM29-1-2 fusion over the N-terminal fusion and (2) higher gains in binding and catalytic efficiency towards galactomannans compared to xyloglucan and galactoglucomannan. Also, the poor performance of CBM29-GaOx could be due to the difference in their targeted saccharides between CBM29-2 and CBM32, where the adjacency of these in CBM29-GaOx may have antagonistic effects. This explains the overall poor binding affinity and the increase of K_M on xyloglucan, where ligands are distantly positioned on the same polysaccharide [2].

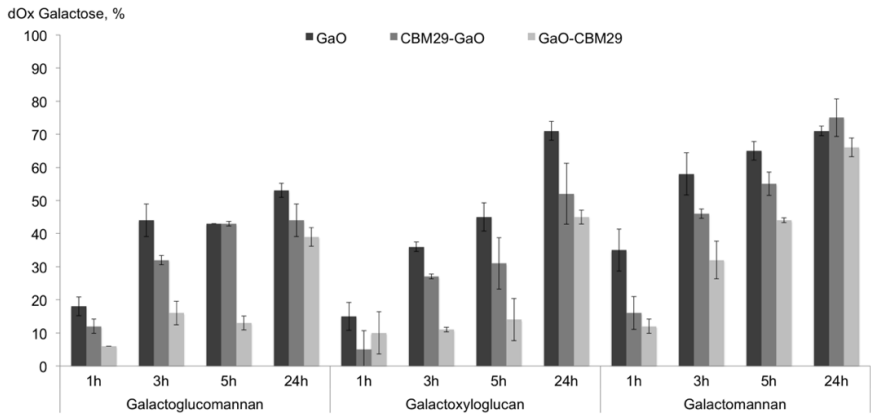


Figure 13. Degree of oxidation of galactose-containing hemicelluloses.
Degree of oxidation (dOx) of galactose in polysaccharides treated with FgrGaOx, CBm29-GaOx and GaOx-CBM29.The degree of oxidized galactose units in each treatment of 0.1% (w/v), spruce galactoglucomannan, tamarind galactoxyloglucan and guar galactomannan. Error bars indicate standard deviation, n=3. This figure was originally published in [2].

The oxidation performance of the fusion proteins was determined by analysing the degree of oxidation soluble polysaccharides (0.1% (w/v)), defined as the % percent of oxidised galactosyl residues relative to total available galactosyls in each polysaccharide. Whereas Michaelis-Menten kinetics evaluates the catalytic rates and affinity to substrates, this analysis is important as it evaluates the performance of the fusion proteins in more practical use. The degree of oxidation was compared after prolonged reaction with *Fgr*GaOx, CBM29-GaOx and GaOx-CBM29 on galactoglucomannan, galactoxylaglucan and guar galactomannan (Figure 13). In contrast

to kinetics, the oxidation performance was lowest for GaOx-CBM29, and the performance of CBM29-GaOx was lower than *Fgr*GaOx. These differences were most apparent in the first 5 hours of the reaction. However, the reduced performance is not surprising when considering that the used substrate concentrations are well above K_M values. At higher substrate concentrations, fusion-protein performance becomes limited by slow dissociation from its substrate to access new oxidation sites. Whereas it is conceivable that the relative performance of GaOx-CBM29 would be better than *Fgr*GaOx at lower substrate concentrations, a substrate concentration of at least 0.1 % (w/v) was necessary to accurately measure the degree of oxidation. In summary, the impacts of fusing *Peq*CBM29-1-2 to *Fgr*GaOx were analogous to impacts of fusion CBM's to glycoside hydrolases, where gains are lost when increasing substrate concentration.

Table 5. Kinetic parameters of GaOx, CBM29-GaOx and GaOx-CBM29

	GaO				GaO–CBM29				CBM29–GaO			
Substrate	k_{cat} (min^{-1})	K_m (mM) [a]	k_{cat}/K_m	K_m (wt%)[b]	k_{cat} (min^{-1})	K_m (mM) [a]	k_{cat}/K_m	K_m (wt%)[b]	k_{cat} (min^{-1})	K_m (mM) [a]	k_{cat}/K_m	K_m (wt%)[b]
D-Galactose	$24\times10^3\pm400$	42.4 ± 3	560 ± 33		$36\times10^3\pm900$	60.4 ± 5	590 ± 36		$31\times10^3\pm900$	56.0 ± 5	560 ± 40	
Galacto-glucomannan	195 ± 28	0.16 ± 0.07	1200 ± 500	28.5 ± 13	200 ± 13	0.074 ± 0.02	2700 ± 700	13.3 ± 4	239 ± 37	0.13 ± 0.08	1800 ± 800	23.4 ± 14
Galacto-xyloglucan	221 ± 7	0.07 ± 0.01	3300 ± 400	12.2 ± 2	274 ± 15	0.076 ± 0.02	3600 ± 700	13.7 ± 4	219 ± 10	0.14 ± 0.02	1500 ± 200	25.2 ± 4
Guar Galactomannan	311 ± 15	0.22 ± 0.03	1400 ± 200	39.6 ± 5	402 ± 11	0.037 ± 0.005	$11\times10^3\pm2000$	6.7 ± 1	313 ± 15	0.081 ± 0.02	3700 ± 800	14.6 ± 4
Locust bean Galactomannan	258 ± 24	0.19 ± 0.06	1400 ± 300	34.2 ± 11	214 ± 9	0.008 ± 0.002	$27\times10^3\pm9000$	1.4 ± 0.4	382 ± 24	0.084 ± 0.02	4600 ± 1000	15.1 ± 4

Since GaOx solely oxidizes the galactosyl units of polysaccharides the kinetic parameters were calculated based on galactose contents in each polysaccharide, although KM values are also presented in wt%. This table was originally published in [2].
a K_M values based on calculated galactose concentration in reactions containing each polysaccharide (mM)
b Apparent KM values based on the final concentration of each polysaccharide (mg/L)

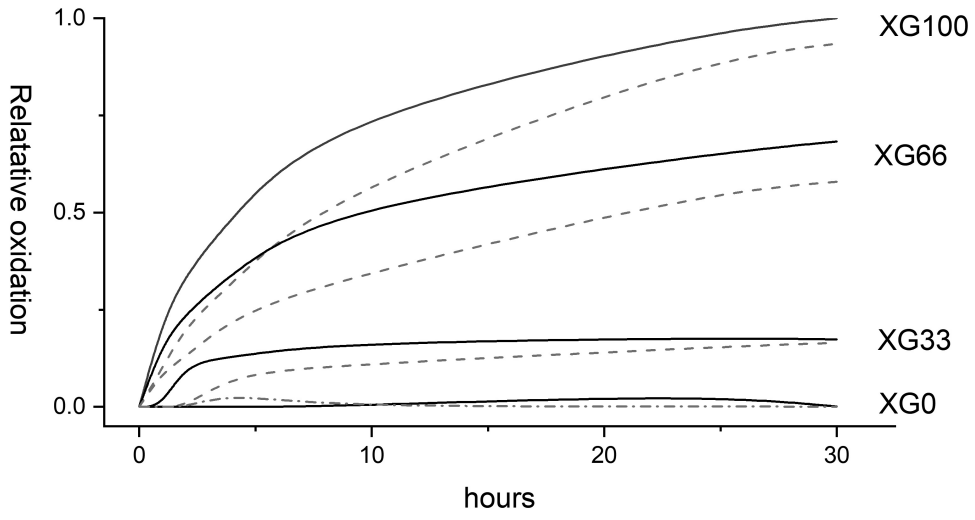


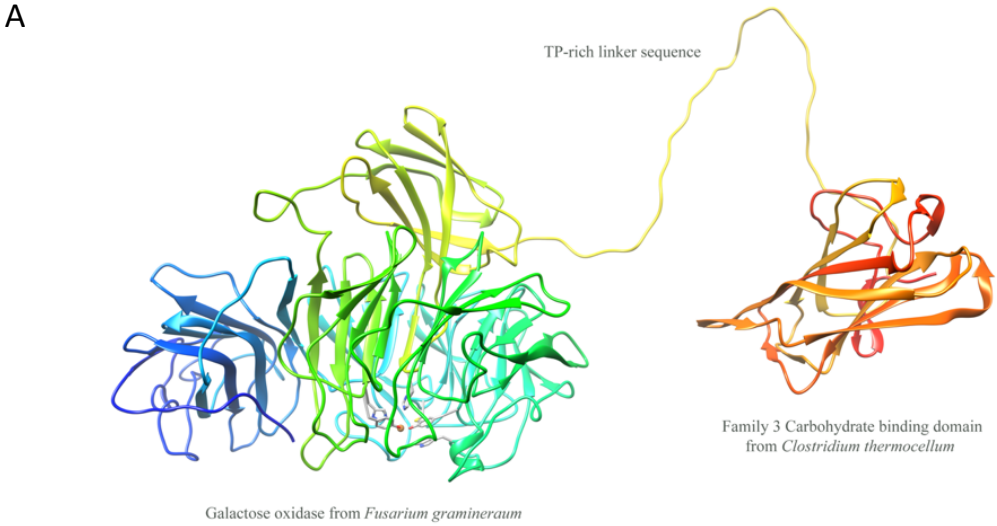
Figure 14. Relative oxidation of galactoxyloglucan immobilized on filter paper cellulose
1 U/mg galactose of *Fgr*GaOx (solid lines) and GaOx-CBM29 (dashed, grey lines) incubated with galactoxyloglucan adsorbed on 350 mg filter papers at XG100 (100% saturation, 14.6 mg XG), XG66 (66% sat., 9.6 mg XG), XG33 (33% sat., 4.9 mg) and XG0 (0% sat., 0 mg XG).

Oxidation of immobilized galactoxyloglucan

To evaluate the impact of *Peq*CBM29-1-2 on *Fgr*GaOx activity on solid substrates, the oxidation was followed on xyloglucan adsorbed on filter paper cellulose (prepared by my co-author Chunlin Xu at Åbo Academy) (Figure 14). Although attempts were made it was not possible to detect aldehyde products on the immobilized surface with a similar method as used for soluble polysaccharides. Instead, the oxidation was followed by monitoring the production of hydrogen peroxide, and thus only calculated estimations of product conversion yields were achieved. These were considerably low, for instance, *Fgr*GaOx yielded a conversion of 0.55 % on XG100 and 0.29% on XG33 as estimated from the generation of hydrogen peroxide. However, these reactions were not performed under optimal conditions for high conversion yields, but rather conditions allowing to follow generation of hydrogen peroxide. Similar to the oxidation of solubilized xyloglucan, *Fgr*GaOx achieved highest degrees of oxidation and oxidation rates at all concentrations of xyloglucan. Accordingly, the impact of CBM29 seemingly also retards action on surface immobilised polysaccharides. These results are not part of article [2] but were presented at the 16th European Conference on Biotechnology (Mollerup et al., 2014).

4.2.4 Glucose and cellulose activity in M3-CBM3

As described in section 1.3.4, Sun and colleagues previously published a mutant (M3; m-RQW) of the galactose oxidase from *F. graminearum* with low activity towards D-glucose (Sun et al., 2002). Similar to the hypotheses giving merit to the fusion of CBM29 to the wild-type GaOx, where fusion enhanced enzyme stability, promoted binding to targeted oligosaccharides and activity at low substrate concentrations, this study evaluated the potential of a family 3 cellulose binding module to enhance the activity of M3 towards glucans and cellulosic substrates. Herein, the family 3 carbohydrate binding domain from *Clostridium thermocellum* was fused to the C-terminal of M3, generating M3-CBM3 (Figure 15). Later, three additional amino acid substitutions (C383S, Y463H and V494A), known to significantly increase the catalytic activity of *Fgr*GaOx such as discussed in section 1.3.3 (Delgrave et al., 2001; Wilkinson et al., 2004), were added to M3-CBM3 to generate the fusion protein M6-CBM3. Both fusion variants were expressed in *P. pastoris* (see 4.1) and subjected to biochemical analysis.



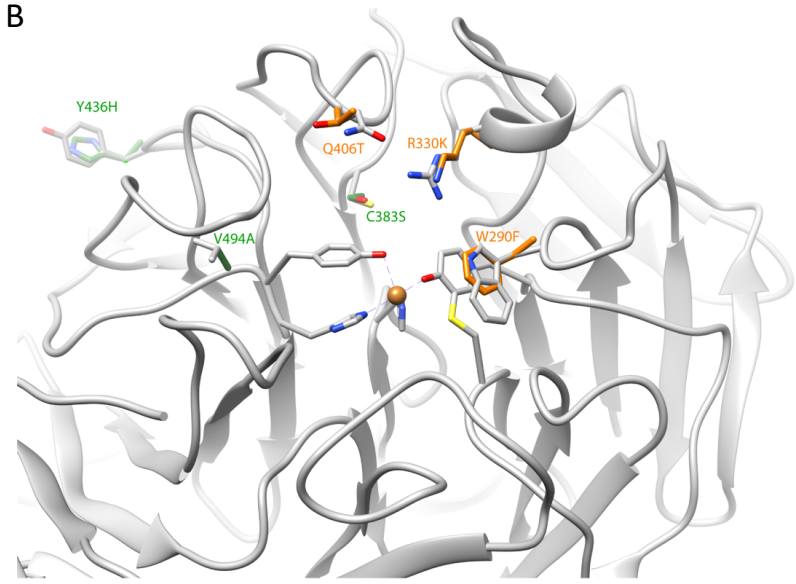


Figure 15. Model structure of M3-CBM3 and M6-CBM3

(A) Model showing the domain structure of M3-CBM3 and M6-CBM3. (B) Structure of the active site of *Fgr*GaOx (pdb code 1gog) showing wild-type amino acids (grey) and the substitutions in M3-CBM3 (orange) and M6-CBM3 (orange and green). The model was constructed from the crystal structures of *Fgr*GaOx (pdb code 1gog) and *Cth*CBM3 (pdb code 1g43).: *FgrGaOx* (1-639, 68 kDa), Tp-linker (640-680, 4.4 kDa), *Cth*CBM3 (681-839, 20.6 kDa)

Characterization of M3-CBM3

The activity of M3-CBM3 on galactose (300 mM) was 228 μmole*min^{-1}*μmole^{-1} (Table 6) which is 114 times lower than *Fgr*GaOx, but approximately 10 times higher than the activity reported for M3 (a.k.a. m-RQW) (Sun et al., 2002). Similarly, the activity of M3-CBM3 on 300 mM glucose was 3 times higher than the relative activity reported for M3, and 15.3% that measured on galactose. Activity on mannose, the C2-epimer of glucose, was 5.4 % that measured for galactose, suggesting that the position of the C2-hydroxyl is of importance to the selectivity of M3-CBM3. A similar activity level was also found on pentoses, including D-xylose (8.5%), as well as D- and L-arabinose (5.6 and 4.8 % relative to galactose), which could be ascribed to the presence of a primary hydroxyl in the D-xylofuranose, D- and L-arabinofuranose forms or

alternatively, oxidation of secondary hydroxyls. Surprisingly, the activity on 1-*O*-methyl-β-D-xylopyranose, which contains no primary hydroxyls, was equal to that of D-xylopyranose, suggesting that a secondary hydroxyl is targeted in xylose. Activity on secondary hydroxyls have also been confirmed for M3 (Sun et al., 2002).

Table 6. Activity of M3-CBM3 and *Fgr*GaOx on oligo- and polysaccharides

	Saccharide composition	M3-CBM3	*Fgr*GaOx
Monosaccharides			
D-Glucose	D-Glc	34.5± 1.3	NA
D-Galactose	D-Gal	228 ± 18	26078 ± 1173
D-Mannose	D-Man	12.4 ± 1	33.6 ± 1.7
D-Arabinose	D-Ara	12.7 ± 1	2.6 ± 0.1
L-Arabinose	L-Ara	10.9 ± 1.2	150 ± 17
D-Xylose	D-Xyl	19.5 ± 1.7	19.4 ± 1
1-O-Me-β-D-Xylopyranose[1]	D-Xyl	20.2 ± 2.7	145 ± 3
D-Fucose	D-Fuc	1.1 ± 0.7	2.5 ± 0.4
Disaccharides			
Cellobiose	D-Glc-β-(1➔4)-D-Glc	10.0 ± 0.9	NA
Maltose	D-Glc-α-(1➔4)-D-Glc	8.4 ± 0.4	NA
iso-Maltose	D-Glc-α-(1➔6)-D-Glc	14.8 ± 0.6	
Melibiose	D-Gal-α(1➔6)-D-Glc	164 ± 49	8224 ± 312
Sucrose	D-Glc-β(1➔2)-D-Fru	16.5 ± 3	1.2 ± 0.2
Lactose	D-Gal-β(1➔4)-D-Glc	75.6 ± 7	2235 ± 114
Trisaccharides			
Maltotriose	α-(1➔4)-D-Glc (x3)	8.8 ± 2	
Raffinose	D-Gal-α(1➔6)-D-Glc-α(1➔4)-D-Fru	146 ± 25	11024 ± 419
Iso-maltotriose	α-(1➔6)-D-Glu (x3)	8.8 ± 1	
Other substrates			
BIM4	Branched α-glucooligo	NA	NA
BIM6	Branched α-glucooligo	NA	NA
Avicell®	Microcrystaline cellulose	minor	NA
Filter paper No. 1	Cellulose filterpaper	Activity	minor
CNC	Cellulose nano-crystals	NA	NA
Cello-oligosaccharides	**concentration**		
D-glucose		0,90	
Cellobiose	6 mM	0,20	
Cellopentaose		0,90	
D-glucose		1,50	
Cellobiose	9 mM	0,20	
Cellopentaose		1,20	

All soluble substrates were tested at 150 mM, with exception of [1] which was tested at 60 mM. 85 ng M3-CBM3 and 10 nG FGRGaOx was incubated with HRP (2U/mL) and ABTS (2mM) in 50 mM sodium phosphate (pH 7.5) for 30 min at 30 ºC prior addition of substrate. Hydrogen peroxide generation as measured continuously for up to 2 hours. NA = no activity Errors indicate standard deviations, n=4

*Fgr*GaOx is known to oxidize a wide range of galactose-containing di-, oligo- and poly-saccharides, where the targeted galactose has a free C4-hydroxyl group and is preferably linked to the neighbouring residue through an α-(1➔6) glycosidic linkage (Parikka et al., 2015; Whittaker, 2005). The activity of *Fgr*GaOx on melibiose and lactose was 31.5% and 8.5% of the activity of galactose (Table 6). By contrast, M3-CBM3 oxidized melibiose and lactose at 72% and 33%, relative to galactose, showing similar preference to α-(1➔6) substituted disaccharides, but with a significantly lower sensitivity than *Fgr*GaOx, and better acceptance of di-sicchacharides. Considering glucose-containing substrates, activity on iso-maltose and cellobiose was 43% and 29% that of glucose, which suggest same preference for α-(1➔6) oligosaccharides when glucose is oxidised.

