Priv.-Doz. Dr. Bernhard Seiboth

DIPLOMARBEIT

Heterologous expression of the thermostable cellulase CEL45A in *Trichoderma reesei*

Ausgeführt am Institut für Verfahrenstechnik, Umwelttechnik und Technische Biowissenschaften
der Technischen Universität Wien

unter der Anleitung von
Priv.-Doz. Dr. Bernhard Seiboth

durch

Philipp Peter Haider, BSc
[redacted]

Datum

Philipp Peter Haider, BSc

Am Ziel deiner Wünsche wirst du jedenfalls eines vermissen:
dein Wandern zum Ziel.

Marie Freifrau von Ebner-Eschenbach

Danksagung

Ich möchte mich auf diesem Weg bei Dr. Bernhard Seiboth bedanken. Vielen Dank, dass ich an einem praxisorientierten Projekt mit Industriebeteiligung arbeiten durfte und für die Unterstützung während des Verfassens dieser Arbeit. Judith Frei danke ich für ihren unermüdlichen Einsatz mich zu motivieren, zu fordern und zu fördern.

Danke auch an die tolle Arbeitsgruppe. Es waren schöne Stunden mit und bei euch. Insbesondere erwähnen möchte ich hier Theresa Radebner, Elisabeth Fitz, Jonas Ramoni und Robert Bischof.

Besonderer Dank gilt meinen Eltern für die finanzielle Unterstützung während des Studiums. Sie waren ein großer Rückhalt und immer mein letzter Hafen, falls ich auf zu stürmischer See unterwegs war.

Meinen lieben Kollegen, Freunden, Verwandten und Bekannten, welche mich immer tatkräftig zum Fortgehen und zu Späßen motiviert haben, sage ich auch ein herzliches Dankeschön. Ohne euch wäre die vergangene Zeit nicht so farbenfroh gewesen.

Danke Julia, dass du mein Leben so viel schöner machst. Ich freue mich noch auf viele gemeinsame Jahre mit dir.

1. Content

2. Abbreviations

mRNA	messenger ribonucleic acid
qPCR	quantitative polymerase chain reaction
CMC	carboxymethyl cellulose
Da	Dalton
DDA	automated data dependent acquisition
DNA	deoxyribonucleic acid
ER	endoplasmic reticulum
ESI	electrospray ionisation
ETD	electron-transfer dissociation
LB	lysogeny broth
LC	liquid chromatography
MS	mass spectrometer
MS/MS	tandem mass spectrometer
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
QTOF	quadrupole time-of-flight mass spectrometer

3. Zusammenfassung

Seit seiner Entdeckung ist der Pilz *Trichoderma reesei* für seine Cellulasen, und damit zusammenhängend dem Abbau von Cellulose bekannt. Der filamentöse Schimmelpilz *T. reesei* wird seit 30 Jahren in der industriellen Produktion von verschiedenen Cellulasen als Expressionshost verwendet. Durch Mutagenese und Selektion wurden *T. reesei* Stämme entwickelt, die beachtliche Cellulasemengen bilden können und durch die Implementierung molekularer Werkzeuge wurden auch beachtliche Erfolge in der Produktion heterologer Cellulasen erzielt. Großes Interesse besteht v.a. in der Produktion thermostabiler Cellulasen, um Lignocellulose mit höheren Reaktionsraten zu fermentierbaren Zuckern umzuwandeln. Das Ziel dieser Arbeit war es eine thermostabile Cellulase in *T. reesei* zu exprimieren. Dazu wurde die kodierende Region der *cel45a* Cellulase unter den homologen Cellobiohydrolase 1 Expressionssignalen von *T. reesei* kloniert. Die Expressionskassette wurde mittels Elektroporation in den *T. reesei* Wildstamm QM6a und den stark Cellulasen produzierenden Stamm Rut C30 eingebracht. Zur Produktion von CEL45A wurden ausgewählte Transformanten auf Cellulose als induzierenden C-Quelle angezüchtet. Im Überstand der Transformanten beider Ausgangsstämme wurde ein 37 kDa Protein gebildet, das mittels MS-MS als CEL45A identifiziert wurde. *Cel45a* Transformanten zeigten eine allgemein erhöhte Endoglucanaseaktivität und die gebildete Cellulase zeigte bei 60 °C erhöhte Thermostabilität. Zusammenfassend konnte mit dieser Arbeit gezeigt werden, dass CEL45A funktionell aktiv und thermostabil in *T. reesei* exprimiert wird.

4. Abstract

Since its discovery the fungus *Trichoderma reesei* is known for its cellulases and therefore for its ability to degrade cellulose. The filamentous ascomycete *T. reesei* is used as expression host for the industrial production of cellulases during the last 30 years. Through mutation and selection strains were generated that are able to produce considerable amounts of cellulases. Based on the development of different molecular tools a great step forward in recombinant strain construction for heterologous production of cellulases could be achieved. There is a great interest especially in producing thermostable cellulases, to hydrolyse lignocellulose with higher yields. The aim of this work was to express a thermostable cellulase in *T. reesei*. For this purpose, the coding region of the *cel45a* cellulase was cloned under the homologous cellobiohydrolase 1 expression signals. The expression cassette was transformed through electroporation into the *T. reesei* wild type strain QM6a and the strong cellulase producer *T. reesei* Rut C30. To produce CEL45A selected transformants were cultivated on cellulose as inducing C-source. In the supernatant of the transformants of both parental strains a 37 kDa protein was secreted which was identified as CEL45A through MS-MS. *Cel45a* transformants showed a general increased endoglucanase activity and the expressed cellulase showed an increased thermostability at 60°C. In summary, it could be shown within this work, that CEL45A could be produced functionally active and thermostable with the host *T. reesei*.

5. Introduction

5.1 *Trichoderma reesei*

T. reesei is a filamentous mesophilic ascomycete, discovered by the US Army during Second World War on the Solomon Islands due to its ability to degrade cotton based army materials. The fungus isolated at Natick was first named *Trichoderma viride*, but later the species was reclassified and renamed *T. reesei* in honour of his discoverer E.T. Reese who worked on cellulase research at the US Army Quartermaster Research and Development Centre at Natick (Mandels and Reese 1957). Filamentous fungi can produce numerous extracellular proteins, and *T. reesei* is especially known for its ability to produce cellulases and hemicellulases. During the oil crises in the 1970s the species gained industrial interest as an alternative to the fossil based industry. But the production of cellulases in such naturally occurring strains is usually too low for industrial application, thus making strain improvement essential. Over decades several methods of random mutagenesis and selection were applied to the *T. reesei* isolate QM6a from the Solomon Islands leading to several strain descendants with higher enzyme productivity (Durand et al. 1988).

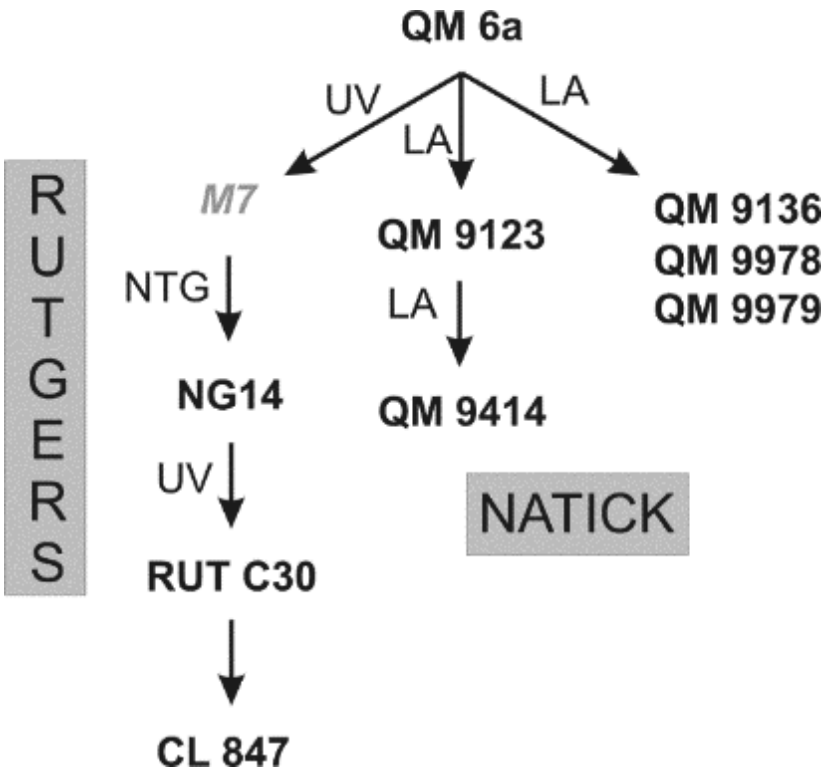


Fig. 1: Strain pedigree of *T. reesei*. Strains of the Rutgers series are shown on the left side, starting with M7. The strains of the NATICK series are presented on the right. 3 rounds of random mutagenesis lead to Rut C30 which can produce about 30 g/l extracellular protein. (Kubicek 2013)

One of them is Rut C30, which produces titers of extracellular protein of 30 g/l (Montenecourt and Eveleigh 1979). When cultivated under optimised conditions, *T. reesei* industrial strains can secrete 100 g/l of homologous protein under cellulase inducing conditions (Cherry and Fidantsef 2003; Peterson and Nevalainen 2012). QM6a has seven chromosomes and its genome is 32.5 Mbp in size (Martinez et al. 2008). Using high-throughput next generation sequencing techniques detailed informations are now available about the genomic changes accompanying cellulase hyperproduction in Rut C30: these changes include large deletions and insertions and single nucleotide variants. Overall it was established that in Rut C30 44 genes are affected by mutations compared to QM6a (Le Crom et al. 2009). Rut C30 shows significant carbon catabolite derepression due to a truncation of *cre1*, a gene encoding the major carbon catabolite repressor (Ilmen et al. 1996b). Regarding the high secretion and synthesis capacity of Rut C30 it could be shown that the volume of the endoplasmic reticulum is increased within this mutant (Ghosh et al. 1982).

5.2 Industrial Importance of *T. reesei*

All *T. reesei* strains used today in industry have been derived from a single isolate *T. reesei* QM6a. Its ability to efficiently degrade lignocellulosic substrates promoted the research on cellulase and xylanase biochemistry and regulation (Foreman et al. 2003; Polizeli et al. 2005). With the possibilities of recombinant gene technology, interest in using the hyperproducing strains to express also heterologous enzymes has been increased. The fungus has GRAS status and is non-pathogenic to humans and therefore safe to use. The strain has been applied within the industry for 30 years now and therefore lots of knowledge has been gained over the past decades. The majority of industrially applied enzymes produced by *T. reesei* are hemicellulases, cellulases, amylases, lipases and proteases (Bischof et al. 2016).

5.2.1 Second Generation Biofuels

Beside the application of enzymes in pulp and paper industry as well as in the textile industry, the cocktails produced by *T. reesei* gained high recognition in second generation bioethanol production. The ecological but also economic interests to hydrolyse cellulose-rich biomass into fermentable sugars lead to increased investigations and high investments in research. By-products from agriculture and forestry are used to produce simple sugars which can be converted to bioethanol (Kocar and Civas 2013; Kumar et al. 2008; Popp et al. 2014).

5.2.2 Heterologous Protein Production in T. reesei

As stated above, the total homologous protein secretion capacity of industrial *T. reesei* strains is about 100 g/l (Cherry and Fidantsef 2003; Peterson and Nevalainen 2012) making this filamentous fungus a preferable production host for extracellular proteins. The yield looks great compared to other microbial cell factories and the protein of interest is secreted. Therefore, the protein of interest is easier to handle within the downstream process. In the secretory pathway the protein is folded, proteolytic processed and glycans are added as main modification. The heterologous protein production in *T. reesei* of mammalian proteins

can often not compete with the yields obtained for homologous proteins and efforts have to be made to make *T. reesei* cost competitive to other expression systems. It is therefore widely accepted that the expression of heterologous fungal proteins have a better chance for improvement in filamentous fungi (Nevalainen and Peterson 2014). Over the years different mammalian proteins were expressed within *T. reesei*, for example calf chymosin (Harkki et al. 1989). The yields for different types of recombinant proteins can vary substantially even within the same host organism (Demain and Vaishnav 2009). Only recently it was shown that the production of certain proteins (therapeutic antibodies, interferon alpha 2b, insulin like growth factor) can considerably be improved by e.g. deleting several proteases in *T. reesei* (Landowski et al. 2015; Landowski et al. 2016).

Several molecular tools were developed for filamentous fungi which allow an efficient molecular manipulation leading to new ways for molecular strain improvement (Bischof et al. 2016). New approaches for increasing the yields of industrial protein production try to engineer whole metabolic pathways (in the field of synthetic biology) and recognising the synergy between fungal enzymes (Gupta et al. 2016b). However, there are still major bottlenecks in heterologous protein production which need to be investigated throughout basic research.

5.2.3 Potential Bottlenecks in Protein Production

The great advantage of filamentous fungi versus other microbial cell factories is their high potential to produce high quantities of extracellular protein. With regard to the overexpression of a homologous proteins lots of molecular tools are established. But concerning heterologous protein expression different bottlenecks were identified (see Fig. 2) and therefore improved strains and methods are highly desirable.

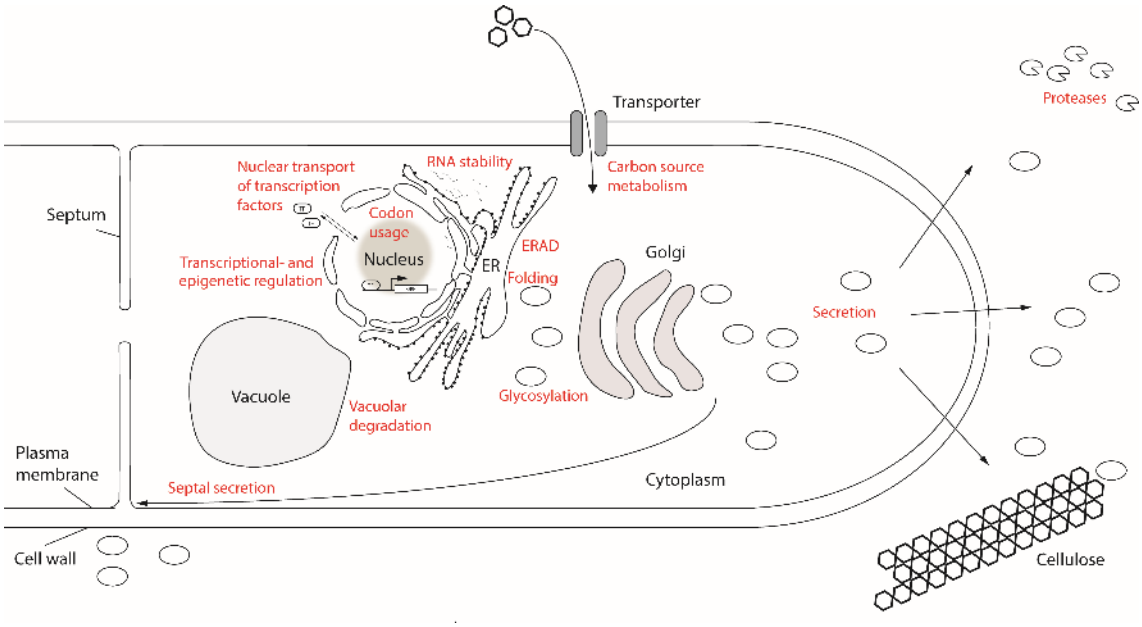


Fig. 2: Bottlenecks in extracellular protein production in filamentous fungi. Starting with the transcription of the gene potential various bottlenecks in fungal protein expression are depicted here in red (Ramoni et al. 2016).

Typical methods to improve the protein yield through genetic modifications and therefore overcome this bottleneck are (Ramoni et al. 2016): (1) optimisation of the codon usage (2) introduction of multiple copies of the gene of interest (3) the use of strong constitutive or inducing promoters (4) efficient secretion signals through gene fusion and (5) construction of protease deficient strains.

The strong promoters used are often derived from genes which encode the most dominant extracellular protein. For *T. reesei* this is the promoter of the gene *cbh1*/*cel7a* which encodes the cellobiohydrolase CBH1/CEL7A (Rahman et al. 2009). It can be induced by cellulose, lactose and sophorose (dos Santos Castro et al. 2014). In general, the main inducing agents for producing high levels of cellulase and hemicellulase are cellulose, xylan (Mach and Zeilinger 2003) and lactose (Seiboth et al. 2005).

Looking at Fig. 2 folding is a concern regarding efficient heterologous protein production. High expression of the gene of interest results in an overload of protein. The folding capacity in the ER is limited and rises soon to its maximum. As a consequence, different proteins are then upregulated within the folding machinery which are associated with the unfolded protein response (UPR). The UPR relieves cellular stress from accumulation of unfolded or misfolded proteins in the ER. This is done by increasing molecular chaperons and extra pro-

tein-folding-associated proteins (Malhotra and Kaufman 2007; Travers et al. 2000).

