Characterization of the *Sorangium cellulosum* endocellulase CelA

Masterarbeit

Vorgelegt zur Erlangung des akademischen Grades eines Diplomingenieurs der Studienrichtung Molekulare Biotechnologie

Monika Friedl

Institut für Molekulare Biotechnologie

Technische Universität Graz

Univ.-Prof. Dipl.-Ing. Dr. techn. Helmut Schwab

Begutachter: Univ.-Prof. Dipl.-Ing. Dr. techn. Helmut Schwab

Betreuerin: Mag.rer.nat. Sandra Nußbaumer

Graz, März 2014

Graz University of Technology

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommene Stellen als solche kenntlich gemacht habe.

Graz, am …… ……..
(Unterschrift)

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

…… ……..
Date (signature)

I. Acknowledgements

First I want to thank Prof. Schwab for giving me the opportunity and support to complete my master thesis in the CD-laboratory group. I also want to thank my supervisor Mag. Sandra Nußbaumer for patiently introducing me to the new topics and whose guidance and knowledge was extremely important for this work. Furthermore I would like to thank Lactosan and the CD-Laboratory for Genetically Engineered Lactic Acid Bacteria for the financing of the master thesis and giving me the opportunity to work in this interesting project.

Thanks to my colleagues, especially Sabrina, Melanie, Simone and Sarah, for lending me an ear and making the time pleasant.

Finally I want to thank my family and friends for supporting me throughout my whole education; thanks for your encouragement whenever I needed you.

II. Abstract

In this work, we characterized the novel endocellulase CelA from *Sorangium cellulosum*. The protein consists of an Alanine, Serine, Threonine and Glycine rich repeat sequence followed by the GH 5 family catalytic enzyme. Since the function of the repeat sequence is not known by now, enzyme variants containing the N-terminal sequence, CelA, and the catalytic core enzyme CelAc, were tagged with N- or C-terminal 6x His Tag. After purification of the enzyme variants, the specific activity towards CMC was determined under optimal conditions (50°C, 1.75% CMC in NaP_i 50 mM, pH 7). Due to the higher activity of CelAc, 400 U/mg, compared to CelA (180 U/mg), no catalytic effect of the N-terminal repeat sequence towards CMC is suggested. Both enzyme variants were highly stable at 5-40°C, while inactivation occurred after 1 h at temperatures higher than 50°C. CelA and CelAc showed high pH stability (pH 4-9) at 4°C, whereas the incubation at 40°C inactivated the enzymes at pH 4 and reduced the activity at pH 5 and 9 to 60-80%. The enzyme variants were not significantly affected by metal ions or EDTA, DTT and DMSO. Urea and glycerin reduced the enzyme activity, while SDS inactivated both variants. Increased enzyme activity was measured after addition of ethanol or pre-incubation with 0.1% SDS.

The core enzyme CelAc was more active towards the insoluble substrate Whatman No. 1. filter paper Thus the N-terminal repeat sequence of CelA has no binding affinity to this filter paper. The synergistic degradation with CelAc and the exocellulase Cex from *Cellulomonas fimi* was determined with CMC or filter paper. The addition of Cex to the enzyme reaction increased the reducing sugar concentration to 160-170%. The degradation of heterogeneous substrates, containing highly crystalline structures, needs to be further optimized.

III. Kurzfassung

In dieser Arbeit wurde die neue GH 5 Endocellulase CelA von *Sorrangium cellulosum* charakterisiert. Das Enzym besteht aus einer Alanin, Serin, Threonin und Glycin reichen, N-terminalen Sequenz deren Funktion nicht bekannt ist. Enzymvarianten mit und ohne der N-terminalen Sequenz und C- oder N- terminalem 6 fach His-Tag wurden exprimiert und aufgereinigt. Die spezifische Enzymaktivität wurde unter optimalen Bedingungen (50°C, 1,75% CMC in NaP_i pH 7) bestimmt. Die höhere Aktivität der verkürzten Enzymvariante CelAc (400 U/mg im Vergleich zu 180 U/mg) deutet darauf hin, dass die N-terminale Repeat Sequenz keine katalytische Funktion aufweist. Beide Enzyme zeigten eine hohe Stabilität bei Temperaturen von 5 bis 40°C während höhere Temperaturen die Proteine nach einer Stunde inaktivierten. CelA und CelAc wiesen eine hohe pH Stabilität bei 4°C auf (pH 4-9), jedoch wurden die Enzyme nach Inkubation bei 40°C und pH 4 inaktiviert, und bei pH 5-9 wurden 60-80% der Aktivität gemessen. Es wurde kein signifikanter Einfluss von Metallen, EDTA, DTT oder DMSO detektiert, aber Harnstoff und Glycerin reduzierten die Enzymaktivität und SDS inaktivierte die Enzyme. Nach Inkubation mit Ethanol oder Vorinkubation mit 0,1% SDS wurde eine erhöhte Enzymaktivität gemessen.

CelAc zeigte auch bei der Degradierung von unlöslichen Substraten (Whatman No.1 Filterpapier) eine höhere Aktivität. Da Cellulose Bindedomänen im Bereich der heterogenen Katalyse, speziell bei kristalliner Cellulose, die Enzymaktivität erhöhen, wird der N-terminalen Repeat- Sequenz keine Affinität zu Cellulose zugeschrieben. Der synergistische Effekt von CelA und der Exocellulase Cex von *Cellulomonas fimi* wurde mittels CMC und Filterpapier als Substrat durchgeführt. Die Zugabe von Cex zur Enzymreaktion erhöhte die Konzentration an reduzierenden Zuckerenden auf 160-170 %. Jedoch muss der Abbau von heterogenen Substraten, die einen hohen Anteil an kristalliner Cellulose enthalten, noch weiter optimiert werden.