M3-CBM3 oxidized galactose and glucose containing oligo-saccharides, such as raffinose and maltotriose, with similar rate as the corresponding di-saccharides (Table 6). Moreover, M3-

CBM3 activity on cellopentaose was similar to that measured for glucose. Notably, however, M3-CBM3 oxidation of branched α-gluco-oligosaccharides was not observed (BIM4 and BIM6).

Gain in M6-CBM3 activity offset by reduced substrate range

The initial activity of purified M6-CBM3 on 150 mM D-galactose 102-fold higher than the activity of M3-CBM3 and nearly same activity as the wild-type (Table 7).

Table 7. Substrate range of M6-CBM3

	Saccharide composition	Activity, U*µmol^{-1} M_6-CBM3
Monosaccharides		
D-Glucose	D-Glu	NA
D-Galactose	D-Gal	23272 ± 1117
D-Mannose	D-Man	NA
D-Arabinose	D-Ara	NA
L-Arabinose	L-Ara	NA
D-Xylose	D-Xyl	165 ± 16.9
Disaccharides		
Cellobiose	β-(1➔4)-D-Glu (x2)	NA
Maltose	α-(1➔4)-D-Glu (x2)	NA
Melibiose	D-Gal-α(1➔6)-D-Glu	1590 ± 174[3]
Lactose	D-Gal-β(1➔4)-D-Glu	6057 ± 296
Trisaccharides		
Raffinose	D-Gal-α(1➔6)-D-Glu-α(1➔4)-D-Fru	876 ± 26,5[3]

All soluble substrates were tested at 150 mM, 15 ng M6-CBM6 was incubated with HRP (2U/mL) and ABTS (2mM) in 50 mM sodium phosphate (pH 7.5) for 30 min at 30 °C prior addition of substrate. Hydrogen peroxide generation as measured continuously for up to 2 hours. NA = no activity Errors indicate standard deviations, n=4

Remarkably, while M6-CBM3 activity on D-galactose reached wild-type levels, activity on D-glucose, cellobiose, and maltose was lost entirely. While M3-CBM3 and *Fgr*GaOx showed low but clearly detectable activity on D-mannose, D- and L-arabinose, M6-CBM3 did not display activity on these substrates (Table 7). Interestingly, M6-CBM3 showed over eight-times higher activity on D-xylose compared to *Fgr*GaOx which could indicate enhanced ability to oxidize secondary hydroxyls. Similarly as reported for other GaOx variants (Escalettes and Turner, 2008). Also, the activity on melibiose and raffinose is relative lower to D-galactose in M6-CBM3 than for *Fgr*GaOx. Instead M6-CBM3 was 2.7-times more active on lactose than *Fgr*GaOx, suggesting that the amino acids substitutions added to M6-CBM3 impacted preference for the glycosylic bond in galactose-containing -di- and oligosaccharides which was otherwise unchanged in M3-CBM6.

4.2.5 Binding and oxidation of cellulose

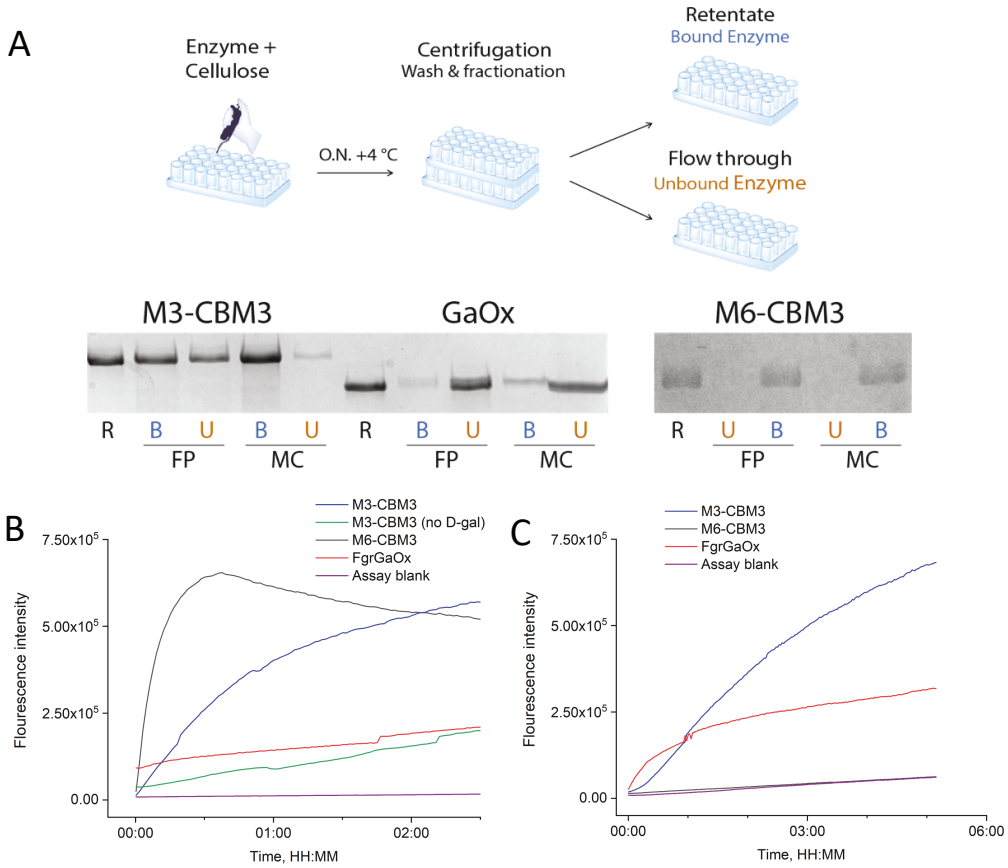


Figure 16. Binding and detection of oxidation of cellulose
(A) Immobilization of M3-CBM3, M6-CBM3 and *Fgr*GaOx on filter paper cellulose (FP, Whatmann #1) and microcrystalline cellulose (MC, Avicel®) detected by SDS-PAGE. B = bound protein, U = unbound protein, R = reference sample with no cellulose collected in flow through. (B) Activity of M3-CBM3, M6-CBM3 and *Fgr*Ox immobilized on microcrystalline cellulose on 300 mM D-galactose. (C) Hydrogen peroxide generation (activity) of M3-CBM3, M6-CBM3 and *Fgr*Ox immobilized on filter paper cellulose (no addition of D-galactose).

Immobilization of M3-CBM3 and M6-CBM3 on cellulose filter paper (FP, Whatmann #1) and microcrystalline cellulose (MC, Avicel®) was higher than that determined for *Fgr*GaOx (Figure 16A). Consistent with earlier studies of *Ct*CBM3 protein fusions, M3-CBM3 and M6-CBM3 bound better to MC than to FP cellulose (Sakka et al., 2011; Voutilainen et al., 2013). Moreover, activity of immobilized M6-CBM3, M3-CBM3 (MC) on 300 mM galactose was confirmed, showing that the immobilized enzymes remain active on soluble carbohydrates (Figure 16B). M6-CBM3 displayed approximately 2.8 times higher activity than M3-CBM3, explained by the higher specific activity on D-galactose. M3-CBM3 showed a slight but detectable oxidation of MC in absence of soluble substrate, putatively owing to oxidation of terminal glycosyls units. This was not observed for M6-CBM3 or *Fgr*GaOx. Although *Fgr*GaOx does not bind specifically to microcrystalline cellulose, a small fraction of non-specific bound *Fgr*GaOx was

detectable by SDS-PAGE and the hydrogen peroxide assay. Hydrogen peroxide generation was detected for both M3-CBM3 and *Fgr*GaOx bound to FP (in absence of D-gal), although it was 2.5 times higher for bound M3-CBM3 than *Fgr*GaOx.

4.2.6 Loss and restored affinity for D-galactose

The k_{cat} and K_M of M3-CBM3 on galactose was 7.12 s^{-1} and 200 mM, respectively, leading to over 300 times reduction in catalytic efficiency relative to *Fgr*GaOx (Table 8). These results are in good agreement with those previously reported for M3 (refer to Sun et al., 2002 and section 1.3.4). Clearly then, although imparting glucose oxidation to *Fgr*GaOx, these amino acid substitutions significantly reduce the catalytic efficiency of the enzyme. Based on earlier studies of *Fgr*GaOx (Delgrave et al., 2001) (1.3.3), the additional amino acid substitutions in M6-CBM3 (C383S, Y436H and V949A) were expected to increase the catalytic efficiency of the enzyme on D-galactose and D-glucose. Indeed, the catalytic efficiency of M6-CBM3 on galactose was increased by 100-times relative to M3-CBM3, where the k_{cat} was 229 s^{-1} and K_M 56.4 mM, and so regained activity levels similar to *Fgr*GaOx. However, activity of M6-CBM3 on glucose was entirely lost.

Table 8. Kinetic parameters of M3-CBM6 and M6-CBM3 on D-galactose

	k_{cat}, s^{-1}	K_M, mM	k_{cat} / K_M, s^{-1} mM^{-1}
M3-CBM3	7.12 ± 0,78	200.9 ± 39,5	0.04 ± 0,012
M6-CBM3	229.03 ± 9,83	56.4 ± 7,5	4.06 ± 0,71
***Fgr*GaOx**	370.35 ± 6,89	29.0 ± 2,3	12.75 ± 1,25

Errors indicate standard deviations, n=4

None of the amino acids introduced to M6-CBM3 were expected to interact directly with the substrate, although C383 and V494 are positioned in close proximity to substrate binding sites within *Fgr*GaOx (Figure 15). As a lone substitution, V494A is only reported to increase k_{cat} while not affecting K_M (Delgrave et al., 2001; Wilkinson et al., 2004). Y436H is positioned far from the active site and likewise only reported to increase k_{cat}. C383S, however, decreases the K_M towards D-galactose by 2 times; this was not observed for the C383A variant, suggesting the hydroxyl group of serine has importance for the affinity to galactose (Wilkinson et al., 2004). Therefore, the increased catalytic efficiency of M6-CBM3 caused by the V494A, Y436H, C383S substitutions was expected; however, the observed impact of these substitutions on the gain in substrate range afforded by W290F, R330K and Q406T was not. From its buried position under Y405 and F441, the sidechain of C383 is not in direct contact with the substituted residues (R330K and Q406T). Nevertheless, C383 evidently indirectly impacts substrate binding by affecting substrate ligands in both *Fgr*GaOx and M6-CBM3, which will be further considered in section 4.3.4.

4.3 New AA5_2 oxidases *Cgr*RaOx and *Pru*AA5_2A

Articles [3] and [4]

At the planning stage of the research, the characterized members of CAZy family AA5 subfamily 2, all being from the *F.* genus, all showed similar substrate profile to the archetypal galactose oxidase from *F. graminearum* (*Fgr*GaOx). The biological function of galactose oxidase has not been determined, but production of extracellular hydrogen peroxide for lignolytic peroxidases or monooxygenases (see section 1.2) been suggested (Kersten and Kirk, 1987). The restricted activity of *Fgr*GaOx to galactose containing carbohydrates, however, seems to conflict with this general function. From an applied perspective, the limited activity of *Fgr*GaOx also narrows its utility for polysaccharide engineering. Accordingly, a bioinformatics analysis of characterized AA5_2 sequences was performed in an effort to identify new carbohydrate oxidases with distinct substrate preference relative to *Fgr*GaOx. This thesis addresses the characterization of two new CAZyme family AA5_2 members selected from *Colletotrichum graminicol*a (*Cgr*RaOx) and *Penicilium rubens Wisconsin* (*Pru*AA5_2A). Therefore, these enzymes will be considered in the following sections as well.

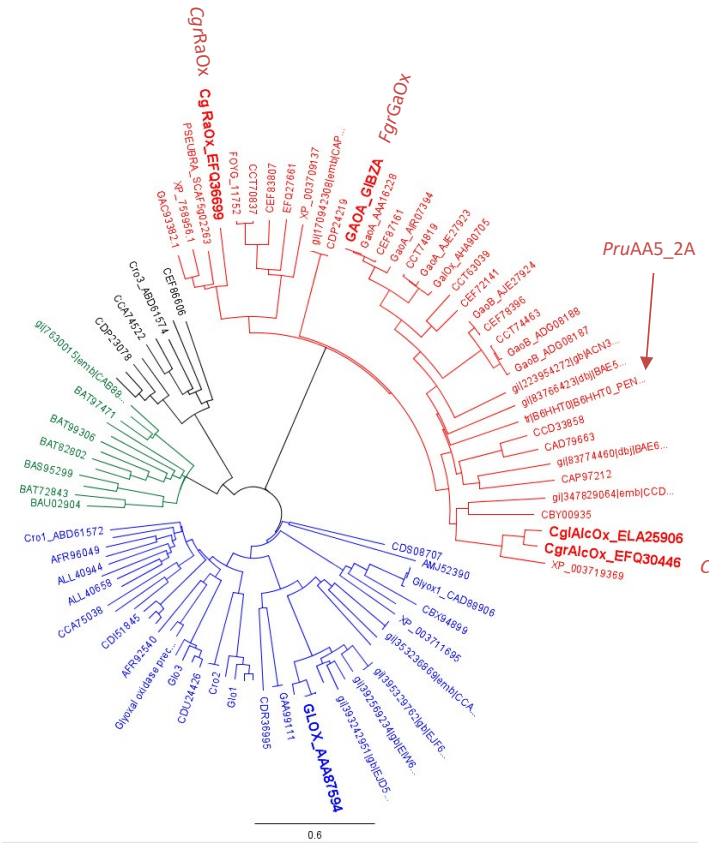


Figure 17. Phylogenetic tree of CAZy family AA5 Subfamily 1, containing glyoxal oxidase is visible in blue, green and black. Subfamily 2 (red), containing galactose oxidase from *F. graminearum* (*Fgr*GaOx) and other members of the *Fusarium* genus, alcohol oxidase from *C. graminicola* (*Cgr*AlcOx). The new AA5_2 members from *C. graminicola* (*Cgr*RaOx) and *P. rubens* (*Pru*AA5_2A). This figure was originally published in [3]

4.3.1 Sequence-function relations in other members of the CAZy family AA5_2

Considering the phylogeny of family AA5_2, sequences from *Fusarium* genus are centred on separate clades with neighboring clades containing uncharacterized sequences (Figure 17). The alcohol oxidases, *Cgr*AlcOx and *Cgl*AlcOx, are isolated within a distinct clade with only one nearby neighbour, leaving large regions of the family uncharacterized. The selection of *Cgr*RaOx and the phylogenetic three of family AA5_2 (Figure 17) was performed by my co-authors [3], through a search in the Pfam database (Finn et al., 2016) for sequences containing the beta-propeller and Duf1929 domains of *Fgr*GaOx. Following statements on the sequence relationships between *Fgr*GaOx and between other AA5_2 members are results of the sequence and structure analysis done by me. These are under the account of this thesis work, and are consequently presented here, in the results, as opposed to previous discussion of *Fgr*GaOx in the introduction. However, following refer back to the introduction and particularly to Figure 2.

Comparative analysis of Fusarium sequences

The sequence analysis described herein included 36 sequences from eukaryotic members of the AA5_2 family (available through the CAZyme database in April 2017) (alignment available in article [4]). Nine of these were from the *Fusarium* genus, including members that have been characterized. Considering the grouping of amino acids around the active site of *Fgr*GaOx, a sequence alignment of the nine sequences from the *Fusarium* genus showed that all amino acids in group 1 are strictly conserved (Table 9) in the *Fusarium* genus and that 10 of the 14 residues lie in sequence stretches containing at least 5 identical amino acids (consensus sequences). Aparecido et. al. (2010) noted the presence of similar sequence stretches in the *Fusarium* genus, six in total, but did not further investigate these.