A potential role during the folding process plays the redox capacity of the ER. The ER provides an oxidizing environment which is well suited to the formation of disulphide bonds in extracellular proteins entering the secretory pathway. Pdi1 (protein disulphide isomerase) is important during formation, isomerization and reduction of disulphide bonds. During disulphide bond formation Pdi1 is reduced by the cysteine residues of the nascent protein. Ero1 (ER oxidoreductin) works as a partner enzyme of Pdi1 and passes the electrons over to molecular oxygen restoring the oxidized form of Pdi1 (Tu et al. 2000), see Fig. 3.

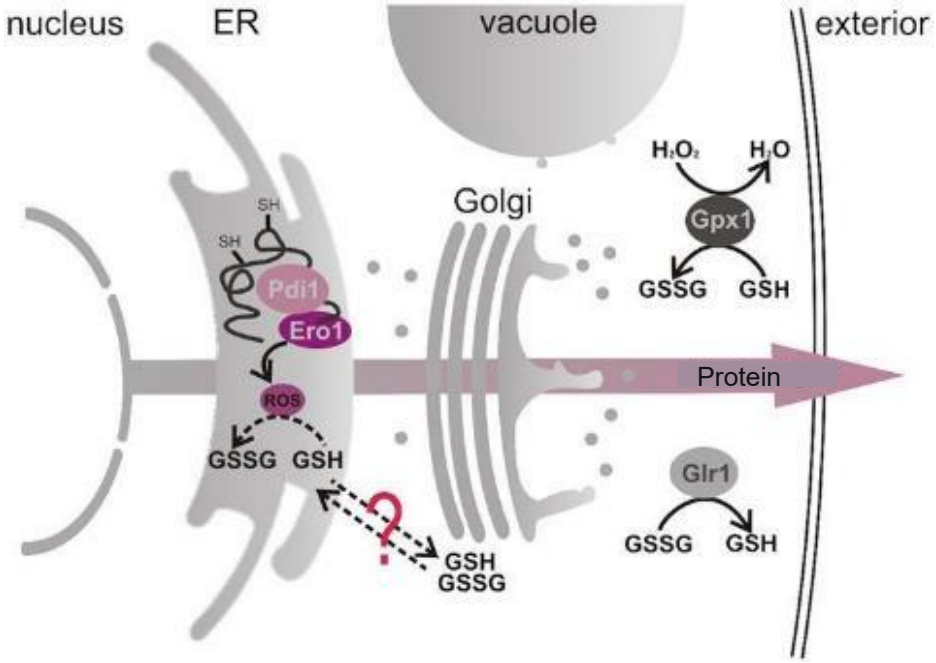


Fig. 3: ROS formation during protein folding. ROS is generated during the protein folding cascade. The tripeptide glutathione works as redox buffer. An overload of protein folding in the ER results in oxidative stress (Delic et al. 2012).

This cascade potentially leads to the generation of ROS (reactive oxygen species). ROS causes oxidative stress which is associated with different diseases due to damage in the cell. Glutathione, a tripeptide, works as redox buffer within cells. In the cell, glutathione is present in its reduced (GSH) and the oxidized form which builds a dimer (GSSG). Within the ER and the cytosol there are typical GSH : GSSG ratios found: in the less oxidizing cytosol the ratio is 100 : 1 and within the ER the ratio is about 10 : 1, depending on the organism (Hwang et al. 1992). It could be shown in *P. pastoris* that protein folding and ER stress

have an impact on the redox conditions in the cytosol. In case of secretion of a disulphide-rich model protein the cytosol was more reduced (Delic et al. 2012).

5.3 Glycoside Hydrolases

T. reesei secrets lots of different enzymes, mainly for degrading plant cell walls which consists of the β-(1,4)-linked glucose polymer cellulose, hemicellulose polysaccharides and lignin. These include as mentioned earlier mostly cellulases and hemicellulases. Most of these enzymes belong to the glycoside hydrolases (GH). There are two ways to classify these enzymes. The first one is the common IUBMB classification, which sorts the enzymes according to their activity. The other classification is done through their structural features within different glycoside hydrolase families. On the internet, a website was created to deliver free access to the knowledge of the currently known carbohydrate active enzymes (CAZymes): www.cazy.org. Additional information as catalytic mechanism of glycoside hydrolases provides the web site www.cazypedia.org. CAZymes in general build and breakdown complex carbohydrates and glycoconjugates. The different families include glycoside hydrolases, glycosyltransferases, polysaccharide lyases, carbohydrate esterase and auxiliary activities. The non-catalytic carbohydrate-binding modules (CBM) which were often appended to the enzyme, referred then as catalytic domain, are also classified on this website. Because of their widespread importance within biotechnology and through biomedical applications the glycoside hydrolases are the best biochemically characterised set of enzymes within the CAZy database. Overall up to today this database contains 136 glycoside hydrolase families (Cantarel et al. 2009; Lombard et al. 2014). Glycoside hydrolases can be classified also by other ways. One characterisation is to specify them via their ability to cleave the substrate at the end (exoenzyme) or within the middle (endoenzyme) of the chain. It also can be distinguished between processive and non-processive enzymes. Processive enzymes attach to the polymeric substrate and perform rounds of catalysis before dissociating. Thereby the processive enzymes are divided into two classes. In one class the enzyme can completely enclosure the substrate. In the other class the substrate is just partially enclosed. Beside these GHs, the

recently discovered lytic polysaccharide monooxygenases (LPMO) cleave the crystalline regions of cellulose by oxidation (presented in Fig. 4).

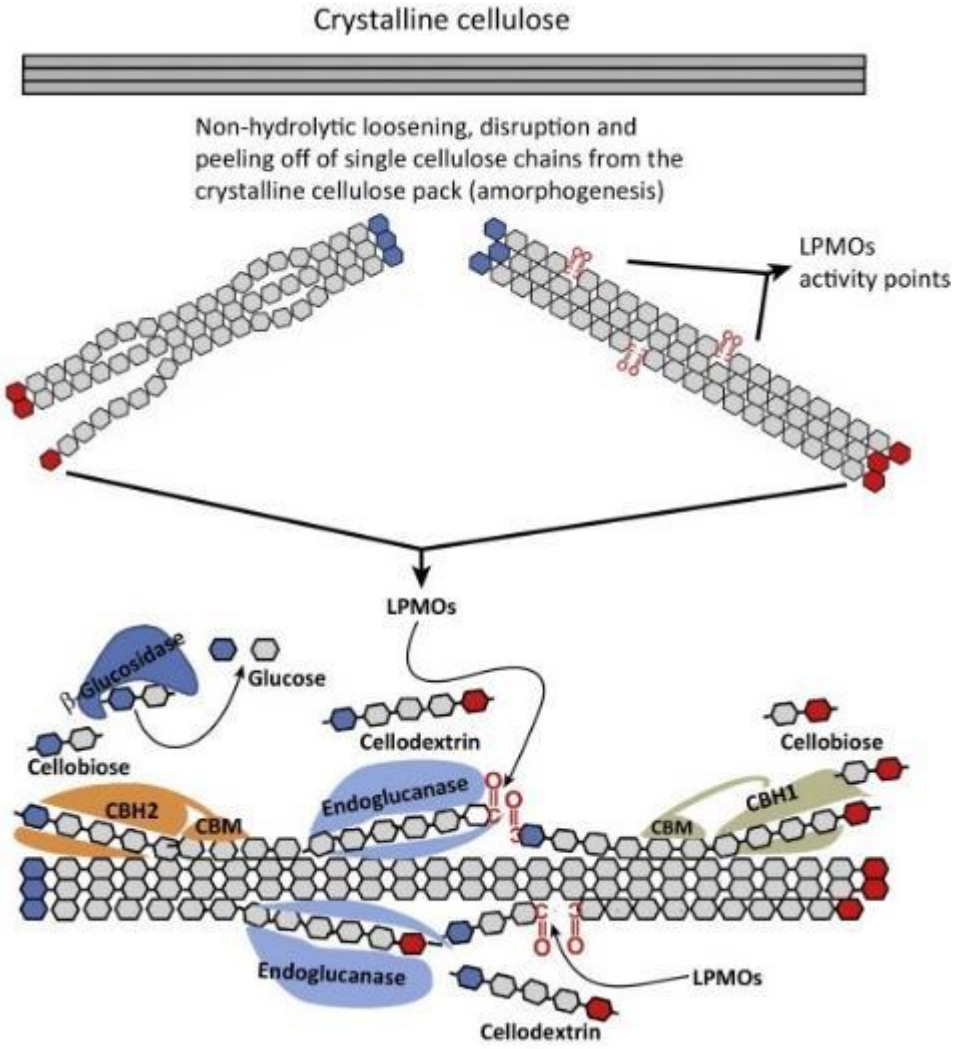


Fig. 4: Oxidative and hydrolytic cleavage of crystalline cellulose to monomeric sugars. A cocktail of enzymes is secreted by fungi to metabolise cellulose. Most of the enzymes belong to the glycoside hydrolases which catalyse the hydrolysis of the glycosidic linkage of glycosides. An important group for cellulose degradation are the LPMOs which render the cellulase more susceptible to hydrolysis by conventional cellulases (Gupta et al. 2016a).

5.3.1 Glycoside Hydrolase Family 45

The protein analysed within this thesis belongs to the glycoside hydrolase family 45. GH 45 enzymes are endoglucanases therefore cleaving the cellulose chains in the middle. Most of the enzymes within these groups consist of two domains, the catalytic and the carbohydrate binding module and they are compared to other cellulases considered as small. This leads to the hypothesis that these enzymes have an evolutionary advantage, because they can penetrate into

smaller pores and therefore have better access to the substrate. *T. reesei* produces one endoglucanase *Tr*CEL45A which belongs to this family (Foreman et al. 2003; Saloheimo et al. 1994). Its molecular mass is of about 23 kDa and possess two domains, the catalytic and carbohydrate binding module which are connected by a linker. CEL45A retains more than 75% of its maximal enzymatic activity in a pH range of 3 to 6 and is thermostable. It shows a maximal activity at 60 °C. The main product after hydrolysis of the cellulose substrate is cellotetraose. It also produces significant amount of cellopentaose and cellotriose (Karlsson et al. 2002).

A main application of GH 45 enzymes is their use within the textile industry. Especially for biostoning GH 45 enzymes of different species are used. The focus of research is to lower the needed dosage, and thus increase the performance of GH 45 enzymes or improving their properties. A future goal is also to save energy and therefore to engineer enzymes which work already within a temperature range of 20 to 40°C (Valtakari et al.).

5.4 Genetic Engineering of Cellulases

The structure and mechanism of the enzymes involved in degrading lignocellulosic biomass, as well as their regulation of expression, the pathways of their formation and secretion was already investigated in detail. One of the present efforts within research is to change the properties of the enzymes involved. The focus has been on e.g. improving thermostability, changing their activity to a broader pH range and relief from inhibition by components (Druzhinina and Kubicek 2017).

5.4.1 Thermostable Cellulases

An important property of cellulases is often their thermostability, because it is required by almost every industrial process. Main advantages of performing processes at high temperature are the increased solubility of many polymeric substrates, the decrease in viscosity, a lower risk of contamination with microbes and a higher reaction rate. Therefore less enzymes are needed (Kuhad et al. 2016). As *T. reesei* is mesophilic, most of its cellulases have their optimal

activity at max. 40 °C, several thermostable cellulases from other fungi (*Acremonium thermophilum*, *Thermoascus aurantiacus, Chaetomium thermophilum, Chaetomium atrobrunneum)* were successfully expressed in *T. reesei* (Colabardini et al. 2016; Voutilainen et al. 2008). But there are also different approaches possible which change the properties of the original cellulases of *T. reesei*. Classical approaches as single amino acid substitutions (Sandgren et al. 2003), as well as methods from synthetic biology, chimeragenesis (Smith et al. 2013), are used to gain cellulases with increased thermostability.

6. Aim

The aim of this work was to transform the low cellulase producer *T. reesei* QM6a and the high cellulase producing strain *T. reesei* Rut C30 with an expression cassette containing a *cel45a* gene under homologous *cbh1/cel7a* expression signals. The obtained transformants will then be tested for the presence of *cel45a* and its functional expression will be tested by determining the endoglucanase activity in the supernatants of the transformants. In addition it will be tested if the protein is produced in a thermostable form in the two *T. reesei* strains.

7. Materials

7.1 Instruments

Table 1: Instruments

Instrument	Manufacturer
AF80 Ice Flaker	Scotsman Ice Ltd., United Kingdom
Balance AW-224	Sartorius AG, Germany
Balance TP-303	Denver Instrument, USA
Biological Safety Cabinet Herasafe™ KS	Thermo Fisher Scientific Inc., USA
Canon EOS 400D incl. lens EF-S18-55	Canon Inc., Japan
Centrifuge 5415 R	Eppendorf AG, Germany
Centrifuge Sigma 3-30KS	SIGMA Laborzentrifugen GmbH, Germany
CertoClav CV-EL 18L	CertoClav Sterilizer GmbH, Austria
Freezer MediLine	Liebherr-International Deutschland GmbH,Germany
Fridge MediLine	Liebherr-International Deutschland GmbH,Germany
Heating Chamber	Binder GmbH, Germany
Helios δ Spectrophotometer	Thermo Fisher Scientific Inc., USA
Heraeus® Incubator	Thermo Fisher Scientific Inc., USA
IKA® Vortex 3	IKA®-Werke GmbH & CO. KG, Germany
Incubator Infors HT Multitron II	Infors AG, Switzerland
Incubator MIR-254	Sanyo Electric Co., Ltd., Japan
Innova® ULT Upright Freezer U725	Eppendorf AG, Germany
inoLab® pH 7110	Xylem Analytics Germany Sales GmbH & Co. KG, Germany
Isotemp™ Wather Bath 220	Fisher Scientific Company LLC, USA
Magnetic Stirrer MR Hei-Standard	Heidolph Instruments GmbH & Co.KG, Germany

Micro Pulser™ Electroporator	Bio-Rad Laboratories, Inc., USA
Mini Centrifuge 3722L	Fisher Scientific Company LLC, USA
Mini Power Pack PS300T	Biometra GmbH, Germany
Mini-Protean® 3 Cell	Bio-Rad Laboratories, Inc., USA
Mini-Sub® Cell GT Cell, #1704467	Bio-Rad Laboratories, Inc., USA
Molecular Imager® Gel Doc™ XR System	Bio-Rad Laboratories, Inc., USA
Multipette® M4	Eppendorf AG, Germany
NanoDrop® ND-1000 Spectrophotometer	Thermo Fisher Scientific Inc., USA
Pipettes, overall range 0.2 – 1000 µl Gilson Pipetman®	Gilson, Inc., USA
Pipettor pipetus®	Hirschmann Laborgeräte GmbH & Co. KG, Germany
PowerPac™ Basic Power Supply	Bio-Rad Laboratories, Inc., USA
Realplex 2 qPCR cycler	Eppendorf AG, Germany
Rotor 11222	SIGMA Laborzentrifugen GmbH, Germany
Rotor 12171	SIGMA Laborzentrifugen GmbH, Germany
Rotor 19776	SIGMA Laborzentrifugen GmbH, Germany
Rotor F45-24-11	Eppendorf AG, Germany
SenTix® 81 precision electrode with temperature sensor	Xylem Analytics Germany Sales GmbH & Co. KG, Germany
Small shaker VXR basic Vibrax® with VX7 dish attachment	IKA®-Werke GmbH & CO. KG, Germany
Thermocycler T3000	Biometra GmbH, Germany
Thermomixer compact	Eppendorf AG, Germany
Vortex-Genie® 2	Scientific Industries, Inc, USA
Wide Mini-Sub® Cell GT Cell, #1704469	Bio-Rad Laboratories, Inc., USA
Zirbus autoclave	ZIRBUS technology GmbH, Germany

7.2 Chemicals

Table 2: Chemicals

Chemical	Article No.	Supplier
$(NH_2)_2SO_4$	3746.3	Carl Roth GmbH + Co. KG, Germany
2-Propanol	1096342511	Merck KGaA, Germany
Acetamide	8223431000	Merck KGaA, Germany
Acrylamide	0189.2	Carl Roth GmbH + Co. KG, Germany
Agar Noble	214230	Becton, Dickinson and Company, USA
Agar-Agar	1016141000	Merck KGaA, Germany
Ampicillin sodium salt	K029.2	Carl Roth GmbH + Co. KG, Germany
APS, SIGMA-ALDRICH®	A3678	Merck KGaA, Germany
Bisacrylamid 4K ul-trapure	A1095	AppliChem GmbH, Germany
$CaCl_2$ * 2 H_2O	5239.2	Carl Roth GmbH + Co. KG, Germany
Cellulose	Arbocel B800	J. RETTENMAIER & SÖHNE GmbH + Co KG, Germany
Chloroform	3313.1	Carl Roth GmbH + Co. KG, Germany
$CoCl_2$ * $6H_2O$, SIGMA-ALDRICH®	203084	Merck KGaA, Germany
CsCl	8627.1	Carl Roth GmbH + Co. KG, Germany
EDTA	2504716	Merck KGaA, Germany
Ethanol	1009712500	Merck KGaA, Germany
$FeSO_4$ * 7 H_2O	1039650100	Merck KGaA, Germany
Glacial acetic acid	3738.1	Carl Roth GmbH + Co. KG, Germany
Glucose monohydrate	6780.2	Carl Roth GmbH + Co. KG, Germany

iQ SYBR Green Super-mix	170-8882	Bio-Rad Laboratories, Inc., USA
Isoamyl alcohol	1009781000	Merck KGaA, Germany
K-Acetate	T874.2	Carl Roth GmbH + Co. KG, Germany
KH_2PO_4	P018.2	Carl Roth GmbH + Co. KG, Germany
Methanol	10600092500	Merck KGaA, Germany
$MgSO_4$ * 7 H_2O	P027.2	Carl Roth GmbH + Co. KG, Germany
$MnSO_4$ * H_2O	1059410250	Merck KGaA, Germany
NaCl	3957.2	Carl Roth GmbH + Co. KG, Germany
Peptone	1022390500	Merck KGaA, Germany
Potato dextrose agar	213200	Merck KGaA, Germany
RNAse	EN0531	Thermo Fisher Scientific Inc., USA
Rotiphorese® 50x TAE buffer	CL86.2	Carl Roth GmbH + Co. KG, Germany
SDS	4360.2	Carl Roth GmbH + Co. KG, Germany
SERVA Blue R	35051	SERVA Electrophoresis GmbH
Sorbitol, CHEMSO-LUTE® bio	81251000	Th. Geyer GmbH & Co. KG, Germany
StarPure Agarose > 1kb	N3101-0500	STARLAB GmbH, Ger-many
TEMED	2367.1	Carl Roth GmbH + Co. KG, Germany
TRIS	5429.3	Carl Roth GmbH + Co. KG, Germany
Triton X-100	1122981001	Merck KGaA, Germany
Tween 80	8221870500	Merck KGaA, Germany
Urea	3941.1	Carl Roth GmbH + Co. KG, Germany
Water for analysis EMSURE®	1167545000	Merck KGaA, Germany