IV. Contents

V. List of abbreviations

ΔS	Without Signal sequence
AA	Amino acid
Amp	Ampicillin
APS	Ammonium persulfat
bla	β Lactamase
CBD	cellulose binding domain
CBM	carbohydrate binding module
CDS	Coding sequence
CelA	*Sorangium cellulosum* CelA
CelAc	*Sorangium cellulosum* CelA core enzyme
celAo	*Sorangium cellulosum celA*, codon optimized gene
celAoc	*Sorangium cellulosum celA*, codon optimized core gene
Cex	*Cellulomonas fimi* Cex
cex2.0	*Cellulomonas fimi cex*, codon optimized sequence
Cf	*Cellulomonas fimi*
cHis	C-terminal 6x His Tag
CL_{IS}	cell lysate insoluble fraction
CL_{LS}	cell lysate lysozyme insoluble fraction
CL_{LT}	cell lysate lysozyme total fraction
CL_{S}	cell lysate soluble fraction
CMC	Carboxymethyl cellulose
CV	Column volume
ddH_2O	Bidistilled water
DNS	3,5-Dinitrosalicylic acid
DOP	degree of polymerization
DS	Degree of substitution
DTT	Dithiothreitol
E_{IP}	eluate fraction impure
E_{P}	eluate fraction pure
FS	fermentation supernatant
GFCC	Gravity flow column chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LAB	lactic acid bacteria
lacI	Lactose repressor
LB	Lennox Broth
MTP	Microtiter plate
NaCit	Sodium citrate buffer
NaP_i	Sodium phosphate

nHis	N-terminal 6x His Tag
OD_{600}	Optical density at 600 nm
ONC	Overnight culture
P	PageRuler Fermentas
PAGE	Polyacrylamide gel electrophoresis
PASC	Phosphoric acid swollen cellulose
pHBAH	4-Hydroxy benzoic acid hydrazide
pUCori	pUC origin of replication
rrnb	rRNA Operon B Terminator
RS	reducing sugar
SDS	Sodium dodecyl sulfate
So ce	*Sorangium cellulosum*
Std.	DNA Ladder Mix
TAE	TRIS-acetate EDTA
TEMED	Tetramethylethylene diamine
WS	Working solution
WSC	Water soluble carbohydrates

VI. List of Figures

VII. List of Tables

1 Introduction

1.1 Cellulase application

Cellulose is the most common organic molecule since it is one of the main compounds in plant cell walls. It has been estimated that 50 to 100 billion tons of carbon are fixed as cellulose by photosynthesis per year (1; 2). Therefore it is of great interest to convert cellulose into shorter oligomers for further industrial applications such as bioenergy, food or agriculture (3).

The production of bioethanol is expected to increase from 31 billion liters in 2001 to 100 billion liters in 2015 (4; 5). In terms of sustainability it is important to use agricultural waste like wheat straw instead of sugarcane or corn for the production of bioethanol. Saccharification of crops and agricultural waste can be catalyzed by cellulases, prior the conversion to bioethanol (6; 4).

Cellulases have a large field of application in food industry. Cellulases, hemicellulases and pectinases, referred to as macerating enzymes, are used to improve the extraction yield of fruits and for clarification of fruit juices, beer and wine. Macerating enzymes hydrolyze the plant cell walls of raw materials like malt, fruits and vegetables. Thus, the viscosity decreases and the texture of fruit juices is maintained (3).

Silage is an important feeding source for livestock, especially for ruminants, since fresh feed is seasonal and not available during winter. Using ensiled stovers, such as corn, grass or cereals, the feed can be conserved. In silage fermentation, aerobic and facultative aerobic microorganisms and yeasts consume residual oxygen (1^{st} phase). After depletion of oxygen, lactic acid bacteria (LAB) become predominant and convert water soluble carbohydrates (WSC) to organic acids (lactic acid), thus decreasing the pH to 3.8-5.0 (2^{nd} phase). As a result, the growth of other microorganisms and fungi is inhibited (3^{rd} phase). The inhibition of *Clostridia* sp. is important since they produce butyric acid, leading to reduced feed intake of the animals. In the 4^{th} phase, the feedout, the silage is exposed to oxygen and aerobic microorganisms (yeasts, acetic acid bacteria and bacilli) are re-activated (7; 8; 9). The quality of the silage process is determined by the concentration of WSC, the buffering capacity and dry matter content (10). LAB express a set of cellulose degrading enzymes. The increase of fibrolytic enzymes (cellulases, hemicellulases and pectinases) increases the concentration of water

soluble carbohydrates and therefore facilitates the growth of LAB, thus decreasing the fibre content and improving the digestibility (3; 11).

1.2 Cellulose and model substrates

Cellulose is a homo-polymer of linear linked β-1,4 D-glucopyranose units and consists of 3 hydroxyl groups per anhydroglucose unit (Figure 1). The two ends of the cellulose chain are different: in the non-reducing end the anomeric carbon atom is involved in a glycosidic linkage, whereas the C1-atom of the reducing end is free (12). After rearrangement of this free hydroxyl groups on the C1-atom an aldehyde group is built, leading to reducing properties.

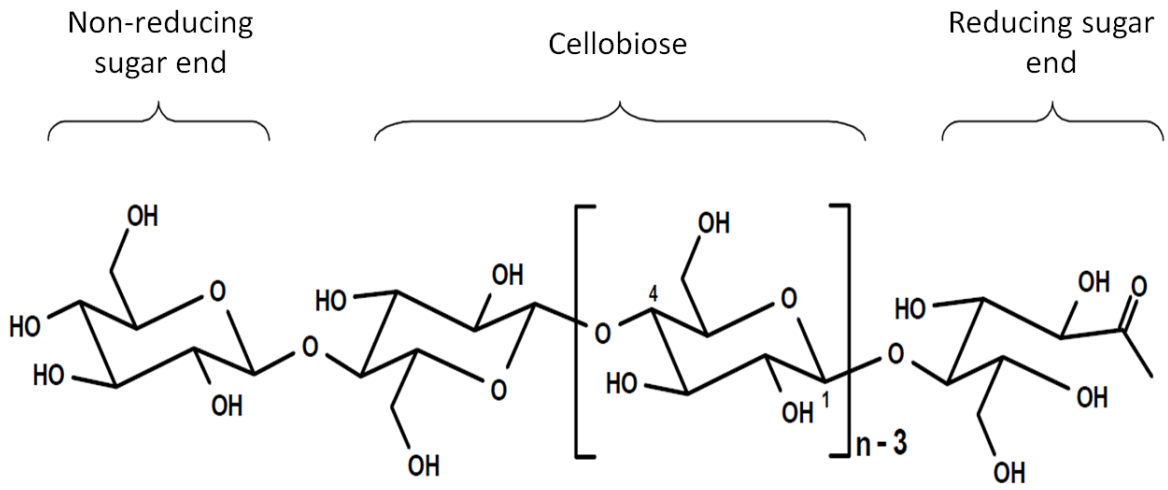

Figure 1: Structure of a cellulose chain. Cellulose consists of a non-reducing, and a reducing end. n indicates the degree of polymerization **(13)**.