In this study, 10 consensus sequences in *Fusarium* AA5_2 members were identified, ranging from 5 to 13 amino acids and containing only members from group 1 (alignment available in [4]). The exceptions were the stacking tryptophan (W290 in *Fgr*GaOx) from group 2 and V494 from group 3 (see Table 9). Nine of the 10 consensus sequences are placed either in or close to the active pocket, where the tenth sequence stretch forms the loop in the C-terminal domain (Figure 2A). No other amino acids from group 2 are placed within the consensus regions. However, many of the amino acids in group 2 are conserved between the *Fusarium* AA5_2's which probably originate in very similar substrate ranges among corresponding enzymes. Higher variability was observed with amino acids in group 3 owing to their position further from the active pocket.

Table 9. Summary of important amino acids in *Fgr*GaOx and AA5_2

Postion in *Fgr* GaOx	*Fgr* GaOx	*Fusarium spp.*	*Pru* AA5_2A	*Cgr* RaOx	*Cgr* AlcOx	Substitutions from litterature	Function/position in *Fgr* GaOx	Relevant references
Group 1: catalytic residues and copper-ligands or positions that are assigned as important for catalytic function								
188	S	Either S or A	S	S	S		Hydrogen bonds to Asn225	
194	F	cons	Y	F	W		Shielding copper(II) from solvent, ring stacking interaction with Phe227	Deacon et. al. 2004
225	D	cons	D	D	D		Hydrogen bonds to Ser188	
227	F	cons	F	F	F		Ring stacking interaction with Tyr272	
228	C	cons	C	C	C		Thioether linkage in Cys228-Tyr272 cofactor	Ito et al. 1991, Whittaker 2005, Rogers et. al. 2007
272	Y	cons	Y	Y	Y		The catalytic tyrosine, copper ligand	Ito et al. 1991, Whittaker 2005, Rogers et. al. 2007
291	S	cons	N	T	S		Positioned in internal cavity	
334	H	cons	H	H	H		Hydrogen bond to Tyr405	
441	F	cons	F	F	F		Ring stacking interaction to Tyr495	Ito et al. 1991
464	F	cons	F	F	F		Potential substrate interaction	Deacon et. al. 2004
495	Y	cons	Y	Y	Y		Copper ligand	Ito et al. 1991, Whittaker 2005
496	H	cons	H	H	H		Copper ligand	Ito et al. 1991, Whittaker 2005
514	L	cons	L	L	L		Positioned in a hydrophobic cavity	
581	H	cons	H	H	H		Copper ligand	Ito et al. 1991, Whittaker 2005
Group 2: positions where amino acid substitutions have altered substrate range of *Fgr* GaOx								
290	W	cons	W	Y	F	F	Ring stacking interaction to Tyr272, substrate ligang, mutation intails great loss of catalytic activity	Rogers et al. 2007, Lippow et al. 2010, Moon et al. 2012, Yin et al. 2015, Ito et. al. 1991, Whittaker 2005
326	Q	Either Q or E	D	A	G	E	Hydrogen bonding to Arg330	Lippow et al. 2010, Sun et al. 2002
329	Y	cons	Y	W	F	K	Substrate ligand and hydrogen bonding to Arg330	Lippow et al. 2010
330	R	cons	R	R	M	K	Substrate ligand, bidentate hydrogen bonds to C3-OH and C4-OH of galactose	Lippow et al. 2010, Rannes et al. 2011, Sun et al. 2002
405	Y	cons	Y	Y	Y	F	Positioned in the active site, hydrogen bonds to Tyr495 and Arg330	Rannes et al. 2011
406	Q	cons (1 D & 1 E)	E	S	T	T,E	Substrate ligand, hydrogen bond to C2-OH of gal	Lippow et al. 2010, Rannes et al. 2011
463	P	cons	P	G	L	I	Mutation considered detrimental to gal or glc activity	Lippow et al. 2010, Rannes et al. 2011
Group 3: positions where amino acid substitions have enhanced the catalytic activity of *Fgr* GaOx								
248	K	var (Q,E,A,S)	K	A	A	E	Surface residue, substitution increases activity towards galactose	Delagrave et al. 2001, Wilkinson et al. 2004
352	T	cons (1 I & 1 K)	K	K	K	S	substitution increases activity of *Fgr* GaOx on gal	Delagrave et al. 2001
366	K	cons	K	K	T	R	Surface residue, increases activity on gal	Delagrave et al. 2001
383	C	var (N, S)	S	C	C	S	Mutation increases V*max* on gal by 1.75x and lowers kM by 3.6	Delagrave et al. 2001, Wilkinson et al. 2004
436	Y	var (I, T, H)	A	K	Y	H, N	Mutation increases V*max* on gal by 2x	Delagrave et al. 2001
494	V	var (T, N)	V	N	N	A	Internal hydrophobic cavity, substitution increases V*max* on Gal by 1,75x	Delagrave et al. 2001, Wilkinson et al. 2004, Sun et. al. 2001

Amino acids from Table 1, identified from sequence and structure analysis of *Fgr*GaOx (Figure 2) and their preservation in the *Fusarium* genus, *Pru*AA5_2A, *Cgr*RaOx and *Cgr*AlcOx.

AA5_2 Alcohol oxidases

While there are currently two known alcohol oxidases (Yin et al., 2015), *Cgr*AlcOx alone was used as a reference for the alcohol oxidase activity, since the sequence identity to *Cgl*AlcOx is high (82%) and they display similar biochemical properties. Considering group 1 amino acids, *Cgr*AlcOx deviates from *Fgr*GaOx by the presence of a tryptophan (W39) at the position of F194 in *Fgr*GaOx. Interestingly, *Cgr*AlcOx contains a phenylalanine (F137) at the stacking position on top of the Cys-Tyr cofactor, whereas *Fgr*GaOx contains a tryptophan (W290). The amino acids of these neighbouring positions are correspondingly swapped in *Cgr*AlcOx relative to *Fgr*GaOx, which could be denotive of family AA5_2 alcohol oxidases (Figure 18 A and E). In fact, of the AA5_2 sequences considered herein, only one other besides *Cgr*AlcOx and *Cgl*AlcOx contains this swap.

*Cgr*AlcOx also differs from *Fgr*GaOx in group 2 residues. Specifically, *Cgr*AlcOx contains a phenylalanine (F171), a methionine (M170) and a threonine (T246) at the corresponding positions of Y329, R330 and Qln406 in *Fgr*GaOx (Table 9, Figure 18A and E). Consequently, *Cgr*AlcOx lacks all the galactose ligands from *Fgr*GaOx, while retaining an aromatic residues at the stacking position (F137) and the shielding residues on both sides of the Cys-Tyr cofactor (W39 and F171). These amino acid deviations to *Fgr*GaOx allegedly explains the inability of *Cgr*AlcOx to oxidize galactose, and instead showing high catalytic efficiency on small aliphatic alcohols. In group 3, *Cgr*AlcOx contains an asparagine (Asp333) corresponding to V494 in *Fgr*GaOx which is comparably larger and more polar and could have impact on the hydrophobic environment between the phenylalanine and the basic tyrosine (F464 and Y495 in *Fgr*GaOx).

Raffinose specificity by CgrRaOx and PruAA5_2

Galactose oxidase from *F. acuminatum* (Alberton et al., 2007) and *F. sambucinum* were the first AA5_2 carbohydrate oxidases reported to have better activity or catalytic efficiency on raffinose than galactose raising attention to these oligosaccharides. In 2017, we reported our similar observation for *Cgr*RaOx [3], although this oxidase shows no significant activity on galactose (as discussed later). The sequence of *Cgr*RaOx contains an N-terminal PAN, a central Kelch_1 and a C-terminal DUF1929 domain, and so differs from *Fgr*RaOx by the N-terminal domain. The larger N-terminal domain of *Cgr*RaOx also accounts for the larger predicted molecular size of *Cgr*RaOx (90 kDa).

The overall *Cgr*RaOx sequence shares 29% identity to *Fgr*GaOx, but 41% identity when removing the N-terminal domains. When considering amino acids important for catalytic function (Group 1 residues; Table 9) *Cgr*RaOx differs from *Fgr*GaOx by a minor deviation, namely the threonine (T351) at the position of Ser291 in *Fgr*GaOx (Figure 18D). Additional sequence

differences were observed when considering substrate ligands (group 2). For example, the stacking tryptophan (w290 in *Fgr*GaOx) is a tyrosine in *Cgr*RaOx (Y350), while the adjacent tyrosine in *Fgr*GaOx (Y329) is replaced by a tryptophan residue (w596), revealing a potential swap between these two positions. *Cgr*RaOx also contains an alanine (A593) at the corresponding position of Q326 in *Fgr*GaOx, a serine (S665) at the Q406 and a glycine (G723) at P464. Considering group 3 amino acids, *Cgr*RaOx contains a lysine (K696) at the position of Y436 and an asparagine (N753) relative to V494. The significance of these deviations is discussed later in 4.3.4.

Like *Cgr*RaOx*, Pru*AA5_2A from *P. rubens (Wisconsin* 54-1255) also displays preference towards raffinose over other tested carbohydrates (section 4.3.3). Similar to *Fgr*GaOx, however, the modular structure of *Pru*AA5_2A from *P. rubens (Wisconsin* 54-1255) consists of an N-terminal CBM32 (F5_F8_type_C) domain, a central 7-bladed ß-propeller (Kelch_1) catalytic domain and a C-terminal DUF1929 domain. *Pru*AA5_2A shares 49.6 % sequence identity to *Fgr*GaOx and 23.3 % to *Cgr*RaOx, and is placed significantly closer to the clades containing the *Fusarium spp*. in the phylogenetic tree (Figure 17). At key amino acid positions (Table 9) *Pru*AA5_2A is one of only four AA5_2 members to contain a tyrosine (Y170) at the position corresponding to F194 in *Fgr*GaOx (Figure 18C), and one of three members to contain an asparagine (N 271) at the position corresponding to S291.

The main indication that *Pru*AA5_2A targets carbohydrates, and potentially galactose, is the presence of the main galactose ligand, arginine (R327) corresponding to R330 in *Fgr*GaOx. However, *Pru*AA5_2A differs from *Fgr*GaOx in group 2 positions by containing an aspartic acid (D323) at the position of Q326 in *Fgr*GaOx, which is involved in coordinating the position of R330 (through hydrogen bonding). Aspartic acid (D323) in *Pru*AA5_2A may affect hydrogen bonding and lead to less strict coordination of R330 allowing more flexible substrate binding. Also, Q406 of *Fgr*GaOx binds to the C2-hydroxyl in galactose, and is glutamic acid (E399) in *Pru*AA5_2A, which could further entail a flexible substrate accommodation. Differences to *Fgr*GaOx amongst positions in group 3 include an alanine (A433) at the position of Y436 in *Fgr*GaOx and, notably, a serine (S376) at the position of C383.

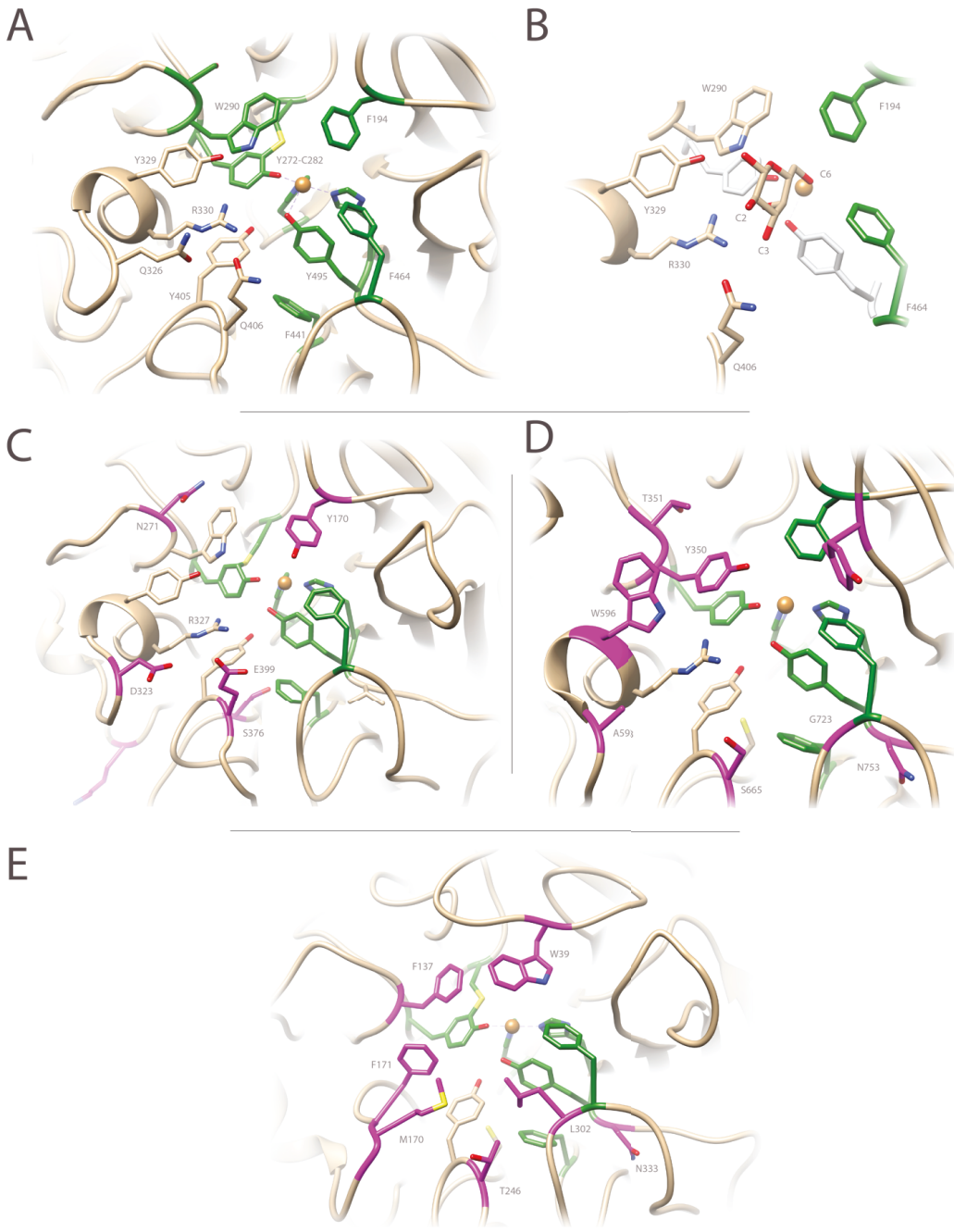


Figure 18. Active site structures if *Fgr*GaOx, *Pru*AA5_2, CgrRaOx and *Cgr*AlcOx
(A) The active site of *Fgr*GaOx (1GOG.pdb) containing residues from group 1, Table 9 (green) and residues from group 2, Table 9 (tan). (B) Docking of D-galactose to the active site of *Fgr*GaOx, depicted here with substrate ligands (W290, Y329, R330 and Q406) as well as F194 and F464 as discussed in section 1.2.1. (C) The structure of *Pru*AA5_2A was created using Phyre2. The active site of *Pru*AA5_2A showing the conserved aromatic residues from group 1 (Table 9) highlighted in green; amino acids that deviate from *Fgr*GaOx are shown in dark purple. (D) The active site of the *Cgr*RaOx was modelled using phyre 2. *Cgr*RaOx contains the arginine corresponding to Arg330 in *Fgr*GaOx, but lacks galactose ligands at positions corresponding to W290 and Q406 in *Fgr*GaOx. (E) The active site of *Cgr*AlcOx (5C92.pdb), which deviate from *Fgr*GaOx by lacking all known galactose ligands.

Consensus sequences for characterized AA5_2

Amino acids in group 1 were highly conserved between *Fgr*GaOx, *Cgr*RaOx, *Pru*AA5_2A and *Cgr*AlcOx. This raised the question whether the consensus sequences identified for *Fusarium* sequences also applied to these members of the AA5_2 family. Indeed, the sequence identity of *Cgr*RaOx, *Pru*AA5_2A and *Cgr*AlcOx compared to *Fgr*GaOx within the 10 consensus sequences are approximately 2-times higher compared to identities across the entire catalytic domain: *Cgr*AlcOx; 46 vs 79 %, *Cgr*RaOx; 29 vs. 77 % and *Pru*AA5_2: 49 vs. 84% (Figure 19). As previously noted, these consensus sequences does not contain galactose ligands other than the stacking tryptophan (W290).