Water, nuclease free	T143.5	Carl Roth GmbH + Co. KG, Germany
Yeast extract	2363.2	Carl Roth GmbH + Co. KG, Germany
$ZnSO_4$ * 7 H_2O	1088830500	Merck KGaA, Germany

7.3 Disposables

Table 3: Disposables

Disposable	Article No.	Supplier
0.1-10 µl pipette tips	37650	Sorenson BioScience, Inc., USA
10-200 µl pipette tips	739290	Greiner Bio-One International GmbH, Austria
200-1000 µl pipette tips	740290	Greiner Bio-One International GmbH, Austria
96 well plates	2239444	Bio-Rad Laboratories, Inc., USA
Conical centrifuge tubes 15 ml	188271	Conical Centrifuge Tubes 50 ml
Conical centrifuge tubes 50 ml	N463.1	Carl Roth GmbH + Co. KG, Germany
Conical centrifuge tubes 50 ml, with skirt	210261	Greiner Bio-One International GmbH, Austria
Eppendorf combitips advanced® 2.5 ml	0030089448	Eppendorf AG, Germany
Filter tips 10µL	771288	Greiner Bio-One International GmbH, Austria
Filter tips 100µL	772288	Greiner Bio-One International GmbH, Austria
Filter tips 1000µL	750288	Greiner Bio-One International GmbH, Austria
Gene Pulser®/MicroPulser™ Electroporation cuvettes, 0.2 cm gap	1652092	Bio-Rad Laboratories, Inc., USA
Glass wool	6574.1	Carl Roth GmbH + Co. KG, Germany

Micro tubes 1.5 ml	72.690.001	Sarstedt AG & Co, Germany
Micro tubes 2 ml	0030120094	Eppendorf AG, Germany
Miracloth Calbiochem®	475855-1R	Merck KGaA, Germany
Optical sealing tapes	2239441	Bio-Rad Laboratories, Inc., USA
PCR SingleCap 8er-SoftStrips 0.2 ml	710970	Biozym Scientific GmbH, Germany
PCR Tubes 0.2 ml	B1402-5500	STARLAB GmbH, Germany
Petri dish 35/10 mm, with vents	627102	Greiner Bio-One International GmbH, Austria
Petri dish 60/15 mm, with vents	628102	Greiner Bio-One International GmbH, Austria
Petri dish 94/16 mm, with vents	633180	Greiner Bio-One International GmbH, Austria
Pipettes 10 ml, sterile	607180	Greiner Bio-One International GmbH, Austria
Pipettes 25 ml, sterile	760180	Greiner Bio-One International GmbH, Austria
Semi-micro cuvettes PMMA 1.5 ml	759115	BRAND GMBH + CO KG, Germany

7.4 Media

Trace Elements (50x):	0.25 g/l $FeSO_4 * 7 H_2O$
	0.085 g/l $MnSO_4 * H_2O$
	0.07 g/l $ZnSO_4 * 7 H_2O$
	0.1 g/l $CoCl_2 * 6H_2O$
Adjusted to pH 2 with concentrated sulphuric acid	

LB Media (+Amp):	10 g/l peptone
	5 g/l yeast extract
	10 g/l NaCl
	1000 µl ampicillin (100 mg/ml) /l
If used as solid medium:	addition of 15 g agar-agar /l

PDA:	39 g/l potato dextrose agar

YPD:	10 g/l yeast extract
	20 g/l peptone
	20 g/l glucose

amdS Medium:	1 g/l $MgSO_4$ * 7 H_2O
	10 g/l KH_2PO_4
	10 g/l glucose
	15 g/l Agar Noble
	20 ml trace elements (50 x) /l
	12.5 ml 1M CsCl per 1 L
	10 ml 1M acetamide per 1 L
If used for spore isolation, add:	0.1 % (v/v) Triton X-100

Mandels-Andreotti Medium:	2 g/l KH_2PO_4
	1.4 g/l $(NH_4)_2SO_4$
	0.3 g/l $MgSO_4$ * 7 H_2O
	0.3 g/l $CaCl_2$ * 2 H_2O
	0.3 g/l urea
	1 g/l peptone
	0.5 g/l Tween 80
	20 ml trace elements (50x) /l
	10 g/l cellulose

YPDS Medium:	4 parts 1.1 M sorbitol
	1.25 parts YPD

SOC Medium:	# 636763, Clontech Laboratories, Inc.

7.5 Buffers and Solutions

Lysis Buffer:	48.5 g/l TRIS-HCl

	17.5 g/l EDTA
	8.8 g/l NaCl
	Adjusted to pH 8 with HCl, then add
	10 g/l SDS

K-Acetate Buffer pH 4.8 (100 ml):	60 ml 5M K-Acetate solution
	11.5 ml glacial acetic acid
	28.5 ml dH_2O

Extraction Buffer pH 7.5 (100 ml):	8.18 g NaCl
	10 ml 1M TRIS (pH 7.5 with HCl)
	1 ml 0.5M EDTA (pH 7.5 with NaOH)
	add 100 ml dH_2O

6x Loading Buffer, 50ml:	7.5 g Ficoll 400
	0.125 g Bromphenolblue
	6 ml 0.5M EDTA, pH 8.0
	add 50 ml dH_2O

Acrylamide Stock Solution 30%:	300 g/l acrylamide
	4 g/l bisacrylamide

Stacking Gel Buffer:	60 g/l TRIS
	4 g/l SDS
	Adjusted to pH 6.8 with HCl

Separating Gel Buffer:	180 g/l TRIS
	4 g/l SDS
	Adjusted to pH 8.8 with HCl

10x Running Buffer:	60 g/l TRIS

	144 g/l glycine
	10 g/l SDS

4x Laemmli Sample Buffer (10 ml):	2.5 ml 1M TRIS-HCl pH 6.8
	0.8 g SDS
	2 ml β-mercaptoethanol
	0.02 g bromphenol blue
	4 ml glycerol
	1.5 ml dH_2O

Colloidal Coomassie Stain, 100 ml:	8 g $(NH_4)_2SO_4$
	78 ml dH_20
	2 ml 85% H_3PO_4
	0.08 g Colloidal Coomassie Brilliant Blue G-250 mixed with 1.6 ml dH_2O in a 2 ml micro tube
	20 ml methanol were added before usage

NaCl-Tween	0.9 % NaCl
	0.05 % Tween 80

7.6 Strains

For transformation and quantitative production of the plasmid pBS_amdS_Cel45a, *E. coli* Stellar™ competent cells (#636763, Clontech Industries, Inc.) were used.

In this work the low cellulase producing wildtype strain *T. reesei* QM6a (ATCC 13631) and the classical mutant strain *T. reesei* Rut C30 (ATCC 56765), a high cellulase producer were used for transformation.

7.7 Plasmid and Primers

The gene of interest, *cel45a*, was provided by Roal Ltd and is integrated into the plasmid pBS_amdS_Cel45a. The plasmid contains also the selection marker *amdS* of *Aspergillus nidulans,* encoding for an acetamidase (Penttilä et al. 1987). *T. reesei* grows only poorly on acetamide and therefore the acetamidase can be used as selection marker. During purification, selection plates were used which contain acetamide as the only nitrogen source. For a strong expression of *cel45a* the expression signals of the cellobiohydrolase 1 (encoded by *cel7a*) were used (Ilmen et al. 1996a). The *cel45a* gene was therefore cloned between the promoter and terminator region of *cel7a*. A scheme of the expression cassette is depicted in Fig. 5.

Fig. 5: Schematic representation of plasmid pBS_amdS_Cel45a. pBS_amdS_Cel45a contains the *A. nidulans amdS* selection marker and the *cel45a* expression cassette. The selection marker gene *amdS* on the plasmid is about 3850 bp long. The *cel45a* expression cassette consists of 1 kb of the cellobiohydrolase 1 promotor region P*cel7a,* followed by the 1076 bp *cel45a* coding region and the 512 bp cellobiohydrolase 1 terminator region T*cel7a*.

For amplification and linearization of plasmid pBS_amdS_Cel45a the M13 primers were applied. Following transformation and purification, the transformants were checked with the two cbh1_genot primers. All primers are listed in Table 4.

Table 4: Primers

Primer name	5' – 3' sequence	Melting temperature
M13 fw	GTAAAACGACGGCCAGT	59 °C
M13 rv	CAGGAAACAGCTATGAC	52 °C
cbh1_genot_fw	ATGAGCTAGTAGGCAAAGTCA	59 °C
cbh1_genot_rv	CATCTCCAGTGAAAGATGAGT	59 °C
B_amdS_qPCR_fw	AAGGCGGTTAGTGAGCTTGA	64 °C
B_amdS_qPCR_rv	TGAGTCTCCGTCCGATAACC	64 °C
qPCR cbh2 fw	ACTACAACGGGTGGAACATTAC	62 °C
qPCR cbh2 rv	CGTGGATGTACAGCTTCTCG	64 °C

8. Methods

8.1 Transformation of *E. coli*

For transformation 5 µl of the plasmid pBS_amdS_Cel45a were inoculated with 50 µl of competent *E. coli* cells on ice for 30 minutes, followed by one minute heat shock at 42 °C. Afterwards 500 µl SOC medium was added and the mixture was incubated for one hour at 37 °C and 400 rpm. The incubated SOC medium was distributed on three LB plates containing 100 µg/l ampicillin using different aliquots (60 µl, 140 µl, 300 µl) and incubated o/n at 37 °C.

8.2 Plasmid Midi-Preparation

For plasmid Midi-Preparation 50 ml liquid LB medium with 100 µg/l ampicillin were inoculated with a single *E. coli* colony and incubated overnight at 37 °C shaking at 180 rpm.

Plasmid Midi-Preparation was done according to the protocol provided with the Promega Kit A2495 “Pure Yield Plasmid Midiprep System”.

8.3 PCR

For PCR amplification the Phusion High-Fidelity DNA Polymerase (2 U/µl) provided with 5x Phusion HF Buffer; (#F530S, Thermo Fisher Scientific Inc.) was employed. However, for diagnostic PCRs Phire Hot Start II DNA Polymerase, provided with 5x Phire reaction buffer; (#F122S, Thermo Fisher Scientific Inc.), was the polymerase of choice. All primer sequences used are listed in chapter 7.7. PCR reactions were mixed and cycling conditions for denaturation and elongation were selected, as suggested in the respective protocols of the polymerases.

8.4 Agarose Gel Electrophoresis

In order to check the sizes of amplified DNA fragments they were loaded onto agarose gels and run in TAE buffer. If not specified differently, 1% agarose gels

were prepared and stained with 3-5 µl SYBR® Safe (#S33102, Thermo Fisher Scientific Inc.). The GeneRuler 1 kb DNA ladder (#SM0313, Thermo Fisher Scientific Inc.) was used as marker. Samples were loaded with 6x loading dye. Gels were run with about 6 Volts/cm if not stated otherwise.

8.5 Purification of DNA by Ethanol Precipitation

For DNA purification 3M NaOAc (1/10 of the sample volume) and 96% ethanol (2x the sample volume) were added to the sample. The micro tube was then stored o/n at -20 °C. On the next day, the samples were centrifuged at 16,000 x g at 4 °C for 30 minutes. Afterwards the supernatant was discarded and the pellet washed with 300 µl of 70 % ethanol, followed by centrifugation at 16,000 x g at 4 °C for 20 minutes. The washing step with 70 % ethanol was done twice. In a final step, after discarding the supernatant, the pellet was reconstituted in 10 µl water (EMSURE®).

8.6 Preparation and Transformation of Electrocompetent Spores of *T. reesei*

Transformation was done as described (Schuster et al. 2012). Therefore the *T. reesei* strain was cultured on PDA for 8 to 14 days in an 8 cm Petri dish. The fully sporulated plate was harvested with a NaCl-Tween solution and then filtered through a 1.5 ml micro tube filled with glass wool into a 15 ml conical centrifuge tube. The tube was vortexed vigorously and the spore solution poured into a sterile 250 ml flask filled with 100 ml YPD. The flask thereafter was incubated at 30 °C at 300 rpm in an orbital shaker for 4 to 6 hours (*T. reesei* QM6a 4 hours, *T. reesei* Rut C30 6 hours). After incubation, the spore suspension was split into two 50 ml conical centrifuge tubes and centrifuged for 5 minutes at 2200 x g at 4 °C. Subsequently the supernatant was discarded and the cell pellet washed with 25 ml 1.1 M cold sorbitol. Following a centrifugation step for 5 minutes at 2200 x g at 4 °C. The supernatant was discarded again and the cells were resuspended in 1 ml 1.1 M cold sorbitol. For the next working steps the suspension was transferred into 2 ml micro tubes and centrifuged for 5 minutes at 1500 x g at 4 °C. After discarding the supernatant 1 ml 1.1 M cold sorbitol

was added and centrifuged again with the same settings as before. Finally, the supernatant was discarded and the pellet resuspended in 150 µl. Out of these, two 75 µl aliquots were transferred into 1.5 ml micro tubes and used for electroporation.

One cooled 75 µl aliquot of electrocompetent spores was gently mixed with 10 µl linearized DNA and incubated on wet ice for 30 minutes. Afterwards the cell suspension was transferred into a sterile electroporation cuvette and electroshocked with 1.8 kVolt for 6 milli-seconds. For recovering, 525 µl YPDS was added to the cuvette and the suspension transferred back to the micro tube, followed by two hours incubation at 28 °C at 800 rpm on a thermomixer. The selective medium for *amdS* selection was MM glucose containing 10mM acetamide ("amdS medium"). After incubation, 150 µl aliquots were plated on amdS medium and incubated at 28 °C in the dark. $AmdS^+$ transformants were single spore isolated on plates containing acetamide and for colony growth restriction (for selective streaking plates) 0.1 % Triton X-100 was added to the medium (Penttilä et al. 1987). After single spore isolation the transformants were transferred to PDA growing at 28 °C in the dark.

8.7 Rapid Genomic DNA Extraction

Some mycelium was scraped off from fully sporulated PDA plates of *T. reesei* and put into a micro tube containing 500 µl lysis buffer. The micro tube was mixed shortly and placed on the bench for ten minutes at room temperature. Afterwards 150 µl K-acetate buffer was added, followed by intense vortexing and one minute centrifugation at 16,000 x g. The supernatant was poured into a new micro tube and centrifuged again with the same settings as before. Then the supernatant was transferred into a new micro tube and about 550 µl 2-propanol was added. For mixing, the micro tube was turned upside down twice and finally centrifuged two minutes at 16,000 x g. After discarding the supernatant, the DNA pellet was washed with 300 µl 70 % ethanol and dried at room temperature with the micro tube cap open. At the end the pellet was dissolved in 50 µl water (EMSURE®).

8.8 Determination of Spore Concentration

A fully sporulated plate *T. reesei* was harvested with NaCl-Tween by filtering the spore solution through a 1.5 ml micro tube filled with glass wool into a 15 ml conical centrifuge tube. Thereafter, 1 ml NaCl-Tween was pipetted in a 1.5 ml micro tube. After vortexing the spores solution, 75 µl of the concentrated conidia suspension were added to the 1 ml NaCl-Tween solution. This diluted suspension was vortexed and the absorbance (Abs) measured in a photometer at 600 nm. A standardised curve was set up by using a hemocytometer.

8.9 Growth Test

T. reesei strains were grown on PDA for around ten days. Then a piece of 4 mm in diameter of each plate was cut out and placed onto a Petri dish containing 20 ml PDA, which was marked with a cross in the middle of the plate. On the following days the strains were grown at 28 °C in the dark and pictures were taken starting from 24 hours (*T. reesei* QM6a and respective transformants) and 72 hours (*T. reesei* Rut C30 and respective transformants) after inoculation with the agar piece. Additionally, the radial growth was measured. The interval of taking photos and measuring the radial growth was 24 hours, until the mycelium reached the edge of the Petri dish. When the plates were fully covered with sporulating mycelium the plates were again photographed and then harvested, as described in chapter 8.8, with the only difference, that the NaCl-Tween amount for harvesting was the same amount for each strain.