Through the 3 hydroxyl groups per glucose molecule, hydrogen bonds are formed intra- or intermolecular, leading to various crystalline structures. These crystalline structures can be divided in 4 allomorph groups, depending on the orientation of the cellulose chains and orientation in an elementary cell of crystal structure. Cellulose I, the native cellulose, can be divided in the subunits cellulose Iα and Iβ. Celluloses from bacteria or algae contain more Iα, whereas higher organisms like plants consist mainly of cellulose Iβ. The difference between those subunits is the crystal structure: Iα belongs to triclinic crystal structure, Iβ to the monoclinic structure. The cellulose chains in cellulose I are oriented parallel, building two inter-, and one intra molecular hydrogen bond (13).

Using mercerization (alkali solubilization) and regeneration (recrystallization), cellulose I can be converted into cellulose II, which is also called regenerated cellulose. Since this conversion is irreversible, cellulose II is suggested to be highly stable. The cellulose chains of regenerated cellulose are orientated antiparallel, building 1 inter-, and 1-intramolecular hydrogen bonds per glucose unit (12).

Cellulose III can be built from Cellulose I or II by treatment with dry liquid ammonia, this crystalline form is reversible. Cellulose IV originates from cellulose I-III through thermal treatment.

The chain length of cellulose, also termed “degree of polymerization” (DOP), strongly depends on the natural source, the isolation and purification method. DOP of hundreds to ten thousands are reported for native cellulose, while processing reduces the chain length (12).

Due to the compact crystalline structure of native and regenerated cellulose, the substrate accessibility is an important factor in enzymatic hydrolysis. The effectiveness of degradation depends on the degree of polymerization, the particle- and pore size (14). The high crystalline structure of cellulose (microfibril) is interrupted by amorphous regions, which are less ordered. Thus, the structure is less compact and more accessible for hydrolysis. It is not known, whether the amorphous regions of cellulose cover the crystalline microfibrils inside of the chains or if amorphous and crystalline structures are alternating (12; 15).

Caused by broad variations of cellulose properties, depending on the natural source, more homogenous model substrates were created to ensure comparability and repeatability for enzymatic degradation in laboratory scale. The most common model substrates for cellulases are Carboxymethyl cellulose (CMC), filter paper (Whatman No. 1 filter paper) and microcrystalline cellulose (Avicel). As shown in Table 1, the properties of those substrates are very diverse. While CMC is a soluble substrate, the content of crystalline cellulose is 45% for Whatman No. 1 filter paper and 50-60% for Avicel. The fraction of β-glycosidic bonds, which are accessible for cellulases, is 100% for CMC, approx. 2% for Whatman No. 1 filter paper and 0.6% for Avicel. The degree of polymerization varies from 150 to 2,800.

Table 1: Physical properties of cellulosic model substrates: Carboxymethyl cellulose (CMC), Whatman No. 1 filter paper and Avicel: Crystallinity Index (CrI), fraction of β-glycosidic bonds accessible for cellulases (F_a) and dregee of polymerization (DOP) **(2)**.

	CrI	**F_a [%]**	**DOP**
CMC	-	100	100-2,000
Whatman No. 1 filter paper	~0.45	1.8	750-2,800
Avicel	0.5-0.6	0.6	150-500

Avicel is prepared by acidic hydrolysis out of wood pulp. In this process most of the amorphous cellulose and the whole hemicellulose is removed. Whatman No.1 filter paper is widely used as a natural- model substrate since its homogenous structure and availability.

CMC consists of β-1,4 linked glucose molecules, where some hydroxyl groups are methylated with carboxyl groups, leading to a higher solubility in water. The Haworth projection of this molecule can be found in Figure 2. When CMC is used as substrate for the enzyme activity assay, the degree of substitution (DS, 0.7 for commercial CMC) has to be considered. Since only non-substituted glucose units are hydrolyzed, the enzyme activity assay should be limited to a hydrolysis rate of 2% (2).

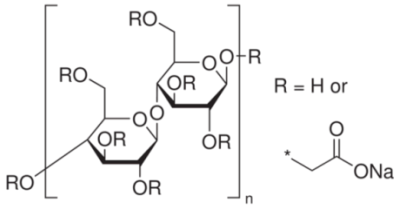

Figure 2: Carboxymethyl cellulose: Haworth projection of the molecule **(16)**

Another model substrate is phosphoric acid swollen cellulose (PASC), which is made from Avicel by dissolution in phosphoric acid, followed by regeneration in water. While Avicel has a highly crystalline structure, PASC is mostly amorphous; therefore it has a higher cellulose accessibility compared to Avicel (17).

1.3 Degradation of cellulose

The hydrolysis of cellulose is catalyzed either by acids or enzymatically. Acidic processes can degrade cellulose at high temperatures and high acid concentrations. The disadvantages of this process are corrosion of the equipment and production of unspecific side products leading to impure products (1). Since 1950, enzymatic hydrolysis was investigated using cellulase producing *Trichoderma* strains, a filamentous fungus. *Trichoderma reesi* produces a variety of enzymes, necessary for a synergistic degradation of cellulolytic material: endocellulases, exocellulases and β-glucosidases (1; 3; 18).

Aerobic fungi (*Trichoderma*) and bacteria (*Cellulomonas*) mainly express the enzymes in dissociable form, while some anaerobic bacteria, i.e. *Clostridium*, assemble the enzymes to a cellulosome (see Figure 3) (19). One assumption for the expression of the cellulosome by anaerobic organisms is energy limitation. Binding of the enzymes enables degradation near the cell and uptake of the products. Soluble enzymes lead to soluble products, which are also available for competing organisms (20).