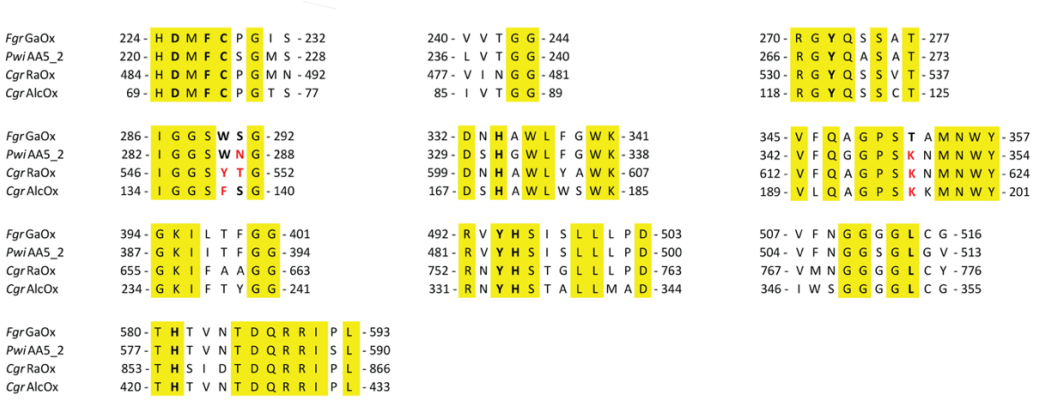


Figure 19. Alignment of consensus sequences from *Fusarium spp.* with *Pru*AA5_2A, *Cgr*RaOx and *Cgr*AlcOx. Residues from Table 10 marked in bold, deviations at these positions are highlighted in red. All are from group 1 with the exception of W290. This figure was originally published in [4]

Although patterns are slightly less clear, these consensus sequences also repeat in all 36 eukaryotic sequences in family AA5_2. As expected, all copper ligands and the cysteine forming the Cys-Tyr cofactor are all strictly conserved, and notably, Y405 is likewise conserved in all 36 family members which substantiates the importance of the H-bond it forms to Y495 in *Fgr*GaOx. Other amino acid positions that were highlighted in section 1.2.1 were either conserved or showed conservative deviations which increase our understanding of sequence-function attributes of the AA5_2 family (Table 10).

Table 10. Summary of selected amino acids in AA5_2 sequences

Amino Acid in *Fgr*GaOx	AA5_2
F194	28 F, 4 Y, 3 W, 1 M
F227	36 F
F441	33 F, 3 Y
F464	35 F, 1 W
W290	28 W, 5 Y, 3 F
Y329	18 Y, 10 W, 3 F, 2 M, 2 Leu, 1 S
R330	26 R, 5 K, 3 F, 1 H, 1 A
Y405	36 Y
Q406	17 Q, 7 T, 6 D, 3 E, 1 S, 1 N

In summary, structural analysis of *Fgr*GaOx (see 1.2.1) combined with the sequence analysis of CAZyme family AA5_2 more generally, show that the active pocket of AA5_2 copper-radical oxidases can be viewed as two distinct regions that form two opposing half-spheres in the active pocket (Figure 2B and C). These regions can be described as follows:

(a) A region of highly conserved amino acids that constitute the functional basis for AA5_2 copper-radical oxidase catalysis, common for all family AA5_2 members (group 1, Table 9, Figure 2B)
(b) A region of more sequence variability, with a few exceptions, where amino acids are either substrate ligands or play important roles in substrate interaction (group 2, Table 9; Figure 2C). Analysis of sequence variability in this region can potentially inform predictions of substrate-preferences for uncharacterized copper-radical oxidases.
(c) In addition, ten consensus sequences, containing most amino acids from group 1, seems to be a characteristic feature of the copper-radical oxidases in family AA5_2

Additionally, the lamost prominet diversity between the *Fgr*GaOx, *Cgr*RaOx, *Cgr*AlcOx and even *Pru*AA5_2A is observed at the N-terminal domains. The N-terminal CBM, otherwise distinctive for the AA5_2 family, is absent in *Cgr*AlcOx, while *Cgr*RaOx contains a significantly larger PAN_1 domain (PF00024). Though PAN_1 is suggested to be involved in protein-protein or protein-carbohydrate interactions, the N-terminal domain of *Cgr*RaOx is not listed in the CAZyme database in contrast to the CBM32 domains of *Fgr*GaOx and, as later described, *Pru*AA5_2A.

4.3.2 *Cgr*RaOx is a raffinose oxidase

As mentioned, the sequence of *Cgr*RaOx from *C. graminicola* is 29% identical to *Fgr*GaOx and distinguished by the substitution of the N-terminal CBM32 domain for a PAN domain. Amino acid deviations identified particularly in group two (Table 9) predicted that *Cgr*RaOx would display a different substrate selectivity relative to *Fgr*GaOx, though still be active on carbohydrates (i.e. prediction of carbohydrate oxidase versus alcohol oxidase functionality). Indeed, *Cgr*RaOx was the first AA5_2 carbohydrate oxidase not from a *Fusarium* species, to display a selectivity towards the tri-saccharide raffinose, and in addition showing notably poor activity on D-galactose (Figure 20). In fact, the substrate range of *Cgr*RaOx is comparably narrow (Substrate range of *Fgr*GaOx available in Table 6). Where *Cgr*RaOx only oxidised melibiose and D-galactose in addition to raffinose. The specific activity on raffinose (150 mM) was 3.0 U/mg followed by melibiose (1.2 U/mg) and D-galactose (0.3 U/mg), and no activity was found on other mono-saccharides, lactose, galactoglucomannan and xyloglucan, which is in contrast to galactose oxidase where activity on lactose and polysaccharides is typical. This activity pattern would suggest that *Cgr*RaOx is a galacto-oligosaccharide oxidase but surprisingly no activity was found on stachyose (a raffinose based oligosaccharide containing a non-reducing galactosyl terminus (see Figure 21 for structures).

NMR and MS analysis showed that *Cgr*RaOx oxidizes the C6-hydroxyl of the galactosyl residue in raffinose to the corresponding aldehyde product, confirming that it generates the same oxidation product as found for *Fgr*GaOx catalysed oxidations [3]. Activity on non-carbohydrates was also tested were *Cgr*RaOx showed activity on glycerol 1.2 (U/mg) but only poor activity towards primary alcohols. Activity was also found on methylglyoxal (0.3 U/mg), which suggests *Cgr*RaOx is capable of oxidizing aldehyde or the corresponding hydrate form (gem-diol). In fact, oxidation of gem-diols was confirmed ESI-MS analysis after oxidation of raffinose in O^{18}-water where the resulting oxidation product is a uronic acid [3]. Notably, the same mechanism was confirmed for *Fgr*GaOx. The exception to the activity profile was oxidation of a fresh solution of glycolaldehyde dimer (glycolaldehyde) where the activity (2.8 U/mg) was similar to that of raffinose. *Fgr*GaOx did not show activity on glycolaldehyde. For this reason, the enzyme was name raffinose oxidase to distinguish it from the broader substrate range of *Fgr*GaOx. The pH optimum of *Cgr*RaOx was pH 8,0 which is high for a fungal enzyme but in range with other known AA5_2 CAZymes.

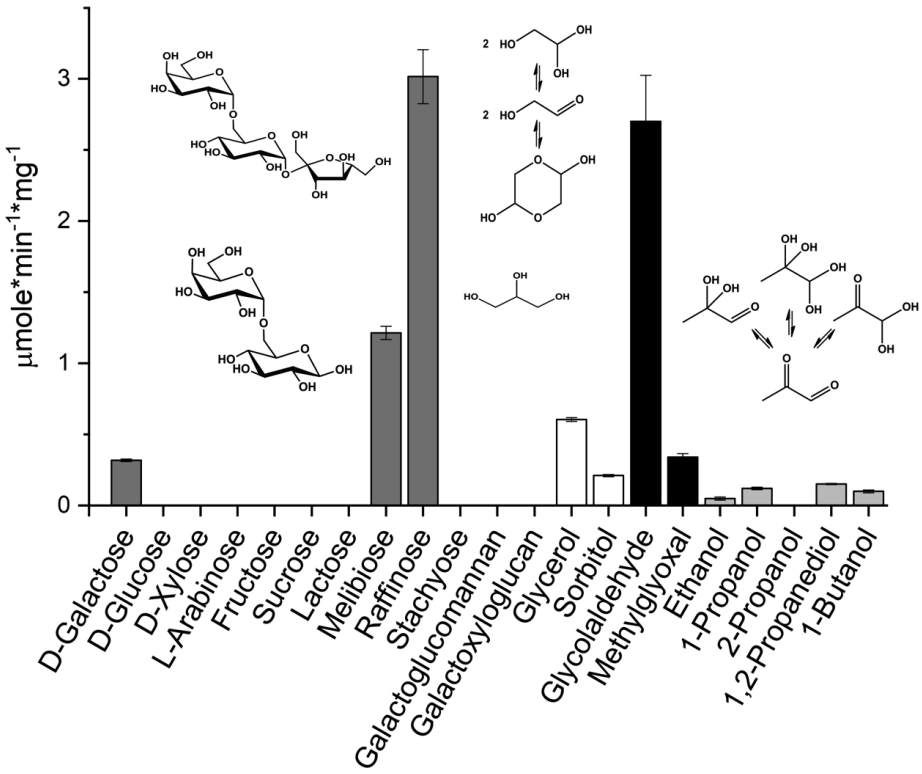


Figure 20. Substrate range of *CgrRaOx*
All substrates were tested at 300 mM with the exception of substrate with limiting solubility: stachyose (75 mM), galactoglucomannan and galactoxyloglucan (2.5 mg/mL). Reactions were performed at pH 7.5 at 30 °C. Error bars indicate standard deviations, n=4. This figure was originally published in [3]

4.3.3 *Pru*AA5_2A shows dual substrate functionality

*Pru*AA5_2A from *P. rubens* is 46 % identical to *Fgr*GaOx and 23% identical to *Cgr*RaOx. Like *Fgr*GaOx, *Pru*AA5_2 includes an N-terminal CBM32 domain. However, as summarized above, *Pru*AA5_2 is distinguished by S291N, Q326D and Q406E substitutions (using *Fgr*Gaox numbering) in the catalytic domain. It was therefore predicted that *Pru*AA5_2A would display activity on D-galactose while also showing activity on non-galactose-containing carbohydrates. Indeed, *Pru*AA5_2A displayed highest activity towards raffinose (specific activity of 30 U/mg), followed by melibiose (17.5 U/mg), D-galactose (14.8 U/mg) and stachyose (9.4 U/mg). Accordingly, *Pru*AA5_2 displayed similar preference for galacto-α-(1→6)-substituted oligosaccharides as *Cgr*RaOx, albeit at 10 times higher specific activity. Also, MS analysis confirmed that *Pru*AA5_2A generates the same C6-galactosyl-aldehyde as *Fgr*GaOx and *Cgr*RacOx [3,4]. Although *Pru*AA5_2A oxidized lactose, the activity was comparably low (2.8 U/mg), and higher activity could be measured against L-arabinose (3.5 U/mg), which shows the importance of the glycosylic bonds in targeted oligosaccharides.

Whereas no significant activity was found on D-glucose and D-xylose, activity was unexpectedly detected on sucrose (Figure 21). Sucrose (α-D-glucopyranosyl-(1→2)-β-D-fructofuranoside) contains three primary hydroxyls; C6-OH on the glucopyranosyl unit, C1-OH and C6-OH on the fructofuranoside unit (see Figure 21 for structure), where especially the terminal C6-OH on fructofuranoside was considered a potential site for oxidation due to the lack of activity on D-glucose. For this reason, the oxidation products were analysed by MS and NMR, where MS identified a minor amount of sucrose containing a single aldehyde, confirming oxidation of sucrose. MS indicated only one oxidation per molecule. In ^{1}H NMR, detection of oxidation products on the fructofuranoside was not possible due to overlapping chemical shifts. Instead the analysis identified a minor amount (5%) of what could correspond to the C6-glycopuranoside-aldehyde. Since MS showed that only a single oxidation of sucrose occurs, it is never the less plausible that *Pru*AA5_2A targets the C6-OH of the glucopyranosyl unit in sucrose. However, much more oxidation product must be generated to confirm this hypothesis. This activity is however interesting since no activity was found on D-glucose; however, the presence of the adjacent fructofuranoside in sucrose could possibly promote activity on glucopyranosyl owing to the inherent selectivity for oligosaccharides of *Pru*AA5_2A. Particularly, oligosaccharides in the raffinose family (RFO's) where sucrose is a precursor.

*Pru*AA5_2A also oxidized a solution of glycolaldehyde dimer with same activity as raffinose. Similar trends as those reported for *Cgr*RaOx were observed, where activity was highest with a fresh solution of glycolaldehyde dimer, diminished after overnight storage of the glycolaldehyde solution, and was completely lost after 48 hours of storage.

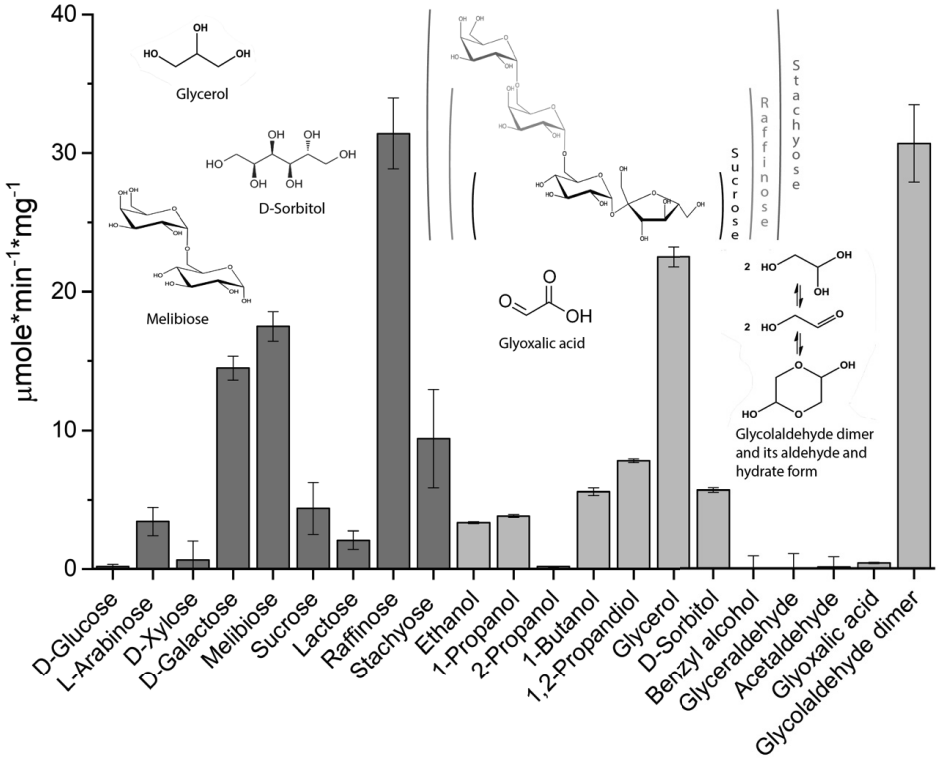


Figure 21. Substrate range of *Pru*AA5_2A
All substrates were tested at 300 mM with the exception of substrate with limiting solubility: stachyose (75 mM), galactoglucomannan and galactoxyloglucan (2.5 mg/mL). Reactions were performed at pH 7.5 at 30 *C. Error bars indicate standard deviations, n=4. This figure was originally published in [4].

4.3.4 Kinetic parameters informed enzyme designations

*Cgr*RaOx and *Pru*AA5_2A showed exceptionally high K_M values for carbohydrates (Table 11). Moreover, *Cgr*RaOx did not display Michaelis-Menten behaviour on D-galactose, and instead showed a strict linear dependency between galactose concentration and activity (Figure 22C). While the K_M of *Pru*AA5_2A on galactose could be calculated (881.5 mM, Table 11), it appears that galactose is unlikely the natural substrate for either *Cgr*RaOx or *Pru*AA5_2A. The catalytic efficiency of *Pru*AA5_2A on raffinose was approximately eight times higher than that of *Cgr*RaOx, which could be explained by the lack of some galactose ligands in *Cgr*RaOx. Consistent with this explanation, both *Fgr*GaOx and *Fsa*GaOx oxidize raffinose, where catalytic efficiencies on raffinose are comparably similar to those measured using galactose (approx. 1.5 times over D-galactose for both).

Whereas *Fgr*GaOx was not active on glycolaldehyde dimer, *Cgr*RaOx showed similar kinetics as raffinose on this substrate. The catalytic efficiency of *Pru*AA5_2 was 6.4 times higher on glycolaldehyde dimer than raffinose (Table 11). Data on the glycolaldehyde dimer solution

showed substrate inhibition at higher concentrations than 120 mM (K_i of 178.5 mM) for PruAA5_2. This is consistent with the fact that oxidation only occurs on fresh solutions of glycolaldehyde dimer, where one of the solubilized derivates of glycolaldehyde dimer potentially inhibits *Pru*AA5_2A or interferes with the HRP/ABTS based activity assay. Although not further investigated, obervations made during the experiments indicate that the absorbance (A_{420nm}) derived from oxidized ABTS decayed significantly faster when solutions of old glycolaldehyde dimer or higher concentrations of fresh solutions were used. Accordingly, it is not possible to draw conclusions on which form of glycolaldehyde that is targeted by *Pru*AA5_2A and *Cgr*RaOx or confirm precise kinetic data on this substrate without further investigation. Yet, kinetic data show a clear selectivity for this substrate solution by PruAA5_2.