8.10 Shake Flask Cultivation for Extracellular Protein Production

For checking the extracellular protein production of the transformants and their parental strains they were cultivated in 500 ml Erlenmeyer flasks filled with 100 ml MA media + 1 % (w/v) cellulose. The medium was inoculated with 10^6 spores/ml and the flasks were incubated at 28 °C, 250 rpm for five days (protected from direct daylight). Sampling was done after 72, 96 and 120 hours following inoculation. About 14 ml of the cultivation broth were collected and stored on wet ice. Afterwards this 14 ml portion was filtered through Miracloth and centrifuged at 10 °C for five minutes at a speed of 5000 x g to pellet re-

maining fungal cells and non-metabolised cellulose. The remaining clear solution was stored on wet ice and aliquoted into micro tubes in different volumes (500 µl, 1000 µl) for upcoming analyses. The aliquots were frozen at -20 °C.

8.11 Chloroform/Methanol Protein Precipitation

Protein precipitation was done as described by Wessel and Flugge (Wessel and Flugge 1984) to obtain a sufficient amount of proteins for the SDS-PAGE gel electrophoresis. Therefore, a 220 µl aliquot of the clear cultivation broth was taken and mixed with 880 µl methanol. Following vortexing, 220 µl chloroform were added. Then the 2 ml micro tube was vortexed and 660 µl water EMSURE® were pipetted into the tube, followed by vigorous vortexing. Now the micro tube was centrifuged five minutes at RT at 16 000 x g. Two phases were visible after centrifugation and in between the desired precipitated proteins were slightly visible. After discarding the upper phase, 660 µl methanol were added. For mixing the solutions, the micro tube was vortexed again, followed by centrifugation with the same settings as before. In the next step the supernatant was discarded. The micro tube was left open on the bench for drying. After the pellet was dried, it was solved in 30 µl of 20 mM TRIS-HCl pH 7.5. The concentrated protein solution was stored at -20 °C pending for concentration analyses, followed by loading onto SDS-PAGE gels.

If more aliquots were necessary to achieve a suitable protein concentration for one sample, the dried proteins of a suitable number of aliquots were dissolved one by one in 30 µl (total volume) of 20 mM TRIS-HCl pH 7.5.

Because only a few µl of the concentrated protein mixture were available, a NanoDrop ND-1000 Spectrophotometer was used to measure the protein concentration. The default setting was “1 Abs” for single measurements. As these values are not directly comparable to other protein measurement methods, the assay is stated here as semi-quantitative.

8.12 Sodium Dodecyl Sulfate - Polyacrylamide Gel Electrophoresis (SDS-PAGE)

A SDS-PAGE was run to separate the proteins according to their molecular weight. Therefore, a SDS-polyacrylamide mini gel was casted. At the top of a polymerized 12 % separating gel, a stacking gel was filled in the gel chamber. Before loading the samples onto the gel, they were mixed with 4x Laemmli buffer and heated up to 95 °C for 5 minutes. Afterwards an appropriate amount of the treated sample corresponding to about 10 µg total protein was loaded into a single well of the gel. A protein ladder (PageRuler™ #26616, Thermo Fisher Scientific, Inc.) was used for a qualitative check. The gel was run at 150 Volts for 60 minutes. Recipe for two mini-gels is shown below in Table 5.

Table 5: Recipe for two 12 % SDS-PAGE Gels

	stacking gel	**separating gel 12 %**
dH_2O	3 ml	4.2 ml
Specific gel buffer	1.25 ml	3 ml
Acrylamide 30 %	0.75 ml	4.8 ml
10 % APS	25 µl	60 µl
TEMED	5 µl	10 µl

8.13 Endo-1,4-β-Glucanase Activity Assay using Azo-CM-Cellulose

The endo-1,4-β-glucanase activity assay was performed according to the booklet provided with the Azo-CM-Cellulose powder (S-ACMC, Megazyme u.c.). For testing thermal stability of the expressed protein, the temperature of incubation was varied. All measurements were executed in a water bath. Details see chapter 9.6.

8.14 In silico Analysis

The gene *cel45a* provided from Roal Ltd was analysed in silico to determine possible glycosylation sites and to calculate the theoretical molecular weight. In a first step the nucleotide sequence was aligned with the *cel45a* gene of *Neurospora crassa* OR74A (GenBank XM_952014.2) to determine the localization of introns and the coding sequence. Based on the coding sequence the gene was translated into its amino acid sequence and the molecular weight was calculated. For all steps listed above the software CLC Main Workbench (https://www.qiagenbioinformatics.com) was used. The signal peptide at the N-terminal end of the protein was predicted with SignalP 4.1 at http://www.cbs.dtu.dk/services/SignalP/ (Petersen et al. 2011). The prediction of the secondary and tertiary structure of the translated protein CEL45A was done by SWISS-MODEL on https://swissmodel.expasy.org/ (Biasini et al. 2014). Potential amino acid residues for glycosylation were analysed with NetNGlyc 1.0 (Technical University of Denmark DTU) and NetOGlyc 4.0 (Steentoft et al. 2013) embedded at http://www.cbs.dtu.dk/services/ in section "Post-translational modifications of proteins". Similarity comparison to other proteins was done using blastp (Altschul et al. 1997); https://blast.ncbi.nlm.nih.gov/Blast.cgi). The protein sequence alignment was performed using T-Coffee [http://tcoffee.crg.cat/, (Di Tommaso et al. 2011)] and the respective secondary structure was generated with ESPript 3.0 [(Robert and Gouet 2014), http://espript.ibcp.fr/ESPript/ESPript/]. For testing potential disulphide bonds, the tools DISULFIND http://disulfind.dsi.unifi.it/ (Ceroni et al. 2006) and DIpro at http://scratch.proteomics.ics.uci.edu/ (Cheng et al. 2006) were employed.

8.15 Isolation of *T. reesei* Chromosomal DNA

DNA was isolated as described (Seiboth et al. 2002). The mycelia were filtered through Miracloth and dried by pressing it between filter papers. The dried mycelia were then frozen in liquid nitrogen followed by grinding with a mortar and pestle in liquid nitrogen. Afterwards the biomass powder was put into a 2 ml tube filled with 800 µl extraction buffer containing 2 µl RNase A (10 mg/ml). The

sample was always kept in liquid nitrogen, preventing the sample from thawing. In a next step the samples were vortexed at 65 °C for 20 minutes at 400 rpm on a thermocycler. After placing the samples on ice every tube was prepared separately. For liquid extraction 400 µl phenol were added and vortexed. Afterwards 400 µl chloroform:isoamyl alcohol (49:1) were added and mixed. The preparation was continued with centrifugation for 15 minutes at 4 °C with 12 000 rpm. The upper phase was then transferred into a new vial containing 750 µl chloroform:isoamylalcohol (49:1). After repeating the centrifugation the supernatant was added to a vial containing 750 µl isopropanol. To pellet the precipitate, the samples were centrifuged again with the same settings as described above. The supernatant was discarded and the pellet washed with 70 % ethanol. After centrifugation at room temperature the pellet was dried for five minutes and afterwards dissolved in 100 µl H_2O. For the digestion of remaining RNA, 2 µl RNase A (10 mg/ml) were added and incubated one hour at 37 °C.

8.16 qPCR Analysis

An Eppendorf Mastercycler® RealPlex^2 was used for quantitative real time PCR (qPCR) to measure the number of gene copies integrated into the *T. reesei* genome. The copy numbers of *cel45a* were determined indirectly by normalizing *amdS* to *cbh2*, which was used as a single copy reference gene. Primers were designed (D'haene et al. 2010) and listed in section 7.7. Each reaction was loaded with a total volume of 25 µl, containing 12.5 µl SYBR® Green Supermix, 10.5 µl nuclease free water and 1 µl each of a primer dilution and the respective DNA sample. All reactions were performed in triplicates and the Relative Expression Software Tool REST® (QIAGEN, USA) was used to analyse the data (Pfaffl et al. 2002).

8.17 MS Analysis

Mass Spectrometer analysis was performed by Dr. Clemens Grünwald-Gruber. The protein band was excised from a SDS-PAGE gel and digested in gel with trypsin (carbamidomethylation of cysteins was done in gel prior digestion). The peptide mixture was loaded on a BioBasic C18 column (BioBasic-18, 150 x 0.32

mm, 5 µm, Thermo Fisher Scientific Inc.) using 65 mM ammonium formate buffer. A gradient from 5 to 75 % acetonitrile was applied over 55 min at a flow rate of 6 µL/min. Detection was performed with a QTOF MS equipped with the standard ESI source (maXis^{TM} 4G ETD, Bruker Corporation) in the positive ion, DDA mode (= switching to MS/MS mode for eluting peaks). Instrument calibration was performed using ESI calibration mixture (Agilent Technologies, Inc.). The analysis files were converted (using Data Analysis, Bruker Corporation) to mgf files, which are suitable for performing a MS/MS ion search with the Global Proteome Machine (GPM). GPM is a web-based, open source user interface for analyzing and displaying protein identification data. The interface creates a series of web browser page views of tandem mass spectrometry data that has been assigned to protein sequences. Additionally, spectra were interpreted manually.

9. Results

9.1 In silico Analysis of *cel45a* and its encoding protein

The sequence of the *cel45a* gene was obtained from Roal Ltd. To determine some of its properties and of its encoded protein, we performed an in silico analysis. The introns in *cel45a* were thereby identified through alignment to similar gene sequences and searching for GT-AG donor/acceptor sites. Two introns of the size of 59 and 126 bp were identified. The exons were then joined and the resulting DNA sequence translated into its amino acid sequence. A model of the translated protein (Fig. 6) was then generated with SWISS-Model (Biasini et al. 2014).

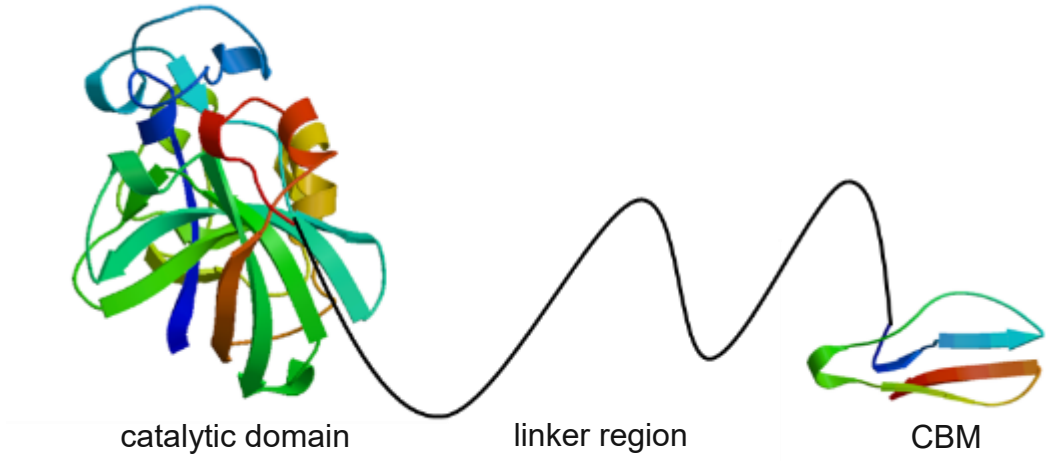


Fig. 6: CEL45A predicted model performed by SWISS-Model. The catalytic domain (CD) of CEL45A is connected via a linker to the family 1 carbohydrate binding modul (CBM) which is located at the C-terminal end of the protein.

The calculated molecular weight of the protein is 30758.87 Da. Removal of the signal peptide results in reduction to 28625.33 Da of the mature protein. Following a blastp of the amino acid sequence of CEL45A against the NCBI database a multiple sequence alignment was performed with similar amino acid sequences. The consensus sequence of CEL45A showed high similarities with other members of the GH45 family.

Since CEL45A is an extracellular enzyme, we also tested if putative recognition sites for post translational modifications such as N- and O-glycosylations which are abundantly found in different fungal cellulases, are present. The presence of

these potential glycosylation sites was tested using the online tools NetNGlyc 1.0 and NetOGlyc 4.0. Only one possible N-glycosylation site was detected, whereas a high number of possible O-glycosylation sites could be found. A high glycosylation potential was predicted for the linker region which is found between amino acid 235 and 260 (Fig. 7).

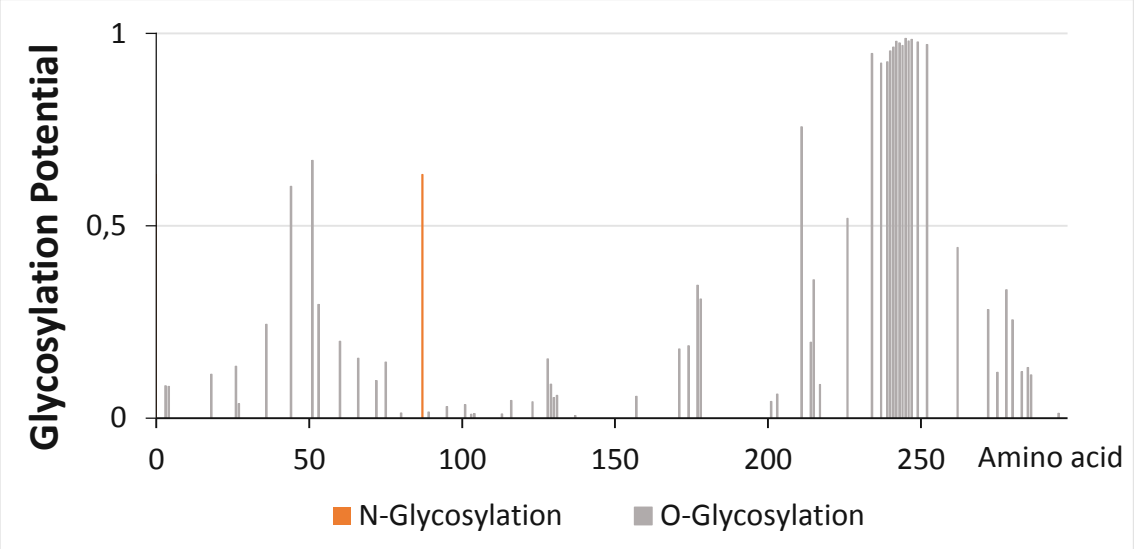


Fig. 7: Glycosylation potential of CEL45A amino acid residues. The glycosylation potential was determined by using NetNGlyc 1.0 and NetOGlyc 4.0. A putatively highly glycosylated area is found in the linker region. O-Glycosylation is marked grey and N-Glycosylation marked orange. The default threshold for glycosylation is 0.5. A value above 0.5 represents therefore a predicted glycosylated site.

The high O-glycosylation in the linker region aids in the binding of the cellulase to the cellulose (Payne et al. 2013). It also increases the MW of the extracellular protein. For that reason, this high glycosylation is a possible explanation, why the band of CEL45A is found at 37 kDa (see below).

In addition, we also investigated if the CEL45A protein contains disulphide bonds found between cysteines using different online tools (DISULFIDE, DIpro). Within the catalytic domain 6 to 7 and within the CBM about three disulphide bonds are possible. This high amount of disulphide bridges (covalent bond) leads to the prediction that the protein could be thermostable (Beeby et al. 2005).

9.2 Transformation of *T. reesei* QM6a and Rut C30 with *cel45a*.

For the expression of *cel45a* in *T. reesei* two different host strains were used. One was the low cellulase producing wild-type strain QM6a, the other strain was the high cellulase producing strain RutC30. To guarantee high expression of *cel45a* the coding region of the gene was cloned under the expression signals of the major cellobiohydrolase 1 resulting in plasmid pBS_amdS_Cel45a. In addition the plasmid contains the *A. nidulans amdS* for selection on acetamide as sole nitrogen source (see Fig. 5). Cloning of *cel45a* in the expression vector was done by Roal Ltd. The final plasmid was then transformed into *E. coli* and extracted and purified through plasmid Midi-Preparation. The plasmid was then linearized through PCR followed by precipitation of the fragment. To introduce the linearized fragment into *T. reesei* strains QM6a and Rut C30, electrocompetent spores were prepared (Schuster et al. 2012) and transformed with the linearized plasmid. Transformants were then selected on amdS medium. Obtained transformants were purified by single spore isolation on amdS selection plates containing Triton-X100 (0.1 %) and then transferred to PDA plates. In total 28 QM6a transformants were selected for the selection process and about 35 % showed stable growth and were selected for further screenings. For Rut C30 45 transformants were selected, with about 45 % left for further analysis.