Cellulosomes consist of Type I and II cohesin and dockerin, the anchoring protein, a cellulose binding domain and the enzymatic subunits. Bacterial type I dockerin consists of 70 amino acids, which contains an EF-hand motif and a calcium binding domain necessary for the dockerin-cohesin interaction. The scaffoldin subunit consists of multiple copies of cohesin I and type II dockerin. Type II dockerin and cohesin (also Ca-dependent) ensure anchoring of the cellulosome. The anchoring protein contains a S-layer homology module (SLH) for interaction with cell-wall components. The size and complexity of the cellulosome depends on the dockerin-cohesin interaction: some dockerins prefer specific interaction, while others bind unspecific to cohesin (21; 22).

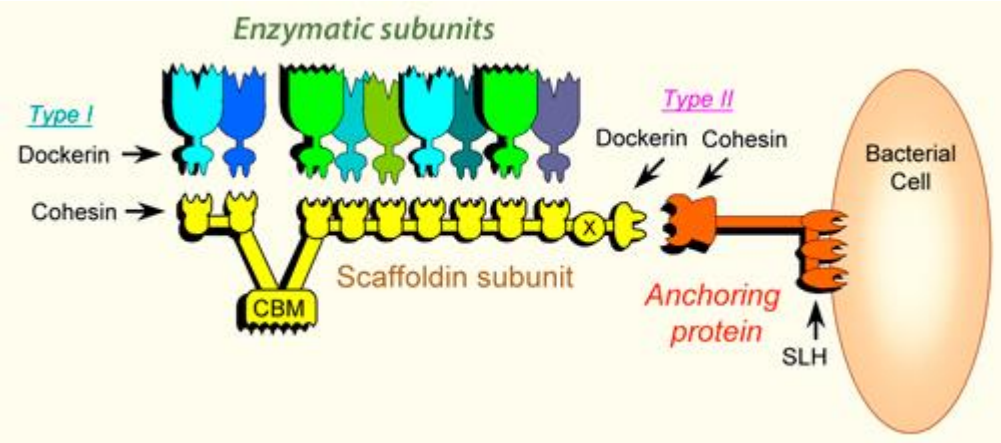

Figure 3: Scheme of a Cellulosome. Cellulosomes consist of various enzymatic subunits, bound to a scaffoldin unit via dockerin-cohesin type I structures. The scaffoldin subunit is bound to the bacterial S-layer cell wall by dockerin-cohesin type II **(21).**

Free cellulases often contain a carbohydrate binding module (CBM) which is not always present in cellulosomes. The CBM is bound through a linker either on the N- or C- terminus of the catalytic domain. By now, 67 families of CBM are identified, and classified according to their protein structure, substrate affinity and amino acids located at the active site (23). Hydrolyzation of crystalline cellulose can be improved by the CBM through loosen of the hydrogen bonds. The CBM bind to the substrate and thereby enhances the processivity of the enzymes (24).

The degradation process can be divided in 2 phases (see Figure 4): the primary, which occurs on solid substrates and the secondary hydrolysis, occurring in the liquid phase. During primary hydrolysis, endocellulases (EC 3.2.1.4) hydrolyze accessible β-1,4-glycosidic bonds producing new chains of variable length since they bind randomly inside the cellulose chain. Therefore, endocellulases decrease the DOP of substrates, whereas exocellulases processively hydrolyze β-1,4-glycosidic bonds. They bind either from the reducing (EC 3.2.1.176) or non-reducing end (EC 3.2.1.91); the products are tetrasaccharides, trisaccharides or disaccharides. Products, with a DOP <6-8 stay insoluble since they have a higher affinity for other cellulose chains than for aqueous solution (25; 20). In the secondary hydrolysis, small oligomers (up to a DOP of 6) are hydrolyzed by β-glucosidases (EC 3.2.1.21) yielding sugar monomers (2). The rate limiting step in hydrolysis is the primary depolymerization by endo- and exocellulases.

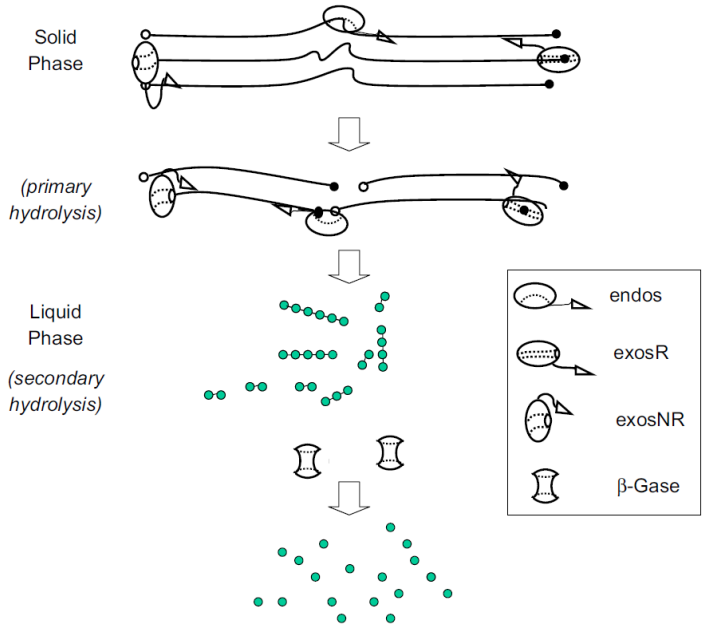

Figure 4: Degradation of cellulose by a cellulases. In primary hydrolysis endocellulases (endos) hydrolyze glycosidic bonds inside the polymer; exocellulases hydrolyze either from the reducing (exosR) or non-reducing end (exosNR). In secondary hydrolysis β-glucosidases (β-Gase) hydrolyze small oligomers to monosaccharides **(2)**.

Cellulases are classified in glycoside hydrolase families according to their tertiary structure, particularly at the active site (23). The tertiary structures of some families are similar, even if the primary sequence is different; those families are classified in GH Clans.