Similar to previous findings from other *Fusarium spp.*, *Fgr*GaOx displays higher catalytic efficiency on raffinose than D-galactose. However, when substrate dependencies are plotted, it is evident that the increase of activity is similar for raffinose and D-galactose in *Fgr*GaOx, where K_M for D-galactose is higher because the reaction saturates at nearly the double turnover rates relative to raffinose (Figure 22). In contrast, the increments of turnover rates for raffinose are significantly higher for both *Pru*AA5_2A and *Cgr*RaOx.

Table 11. Kinetic parameters of *Cgr*RaOx, *Pru*AA5_2A and *Fgr*GaOx
Michaelis-Menten and Linear regression kinetics on raffinose, D-galactose and glycolaldehyde dimer

	k_{cat}, S^{-1}	K_M, mM	k_{cat}/K_M, $S^{-1}*mM^{-1}$	*Linear regression* k_{cat}/K_M, $S^{-1}*mM^{-1}$	K_i, mM
***Cgr* RaOx**					
Raffinose	13.8 ± 1.2	481 ± 49	0.0293 ± 0.013	0.0474 ± 0.0033	
D-Galactose	na	na	na	0.00313 ± 0.0082	
Glycolaldehyde dimer	8.81 ± 0.16	184 ± 13	0.0481 ± 0.0042	0.0611 ± 0.013	
***Pru* AA5_2A**					
Raffinose	57.8 ± 1.5	248 ± 13	0.233 ± 0,018		
D-Galactose	75.4 ± 4.2	881 ± 106	0.0864 ± 0.015		
Glycolaldehyde dimer	78.8 ± 21	52.7 ± 24	1.49 ± 1.1		178 ± 82
FgrGaOx					
Raffinose	436 ± 11	16.5 ± 2.4	30.0 ± 5.1		
D-Galactose	795 ± 27	39.8 ± 5.0	19.9 ± 3.2		
Glycolaldehyde dimer	*No activity*				

Errors indicate standard deviations, n = 4.

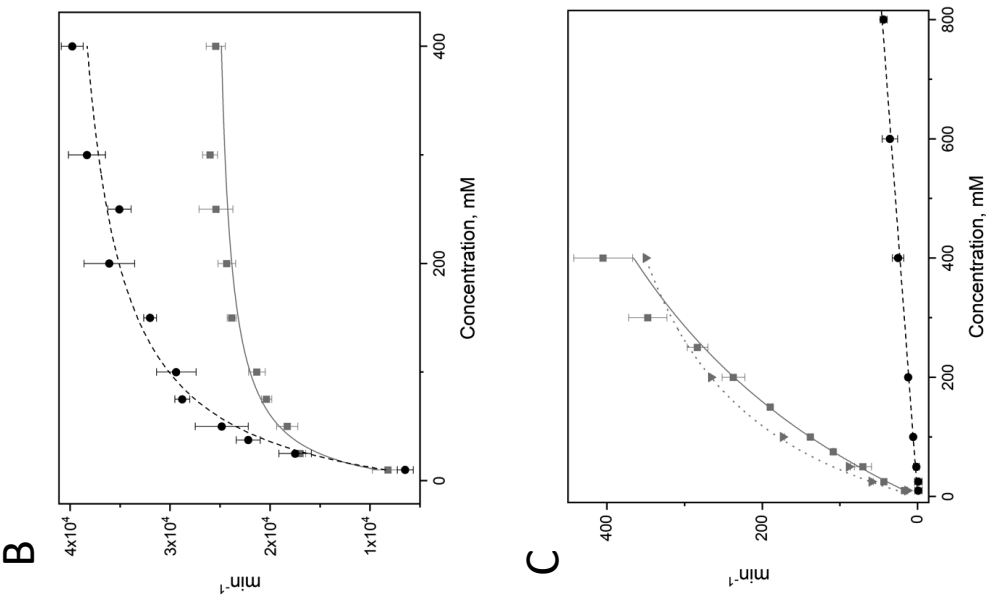


Figure 22. Kinetic plots for A: *Pru*AA5_2A B: *Fgr*GaOx and C: *Cgr*RaOx
Activity dependence of substrate concentration on glycolaldehyde dimer (▶ and dotted line), raffinose (■ solid line) and D–Galactose (● dashed line). Error bars indicate standard deviations, n = 4).

Despite the activity of *Fgr*GaOx and *Fsa*GaOx (Paukner et al., 2015) on raffinose, the high selectivity of *Cgr*RaOx and *Pru*AA5_2A towards raffinose and oxidation of glycolaldehyde dimer raised questions about how to best name these new AA5_2 oxidases. Certainly, both *Cgr*RaOx and *Pru*AA5_2A target the terminal galactosyl unit in raffinose which would argue for naming these enzymes as a galactose oxidase. However, the distant phylogeny and the specific kinetic profiles significantly distinguish these enzymes from *Fgr*GaOx and other known galactose oxidases. Accordingly, when we first published the characterization of the raffinose oxidase from *C. graminicola*, we used raffinose oxidase in order to substantiate the specific

oxidation of this substrate. Later, when characterizing *Pru*AA5_2, we left the functional assignment open, given the comparatively high activity of this enzyme on the glycolaldehyde dimer.

Overall, *Pru*AA5_2A and *Cgr*RaOx differ from *Fgr*GaOx by (1) showing highest catalytic efficiency on a glycolaldehyde dimer solution where *Fgr*GaOx is inactive, (2) their remarkably high K_m values for raffinose and D-galactose, and (3) their lack of activity on galactose containing polysaccharides. In fact, *Fgr*GaOx is the only AA5_2 member reported to have high performance on galactose-containing polysaccharides (section 4.2.3), and since *Cgr*RaOx and *Pru*AA5_2A are inactive on these substrates, it is conceivable that *Fgr*GaOx has evolved to target different substrates than these.

Whether polysaccharides are oxidized by other galactose oxidases from the *Fusarium* genus, or this trait is unique to *Fgr*GaOx, remains unknown since no other *Fusarium spp*. have been tested on polysaccharides so far. R330 in *Fgr*GaOx appears to be particularly critical for activity on galactose and is present in *Cgr*RaOx and *Pru*AA5_2A. However, *Cgr*RaOx deviates from *Fgr*GaOx at the other substrate ligands by containing (1) a tyrosine at the stacking position instead of a tryptophan like *Pru*AA5_2A and *Fgr*GaOx (W290 in *Fgr*GaOx), (2) a tryptophan where *Fgr*GaOx contains Y329, accordingly these two neighbouring positions seems swapped, and (3) a serine instead of Q406, which together explains the exceptionally poor catalytic performance of *Cgr*RaOx on galactose. Despite the lack of the three putative D-galactose ligands, catalytic activity is found on the galactosyl end of raffinose. This activity is highly specific for α(1→6) linkages between galactosyl and glucosyl units. Yet, *Cgr*RaOx is inactive on stachyose and lactose which supports that *Cgr*RaOx interacts with glucosyl and fructosyl units of raffinose to facilitate activity. A similar case can be argued for *Pru*AA5_2A although it only deviates from *Fgr*GaOx at two positions in group 2. In addition, *Pru*AA5_2A contains a tyrosine rather than F194 in *Fgr*GaOx and an asparagine rather than S291, which may be implicated in coordination of the glucosyl unit. Accordingly,she sequence analysis herein and the substrate preference of known AA5_2 oxidases begin to suggest that the R330-Y329 pair, as well as F194-W290 pair, may delineate carbohydrate versus alcohol oxidase functionally within this enzyme subfamily.

4.4 Impact of buffer, pH, and substrate on lag-phase

The characterization of both *Pru*AA5_2 and the M6-CBM3 variant of *Fgr*GaOx uncovered a similar and unexpected attribute. Specifically, both of these enzymes displayed a distinctive lag-phase prior to catalysis (Figure 23 and Figure 24), however, the lag phase was not alleviated by the routine activation step to confirm the enzyme has acquired is fully active state (i.e.,

treatment with $CuSO_4$, $K_3Fe(CN)_6$ or horseradish peroxidase) (see 4.1). Lag-phases and buffer inhibition have not previously been reported for AA5_2 oxidases but impacts of phosphate and Tris buffers on other metal-containing oxidoreductases are known. For example, impact of phosphate and Tris buffers have been observed for iron-lipoxygenases (Schilstra et al., 1994; Wang et al., 1994), copper-tyrosinases (Molina et al., 2007) and copper and zinc dismutases (Winterbourn et al., 2002).

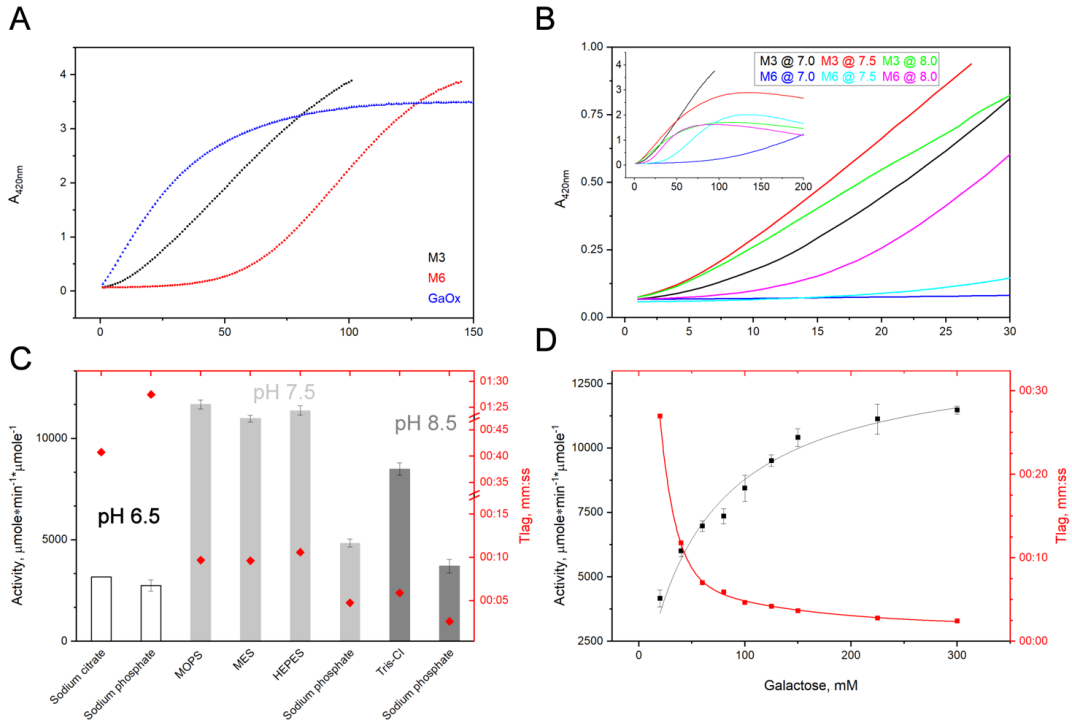


Figure 23. Lag-phase properties in M6-CBM3
The lag-phase (Tlag) is defined as the time from initiation of the reaction (T = 00:00) to where the maximum slope crosses the x-axis. All reactions were with 300 mM D-galactose. (A) Oxidation reactions of M3-CBM3 (M3), M6-CBM3 (M6) and *Fgr*GaOx (GaOx) loaded with equal enzyme units (50 mM sodium phosphate, pH 7.5). (B) Impact of pH on reaction rate and Tlag of M3 and M6, where the pH is established using 25 mM sodium phosphate buffer. (C) Impact of buffer and pH reaction rate and Tlag on M6. (D) Impact of substrate concentration on reaction rate and Tlag, where D-galactose was prepared in 20 mM MOPS (pH 7.5). n=4 error bars indicate standard deviation.

Lag-phase (T_{lag}) was observed in all tested conditions for *Pru*AA5_2A and M6-CBM3, but also the activity was affected by pH and buffer type (Figure 23 for M6-CBM3 and Figure 24 for *Pru*AA5_2A) where the impact was highest for citrate buffer, followed by sodium phosphate, Tris-Cl, and MOPS buffers (exemplified in Figure 23C for M6-CBM3). For *Pru*AA5_2A the impact of phosphate buffer (pH 7.5) was shown by the increase of T_{lag} with increasing buffer concentration (Figure 24A). Also, T_{lag} was reduced by over 90% when shifting from pH 6.0 to pH 8.0 (Figure 24B). M6-CBM3 also showed dependency of increased pH relative to M3-CBM3 (Figure 23B). Notably, for *Pru*AA5_2A T_{lag} was shortest for most preferred substrates (Figure

24C) and decreased with increasing substrate concentrations (Figure 24D), which was also observed for M6-CBM3 (Figure 23D).

Whereas *Pru*AA5_2A is the first wild-type AA5_2 member reported to display this phenotype, this behaviour was also observed for the M3 variant of *Fgr*GaOx (Moon et al., 2012), albeit to a significantly lesser extent compared to M6-CBM3 (Figure 23A) and *Pru*AA5_2A (Figure 24A). In this study the lag-phase was attributed to the W290F substitution, where substrate-induced alteration of the enzyme structure leads to a more active form. The substrate dependence of the lag phase observed for M6-CBM3 and *Pru*AA5_2A is consistent with this model, but suggests it is not only attributed to the W290F substitution, given that *Pru*AA5_2A contains a tryptophan at the corresponding position.

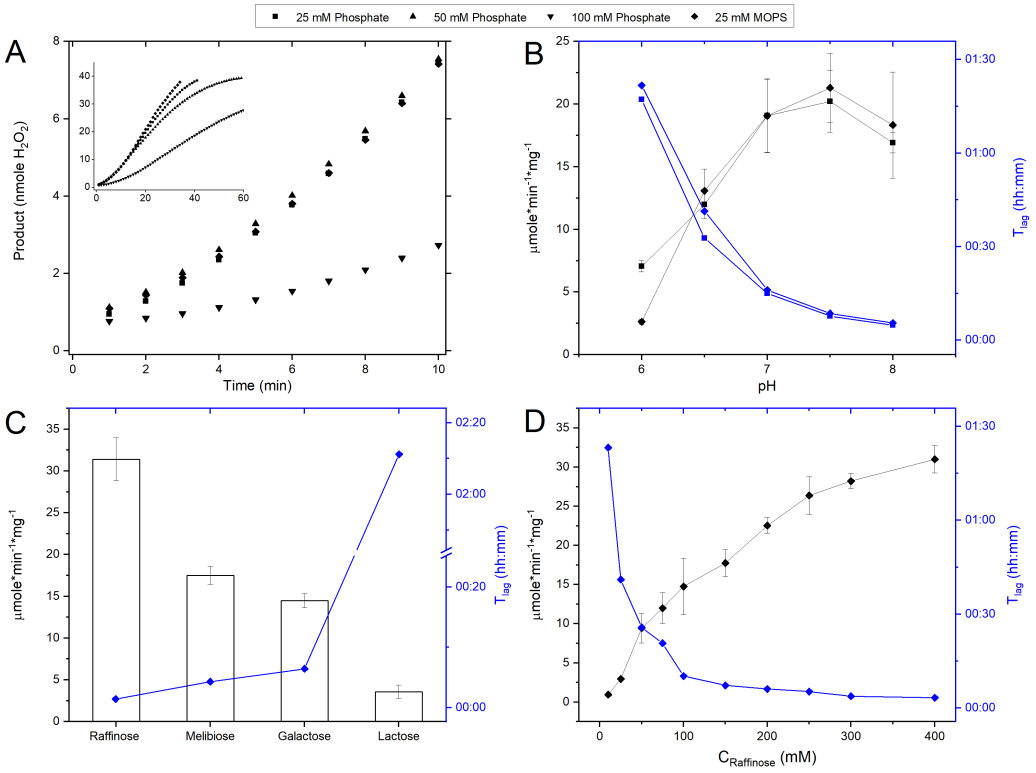


Figure 24. Lag-Phase properties in *Pru*AA5_2
The lag-phase (Tlag) is defined as the time from initiation of the reaction (T = 00:00) to where the maximum slope crosses the x-axis. (A) Impact of buffer type and concentration on rate of product formation (pH 7.5). (B) Impact of pH on reaction rate and Tlag, where the pH is established using 25 mM sodium phosphate buffer and 25 mM MOPS. (C) Impact of substrate on reaction rate and Tlag, where each substrate was prepared to 300 mM in 20 mM MOPS (pH 7.5). (D) Impact of substrate concentration on reaction rate and Tlag, where raffinose was prepared in 20 mM MOPS (pH 7.5). n=4 error bars indicate standard deviation. This figure was originally published in [4]

Common to both M6 and *Pru*AA5_2A is that they contain a serine at the position of C383 in *Fgr*GaOx. In *Fgr*GaOx, C383 is positioned just beneath Y405 and between H334 and F464 which are all highly conserved residues in subfamily AA5_2 (Table 10). Furthermore, Y405 is positioned directly below the main substrate ligand, R330, and forms hydrogen bonds to H334 and the catalytic Y495 (Figure 25). Changes in the environment beneath Y405 can thus have indirect effects on Y495 orientation. An altered orientation of Y495 may well affect its ability to acquire a proton from the substrate hydroxyl causing failure of catalysis (see Figure 5). Although there are no reports of other C383S variants of *Fgr*GaOx displaying a catalytic lag-phase, and there were no observable changes in the crystal structure of the C383S variant relative to the wild-type *Fgr*GaOx (Wilkinson et al., 2004), the additional differences in M6 and *Pru*AA5_2 relative to *Fgr*GaOx could impart this catalytic behaviour. For example, the presence of the hydroxyl group instead of a thiol (C383S in M6-CBM3 or S379 in *Pru*AA5_2A) could have a relay effect through Y405 on the orientation of the basic tyrosine copper-ligand (Y495 in *Fgr*GaOx), causing a suboptimal orientation that is corrected upon substrate binding. It must be assumed then, that the resting forms of M6 and *Pru*AA5_2A have low substrate affinity, leading to the lag phase.