The purified acetamide utilizing transformants of strains QM6a and RutC30 were then checked for integration of the *cel45a* expression cassette by diagnostic PCR. For this analysis primers were used which bind in the *cel7a* (*cbh1*) promoter and terminator region. Amplification of the endogenous *cel7a* results therefore in a band of 2 kb. In the case the expression cassette of pBS_amdS_Cel45a is present in the transformants, an additional band of around 1400 bp representing the *cel45a* expression cassette will be amplified. Results of the diagnostic PCR for the respective transformants are presented in Fig. 8. Following this analysis 6 positive transformants were identified for strain QM6a (designated Q2.1, Q2.4, Q2.3, Q2.6, Q2.9, Q2.21) and 18 positive transformants for RutC30 (designated R1.2, R1.3, R2.4, R2.5, R2.7, R2.8, R2.9, R2.13, R2.17, R2.18, R2.20, R2.24, R2.26, R2.27, R2.32, R2.34, R2.35, R2.36).

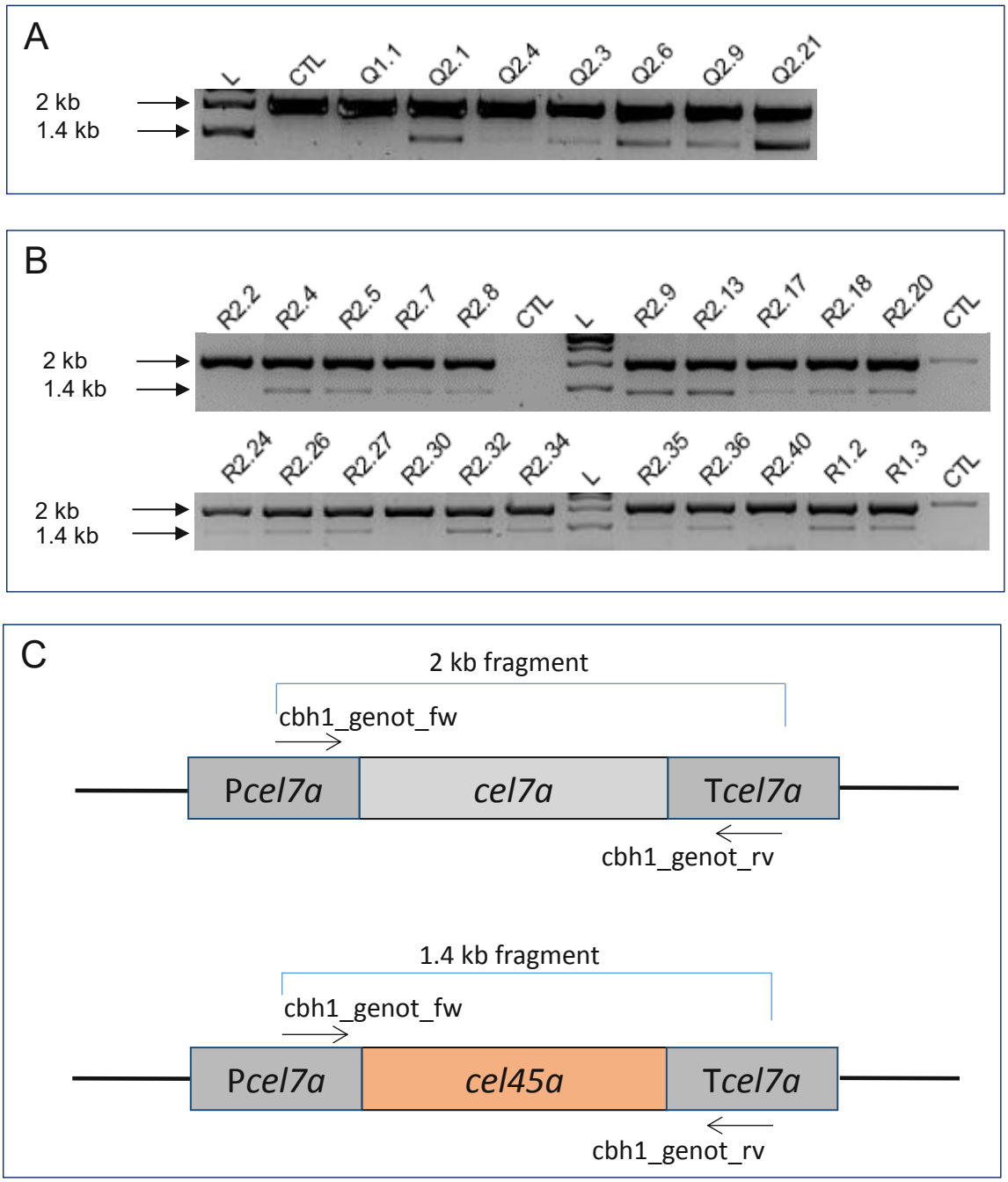


Fig. 8: Diagnostic PCR of selected *T. reesei cel45a* transformants. To test if the expression cassette of *cel45a* is present in the different *T. reesei* transformants primers were used which bind in the *cel7a* promoter and terminator region. Amplification with these primers results in a 2 kb band for the endogenous *cel7a* in *T. reesei* QM6a **(A)** and RutC30 **(B)**. Presence of the *cel45a* expression cassette results in a band of 1.4 kb in *cel45a* transformants of *T. reesei* strain QM6a **(A)** and Rut C30 **(B)**. The forward primer cbh1_genot_fw binds 162 bp upstream of the start codon and therefore within the promotor region P*cel7a*. The reverse primer cbh1_genot_rv binds 154 bp following the stop codon within the terminator region T*cel7a* **(C)**.

In the next step the copy number of the positive transformants was quantified by qPCR. Within this analysis the *cel45a* gene was detected indirectly (details see chapter 8.16) and quantified within the genomes of the transformants. Com-

pared to the expression of the single copy reference gene *cbh2*, the different expression and therefore the gene copies integrated into the genome can be calculated. The detailed results of the qPCR are found in Appendix 12.1. In Table 6 the gene copy numbers of the selected *cel45a* transformants are depicted.

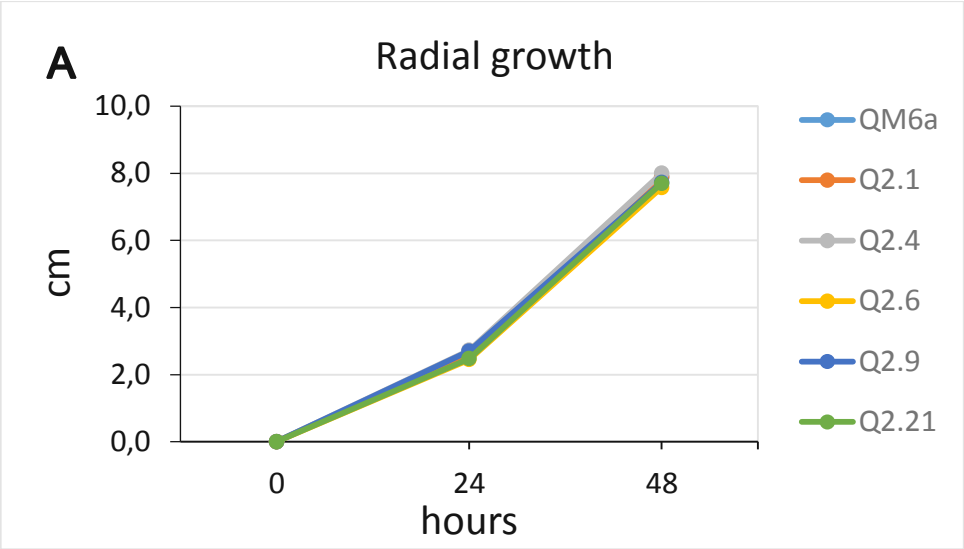
Table 6: Estimated *cel45a* genome copy number of selected *T. reesei* QM6a and RutC30 transformants

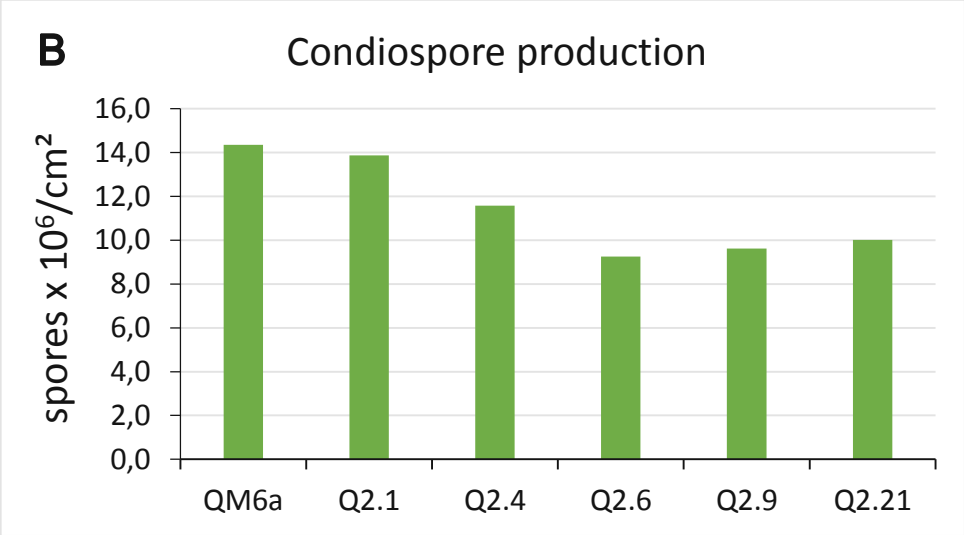
QM6a *cel45a* transformants	*cel45a* copies	Rut C30 *cel45a* transformants	*cel45a* copies
Q2.1	2	R2.4	2
Q2.4	1	R2.5	2
Q2.6	1	R2.9	3
Q2.9	1	R2.32	1
Q2.21	4	R2.34	1

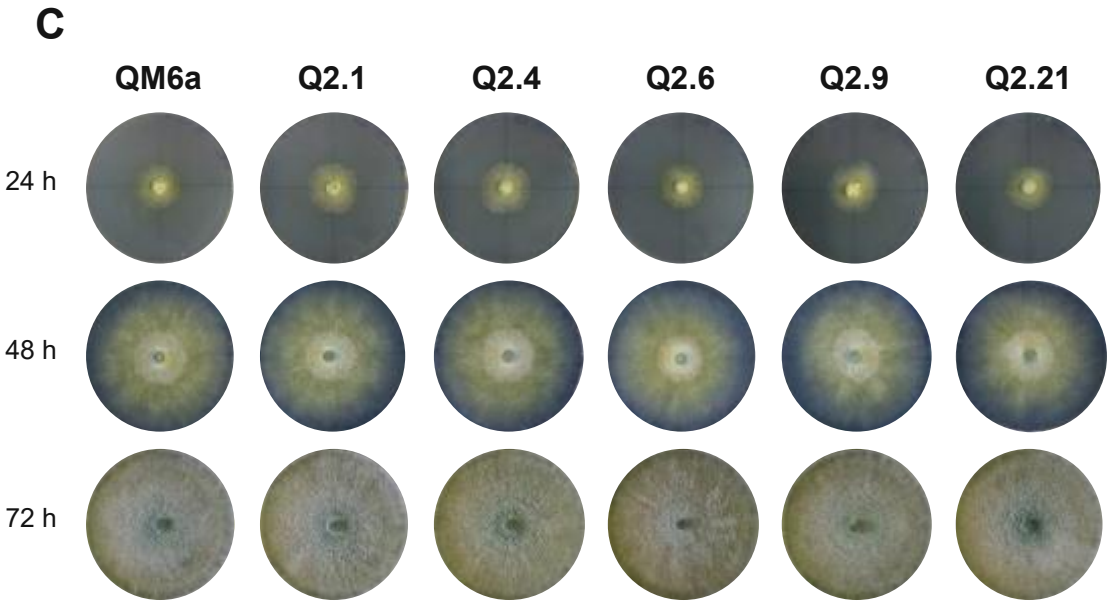
9.3 Growth and spore production of the *cel45a* transfomants on PDA

Growth tests on PDA were performed to check if any of the *cel45a* transformants shows a different phenotype compared to the parental strain. Therefore, QM6a and the *cel45a* transformants were incubated for 8 days, whereas Rut C30 and its *cel45a* transformants strains were incubated for 10 days.

The QM6a *cel45a* strains did not show any growth defect following the transformation and purification procedure when compared to the parental strain QM6a. These strains showed a similar growth rate as their parental strain. *Cel45a* transformants produced about 30 % less spores per cm² than the parental strain QM6a, with the only exception of Q2.1, which generated about the same amount of spores than the parental strain *T. reesei* QM6a (Fig. 9).







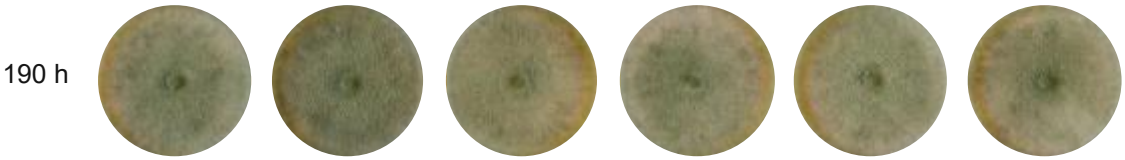
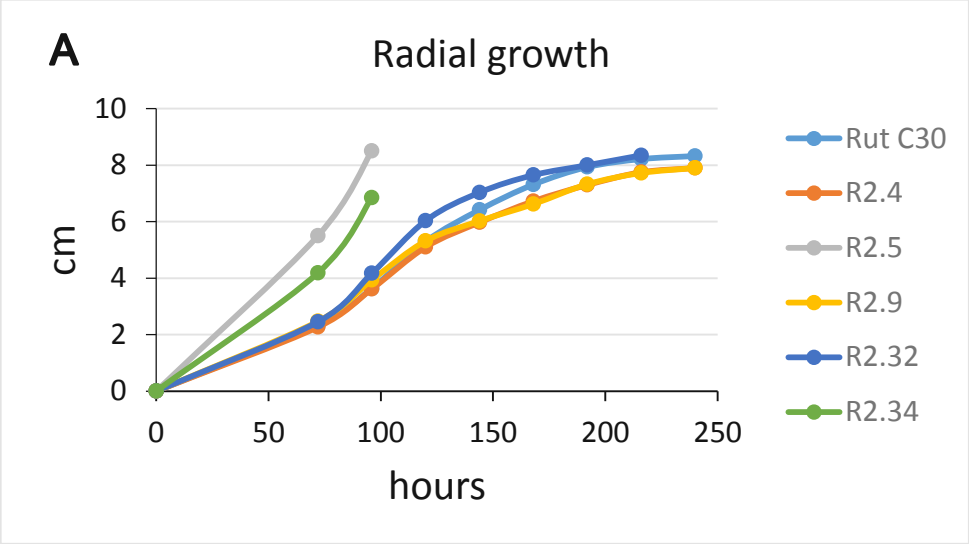
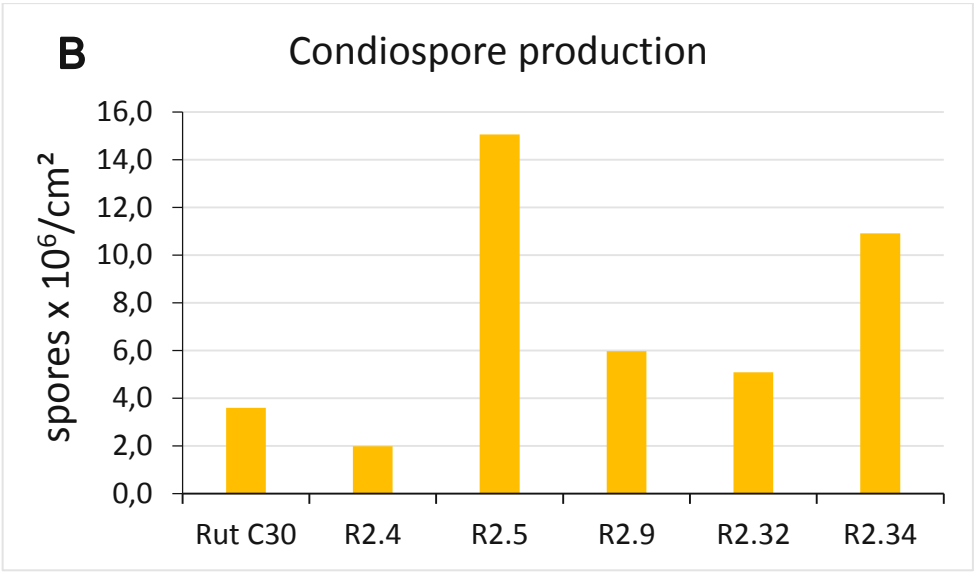
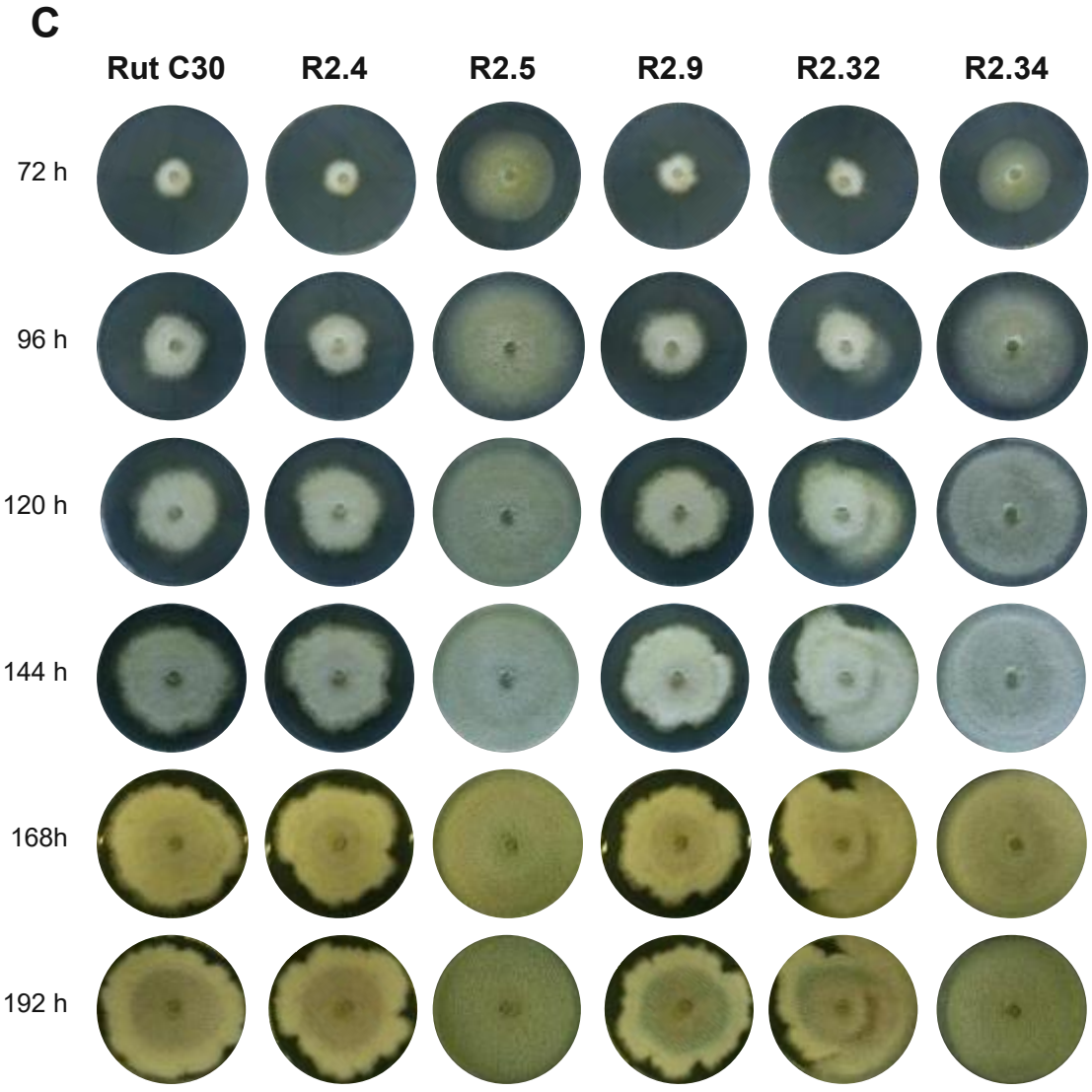