Two glutamic or aspartic acids are necessary as proton donor and nucleophile. The glycosidic bond is hydrolyzed either by retaining or inverting the anomeric configuration (see Figure 5). During the inverting mechanism water directly binds to the glycosidic group, replacing the leaving group. The retaining mechanism involves an enzyme-substrate intermediate. A glutamic or aspartic acid of the enzyme binds the glycosidic bond before water hydrolyzes it again, releasing the free glycosidic bond and enzyme (26; 20).

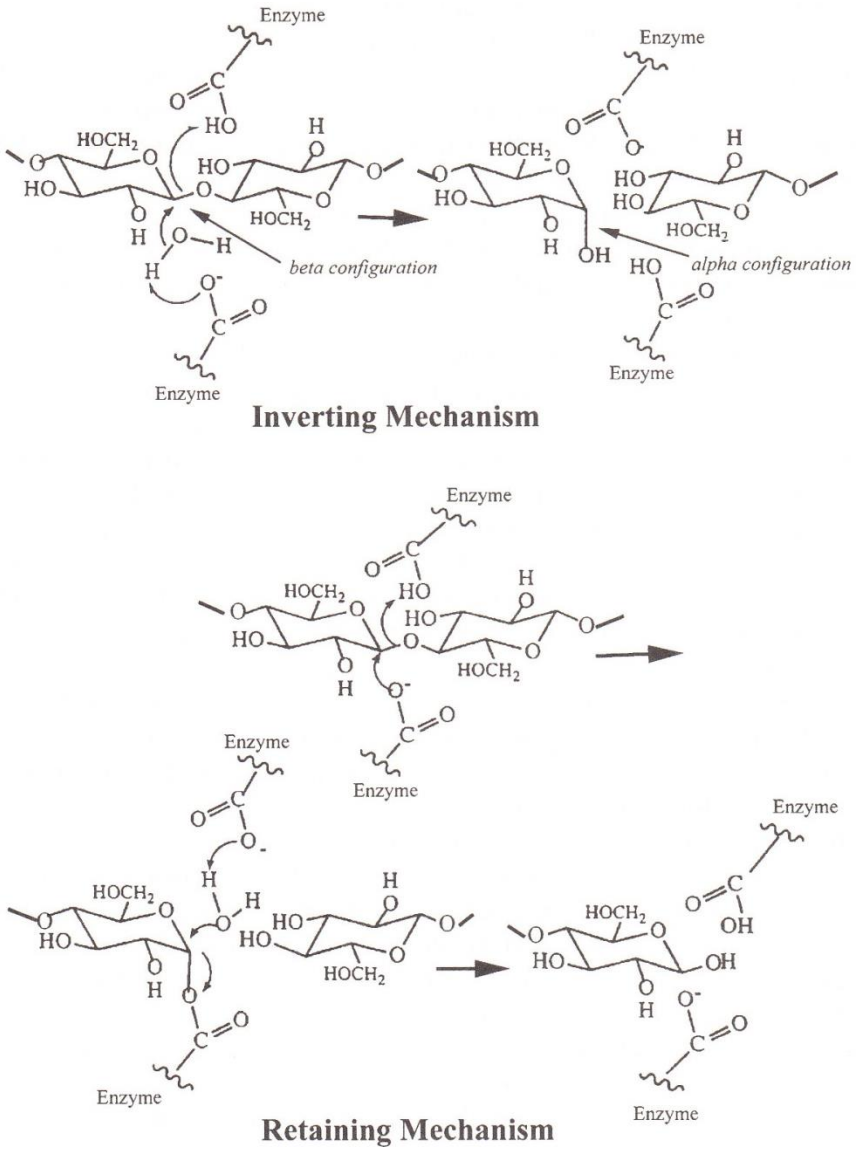

Figure 5: Mechanism of enzymatic hydrolysis of glycosidic bonds. The cleavage of the glycosidic bond can be hydrolyzed either by retaining or inverting mechanism **(20)**.

1.4 *Sorangium cellulosum* endocellulase CelA

*Sorangium cellulosum So ce*56 is a ubiquitous gram-negative myxobacterium and was isolated from a soil sample in 1985 in Indonesia (27). The circular genome of *So ce*56 (13 Mb) encodes for approx. 10,000 predicted protein coding sequences (CDS) and has an average GC-content of 71.4% (28). *Sorangium* sp. are known for their complexity, involving swarm-formation, cell-to-cell signaling and the ability to produce pharmaceutical, bioactive metabolites such as antitumor components (29).

As they occur on wood and degrade organic material, they are able to decompose complex plant material through their secreted hydrolytic enzymes (30). Therefore, this bacterium was chosen for characterization of its novel endocellulase CelA (28). CelA is a protein of 409 amino acids and it contains a putative signal sequence of approx. 35 amino acids (31) followed by repetitions of Serine, Alanine, Threonine and Glycine (80 amino acids) and the functional GH5 core enzyme. The functionality of this repetitive sequence is yet unknown, but no cellulose binding affinity is suggested. Blast search of the repeat sequence resulted in 1 similar sequence: Endocellulase 5 from *Verticillium alfalfae* VaMs. 102 (61% sequence identity, NCBI Reference XP_002999920.1) (32). The similar sequence is located at the C-terminal region, followed by fungal cellulose binding domain (CBM 1). The function of this structure is also not known by now.

CelA is annotated as endo-β-1,4-cellulase (EC 3.2.1.4) and searching of conserved domains resulted in glycoside hydrolase family 5 (GH5) (32). This glycoside hydrolase family has a $(\beta/\alpha)_8$ tertiary structure. This structure was predicted by "Psipred" (Figure 6) and "M4T server" (33; 34). Due to the lack of homologous sequences of the N-terminal repetitions, the predicted protein structure begins at V117. Template for the structure prediction is the endocellulase Cel5A from *Bacillus agaradhaerens* (PBD code 7A3H) (35). The active site of CelA is depicted in Figure 6. The two necessary glutamic acids (E132, E219) for CelA are colored in blue. The distances between the glutamic acids is 3- 5 Å. The enzymatic mechanism is proposed to be retaining, which involves an enzyme-substrate-intermediate, whereby one glutamic acid is involved in substrate binding (see Figure 5) (23).