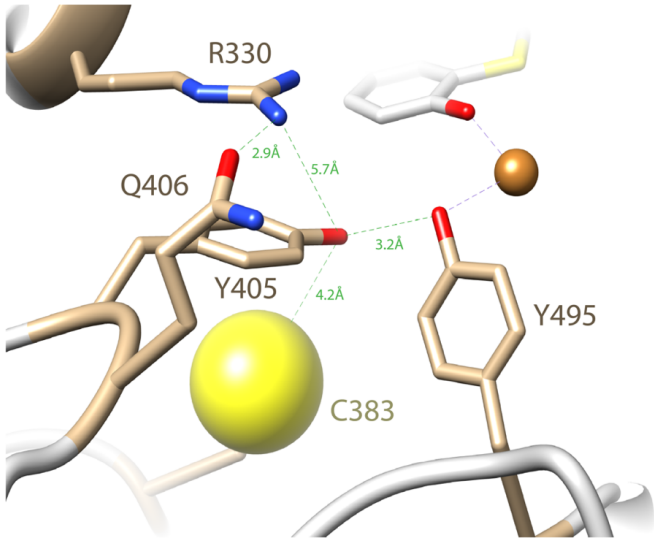


Figure 25. Proximity of C383, Y405 to Y495 and R330
C393 is depicted as a sphere to illustrate the size of the sulphur atom. A corresponding oxygen would is smaller and more hydrophilic.

Based on this interpretation of the lag phase, M6 was tested on glucose after a preincubation with galactose. However, gain in glucose activity was not observed (results not shown).
Certainly, the C383S substitution, through Y405, could also affect the orientation of the R330K and Q406T substitutions in M6-CBM3, leading to the observed loss of activity on glucose and

gain in activity on galactose. The low of activity in M3-CBM3 and re-gained activity achieved by addition of C383S, V494A and Y436H in M6 supports that the W290F substitution does not significant impact the redox function of the Cys-Tyr cofactor. Rather, the substitution to phenylalanine impaired carbohydrate oxidation by disturbing a hyperfine coordination of the galactose primary hydroxyl group to the copper and tyrosinate ligands. A coordination loss that seems possible to regain through additional mutagenesis, such as the example of M6-CBM3.

Similar to the trends seen for M6-CBM6 and *Pru*AA5_2A (Figure 23 and Figure 24), all AA5_2 members characterized to date have reported pH optimum from 7.0 to 8.5 which is surprisingly high. Due to the functionality as a base, Y495 may be prone to unintended protonation in the acidic pH range which would prevent its catalytic function (see Figure 5).

5. Conclusions and future aspects

Main research conclusions

The aims of this study were 1: to investigate the potential of enhancing the performance of *Fgr*GaOx and M3 on polysaccharides through fusion with *Pe*CBM29-1-2 and *Ct*CBM3 respectively. 2: to increase the catalytic activity of M3-CBM3 by generating M6-CBM3 containing amino acid substitutions known to increase the activity of wild-type *Fgr*GaOx. 3: to discover and analyze new carbohydrate oxidases from CAZyme family AA5_2 aiming at broadening the diversity of the family and find novel candidates for chemo-enzymatic carbohydrate engineering.

Fusion of *Ct*CBM29-1-2 to *Fgr*GaOx did not appear to impact enzyme maturation or activity on D-galactose, but affected expression yields, which was further affected by the position of *Pe*CBM29-1-2. Clearly, N-terminal fusion impacted expression and also functionality of the involved carbohydrate binding modules. C-terminal fusion lead to greatest binding affinity and thereby highest gains in catalytic efficiency on polysaccharides. Together with the QCM-D studies, the inferior performance of CBM29-GaOx could be explained by potential antagonism between neighboring CBM29-2 and CBM32 domains, or disruption of cooperative function between CBM29-1 and CBM29-2. Gains of catalytic activity at low substrate concentrations did not benefit oxidation of polysaccharides at higher concentrations. At high substrate concentrations, slow enzyme detachment became limiting to the oxidation rate, similar to reports from fusing cellulose binding domains to glycoside hydrolases. Neither did fusion of *Pe*CBM29-1-2 increase performance of immobilized galactoxyloglucan, suggesting that the *Fgr*GaOx is already optimized for oxidation of polysaccharides.

The fusion of *Ct*CBM3 to the M3 did not significantly affect the catalytic properties on mono- and oligosaccharides, similar to that of observed for GaOx-CBM29. As previously reported for M3 and other glucose active variants, amino acids substitutions broadening the substrate range of *Fgr*GaOx leads to significant loss of activity and affinity to galactose. C-terminal fusion of *Ct*CBM3 provided convenient immobilization on cellulose, while the activity on soluble carbohydrates was maintained. M3-CBM3 also displayed increased oxidation over *Fgr*GaOx on cellulose filter paper after immobilization. The reactions where compared on a mass dosage,

suggesting that M3-CBM3 is more efficient in oxidation of solid polysaccharides when accounting the of molecular weight of M3-CBM3 and its reduced activity levels. Efforts to increase to activity of M3-CBM3 by generating M6-CBM3, gained wild-type level catalytic efficiency on D-galactose entailed a complete loss of activity on other carbohydrates, suggesting that V494A and particularly C383S have unexpected impacts on substrate coordination. These mutations having comparably higher effects on M3-CBM3 than *Fgr*GaOx suggest that the loss of activity on D-galactose is caused by reduced substrate affinity, which can be regained, rather than implicating CRO-activity per se.

Combinatorial sequence analysis, a detailed review of literature describing *Fgr*GaOx variants and the characterization of new AA5_2 members, revealed important amino acid positions and stretches of consensus sequences in the AA5_2 family that can be used to predict substrate preference of homologous copper-radical oxidases. These analyses further showed the distinct division of the active pocket in a region of conserved residues important for copper-radical catalysis and a more variable region important for the coordination of the D-galactose. The dual activity on glycolaldehyde dimer and raffinose by *Cgr*RaOx and *Pru*AA5_2A builds a bridge between subfamily one and two in AA5 as some AA5_1 members display activity on glycolaldehyde dimer (Wymelenberg et al., 2006). Interestingly, both *Cgr*RaOx and *Pru*AA5_2A show remarkably poor affinity to both raffinose and D-galactose; however, the comparatively higher performance of *Pru*AA5_2A on these substrates coincides with the type of N-terminal domain and presence of residues that coordinate D-galactose (W290, R330, Q406 in FgrGaOx) or play an important role in this region (Q326, Y329 and P463). Taken together with recent reports of alcohol oxidases within family AA5_2, this pattern suggests that incorporation of the N-terminal CBM32 could have evolved in association with activity on carbohydrates (i.e. substitutions of amino acids listed above).

Considering the thesis conclusions summarized above, the main contributions of my thesis can be listed as follows:

1. C-terminal protein fusion to *Fgr*GaOx ensures a functional fusion construct, whereas N-terminal fusion may have detrimental effects on functionality.
2. Immobilization of functional galactose oxidase can be facilitated through fusion of carbohydrate domains.
3. Amino acids substitutions broadening substrate range does not implicate CRO- functionality, suggesting a potential to optimize variants.
4. Identification of consensus sequence stretches and relevant amino acid positions in the AA5_2 family that help identify CAZyme AA5_2 copper-radical oxidases.
5. Use of point four in the discovery of two new AA5_2 oxidases that connect the functional space between AA5_1 and AA5_2.

5.1 Future research aspects

Raffinose is a plant defence molecule

The biological role of *Fgr*GaOx has often been explained by its production of hydrogen peroxide. However this argument seems insufficient to explain the specificity towards D-galactose and particularly raffinose which seems the preferred substrate within many AA5_2 oxidases. Specific generation of the galactosyl-aldehyde in raffinose must be of significance to substantiate their selectivity. Synthesis of stachyose, verbascose and longer raffinose family oligosaccharides has been connected to biotic stress response in plants (Ende, 2013). These oligosaccharides, which are formed from raffinose by stachyose synthase (Nishizawa et al., 2008), have reactive oxygen species scavenging properties and thus function as stabilizers in the plant cell wall during stress responses. Oxidation of raffinose implies a role in inhibiting oxidative stress response pathways in plants. Specifically, since raffinose is an intermediate in this pathway and a substrate for stachyose synthase, oxidation of raffinose could diminish biotic stress response during colonization by plant pathogenic fungi like *C. graminicola* and *F. graminearum*.

Glycolaldehyde active oxidases are implicated filamentous growth

The activity of *Pru*AA5_2A and *Cgr*RaOx on glycolaldehyde dimer solution recalls the discovery of the same activity in a newly discovered copper-radical oxidase, GlxA, from *Streptomyces lividans* (Chaplin et al., 2015). As a representative archetype, *Fgr*GAOx galactose oxidase was compared to GlxA revealing that these enzymes only share similar structural and biochemical traits (CRO-functionality). The kinetic behaviour of GlxA is, however, more reminiscent of *Cgr*RaOx where both enzymes show a linear relation between activity and increasing concentrations of D-galactose but approximate Michealis-Menten kinetics on glycolaldehyde. Genetic deletions suggested that GlxA plays a role in glycan deposition during aerial hyphae development under osmotic growth by this bacterium. Similar functionality was suggested for galactose oxidase-like genes in other actinobacteria (Liman et al., 2013; Silakowski et al., 1998). Glycolaldehyde is also a known substrate for glyoxal oxidases in CAZyme family AA5_1, which generate hydrogen peroxide utilized by ligninolytic enzymes (Kersten and Cullen, 2014). Glyoxal oxidase have similarly been suggested to play a role in filamentous growth of phytopathogentic fungus *Ustilago maydis* (Leuthner et al., 2005). However, glyoxal oxidase from subfamily AA5_1 are not reported to act on alcohols.

Further studies to address the biological function of AA5_2 oxidases

Studies that investigate inhibition of the stachyose synthase by the raffinose-aldehyde generated by AA5_2 oxidases, or generate and test *Δgao*-strains of *F. graminearum*, *P. rubens* or *C. graminicola* for impaired osmotic stress tolerance or morphogenesis, would directly test the

biological functions proposed above. Although genetic engineering of fungal organisms has previously been challenging, new CRISPR gene editing technologies might facilitate gene editing in fungi allowing more detailed studies on these microorganisms. Notably, CAZyme family AA5_2 contains genes from the proteobacteria *Burkholderia genus* where *B. pseudomallei* is also known to infect plant hosts (Lee et al., 2010). Since bacterial members of family AA5_2 have not yet been described, the expression and characterization of these AA5_2 members would expand the knowledge of copper-radical oxidase activity and likely help illuminate their biological functions. Related to this aspect, evolutionary comparison between the actinobacteria *glxA* and fungal *gao* genes as well as subfamily AA5_1 versus AA5_2 could elucidate whether the microbial copper-radical oxidase originally evolved with affinity for galactose and raffinose or this function has evolved specifically in plant pathogenic fungi from other copper-radical oxidases. Intriguingly, galactose-ligands in *Fgr*GaOx are of higher variability in family AA5_2 and seems specifically placed so they do not impact the position and function of more conserved amino acids commonly important for CRO-functionality.

Certainly, the determination of the crystal structure of *Pru*AA5_2A and particularly *Cgr*RaOx would shed new light on structure-function relationships within AA5_2. In fact, crystallization was discussed for *Cgr*RaOx and attempted for *Pru*AA5_2A, but due to challenges relating to removal of glycans from *Pru*AA5_2A, efforts to date have not been successful.

Overcoming the Lag-phase

Hypothesizing that the lag phase in M6-CBM3 and *Pru*AA5_2 is caused by positional shifts of the Y495 sidechain (stated in 4.4), a change in the copper-coordination sphere of *Fgr*GaOx variants would be detectable by Electron Paramagnetic Resonance (EPR). EPR spectra of *Fgr*GaOx, M3-CBM3 and M6-CBM3 would reveal gradual changes to the copper-coordination chemistry, which could be correlated to the severity of the lag-phase. Also, time-resolved optical spectroscopy has been used to measure the radical decay rate in *Fgr*GaOx (Rogers et al., 2007), indicating that the W290F substitution does not lead to significant de-stabilization (see 1.2.1). The authors suggest that a radical-decay pathway through the copper-ion via Y495 is abolished when the Y495-Cu interaction is broken, for example by protonation of Y495 in acidic pH. Similarly, if the Y495-Cu bond is distorted, M6-CBM3 could display higher radical stability compared to M3-CBM3. To test this hypothesis, S383C in M6 and S379C in *Pru*AA5_2A substitutions could be introduced to alleviate lag phase, and impacts could then be tested by same methods as describe above.

Engineering of new competent glucose oxidase variants

The activity gain from M3-CBM3 to M6-CBM3 shows that is possible to regain affinity towards carbohydrates in R330 and W290 variants of *Fgr*GaOx. Further attempts to engineer a *Fgr*GaOx variant with good activity towards other carbohydrates should try other combinations of amino acids substitutions from group two and three in Table 9. Also, the inherent activity on sucrose by *Pru*AA5_2A makes it a novel candidate for engineering and optimization.

References

Alberton, D., De Oliveira, L.S., Peralta, R.M., Barbosa-Tessmann, I.P., 2007. Production, purification, and characterization of a novel galactose oxidase from *Fusarium acuminatum*. J. Basic Microbiol. 47, 203–212. https://doi.org/10.1002/jobm.200610290

Amaral, D., Bernstein, L., Morse, D., Horecker, B.L., 1963. Galactose oxidase of *Polyporus circinatus*: a copper enzyme. J. Biol. Chem. 238, 2281–4.

Anasontzis, G.E., Salazar Penã, M., Spadiut, O., Brumer, H., Olsson, L., 2014. Effects of temperature and glycerol and methanol-feeding profiles on the production of recombinant galactose oxidase in *Pichia pastoris*. Biotechnol. Prog. 30, 728–35. https://doi.org/10.1002/btpr.1878

Aparecido Cordeiro, F., Bertechini Faria, C., Parra Barbosa-Tessmann, I., 2010. Identification of new galactose oxidase genes in *Fusarium spp*. J. Basic Microbiol. 50, 527–537. https://doi.org/10.1002/jobm.201000078

Arndt, T., 2003. Asialotransferrin--an alternative to carbohydrate-deficient transferrin? Clin. Chem. 49, 1022–3. https://doi.org/10.1373/49.6.1022

Arnold, F., Ioanna, P., Sun, L., 2000. Directed evolution of galactose oxidase enzymes.

Baird, J.K., Smith, W.W., 1989. A simple colorimetric method for the specific analysis of food-grade galactomannans. Food Hydrocoll. 3, 413–416. https://doi.org/10.1016/S0268-005X(89)80016-5

Baron, A.J., Stevens, C., Wilmot, C., Seneviratne, K.D., Blakeley, V., Dooley, D.M., Phillips, S.E. V, Knowles, P.F., McPherson, M.J., 1994. Structure and mechanism of galactose oxidase: The free radical site. J. Biol. Chem. 269, 25095–25105.

Behrens, C., Garibay, P.W., Zundel, M., Klausen, N.K., Bjorn, S.E., 2005. Use of galactose oxidase for selective chemical conjugation of protractor molecules to glycoproteins of therapeutic or diagnostic interest. PCT Int. Appl.