Fig. 9: Radial growth and conidiospore production of QM6a and its *cel45a* transformants. **(A)** Radial growth was measured by inoculation of a 4 mm agar piece on a 20 ml PDA Petri dish. Strains were grown for 190 h at 28 °C. The diameter of each strain was measured every 24 h until the mycelium reached the edge of the Petri dish. **(B)** After 190 h the conidiospore production of *cel45a* strains was measured through harvesting the spores and determination of the spore concentration. **(C)** Phenotypes of QM6a and the respective *cel45a* transformants on PDA plates. Incubation time is indicated at the left side.

Rut C30 *cel45a* transformants showed a different picture regarding their growth rate and conidia production. Three strains (R2.4, R2.9 and R2.32) were growing identical compared to Rut C30 whereas two strains (R2.5 and R2.34) were growing faster than the parental strain and were producing about 250 % more spores. One strain had a different colony phenotype and produced sectors which showed faster growth. This indicates that this strain seems to be unstable despite the single spore purification (R2.32). The spore production was not affected in R2.32 (Fig. 10).







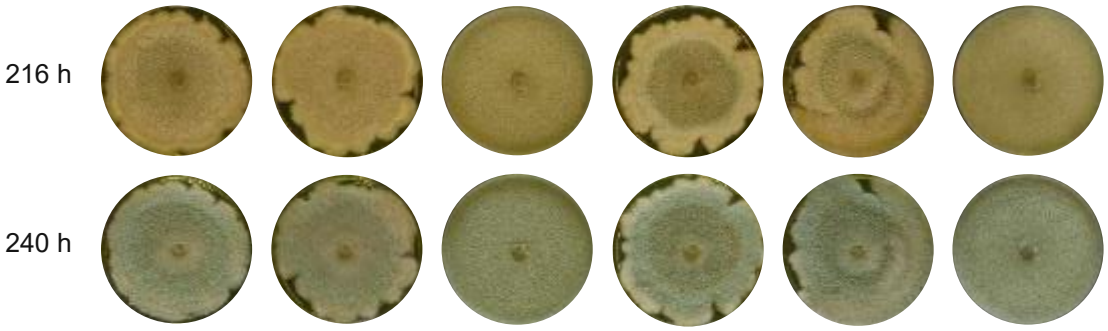
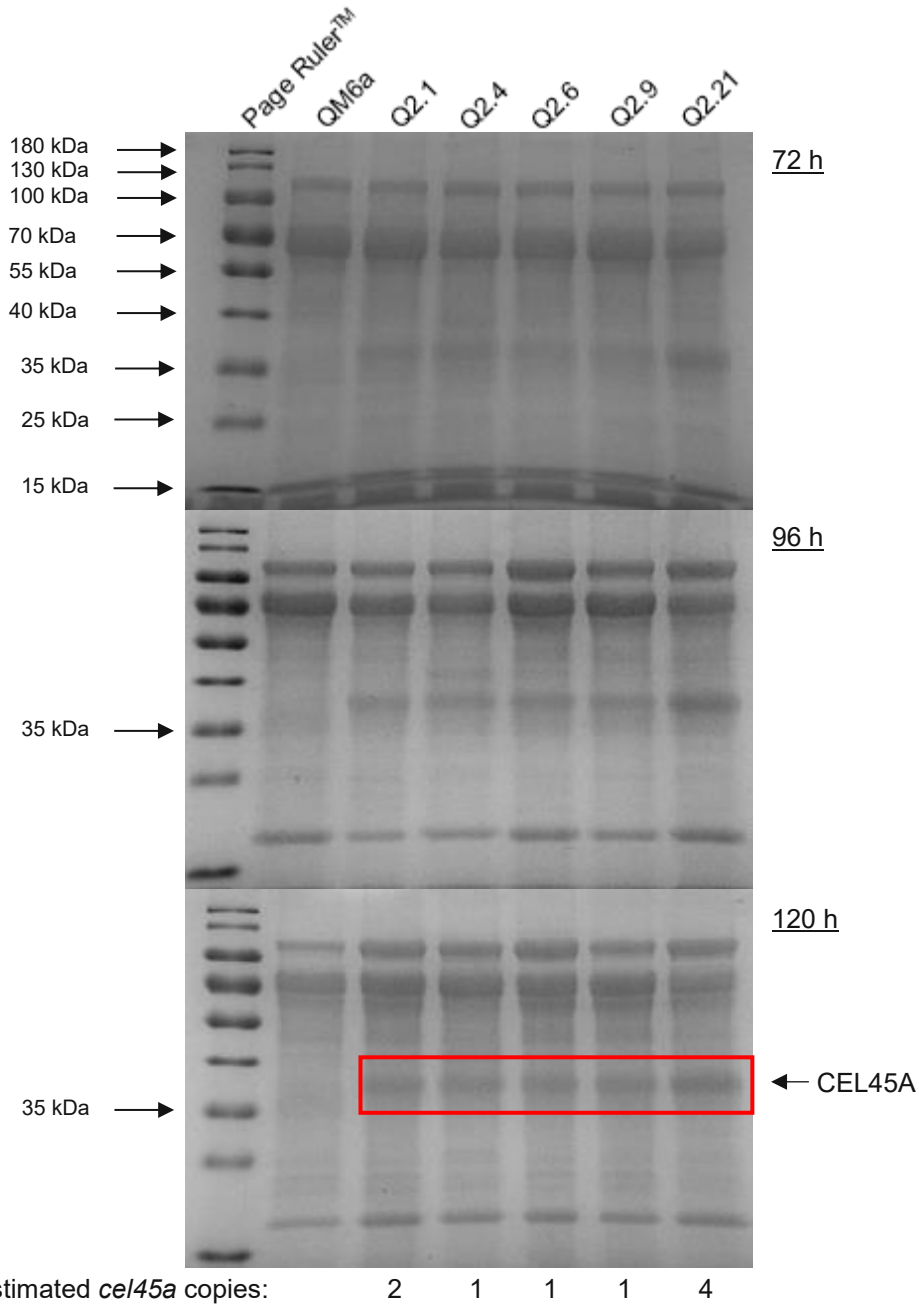


Fig. 10: Radial growth and conidiospore production of Rut C30 and its *cel45a* transformants. (A) Radial growth was measured by inoculation of a 4 mm agar piece on a 20 ml PDA Petri dish. Strains were then grown for 240 h at 28 °C. The diameter of each strain was measured every 24 hours until the mycelium reached the edge of the Petri dish. **(B)** After 240 h the conidiospore production of *cel45a* transformants was measured through harvesting the spores and determination of the spore concentration. **(C)** Phenotypes of Rut C30 and the respective *cel45a* transformants on PDA plates. Incubation time is indicated at the left side.

9.4 Analysis of CEL45A production in *T. reesei* QM6a and RutC30 transformants

Following the copy number determination of the *T. reesei cel45a* transformants, we tested if the strains produce the CEL45A protein. Therefore strains with different copy numbers were chosen and shake flask cultivations were performed for 120 h using cellulose as carbon source to induce expression of *cel45a* under the *cel7a* expression signals. Samples of the supernatant were taken after 72, 96 and 120 h, the proteins precipitated and then separated by their molecular weight in an SDS-PAGE. Following Coomassie blue staining of the protein gels of the two parental strains QM6a and RutC30 and the different *cel45a* transformants, an additional band became visible in the *cel45a* transformants compared to their parental strains. This band is found at around 37 kDa and increased in intensity over time (Fig. 11). The calculated molecular weight of CEL45A without the signal peptide is about 29 kDa (see chapter 9.1). The difference between calculated molecular weight and the result from the SDS-PAGE is therefore about 8 kDa and is most likely caused by N- and especially O-glycosylation (see Fig. 7).

A



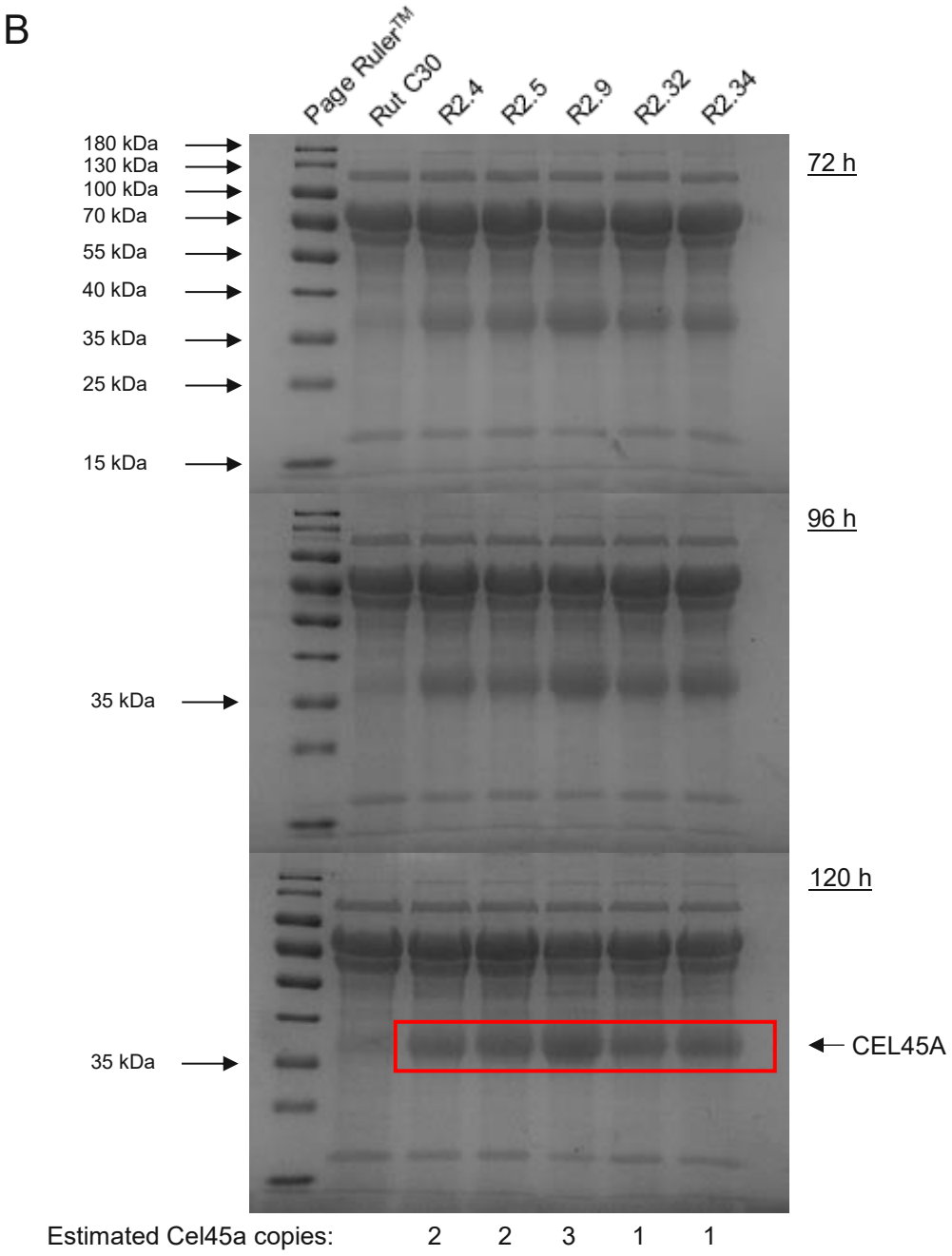


Fig. 11: SDS-PAGE of the supernatant of CEL45A expressing *T. reesei* QM6a and RutC30 transformants. The parental *T. reesei* strains and their respective *cel45a* transformants were cultivated in Mandels-Andreotti medium with 1 % cellulose as sole carbon source for 120 h and samples were taken at the indicated time points. For the SDS PAGE about 10 µg of extracellular protein was loaded on the SDS-PAGE per sample. **(A)** *T. reesei* QM6a and selected *cel45a* QM6a transformants. **(B)** *T. reesei* Rut C30 and Rut C30 *cel45a* transformants. A band of around 37 kDa representing the putative CEL45A protein is highlighted. As size marker PAGE Ruler™ (Thermo Fisher Scientific, Inc.) was used.

9.5 Identification of the 37 kDa band expressed in *T. reesei* transfomants as CEL45A

To proof the assumption, that the 37 kDa band corresponds to CEL45A, the CEL45A protein band of the transformant Q2.21 was analysed by MS analysis. After a tryptic digestion of the 37 kDa protein band, the sample was measured by LC-ESI-MS. The data generated were analysed online with the Global Proteome Machine. The MS analysis verified that the protein at the 37 kDa band corresponds to the CEL45A protein. The log(e) value was -91.4, showing that this result is not likely to be randomly generated.

9.6 Endoglucanase activity of the *T. reesei cel45a* transformants

Having established that *cel45a* was expressed in the different QM6a and RutC30 transformants, the next step was to test if the protein was functionally produced. Therefore, the culture supernatants of the shake flask cultivation were analysed for endo-1,4-β-glucanase activity using the Azo CMC assay. Default conditions for this assay were 40 °C and pH 4.5. In the first set of experiments the increase in endoglucanase activity of each parental strain and their respective transformants was checked over the cultivation period. The overall endoglucanase activity produced was much higher for RutC30 compared to QM6a and the endoglucanase activity increased over time in all strains. For the respective QM6a transformants an increase in endo-1,4-β-glucanase activity of about 150 % compared to the parental strain could be observed, whereas for the Rut C30 transformants an increase of about 50 % on average was found under these conditions (Fig. 12).