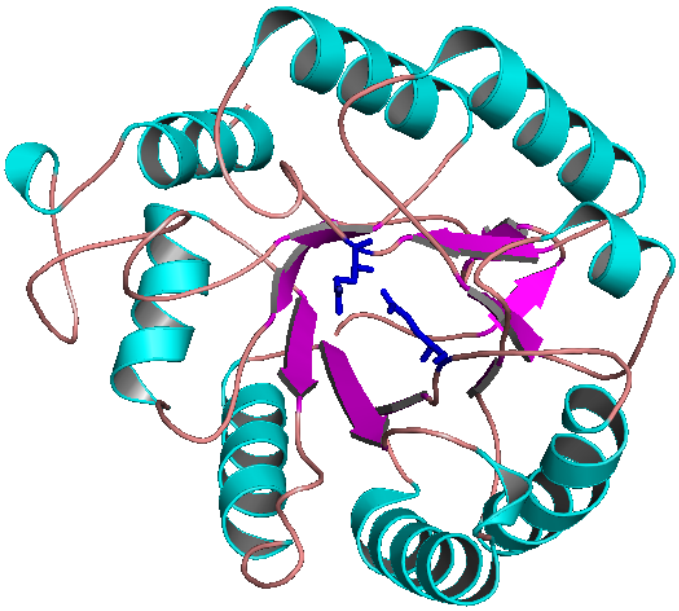

Figure 6: Structure prediction of CelA. The tertiary structure $(\beta/\alpha)_8$ of CelA after alignment with Cel5A (7a3h). The two glutamic acids in the active site of CelA are shown in dark blue **(33)**.

The active site of endocellulases consists of 4-7 amino acids, arranged in a cleft of the protein (see Figure 7). The substrate is bound via hydrogen bond and π-stacking of carbohydrate ring with aromatic amino acids. According to the homologous sequence Cel5A from *Bacillus agaradhaerens*, amino acids interacting with Cellotriose (PDB code 1HF6) were highlighted in the proposed structure of CelA. Similar interactions were determined in the GH5 endoglucanase EngD of Clostridium cellulovorans. The crystal structure of the protein was determined and amino acids, interacting with Cellotriose in the active site identified (Arg63, His107, Asn151, Glu151, His227, Tyr229, Glu275 and Trp308) (36).

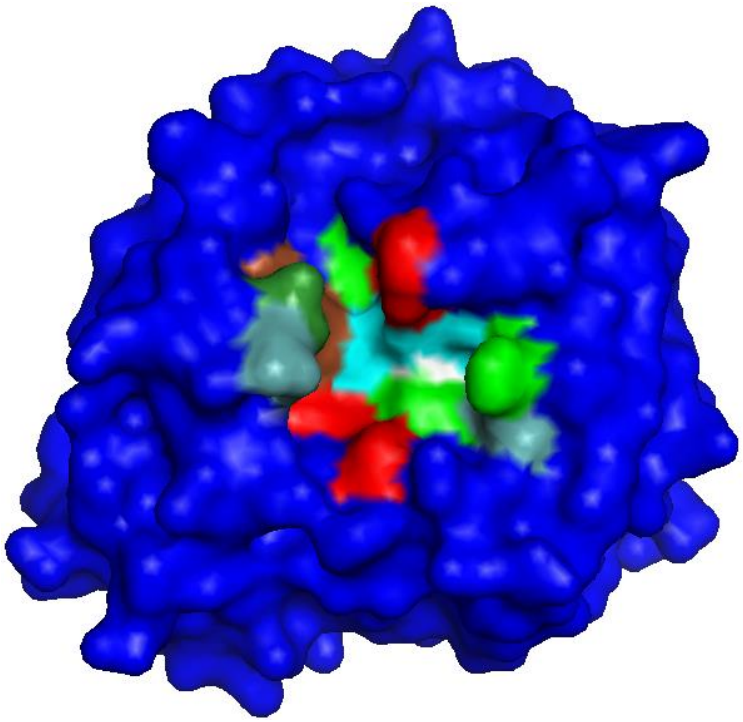

Figure 7: Structure prediction of CelA. The active site (E132 and E219) in the cleft is colored in cyan. AAs, possibly interactive with the substrate are highlighted. Green: His (97, 99, 191), dark green Ala225; red: Trp (31, 169, 235); grey: Asp (98, 226); gold: Tyr 193; white: Asn131.

1.5 *Cellulomonas fimi* endoxylanase / exocellulase Cex

Cellulomonas fimi is a gram-positive actinobacteria and its genome is GC rich (72 mol%) (37). Cex belongs to GH 10 (clan A, $(\beta/\alpha)_8$ configuration). The enzyme is predicted to have xylanase and exocellulase activity. A CBM is located at the C-terminal end of the catalytic domain. These 100 AA are classified as CBM family 2 with affinity to cellulose, chitin and xylan. Xylan is a hemicellulose and consists of β-1,4- linked D-xylose molecules. The enzyme belongs to class 3.2.1.91; therefore it hydrolyzes from the non-reducing end of the cellulose chain. The crystal structure of Cex was determined in complex with thio-linked xylotetraose (PDB code 3CUI). The structure is shown in Figure 8: the predicted active site is shown in dark blue (Glu43, Glu127, Gln203, Trp273, Asn44) and the bound thio-linked xylotetraose is colored black.