Birmingham, W., Turner, N.J., 2018. A single enzyme oxidative 'cascade' via a dual functional galactose oxidase. ACS Catal. acscatal.8b00043. https://doi.org/10.1021/acscatal.8b00043

Boraston, A.B., Bolam, D.N., Gilbert, H.J., Davies, G.J., 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem. J. 382, 769–781. https://doi.org/10.1042/BJ20040892

Bragd, P.L., van Bekkum, H., Besemer, A.C., 2004. TEMPO-mediated oxidation of polysaccharides: survey of methods and applications. Top. Catal. 27, 49–66. https://doi.org/10.1023/B:TOCA.0000013540.69309.46

Campia, P., Ponzini, E., Rossi, B., Farris, S., Silvetti, T., Merlini, L., Brasca, M., Grandori, R., Galante, Y.M., 2017. "Aerogels of enzymatically oxidized galactomannans from leguminous plants: Versatile delivery systems of antimicrobial peptides and enzymes." Carbohydr. Polym. 158, 102–111. https://doi.org/10.1016/j.carbpol.2016.11.089

Chaplin, A.K., Petrus, M.L.C., Mangiameli, G., Hough, M.A., Svistunenko, D.A., Nicholls, P., Claessen, D., Vijgenboom, E., Worrall, J.A.R., 2015. GlxA is a new structural member of the radical copper oxidase family and is required for glycan deposition at hyphal tips and morphogenesis of *Streptomyces lividans*. Biochem. J. 469, 433–444. https://doi.org/10.1042/BJ20150190

Charmantray, F., Touisni, N., Hecquet, L., Mousty, C., 2013. Amperometric Biosensor Based on Galactose Oxidase Immobilized in Clay Matrix. Electroanalysis 25, 630–635. https://doi.org/10.1002/elan.201200274

Charnock, S.J., Bolam, D.N., Nurizzo, D., Szabó, L., McKie, V.A., Gilbert, H.J., Davies, G.J., 2002. Promiscuity in ligand-binding: The three-dimensional structure of a *Piromyces* carbohydrate-binding module, CBM29-2, in complex with cello- and mannohexaose. Proc. Natl. Acad. Sci. U. S. A. 99, 14077–82. https://doi.org/10.1073/pnas.212516199

Chen, C., Chen, J., Geng, Z., Wang, M., Liu, N., Li, D., 2018. Regioselectivity of oxidation by a polysaccharide monooxygenase from *Chaetomium thermophilum*. Biotechnol. Biofuels 11, 155. https://doi.org/10.1186/s13068-018-1156-2

Chhabra, S.R., Shockley, K.R., Conners, S.B., Scott, K.L., Wolfinger, R.D., Kelly, R.M., 2003. Carbohydrate-induced differential gene expression patterns in the hyperthermophilic bacterium *Thermotoga maritima*. J. Biol. Chem. 278, 7540–52. https://doi.org/10.1074/jbc.M211748200

Cooper, J.A.D., Smith, W., Bacila, M., Medina, H., 1959. Galactose oxidase from *Polyporus circinatus*, FL. J. Biol. Chem. 445–448.
Cowley, R.E., Cirera, J., Qayyum, M.F., Rokhsana, D., Hedman, B., Hodgson, K.O., Dooley, D.M., Solomon, E.I., 2016. Structure of the reduced copper active site in preprocessed galactose oxidase: ligand tuning for one-electron O_2 activation in cofactor Biogenesis. J. Am. Chem. Soc. 138, 13219–13229. https://doi.org/10.1021/jacs.6b05792
Deacon, S.E., Mahmoud, K., Spooner, R.K., Firbank, S.J., Knowles, P.F., Phillips, S.E. V, McPherson, M.J., 2004. Enhanced fructose oxidase activity in a galactose oxidase variant. ChemBioChem 5, 972–979. https://doi.org/10.1002/cbic.200300810
Deacon, S.E., McPherson, M.J., 2011. Enhanced expression and purification of fungal galactose oxidase in *Escherichia coli* and use for analysis of a saturation mutagenesis library. ChemBioChem 12, 593–601. https://doi.org/10.1002/cbic.201000634
Delgrave, S., Murphy, D.J., Pruss, J.L., Maffia, A.M., Marrs, B.L., Bylina, E.J., Coleman, W.J., Grek, C.L., Dilworth, M.R., Yang, M.M., Youvan, D.C., 2001. Application of a very high-throughput digital imaging screen to evolve the enzyme galactose oxidase. Protein Eng. 14, 261–267.
Díaz-Rodríguez, A., Martínez-Montero, L., Lavandera, I., Gotor, V., Gotor-Fernández, V., 2014. Laccase/2,2,6,6-Tetramethylpiperidinoxyl Radical (TEMPO): An efficient catalytic system for selective oxidations of primary hydroxy and amino groups in aqueous and biphasic media. Adv. Synth. Catal. 356, 2321–2329. https://doi.org/10.1002/adsc.201400260
Ende, W. Van den, 2013. Multifunctional fructans and raffinose family oligosaccharides. Front. Plant Sci. 4, 247. https://doi.org/10.3389/fpls.2013.00247
Escalettes, F., Turner, N.J., 2008. Directed evolution of galactose oxidase: generation of enantioselective secondary alcohol oxidases. ChemBioChem 9, 857–860. https://doi.org/10.1002/cbic.200700689
Evik, E., Şenel, M., Fatih Abasyank, M., 2010. Construction of biosensor for determination of galactose with galactose oxidase immobilized on polymeric mediator contains ferrocene. Curr. Appl. Phys. 10, 1313–1316. https://doi.org/10.1016/j.cap.2010.03.014
Ficko-Bean, E., Boraston, A.B., 2018. Carbohydrate Binding Module Family 32 in CAZypedia [WWW Document]. URL: http://www.cazypedia.org/index.php/Carbohydrate_Binding_Module_Family_32 (accessed 4.16.18).
Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., Bateman, A., 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 44, D279–D285. https://doi.org/10.1093/nar/gkv1344
Firbank, S., Rogers, M., 2001. Crystal structure of the precursor of galactose oxidase: An unusual self-processing enzyme. PNAS. 98 (23). https://doi.org/10.1073/pnas.2314637998
Flint, J., Bolam, D.N., Nurizzo, D., Taylor, E.J., Williamson, M.P., Walters, C., Davies, G.J., Gilbert, H.J., 2005. Probing the mechanism of ligand recognition in family 29 carbohydrate-binding modules. J. Biol. Chem. 280, 23718–23726. https://doi.org/10.1074/jbc.M501551200
Flint, J., Nurizzo, D., Harding, S.E., Longman, E., Davies, G.J., Gilbert, H.J., Bolam, D.N., 2004. Ligand-mediated dimerization of a carbohydrate-binding module reveals a novel mechanism for protein–carbohydrate Recognition. J. Mol. Biol. 337, 417–426. https://doi.org/10.1016/j.jmb.2003.12.081
Foumani, M., 2011. Altered substrate specificity of the gluco-oligosaccharide oxidase from *Acremonium strictum*. Biotechnol. Bioeng. 108, 2261–2269. https://doi.org/10.1002/bit.23149
Freelove, A.C., Bolam, D.N., White, P., Hazlewood, G.P., Gilbert, H.J., 2001. A novel carbohydrate-binding protein is a component of the plant cell wall-degrading complex of *Piromyces equi*. J. Biol. Chem. 276, 43010–7. https://doi.org/10.1074/jbc.M107143200
Galante, Y.M., Merlini, L., Silvetti, T., Campia, P., Rossi, B., Viani, F., Brasca, M., 2018. Enzyme oxidation of plant galactomannans yielding biomaterials with novel properties and applications, including as delivery systems. Appl. Microbiol. Biotechnol. https://doi.org/10.1007/s00253-018-9028-z
Ghafar, A., Gurikov, P., Subrahmanyam, R., Parikka, K., Tenkanen, M., Smirnova, I., Mikkonen, K.S., 2017. Mesoporous guar galactomannan based biocomposite aerogels through enzymatic crosslinking. Compos. Part A Appl. Sci. Manuf. 94, 93–103. https://doi.org/10.1016/j.compositesa.2016.12.013
Ghafar, A., Parikka, K., Sontag-Strohm, T., Österberg, M., Tenkanen, M., Mikkonen, K.S., 2015. Strengthening effect of nanofibrillated cellulose is dependent on enzymatically oxidized polysaccharide gel matrices. Eur. Polym. J. 71, 171–184. https://doi.org/10.1016/j.eurpolymj.2015.07.046
Goswami, P., Chinnadayyala, S.S.R., Chakraborty, M., Kumar, A.K., Kakoti, A., 2013. An overview on alcohol oxidases and their potential applications. Appl. Microbiol. Biotechnol.

https://doi.org/10.1007/s00253-013-4842-9
Hallahan, T., Gilbert, C., 1993. Factor IX- polymeric conjugates. US patent 5621039A.
Hammel, K.E., Mozuch, M.D., Jensen, K.A., Kersten, P.J., 1994. H_2O_2 Recycling during Oxidation of the Arylglycerol β-Aryl Ether Lignin Structure by Lignin Peroxidase and Glyoxal Oxidase. Biochemistry 33, 13349–13354. https://doi.org/10.1021/bi00249a022
Hartmans, S., de Vries, H.T., Beijer, P., Brady, R.L., Hofbauer, M., Haandrikman, A.J., 2004. Production of oxidized guar galactomannan and its applications in the paper industry. ACS Symp. Ser. 864, 360–371. https://doi.org/10.1021/bk-2004-0864.ch023
Herter, S., McKenna, S.M., Frazer, A.R., Leimkühler, S., Carnell, A.J., Turner, N.J., 2015. Galactose oxidase variants for the oxidation of amino alcohols in enzyme cascade synthesis. ChemCatChem 7, 2313–2317. https://doi.org/10.1002/cctc.201500218
Isobe, K., Kataoka, M., Ogawa, J., Hasegawa, J., Shimizu, S., 2012. Microbial oxidases catalyzing conversion of glycolaldehyde into glyoxal. N. Biotechnol. 29, 177–182. https://doi.org/10.1016/j.nbt.2011.05.001
Ito, N., Phillips, S.E.V., Yadav, K.D.S., Knowles, P.F., 1994. Crystal structure of a free radical enzyme, galactose oxidase. J. Mol. Biol. 238, 704–814. https://doi.org/10.1006/jmbi.1994.1335
Ito, N., Phillips, S.E. V., Stevens, C., Ogel, Z.B., McPherson, M.J., Keen, J.N., Yadav, K.D.S., Knowles, P.F., 1991. Novel thioether bond revealed by a 1.7 Å crystal structure of galactose oxidase. Nature 350, 87–90. https://doi.org/10.1038/350087a0
Jeffrey, G.A., 1995. Hydrogen-Bonding: An Update. Crystallogr. Rev. 4, 213–254. https://doi.org/10.1080/08893119508039923
Johansen, K.S., 2016. Lytic Polysaccharide Monooxygenases: the microbial power tool for lignocellulose degradation. Trends Plant Sci. 21, 926–936. https://doi.org/10.1016/j.tplants.2016.07.012
Johansen, K.S., 2016. Discovery and industrial applications of lytic polysaccharide mono-oxygenases. Biochem. Soc. Trans. 44. https://doi.org/10.1042/BST20150204
Johnson, J.M., Halsall, H.B., Heineman, W.R., 1985. Redox activation of galactose oxidase: thin-layer electrochemical study. Biochemistry 24, 1579–1585. https://doi.org/10.1021/bi00328a001
Kanyong, P., Krampa, F.D., Aniweh, Y., Awandare, G.A., 2017. Enzyme-based amperometric galactose biosensors: a review. Microchim. Acta. https://doi.org/10.1007/s00604-017-2465-z
Kempner, E.S., Whittaker, J.W., Miller, J.H., 2010. Radiation inactivation of galactose oxidase, a monomeric enzyme with a stable free radical. Protein Sci. 19, 236–241. https://doi.org/10.1002/pro.302
Kersten, P., Cullen, D., 2014. Copper radical oxidases and related extracellular oxidoreductases of wood-decay *Agaricomycetes*. Fungal Genet. Biol. 72, 124–130. https://doi.org/10.1016/j.fgb.2014.05.011
Kersten, P.J., Kirk, T.K., 1987. Involvement of a new enzyme, glyoxal oxidase, in extracellular H_2O_2 production by *Phanerochaete chrysosporium*. J. Bacteriol. 169, 2195–2201. https://doi.org/10.1128/JB.169.5.2195-2201.1987
Kracher, D., Scheiblbrandner, S., Felice, A.K.G., Breslmayr, E., Preims, M., Ludwicka, K., Haltrich, D., Eijsink, V.G.H., Ludwig, R., 2016. Extracellular electron transfer systems fuel cellulose oxidative degradation. Science. https://doi.org/10.1126/science.aaf3165
Kupper, C.E., Rosencrantz, R.R., Henßen, B., Pelantová, H., Thönes, S., Drozdová, A., Křen, V., Elling, L., 2012. Chemo-enzymatic modification of poly-N-acetyllactosamine (LacNAc) oligomers and N,N-diacetyllactosamine (LacDiNAc) based on galactose oxidase treatment. Beilstein J. Org. Chem. 8, 712–25. https://doi.org/10.3762/bjoc.8.80
Kwiatkowski, L.D., Kosman, D.J., 1973. On the role of superoxide radical in the mechanism of action of galactose oxidase. Biochem. Biophys. Res. Commun. 53, 715–721. https://doi.org/10.1016/0006-291X(73)90152-6
Lee, Y.-K., Whittaker, M.M., Whittaker, J.W., 2008. The electronic structure of the Cys-Tyr free radical in galactose oxidase determined by EPR spectroscopy. Biochemistry 47, 6637–6649. https://doi.org/10.1021/bi800305d
Lee, Y.H., Chen, Y., Ouyang, X., Gan, Y.-H., 2010. Identification of tomato plant as a novel host model for Burkholderia pseudomallei. BMC Microbiol. 10, 28. https://doi.org/10.1186/1471-2180-10-28
Leppänen, A.-S., 2013. Regioselective modification of galactose-containing polysaccharides in aqueous media. PhD. thesis. Painosalama Oy - Turku, Finland. ISBN 978-952-12-2942-8
Leppänen, A.-S., Niittymäki, O., Parikka, K., Tenkanen, M., Eklund, P., Sjöholm, R., Willför, S., 2010. Metal-mediated allylation of enzymatically oxidized methyl α-d-galactopyranoside. Carbohydr. Res. 345, 2610–2615. https://doi.org/10.1016/j.carres.2010.09.026
Leppänen, A.-S., Xu, C., Parikka, K., Eklund, P., Sjöholm, R., Brumer, H., Tenkanen, M., Willför, S., 2014. Targeted allylation and propargylation of galactose-containing polysaccharides in water. Carbohydr. Polym. 100, 46–54. https://doi.org/10.1016/j.carbpol.2012.11.053
Leuthner, B., Aichinger, C., Oehmen, E., Koopmann, E., Møller, O., Møller, P., Kahmann, R., Bølker,

M., Schreier, P.H., 2005. A H_2O_2-producing glyoxal oxidase is required for filamentous growth and pathogenicity in *Ustilago maydis*. Mol. Genet. Genomics 272, 639–650. https://doi.org/10.1007/s00438-004-1085-6
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. Biotechnol. Biofuels 6, 41. https://doi.org/10.1186/1754-6834-6-41
Liman, R., Facey, P.D., van Keulen, G., Dyson, P.J., Del Sol, R., 2013. A laterally acquired galactose oxidase-like gene is required for aerial development during osmotic stress in *Streptomyces coelicolor*. PLoS One 8, e54112. https://doi.org/10.1371/journal.pone.0054112
Lippow, S.M., Moon, T.S., Basu, S., Yoon, S.H., Li, X., Chapman, B.A., Robison, K., Lipovšek, D., Prather, K.L.J., 2010. Engineering enzyme specificity using computational design of a defined-sequence library. Chem. Biol. 17, 1306–1315. https://doi.org/10.1016/j.chembiol.2010.10.012
Mamo, G., Hatti-Kaul, R., Mattiasson, B., 2007. Fusion of carbohydrate binding modules from *Thermotoga neapolitana* with a family 10 xylanase from *Bacillus halodurans S7*. Extremophiles 11, 169–77. https://doi.org/10.1007/s00792-006-0023-4
Maradufu, A., Perlin, A.S., 1974. A non-hydrogen-bonding role for the 4-hydroxyl group of D-dalactose in its reaction with D-Galactose oxidase. Carbohydr. Res. 32, 93–99. https://doi.org/10.1016/S0008-6215(00)82466-0
McPherson, M., Stevens, C., Baron, A., 1993. Galactose oxidase: Molecular analysis and mutagenesis studies. Biochem. Soc. Trans. 21.752 - 756. PMID: 8224504
Mikkonen, K.S., Parikka, K., Ghafar, A., Tenkanen, M., 2013. Prospects of polysaccharide aerogels as modern advanced food materials. Trends Food Sci. Technol. 34, 124–136. https://doi.org/10.1016/j.tifs.2013.10.003
Mikkonen, K.S., Parikka, K., Suuronen, J.-P., Ghafar, A., Serimaa, R., Tenkanen, M., 2014. Enzymatic oxidation as a potential new route to produce polysaccharide aerogels. RSC Adv. 4, 11884–11892. https://doi.org/10.1039/c3ra47440b
Minasian, S.G., Whittaker, M.M., Whittaker, J.W., 2004. Stereoselective hydrogen abstraction by galactose oxidase. Biochemistry 43, 13683–13693. https://doi.org/10.1021/bi048554s
Mizutani, K., Sakka, M., Kimura, T., Sakka, K., 2014. Essential role of a family-32 carbohydrate-binding module in substrate recognition by *Clostridium thermocellum* mannanase CtMan5A. FEBS Lett. 588, 1726–30. https://doi.org/10.1016/j.febslet.2014.03.022
Molina, F.G., Muñoz, J.L., Varón, R., López, J.N.R., Cánovas, F.G., Tudela, J., 2007. An approximate analytical solution to the lag period of monophenolase activity of tyrosinase. Int. J. Biochem. Cell Biol. 39, 238–252. https://doi.org/10.1016/j.biocel.2006.08.007
Mollerup, F., Haara, M., Parikka, K., Xu, C., Tenkanen, M., Master, E., 2014. Enzymatic oxidation of plant polysaccharides adsorbed to cellulose surfaces. N. Biotechnol. 31, S7--S8. https://doi.org/10.1016/j.nbt.2014.05.1631
Moon, T.S., Nielsen, D.R., Prather, K.L.J., 2012. Sensitivity analysis of a proposed model mechanism for newly created glucose-6-oxidases. AIChE J. 58, 2303–2308. https://doi.org/10.1002/aic.12762
Mottiar, Y., 2012. M.Sc. thesis. Univ. Toronto, Toronto, Ontario, Canada 15–22.
Nishizawa, A., Yabuta, Y., Shigeoka, S., 2008. Galactinol and raffinose constitute a novel function to protect plants from oxidative damage. Plant Physiol. 147, 1251–63. https://doi.org/10.1104/pp.108.122465
Ögel, Z.B., Brayford, D., McPherson, M.J., 1994. Cellulose-triggered sporulation in the galactose oxidase-producing fungus *Cladobotryum (Dactylium) dendroides* NRRL 2903 and its re-identification as a species of *Fusarium*. Mycol. Res. 98, 474–480. https://doi.org/10.1016/S0953-7562(09)81207-0
Oliveira, C., Carvalho, V., Domingues, L., Gama, F.M., 2015. Recombinant CBM-fusion technology — Applications overview. Biotechnol. Adv. 33, 358–369. https://doi.org/10.1016/J.BIOTECHADV.2015.02.006
Parikka, K., Ansari, F., Hietala, S., Tenkanen, M., 2012a. Thermally stable hydrogels from enzymatically oxidized polysaccharides. Food Hydrocoll. 26, 212–220. https://doi.org/10.1016/j.foodhyd.2011.05.012
Parikka, K., Leppänen, A.-S.S., Pitkänen, L., Reunanen, M., Willför, S., Tenkanen, M., 2010. Oxidation of polysaccharides by galactose oxidase. J. Agric. Food Chem. 58, 262–71. https://doi.org/10.1021/jf902930t
Parikka, K., Leppänen, A.S., Xu, C., Pitkänen, L., Eronen, P., Österberg, M., Brumer, H., Willför, S., Tenkanen, M., 2012b. Functional and anionic cellulose-interacting polymers by selective chemo-enzymatic carboxylation of galactose-containing polysaccharides. Biomacromolecules 13, 2418–2428. https://doi.org/10.1021/bm300679a
Parikka, K., Master, E., Tenkanen, M., 2015. Oxidation with galactose oxidase: Multifunctional enzymatic catalysis. J. Mol. Catal. B Enzym. 120, 47–59.