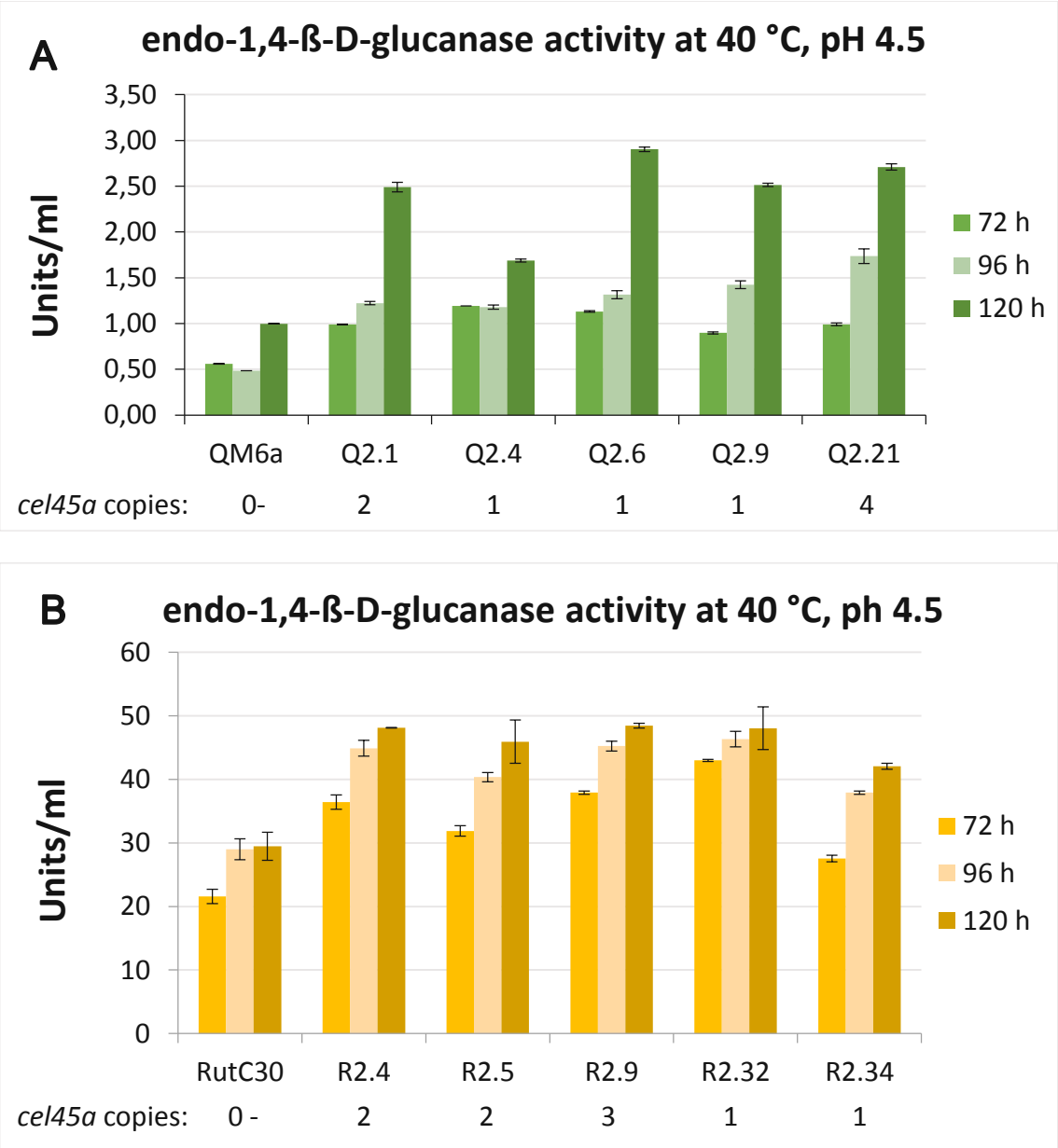


Fig. 12: Endoglucanase activity produced by different *T. reesei cel45a* transformants. All strains were grown for 120 hours in shake flasks using cellulose as carbon source. Endoglucanase activity was measured after 72, 96 and 120 h of cultivation. *Cel45a* copy numbers of the strains is indicated below. **(A)** QM6a and respective *cel45a* transformants. **(B)** Rut C30 and respective *cel45a* transformants.

Comparison between *cel45a* copy number in the *T. reesei* QM6a and Rut C30 transformants showed that no correlation between endoglucanase activity and gene copy number was found.

CEL45A is predicted to be a thermostable cellulase and it was tested now in the next step if the enzyme could be produced in a thermostable form by *T. reesei*. Therefore the temperature for the Azo-CMC assay was raised in 10 °C steps until 80 °C and the culture supernatants of QM6a and selected *cel45a* trans-

formants compared. Endoglucanase activity of the QM6a supernatant decreased already at 50°C and no activity was found at 60°C. In contrast, endoglucanase activity was still high at 50 and 60°C in the QM6a *cel45a* transformants and some residual activity was still found at 70 and 80°C (Fig. 13). Highest activity could be observed for transformants Q2.21 which harbours four copies of the *cel45a* gene. The results indicate that CEL45A could be functionally produced in *T. reesei* and shows that the thermostability of CEL45A could also be found for the *T. reesei* produced cellulase.

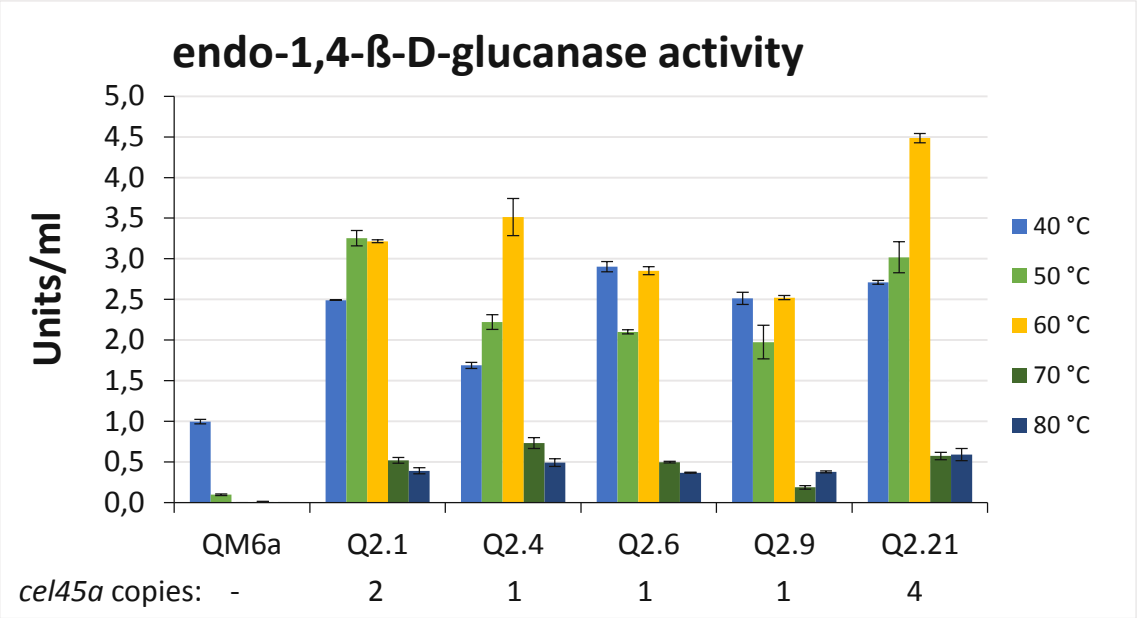


Fig. 13: Thermostability of endoglucanases produced by different *T. reesei* QM6a *cel45a* transformants. All strains were grown for 120 hours in shake flasks using cellulose as carbon source. Endoglucanase activity was measured after 120 h of cultivation. *Cel45a* gene numbers of the strains is indicated below.

In the next step we determined the thermostability of the endoglucanases of *T. reesei* RutC30 and its *cel45a* transformants. When we measured the enzymatic activity of the cellulases at 60 °C, the cellulases produced by strain Rut C30 turned out to be more thermostable than that of QM6a. Our data are in accordance with earlier reports which showed that RutC30 produced a more thermostable endoglucanase activity (Zou et al. 2012). When we measured the endoglucanase activity following a 24 hours preincubation in a thermomixer at 60 °C. The enzymatic activity of the Rut C30 sample decreased after this preincubation for about 40 %, whereas that of the *cel45a* transformants decreased around 20 % indicating, that a more thermostable cellulase activity was produced by the *cel45a* expressing strains (Fig. 14).

Overall the thermostability of CEL45A until 60 °C could be shown which declined at higher temperatures. This observation is similar to the related CEL45A protein of *N. crassa* (Kadowaki et al. 2015) which also showed a reduction at 60 °C.

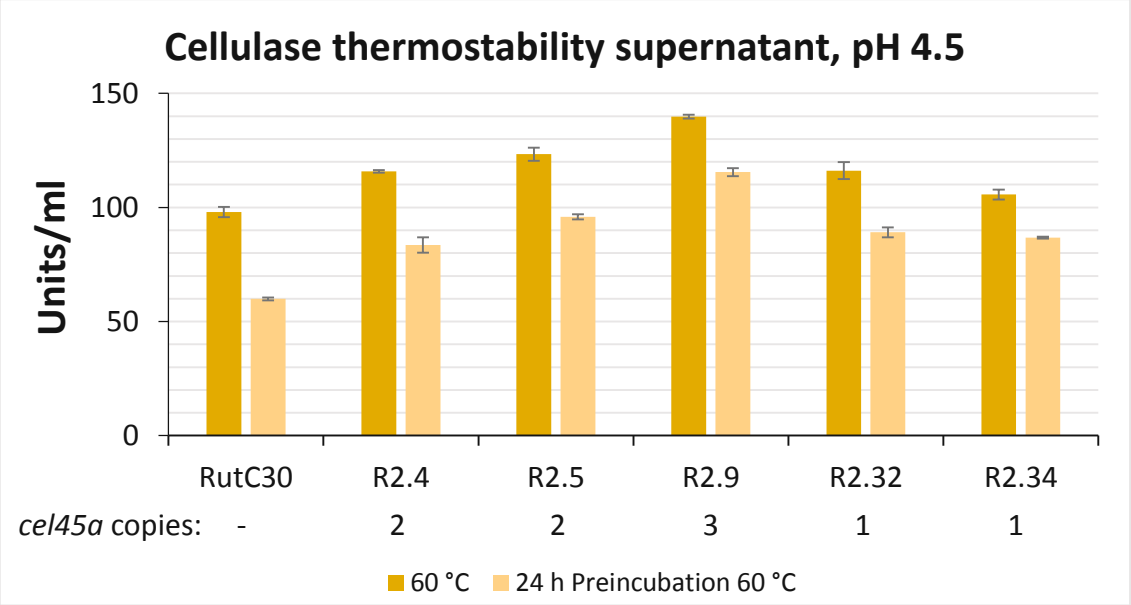


Fig. 14: Thermostabilty of *T. reesei* Rut C30 cellulases. All strains were grown for 120 h in shake flasks using cellulose as carbon source. Endoglucanase activity was measured after 120 h of cultivation. As Rut C30 produces thermostable cellulases, the supernatant of all Rut C30 strains were preincubated for 24 hours in a thermomixer at 60 °C and then measured at 60 °C in a water bath. *Cel45a* gene numbers of the strains is indicated below.

10. Discussion

T. reesei has been used as an expression host for industrial enzyme production for over 30 years now. Through different mutagenesis and selection regimes high cellulase producing strains were generated by classical strain engineering. Several molecular tools were also developed to improve homologous and heterologous protein expression for recombinant strain engineering (Bischof and Seiboth 2014). Whereas homologous protein overexpression is well established the production of heterologous proteins often needs further improvements to become cost competitive to other expression systems.

The aim of the present study was to express a thermostable endoglucanase of the CEL45A GH family in *T. reesei*. Analysis of CEL45A sequence shows that it is composed of a carbohydrate binding domain belonging to family 1 and the catalytic domain connected by a linker. Analysis of the amino acid sequence shows that the protein contains a high number of disulphide bridges which are necessary to confer the thermostability to the protein. In addition the protein sequence exhibits a number of putative O-glycosylation sites and only a single N-glycosylation site.

For the expression of *cel45a* two different *T. reesei* strains were selected. One is the wild-type strain QM6a, a low cellulase producer. The other one is strain Rut C30 which is the highest cellulase producer publically available (Peterson and Nevalainen 2012). Both were transformed with an expression cassette based on the *cbh1*/*cel7A* expression signals and the *cel45a* coding region using *amdS* as selection marker. Following the transformation of the expression cassette, the presence and copy number of the expression cassette was determined in the transformants. The copy number of the expression cassette had to be determined indirectly through the quantification of the *amdS* marker gene. A direct determination of the *cel45a* sequence failed as no suitable primers could be identified for the *cel45a* coding region. Further efforts will therefore be necessary to find suitable primers for qPCR to determine the *cel45a* copy number. Additionally, it will also be necessary in follow up experiments to determine the transcript levels of *cel45a* in the different transformants with the different copy numbers and to relate this level to the endoglucanase activity produced.

Transformants with a different number of *cel45a* gene copies were then chosen and cultivated on minimal-media containing the *cel7a* inducing carbon source cellulose. A band at 37 kDa could be detected in the protein mix of the supernatant of the transformants. In comparison, the calculated molecular weight of CEL45A without the signal peptide is about 29 kDa. To clarify if this band corresponds to CEL45A, an MS analysis was performed, which revealed that the band at 37 kDa is CEL45A. A possible explanation for the increased molecular weight is the above mentioned O-glycosylation, predicted by NetOGlyc 4.0 for CEL45A, which occurs often in the linker region of cellulases connecting the catalytic and the binding domain.

The activity of the overexpressed CEL45A was tested without any further purification in the supernatant of these strains. An improved endo-1,4-β-D-glucanase activity in the strains could be proven. Regarding the measurements at 40 °C for QM6a and Rut C30 *cel45a* transformants no obvious relationship regarding gene copy numbers and the produced Units/ml could be shown. One possible explanation could be that the indirect determination of the copy number described above did not accurately estimate the *cel45a* copy number. Another reason might be that the expression of this protein leads to an overloading of the redox capacity of the protein folding machinery and that only a certain level of the protein could be produced functionally. Again, this will be subject of a further deeper analysis of the strains. Nevertheless, the aim of this work was to overexpress the protein and prove its functionality. As the protein is expected to be thermostable, enzyme activity assays were performed at different temperatures. The supernatant samples of the *T. reesei* QM6a *cel45a* transformants producing CEL45A showed an activity until 60°C whereas the secreted proteins of the parental strain showed nearly no activity after incubation at 50°C. Due to the reason that Rut C30 produces also more thermostable cellulases (Zou et al. 2012) the supernatants were measured at 60°C and for comparison after 24 hours of preincubation at 60 °C. The assumption was that the homologous cellulases produced by the Rut C30 parental strain are not as thermostable as the heterologous CEL45A cellulase of the transformants. Indeed, there was a difference detectable between the RutC30 transformants and the parental strain. Its enzymatic activity after 24 hours of preincubation decreased 40 %, whereas

the enzymatic activity of the Rut C30 transformants decreased just for 20 %. This indicates that the transformants produce a more thermostable cellulase. Regarding the phenotype of the transformants the QM6a *cel45a* transformants showed homogeneity towards colony extension. Total spore production was decreased about 30 % within four transformants regardless of the gene copy number which would indicate that the CEL45A protein affects sporulation. For the Rut C30 *cel45a* transformants a difference in colony extension was observed for two strains. They grew faster than the parental strain and did not exhibit the typical compact RutC30 phenotype. One strain (R2.32) showed a very asymmetric growth pattern on PDA which indicates that the strain is still unstable and further purification rounds might therefore be necessary to obtain stable strains. Regarding spore production, the two strains R2.5 and R2.34 were producing about 250 % more spores than the parental strain. This leads to the assumption, that after successful integration of *cel45a* into the genome of *T. reesei* QM6a and Rut C30, QM6a can handle this integration better than Rut C30. A reason for this impaired growth after the integration might be the higher protein production of the strain which might already stress the metabolism of the fungus.

In summary, the aim of this thesis to overexpress *cel45a* in *T. reesei* was achieved and the protein was expressed in both strains functionally and thermostable. A further characterization of the full biochemical properties of the protein CEL45A was beyond this work and its activity could be further tested under different pH and temperature values. Additionally the saccharification potential of the enzyme mixture with respect to different lignocellulosic substrates could be another possibility for subsequent studies.

11. Conclusion

The low cellulase producer *T. reesei* QM6a and the high cellulase producer *T. reesei* Rut C30 could successfully be transformed with the provided *cel45a* expression cassette. Different transformants with various gene copy numbers of *cel45a* were generated expressing CEL45A. The enzyme exhibit thermostability and can be used now to improve cellulose degradation at higher temperatures.