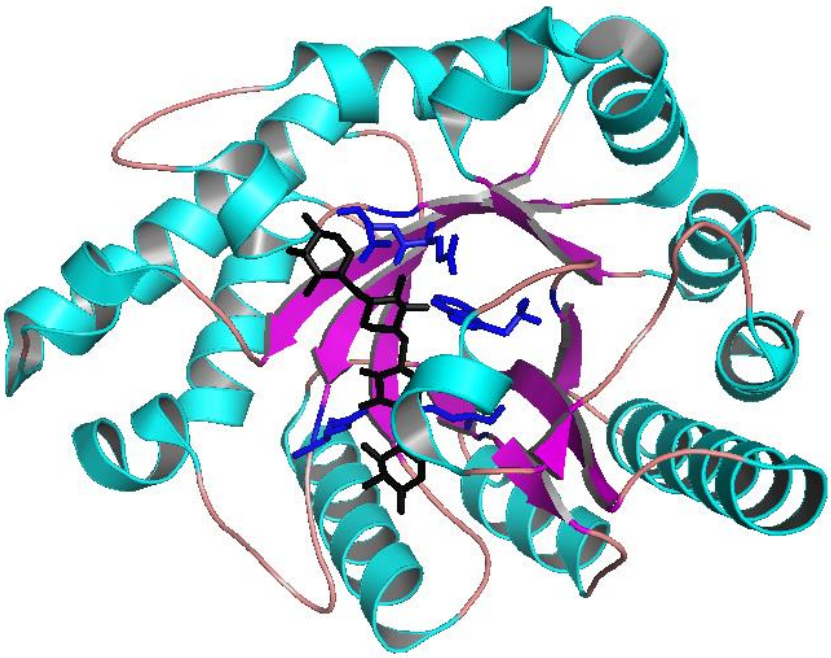

Figure 8: Crystal structure of Cex (3CUI). The active site is highlighted in dark blue: E43, E127 and N44, Q203, W273; black: thio-linked xylotetraose.

1.6 Enzyme activity assays

For the determination of enzyme activity while enzyme purification, two agar plate assays were used. Endocellulase activity was monitored using Congo red staining which was routinely done on CMC agar plates. After destaining of the CMC agar plate, orange halos are visible against a red background. The orange halos occur if endocellulases degraded CMC to oligosaccharides with a maximal chain length of four sugar molecules and Congo red cannot intercalate. The size and intensity of the halos on the agar plates correspond to the enzyme activity and thus one can conclude the amount of active enzyme.

Exocellulase activity was monitored using 4-Methylbelliferyl β-D-cellobioside (MUC). This model substrate consists of a fluorescent dye bound on the reducing end of cellobiose. Agar plates containing MUC are incubated with the protein fractions, the exocellulase hydrolyzes the glycosidic bond, and the enzyme activity can be detected by fluorescence under UV light excitation.

Since those plate assays are qualitative, reducing sugar assays were used for quantitative determination of the enzyme activity.

The cellulase activity can be measured by incubation of single enzymes or by determination of the synergism of several enzymes. As described previously, several model substrates are available. For the determination of endocellulase activity, CMC was chosen as soluble substrate. For synergistic degradation, soluble and insoluble substrates (filter paper) were incubated with CelA variants and Cex. Using those substrates, the enzyme activity can be calculated after determination of the reducing sugar amount, released in the enzyme reaction. Reducing sugars, released from the soluble substrate CMC, were determined using 4-Hydroxybenzoic acid hydrazide (pHBAH) reagent (38). Two pHBAH molecules bind to one free reducing sugar end forming osazone, a yellow derivate, and leading to a measurable absorbance shift. After incubation of insoluble substrates, the reducing sugar concentration was measured using the 3,5-Dinitrosalicylic acid (DNS) assay (39). 1 mol of reducing sugar reduces 1 mol of 3,5-Dinitrosalicylic acid, forming 3-Amino-nitrosalicyclic acid and leading again to an absorbance shift.

Both reducing sugar assays were calibrated with glucose. Therefore, the amount of reducing sugar is expressed as glucose concentration. The formula depicted in Figure 9 was used for the calculation of the specific enzyme activity of the enzymes.

$$\frac{Reducing\ sugar\ concentration\left[\frac{mg}{ml}\right] * 1000}{Molecular\ weight\ glucose\left[\frac{g}{mol}\right] * Time\ [min] * Protein\ concentration\ in\ enzyme\ reaction\left[\frac{mg}{ml}\right]} = IU\left[\frac{\mu mol}{min * mg}\right]$$

Figure 9: Formula for the calculation of the specific enzyme activity.

1.7 Objective of this work

This work was done in the Christian-Doppler laboratory for genetically engineered lactic acid bacteria, module 2: enzyme engineering. The aim of the project is to find new cellulases, which can be expressed in lactic acid bacteria, mainly in *Lactobacillus plantarum*. Integration of cellulases into the *L. plantarum* genome enables enhanced degradation of cellulose, thus improving the silage process.

In this diploma thesis the novel endocellulase CelA from *Sorangium cellulosum* was characterized. The optimal conditions for the measurement of the specific enzyme activity were determined as well as the enzyme stability. Moreover, the influence of metals, reagents, chemicals and inhibitors towards the enzyme activity was tested.

For the determination of the enzyme activity, 3 model substrates were chosen: the soluble Carboxymethyl cellulose (CMC) and the insoluble substrates filter paper (Whatman No. 1 and Macherey Nagel MN 619 de) and microcrystalline cellulose (Avicel).

The synergistic effect of CelA with the exocellulase *Cellulomonas fimi* Cex was determined on several substrates.

2 Materials and Methods

2.1 Materials

Reagents and chemicals were ordered at Roth (40) or Sigma Aldrich (16) in chemical purity or per analysis. Reagents, ordered from different companies, are mentioned separately.

2.1.1 Cells

Table 2: *E. coli* cells used for cloning and expression of CelA and Cex variants (41)**.**

	Genotype
***E. coli* XL1**	*Tc-r; recB, recI, sbcC201, uvrC, umuC::Tn5(Km-r), mcpA, mcrB, mrr, lac, Δ(hsdRMS), endA1, gyrA96, thi, relA1, supE44(F´), proAB, lacI9ZΔM15, Tn10(Tc-r)*
***E. coli* BL21**	*F- dcm ompT hsdS(r_B- m_B-) gal [malB$^+$]$_{K-12}$(λ^S)*
***E. coli* BL21 (DE3)**	*F- dcm ompT hsdS(r_B- m_B-) gal λ(DE3)*

2.1.2 Primer

All primers were ordered from Integrated DNA Technology. The primer sequences and melting temperatures are shown in Table 3 (42).