https://doi.org/10.1016/j.molcatb.2015.06.006
Parikka, K., Tenkanen, M., 2009. Oxidation of methyl α-d-galactopyranoside by galactose oxidase: products formed and optimization of reaction conditions for production of aldehyde. Carbohydr. Res. 344, 14–20. https://doi.org/10.1016/j.carres.2008.08.020
Paukner, R., Staudigl, P., Choosri, W., Haltrich, D., Leitner, C., 2015. Expression, purification, and characterization of galactose oxidase of *Fusarium sambucinum* in *E. coli*. Protein Expr. Purif. 108, 73–79. https://doi.org/10.1016/j.pep.2014.12.010
Paukner, R., Staudigl, P., Choosri, W., Sygmund, C., Halada, P., Haltrich, D., Leitner, C., 2014. Galactose oxidase from *Fusarium oxysporum* - Expression in *E. coli* and *P. pastoris* and biochemical characterization. PLoS One 9, e100116. https://doi.org/10.1371/journal.pone.0100116
Pedersen, A., Birmingham, W.R., Rehn, G., Charnock, S.J., Turner, N.J., Woodley, J.M., 2015. Process requirements of galactose oxidase catalyzed oxidation of alcohols. Org. Process Res. Dev. 19, 1580–1589. https://doi.org/10.1021/acs.oprd.5b00278
Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8, 785–786. https://doi.org/10.1038/nmeth.1701
Pierre, G., Punta, C., Delattre, C., Melone, L., Dubessay, P., Fiorati, A., Pastori, N., Galante, Y.M., Michaud, P., 2017. TEMPO-mediated oxidation of polysaccharides: An ongoing story. Carbohydr. Polym. 165, 71–85. https://doi.org/10.1016/J.CARBPOL.2017.02.028
Pizarro, L.F., 2016. Optimization and modeling of recombinant protein production in *Pichia Pastoris*. Aalto University. Master Thesis.
Rannes, J.B., Ioannou, A., Willies, S.C., Grogan, G., Behrens, C., Flitsch, S.L., Turner, N.J., 2011. Glycoprotein labeling using Engineered variants of galactose oxidase obtained by directed evolution. J. Am. Chem. Soc. 133, 8436–8439. https://doi.org/10.1021/ja2018477
Reynolds, M.P., Baron, A.J., Wilmot, C.M., Vinecombe, E., Stevens, C., Phillips, S.E. V., Knowles, P.F., McPherson, M.J., 1997. Structure and mechanism of galactose oxidase: catalytic role of tyrosine 495. JBIC J. Biol. Inorg. Chem. 2, 327–335. https://doi.org/10.1007/s007750050139
Rogers, M., Baron, A., McPherson, M.J., Knowles, P.F., Dooley, D.M., 2000. Galactose oxidase pro-sequence cleavage and cofactor assembly are self-processing reactions. J. Amercan Chem. Soc. 122, 990–991. https://doi.org/10.1021/ja993385y
Rogers, M., Tyler, E., Akyumani, N., 2007. The stacking tryptophan of galactose oxidase: a second-coordination sphere residue that has profound effects on tyrosyl radical behavior and enzyme catalysis. Biochemistry 46, 4606–4618. https://doi.org/10.1021/bi062139d
Rogers, M.S., Dooley, D.M., 2001. Posttranslationally modified tyrosines from galactose oxidase and cytochrome C oxidase. Adv. Protein Chem. https://doi.org/10.1016/S0065-3233(01)58009-2
Rossi, B., Campia, P., Merlini, L., Brasca, M., Pastori, N., Farris, S., Melone, L., Punta, C., Galante, Y.M., 2016. An aerogel obtained from chemo-enzymatically oxidized fenugreek galactomannans as a versatile delivery system. Carbohydr. Polym. 144, 353–361. https://doi.org/10.1016/j.carbpol.2016.02.007
Rossi, B., Ponzini, E., Merlini, L., Grandori, R., Galante, Y.M., 2017. Characterization of aerogels from chemo-enzymatically oxidized galactomannans as novel polymeric biomaterials. Eur. Polym. J. 93, 347–357. https://doi.org/10.1016/j.eurpolymj.2017.06.016
Rouau, X., Schrøder, M., Søe, J., 1999. A composition comprising an enzyme having galactose oxidase activity and use thereof. US patent WO1999003351A1.
Sakka, M., Higashi, Y., Kimura, T., Ratanakhanokchai, K., Sakka, K., 2011. Characterization of *Paenibacillus curdlanolyticus B-6* Xyn10D, a xylanase that contains a family 3 carbohydrate-binding module. Appl. Environ. Microbiol. 77, 4260–4263. https://doi.org/10.1128/AEM.00226-11
Schilstra, M.J., Veldink, G.A., Vliegenthart, J.F.G., 1994. The dioxygenation rate in lipoxygenase catalysis is determined by the amount of Iron(III) lipoxygenase in solution. Biochemistry 33, 3974–3979. https://doi.org/10.1021/bi00179a025
Schoevaart, R., Kieboom, T., 2002. Galactose dialdehyde as potential protein cross-linker: proof of principle. Carbohydr. Res. 337, 899–904. https://doi.org/10.1016/S0008-6215(02)00051-4
Selanere, M.L., Andersson, R., 2002. Cell wall composition of 1B/1R translocation wheat grains. J. Sci. Food Agric. 82, 538–545. https://doi.org/10.1002/jsfa.1070
Sharma, S.K., Suman, Pundir, C.S., Sehgal, N., Kumar, A., 2006. Galactose sensor based on galactose oxidase immobilized in polyvinyl formal. Sensors Actuators, B Chem. 119, 15–19. https://doi.org/10.1016/j.snb.2005.11.046
Shoseyov, O., Shani, Z., Levy, I., 2006. Carbohydrate binding modules: biochemical properties and novel applications. Microbiol. Mol. Biol. Rev. 70, 283–95. https://doi.org/10.1128/MMBR.00028-05
Siebum, A., van Wijk, A., Schoevaart, R., Kieboom, T., 2006. Galactose oxidase and alcohol oxidase:

Scope and limitations for the enzymatic synthesis of aldehydes. J. Mol. Catal. B Enzym. 41, 141–145. https://doi.org/10.1016/j.molcatb.2006.04.003
Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539. https://doi.org/10.1038/msb.2011.75
Silakowski, B., Ehret, H., Schairer, H., 1998. fbfB, a gene encoding a putative galactose oxidase, is involved in *Stigmatella aurantiaca* fruiting body formation. J. Bacteriol. 180, 1241–1247.
Silvetti, T., Merlini, L., Brasca, M., Galante, Y.M., 2018. Aerogel from chemo-enzymatically oxidized fenugreek gum: an innovative delivery system of isothiazolinones biocides. Appl. Microbiol. Biotechnol. 102, 2683–2692. https://doi.org/10.1007/s00253-018-8804-0
Spadiut, O., Olsson, L., Brumer, H., 2010. A comparative summary of expression systems for the recombinant production of galactose oxidase. Microb. Cell Fact. 9, 1–13. https://doi.org/10.1186/1475-2859-9-68
Stefano, J., 2006. Targeting of glycoprotein therapeutics. US patent 20170042979A1.
Sun, L., Bulter, T., Alcalde, M., Petrounia, I.P., Arnold, F.H., 2002. Modification of galactose oxidase to introduce glucose 6-oxidase activity. ChemBioChem 3, 781. https://doi.org/10.1002/1439-7633(20020802)3:8<781::AID-CBIC781>3.0.CO;2-8
Sun, L., Petrounia, I., Yagasaki, M., 2001. Expression and stabilization of galactose oxidase in *Escherichia coli* by directed evolution. Prot Eng. Des and Sel. 14, 699–704. https://doi.org/10.1093/protein/14.9.699
Tkac, J., Sturdik, E., Gemeiner, P., 2000. Novel glucose non-interference biosensor for lactose detection based on galactose oxidase-peroxidase with and without co-immobilised β-galactosidase. Analyst 125, 1285–1289. https://doi.org/10.1039/b001432j
Tkac, J., Whittaker, J.W., Ruzgas, T., 2007. The use of single walled carbon nanotubes dispersed in a chitosan matrix for preparation of a galactose biosensor. Biosens. Bioelectron. 22, 1820–1824. https://doi.org/10.1016/j.bios.2006.08.014
Várnai, A., Mäkelä, M.R., Djajadi, D.T., Rahikainen, J., Hatakka, A., Viikari, L., 2014. Carbohydrate-binding modules of fungal cellulases. occurrence in nature, function, and relevance in industrial biomass conversion. Adv. Appl. Microbiol. 88, 103–165. https://doi.org/10.1016/B978-0-12-800260-5.00004-8
Voutilainen, S.P., Nurmi-Rantala, S., Penttila, M., Koivula, A., 2013. Engineering chimeric thermostable GH7 cellobiohydrolases in *Saccharomyces cerevisiae*. Appl. Microbiol. Biotechnol. 1–11. https://doi.org/10.1007/s00253-013-5177-2
Vuong, T. V., Master, E.R., 2014. Fusion of a xylan-binding module to gluco- oligosaccharide oxidase increases activity and promotes stable immobilization. PLoS One. https://doi.org/10.1371/journal.pone.0095170
Wang, W.H., Takano, T., Shibata, D., Kitamura, K., Takeda, G., 1994. Molecular basis of a null mutation in soybean lipoxygenase 2: substitution of glutamine for an iron-ligand histidine. Proc. Natl. Acad. Sci. U. S. A. 91, 5828–32. https://doi.org/10.1073/pnas.91.13.5828
Whittaker, J., 2005. The radical chemistry of galactose oxidase. Arch. Biochem. Biophys. 433, 227–239. https://doi.org/10.1016/j.abb.2004.08.034
Whittaker, J.W., 2003. Free radical catalysis by galactose oxidase. Chem. Rev. 103, 2347–2364. https://doi.org/10.1021/cr020425z
Whittaker, M.M., DeVito, V.L., Asher, S.A., Whittaker, J.W., 1989. Resonance Raman evidence for tyrosine involvement in the radical site of galactose oxidase. J. Biol. Chem. 264, 7104–7106.
Whittaker, M.M., Kersten, P.J., Nakamura, N., Sanders-Loehr, J., Schweizer, E.S., Whittaker, J.W., 1996. Glyoxal Oxidase from *Phanerochaete chrysosporium* is a new radical-copper oxidase. J. Biol. Chem. 271, 681–687. https://doi.org/10.1074/jbc.271.2.681
Whittaker, M.M., Whittaker, J.W., 2003. Cu(I)-dependent biogenesis of the galactose oxidase redox cofactor. J. Biol. Chem. 278, 22090–22101. https://doi.org/10.1074/jbc.M300112200
Whittaker, M.M., Whittaker, J.W., 2001. Catalytic reaction profile for alcohol oxidation by galactose oxidase. Biochemistry 40, 7140–7148. https://doi.org/10.1021/bi010303l
Whittaker, M.M., Whittaker, J.W., 2000. Expression of recombinant galactose oxidase by *Pichia pastoris*. Protein Expr. Purif. 20, 105–11. https://doi.org/10.1006/prep.2000.1287
Whittaker, M.M., Whittaker, J.W., 1993. Ligand interactions with galactose oxidase: mechanistic insights. Biophys. J. 64, 762–72. https://doi.org/10.1016/S0006-3495(93)81437-1
Wilkinson, D., Akumanyi, N., Hurtado-Guerrero, R., Dawkes, H., Knowles, P.F., Phillips, S.E. V, McPherson, M.J., 2004. Structural and kinetic studies of a series of mutants of galactose oxidase identified by directed evolution. Protein Eng. Des. Sel. 17, 141–8. https://doi.org/10.1093/protein/gzh018
Willför, S., Sjöholm, R., Laine, C., Roslund, M., Hemming, J., Holmbom, B., 2003. Characterisation of water-soluble galactoglucomannans from *Norway spruce* wood and thermomechanical pulp.

Carbohydr. Polym. 52, 175–187. https://doi.org/10.1016/S0144-8617(02)00288-6
Winterbourn, C.C., Peskin, A. V., Parsons-Mair, H.N., 2002. Thiol oxidase activity of copper,zinc superoxide dismutase. J. Biol. Chem. 277, 1906–1911. https://doi.org/10.1074/jbc.M107256200
Wymelenberg, A. Vanden, Sabat, G., Mozuch, M., Kersten, P.J., Cullen, D., Blanchette, R.A., 2006. Structure, organization, and transcriptional regulation of a family of copper radical oxidase genes in the lignin-degrading basidiomycete *Phanerochaete chrysosporium*. Appl. Environ. Microbiol. 72, 4871–4877. https://doi.org/10.1128/AEM.00375-06
Xie, J., Chen, C., Zhou, Y., Fei, J., Ding, Y., Zhao, J., 2016. A galactose oxidase biosensor based on graphene composite film for the determination of galactose and dihydroxyacetone. Electroanalysis 28, 183–188. https://doi.org/10.1002/elan.201500486
Xu, C., Spadiut, O., Araújo, A.C., Nakhai, A., Brumer, H., 2012. Chemo-enzymatic assembly of clickable cellulose surfaces via multivalent polysaccharides. ChemSusChem 5, 661–665. https://doi.org/10.1002/cssc.201100522
Yin, D.T., Urresti, S., Lafond, M., Johnston, E.M., Derikvand, F., Ciano, L., Berrin, J.-G., Henrissat, B., Walton, P.H., Davies, G.J., Brumer, H., 2015. Structure-function characterization reveals new catalytic diversity in the galactose oxidase and glyoxal oxidase family. Nat. Commun. 6, 10197. https://doi.org/10.1038/ncomms10197
Zhu, Y., 2002. Methods for introducing mannose 6-phosphate and other oligosaccharides onto glycoproteins. US patent 7001994B2.

This doctoral thesis is an account for my research on the galactose oxidase (GaOx) and its family members in the Carbohydrate-Active enZYmes database. GaOx converts the primary alcohol of the carbohydrate (a.k.a. sugars) galactose into the corresponding aldehyde using only atmospheric air, eliminating the need for additional electron-accepting molecules. This copper-containing oxidase is particularly known for its ability to act on polymeric carbohydrate chains, which presents a unique potential for biocatalytic engineering of plant-based fibrous materials. However, GaOx is strictly sel-ective to galactose limiting this potential to include only galactose-containing carbohydrates.

In this research, protein engineering technology was applied in an attempt to improve the performance of GaOx on galactose- and also glucose-containing carbohydrates. In the second half, enzyme discovery aimed to explore the functional space of the AA5_2 family, leading to increased diversity.

Aalto-DD 232/2018

ISBN 978-952-60-8309-4 (printed)
ISBN 978-952-60-8310-0 (pdf)
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)

Aalto University
School of Chemical Engineering
Department of Bioproducts and Biosystems
www.aalto.fi

BUSINESS + ECONOMY

ART + DESIGN + ARCHITECTURE

SCIENCE + TECHNOLOGY

CROSSOVER

DOCTORAL DISSERTATIONS