12. Appendix

12.1 qPCR results of *amdS* quantification

Strain		Gene	Type	Efficiency	Expression	Std. Error
QM6a	2.1 *cel45a*	*amdS*	TRG	0,94	1,596	1,527 - 1,666
		cbh2	REF	0,84	1	
QM6a	2.4 *cel45a*	*amdS*	TRG	0,94	1,42	1,283 - 1,526
		cbh2	REF	0,84	1	
QM6a	2.6 *cel45a*	*amdS*	TRG	0,94	0,714	0,670 - 0,754
		cbh2	REF	0,84	1	
QM6a	2.9 *cel45a*	*amdS*	TRG	0,94	0,642	0,556 - 0,706
		cbh2	REF	0,84	1	
QM6a	2.21 *cel45a*	*amdS*	TRG	0,94	4,102	3,742 - 4,404
		cbh2	REF	0,84	1	

Strain		Gene	Type	Efficiency	Expression	Std. Error
Rut C30	2.4 *cel45a*	*amdS*	TRG	0,94	2,22	2,086 - 2,364
		cbh2	REF	0,84	1	
Rut C30	2.5 *cel45a*	*amdS*	TRG	0,94	1,93	1,192 - 4,572
		cbh2	REF	0,84	1	
Rut C30	2.9 cel45a	*amdS*	TRG	0,94	3,477	3,323 - 3,749
		cbh2	REF	0,84	1	
Rut C30	2.32 *cel45a*	*amdS*	TRG	0,94	1,06	0,986 - 1,149
		cbh2	REF	0,84	1	
Rut C30	2.34 *cel45a*	*amdS*	TRG	0,94	1,227	1,129 - 1,317
		cbh2	REF	0,84	1	

13. References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 25(17):3389-402

Beeby M, O'Connor BD, Ryttersgaard C, Boutz DR, Perry LJ, Yeates TO (2005) The genomics of disulfide bonding and protein stabilization in thermophiles. PLoS Biology 3(9):e309 doi:10.1371/journal.pbio.0030309

Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, Schwede T (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Research 42(W1):W252-W258 doi:10.1093/nar/gku340

Bischof R, Seiboth B (2014) Molecular Tools for Strain Improvement of *Trichoderma* spp. In: Gupta VK (ed) Biotechnology and Biology of Trichoderma. pp 179-191

Bischof RH, Ramoni J, Seiboth B (2016) Cellulases and beyond: the first 70 years of the enzyme producer *Trichoderma reesei*. Microbial Cell Factories 15(1):106 doi:10.1186/s12934-016-0507-6

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Research 37:D233-D238 doi:10.1093/nar/gkn663

Ceroni A, Passerini A, Vullo A, Frasconi P (2006) DISULFIND: a disulfide bonding state and cysteine connectivity prediction server. Nucleic Acids Research 34(Web Server issue):W177-81 doi:10.1093/nar/gkl266

Cheng J, Saigo H, Baldi P (2006) Large-scale prediction of disulphide bridges using kernel methods, two-dimensional recursive neural networks, and weighted graph matching. Proteins 62(3):617-29 doi:10.1002/prot.20787

Cherry JR, Fidantsef AL (2003) Directed evolution of industrial enzymes: an update. Current Opinion in Biotechnology 14(4):438-443 doi:10.1016/s0958-1669(03)00099-5

Colabardini AC, Valkonen M, Huuskonen A, Siika-aho M, Koivula A, Goldman GH, Saloheimo M (2016) Expression of Two Novel beta-Glucosidases from Chaetomium atrobrunneum in Trichoderma reesei and Characterization of the Heterologous Protein Products. Molecular Biotechnology 58(12):821-831 doi:10.1007/s12033-016-9981-7

D'haene B, Vandesompele J, Hellemans J (2010) Accurate and objective copy number profiling using real-time quantitative PCR. Methods 50(4):262-270 doi:10.1016/j.ymeth.2009.12.007

Delic M, Rebnegger C, Wanka F, Puxbaum V, Haberhauer-Troyer C, Hann S, Kollensperger G, Mattanovich D, Gasser B (2012) Oxidative protein folding and unfolded protein response elicit differing redox regulation in endoplasmic reticulum and cytosol of yeast. Free Radical Biology and Medicine 52(9):2000-12 doi:10.1016/j.freeradbiomed.2012.02.048

Demain AL, Vaishnav P (2009) Production of recombinant proteins by microbes and higher organisms. Biotechnology Advances 27(3):297-306 doi:10.1016/j.biotechadv.2009.01.008

Di Tommaso P, Moretti S, Xenarios I, Orobitg M, Montanyola A, Chang JM, Taly JF, Notredame C (2011) T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. Nucleic Acids Research 39(Web Server issue):W13-7 doi:10.1093/nar/gkr245

dos Santos Castro L, Pedersoli WR, Antoniêto ACC, Steindorff AS, Silva-Rocha R, Martinez-Rossi NM, Rossi A, Brown NA, Goldman GH, Faça VM, Persinoti GF, Silva RN (2014) Comparative metabolism of cellulose, sophorose and glucose in *Trichoderma*

reeseiusing high-throughput genomic and proteomic analyses. Biotechnology for Biofuels 7(1):41 doi:10.1186/1754-6834-7-41

Druzhinina IS, Kubicek CP (2017) Genetic engineering of Trichoderma reesei cellulases and their production. Microbial Biotechnology doi:10.1111/1751-7915.12726

Durand H, Clanet M, Tiraby G (1988) Genetic improvement of *Trichoderma reesei* for large scale cellulase production. Enzyme and Microbial Technology 10(6):341-346 doi:http://dx.doi.org/10.1016/0141-0229(88)90012-9

Foreman PK, Brown D, Dankmeyer L, Dean R, Diener S, Dunn-Coleman NS, Goedegebuur F, Houfek TD, England GJ, Kelley AS, Meerman HJ, Mitchell T, Mitchinson C, Olivares HA, Teunissen PJM, Yao J, Ward M (2003) Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei*. Journal of Biological Chemistry 278(34):31988-31997 doi:10.1074/jbc.M304750200

Ghosh A, Al-Rabiai S, Ghosh BK, Trimiño-Vazquez H, Eveleigh DE, Montenecourt BS (1982) Increased endoplasmic reticulum content of a mutant of *Trichoderma reesei* (RUT-C30) in relation to cellulase synthesis. Enzyme and Microbial Technology 4(2):110-113 doi:http://dx.doi.org/10.1016/0141-0229(82)90093-X

Gupta VK, Kubicek CP, Berrin JG, Wilson DW, Couturier M, Berlin A, Filho EXF, Ezeji T (2016a) Fungal Enzymes for Bio-Products from Sustainable and Waste Biomass. Trends in Biochemical Sciences 41(7):633-645 doi:10.1016/j.tibs.2016.04.006

Gupta VK, Steindorff AS, de Paula RG, Silva-Rocha R, Mach-Aigner AR, Mach RL, Silva RN (2016b) The Post-genomic Era of *Trichoderma reesei*: What's Next? Trends in Biotechnology 34(12):970-982 doi:10.1016/j.tibtech.2016.06.003

Harkki A, Uusitalo J, Bailey M, Penttila M, Knowles JKC (1989) A novel fungal expression system - secretion of active calf chymosin from the filamentous fungus *Trichoderma reesei*. Nature Biotechnology 7(6):596-& doi:DOI 10.1038/nbt0689-596

Hwang C, Sinskey AJ, Lodish HF (1992) Oxidized redox state of glutathione in the endoplasmic reticulum. Science 257(5076):1496-502 doi:10.1126/science.1523409

Ilmen M, Onnela ML, Klemsdal S, Keranen S, Penttila M (1996a) Functional analysis of the cellobiohydrolase I promoter of the filamentous fungus *Trichoderma reesei*. Molecular and General Genetics 253(3):303-14

Ilmen M, Thrane C, Penttila M (1996b) The glucose repressor gene *cre1* of *Trichoderma*: isolation and expression of a full-length and a truncated mutant form. Molecular and General Genetics 251(4):451-60

Kadowaki MAS, Camilo CM, Muniz AB, Polikarpov I (2015) Functional characterization and low-resolution structure of an endoglucanase CEL45A from the filamentous fungus *Neurospora crassa* OR74A: thermostable enzyme with high activity toward lichenan and beta-glucan. Molecular Biotechnology 57(6):574-588 doi:10.1007/s12033-015-9851-8

Karlsson J, Siika-aho M, Tenkanen M, Tjerneld F (2002) Enzymatic properties of the low molecular mass endoglucanases Cel12A (EG III) and Cel45A (EG V) of Trichoderma reesei. Journal of Biotechnology 99(1):63-78 doi:10.1016/s0168-1656(02)00156-6

Kocar G, Civas N (2013) An overview of biofuels from energy crops: Current status and future prospects. Renewable & Sustainable Energy Reviews 28:900-916 doi:10.1016/j.rser.2013.08.022

Kubicek CP (2013) Systems biological approaches towards understanding cellulase production by *Trichoderma reesei*. Journal of Biotechnology 163(2):133-142 doi:http://dx.doi.org/10.1016/j.jbiotec.2012.05.020

Kuhad RC, Deswal D, Sharma S, Bhattacharya A, Jain KK, Kaur A, Pletschke BI, Singh A, Karp M (2016) Revisiting cellulase production and redefining current strategies based on major challenges. Renewable & Sustainable Energy Reviews 55:249-272 doi:10.1016/j.rser.2015.10.132

Kumar R, Singh S, Singh OV (2008) Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. Journal of Industrial Microbiology & Biotechnology 35(5):377-391 doi:10.1007/s10295-008-0327-8

Landowski CP, Huuskonen A, Wahl R, Westerholm-Parvinen A, Kanerva A, Hanninen AL, Salovuori N, Penttila M, Natunen J, Ostermeier C, Helk B, Saarinen J, Saloheimo M (2015) Enabling low cost biopharmaceuticals: a systematic approach to delete proteases from a well-known protein production host *Trichoderma reesei*. PLoS One 10(8) doi:10.1371/journal.pone.0134723

Landowski CP, Mustalahti E, Wahl R, Croute L, Sivasiddarthan D, Westerholm-Parvinen A, Sommer B, Ostermeier C, Helk B, Saarinen J, Saloheimo M (2016) Enabling low cost biopharmaceuticals: high level interferon alpha-2b production in *Trichoderma reesei*. Microbial Cell Factories 15 doi:10.1186/s12934-016-0508-5

Le Crom S, Schackwitz W, Pennacchio L, Magnuson JK, Culley DE, Collett JR, Martin J, Druzhinina IS, Mathis H, Monot F, Seiboth B, Cherry B, Rey M, Berka R, Kubicek CP, Baker SE, Margeot A (2009) Tracking the roots of cellulase hyperproduction by the fungus *Trichoderma reesei* using massively parallel DNA sequencing. Proceedings of the National Academy of Sciences of the United States of America 106(38):16151-6 doi:10.1073/pnas.0905848106

Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Research 42(D1):D490-D495 doi:10.1093/nar/gkt1178

Mach RL, Zeilinger S (2003) Regulation of gene expression in industrial fungi: Trichoderma. Applied Microbiology and Biotechnology 60(5):515-522 doi:10.1007/s00253-002-1162-x

Malhotra JD, Kaufman RJ (2007) The endoplasmic reticulum and the unfolded protein response. Seminars in Cell & Developmental Biology 18(6):716-731 doi:http://dx.doi.org/10.1016/j.semcdb.2007.09.003

Mandels M, Reese ET (1957) Induction of cellulase in *Trichoderma viride* as influenced by carbon sources and metals. Journal of Bacteriology 73(2):269-78

Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, Chapman J, Chertkov O, Coutinho PM, Cullen D, Danchin EGJ, Grigoriev IV, Harris P, Jackson M, Kubicek CP, Han CS, Ho I, Larrondo LF, de Leon AL, Magnuson JK, Merino S, Misra M, Nelson B, Putnam N, Robbertse B, Salamov AA, Schmoll M, Terry A, Thayer N, Westerholm-Parvinen A, Schoch CL, Yao J, Barbote R, Nelson MA, Detter C, Bruce D, Kuske CR, Xie G, Richardson P, Rokhsar DS, Lucas SM, Rubin EM, Dunn-Coleman N, Ward M, Brettin TS (2008) Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). Nature Biotechnology 26(5):553-560 doi:10.1038/nbt1403

Montenecourt BS, Eveleigh DE (1979) Selective screening methods for the isolation of high yielding cellulase mutants of *Trichoderma reesei* Hydrolysis of Cellulose: Mechanisms of Enzymatic and Acid Catalysis. Advances in Chemistry, vol 181. American Chemical Society, pp 289-301

Nevalainen H, Peterson R (2014) Making recombinant proteins in filamentous fungi- are we expecting too much? Frontiers in Microbiology 5:75 doi:10.3389/fmicb.2014.00075

Payne CM, Resch MG, Chen LQ, Crowley MF, Himmel ME, Taylor LE, Sandgren M, Stahlberg J, Stals I, Tan ZP, Beckham GT (2013) Glycosylated linkers in multimodular lignocellulose-degrading enzymes dynamically bind to cellulose. Proceedings of the National Academy of Sciences of the United States of America 110(36):14646-14651 doi:10.1073/pnas.1309106110

Penttilä M, Nevalainen H, Rättö M, Salminen E, Knowles J (1987) A versatile transformation system for the cellulolytic filamentous fungus *Trichoderma reesei*. Gene 61(2):155-164 doi:http://dx.doi.org/10.1016/0378-1119(87)90110-7

Petersen TN, Brunak S, Von Heijne G, Nielsen H (2011) SignalP 4.0: Discriminating signal peptides from transmembrane regions. Nature Methods 8(10):785-786 doi:10.1038/nmeth.1701

Peterson R, Nevalainen H (2012) *Trichoderma reesei* RUT-C30--thirty years of strain improvement. Microbiology 158(Pt 1):58-68 doi:10.1099/mic.0.054031-0

Pfaffl MW, Horgan GW, Dempfle L (2002) Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Research 30(9):e36

Polizeli M, Rizzatti ACS, Monti R, Terenzi HF, Jorge JA, Amorim DS (2005) Xylanases from fungi: properties and industrial applications. Applied Microbiology and Biotechnology 67(5):577-591 doi:10.1007/s00253-005-1904-7

Popp J, Lakner Z, Harangi-Rakos M, Fari M (2014) The effect of bioenergy expansion: Food, energy, and environment. Renewable & Sustainable Energy Reviews 32:559-578 doi:10.1016/j.rser.2014.01.056

Rahman Z, Shida Y, Furukawa T, Suzuki Y, Okada H, Ogasawara W, Morikawa Y (2009) Evaluation and characterization of *Trichoderma reesei* cellulase and xylanase promoters. Applied Microbiology and Biotechnology 82(5):899-908 doi:10.1007/s00253-008-1841-3

Ramoni J, Seidl-Seiboth V, Bischof RH, Seiboth B (2016) Gene Expression Systems in Industrial Ascomycetes: Advancements and Applications. In: Schmoll M, Dattenböck C (eds) Gene Expression Systems in Fungi: Advancements and Applications. Springer International Publishing, Cham, pp 3-22

Robert X, Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Research 42(W1):W320-W324 doi:10.1093/nar/gku316

Saloheimo A, Henrissat B, Hoffrén A-M, Teleman O, Penttilä M (1994) A novel, small endoglucanase gene, egl5, from *Trichoderma reesei* isolated by expression in yeast. Molecular Microbiology 13(2):219-228 doi:10.1111/j.1365-2958.1994.tb00417.x

Sandgren M, Gualfetti PJ, Shaw A, Gross LS, Saldajeno M, Day AG, Jones TA, Mitchinson C (2003) Comparison of family 12 glycoside hydrolases and recruited substitutions important for thermal stability. Protein Science 12(4):848-60 doi:10.1110/ps.0237703

Schuster A, Bruno KS, Collett JR, Baker SE, Seiboth B, Kubicek CP, Schmoll M (2012) A versatile toolkit for high throughput functional genomics with *Trichoderma reesei*. Biotechnology for Biofuels 5(1):1 doi:10.1186/1754-6834-5-1

Seiboth B, Hartl L, Salovuori N, Lanthaler K, Robson GD, Vehmaanpera J, Penttila ME, Kubicek CP (2005) Role of the *bga1*-encoded extracellular beta-galactosidase of *Hypocrea jecorin*a in cellulase induction by lactose. Applied and Environmental Microbiology 71(2):851-857 doi:10.1128/Aem.71.2.851-857.2005

Seiboth B, Hofmann G, Kubicek CP (2002) Lactose metabolism and cellulase production in *Hypocrea jecorina*: the *gal7* gene, encoding galactose-1-phosphate uridylyltransferase, is essential for growth on galactose but not for cellulase induction. Molecular Genetics and Genomics 267(1):124-32 doi:10.1007/s00438-002-0654-9

Smith MA, Romero PA, Wu T, Brustad EM, Arnold FH (2013) Chimeragenesis of distantly-related proteins by noncontiguous recombination. Protein Science 22(2):231-8 doi:10.1002/pro.2202

Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager K, Lavrsen K, Dabelsteen S, Pedersen NB, Marcos-Silva L, Gupta R, Bennett EP, Mandel U, Brunak S, Wandall HH, Levery SB, Clausen H (2013) Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO Journal 32(10):1478-1488 doi:10.1038/emboj.2013.79

Travers KJ, Patil CK, Wodicka L, Lockhart DJ, Weissman JS, Walter P (2000) Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. Cell 101(3):249-58

Tu BP, Ho-Schleyer SC, Travers KJ, Weissman JS (2000) Biochemical basis of oxidative protein folding in the endoplasmic reticulum. Science 290(5496):1571-4

Valtakari L, Alapuranen M, Szakacs G, Kallio J, Ojapalo P, Vehmaanperä J, Puranen T Fungal endoglucanases, their production and use. WO2010076388 A1, 8. July 2010

Voutilainen SP, Puranen T, Siika-Aho M, Lappalainen A, Alapuranen M, Kallio J, Hooman S, Viikari L, Vehmaanpera J, Koivula A (2008) Cloning, expression, and characterization of novel thermostable family 7 cellobiohydrolases. Biotechnology and Bioengineering 101(3):515-28 doi:10.1002/bit.21940

Wessel D, Flugge UI (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. Analytical Biochemistry 138(1):141-3

Zou G, Shi S, Jiang Y, van den Brink J, de Vries RP, Chen L, Zhang J, Ma L, Wang C, Zhou Z (2012) Construction of a cellulase hyper-expression system in *Trichoderma reesei* by promoter and enzyme engineering. Microbial Cell Factories 11(1):21 doi:10.1186/1475-2859-11-21