Table 3: Primers for cloning and sequencing: the restriction sites are underlined in the primer sequence.[1] Sequencing primers

No.	**Name**	**Sequence**	**Annealing Temp.**
173	cexDNA2.0ΔS_*Nde*Ifor	5'-ATCTCATATGGCCACCACCCTGAAAGAAGC-3'	63°C
174	cexDNA2.0_Δstop_*Hin* *d*III	5'-ATATAAGCTTGCCGACGGTGCAGGGTGTTC-3'	65°C
175	cexDNA2.0_*Nde*Ifor	5'-ATCTCATATGCCTCGTACCACCCCAGCTCC-3'	66°C
176	cexDNA2.0_*Hind*III	5'-ATATAAGCTTTTAGCCGACGGTGCAGGGTG-3'	63°C
168	cexΔS_*Nde*Ifor	5´-ATCTCATATGGCGACCACGCTCAAGG-3´	62°C
130	cexΔstop_*Hind*III	5'-ACATAAGCTTGCCGACCGTGCAGG-3'	66°C
133	pMS470fwd[1]	5'-GCATAATTCGTGTCGCTCAAGG-3'	57°C
134	pMS470rev[1]	5'-GCAAATTCTGTTTTATCAGACC-3'	50°C

2.1.3 Plasmid

Expression of the enzymes in *E. coli* cells was fulfilled using the expression vector pMS470. The copy number is regulated by the high copy origin of replication pUC leading to 10-100 copies per cell. The expression is regulated through the IPTG-inducible tac promoter. The transcription is terminated using the antiterminator rrnB. The resistance gene for β-Lactamase is encoded for the later cleavage of Ampicillin. For later purification of the enzymes vector variants encoding N- or C- terminal 6x His Tag were used for expression. The vector pMS470nHis and pMS470cHis were constructed in CD-laboratory of IMBT, vector cards of pMS470cHis, pMS470nHis and variants with inserts are shown in Figure 10 A-D.

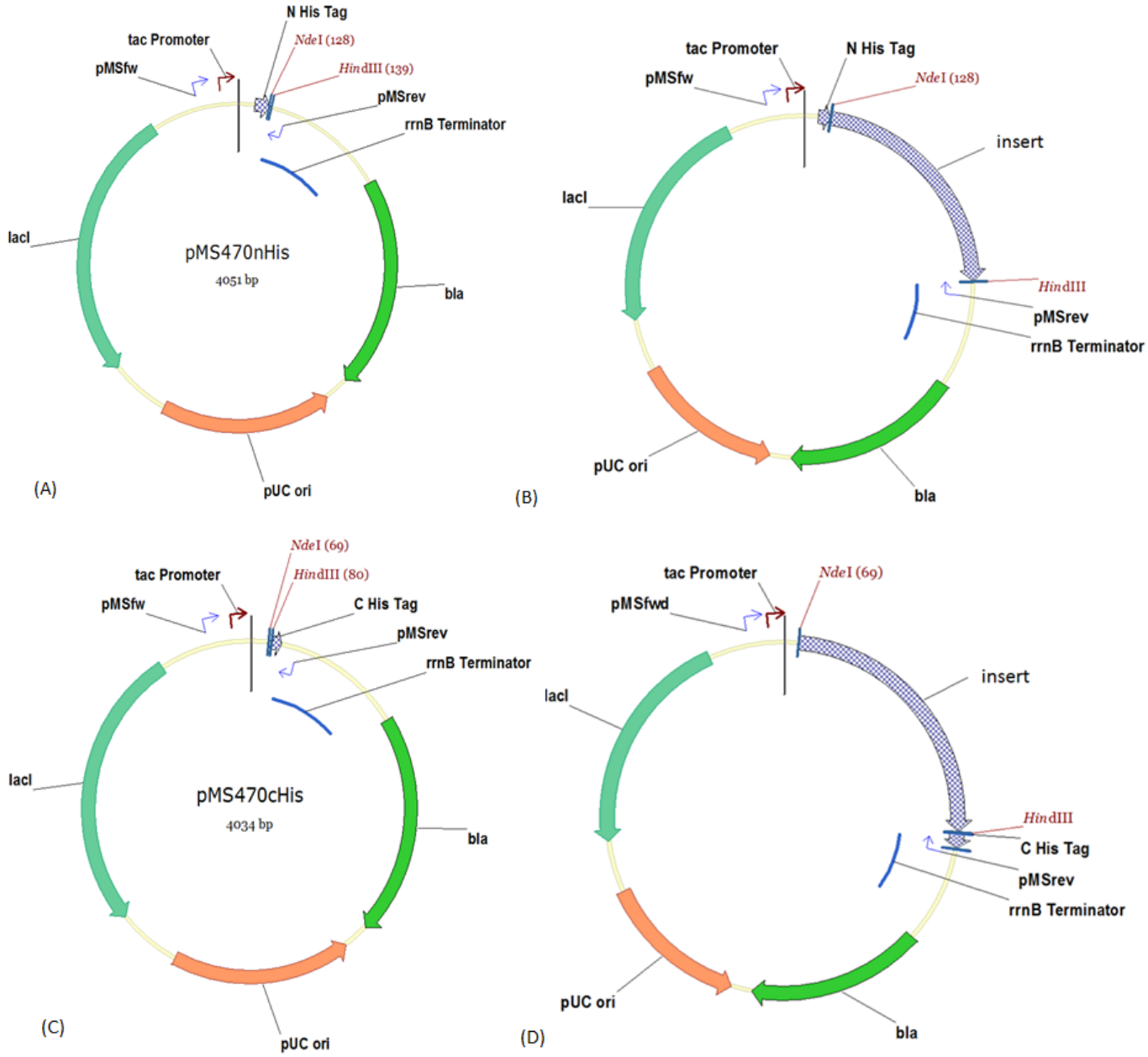

Figure 10: Vector maps. (A) pMS470*nHis*, (B) pMS470*nHis* with insert, (C) pMS470*cHis*, (D) pMS470*cHis* with insert

Schemes of the wild type and the truncated CelA and Cex variants are shown in Figure 11; whereas the corresponding gene- and amino acid sequences can be found in the Appendix (7.2 Gene and amino acid sequences). For a successful expression of the genes in *E. coli* strains, the codon optimized DNA sequence was cloned into pMS470*nHis* and pMS470*cHis* vector without signal sequence.

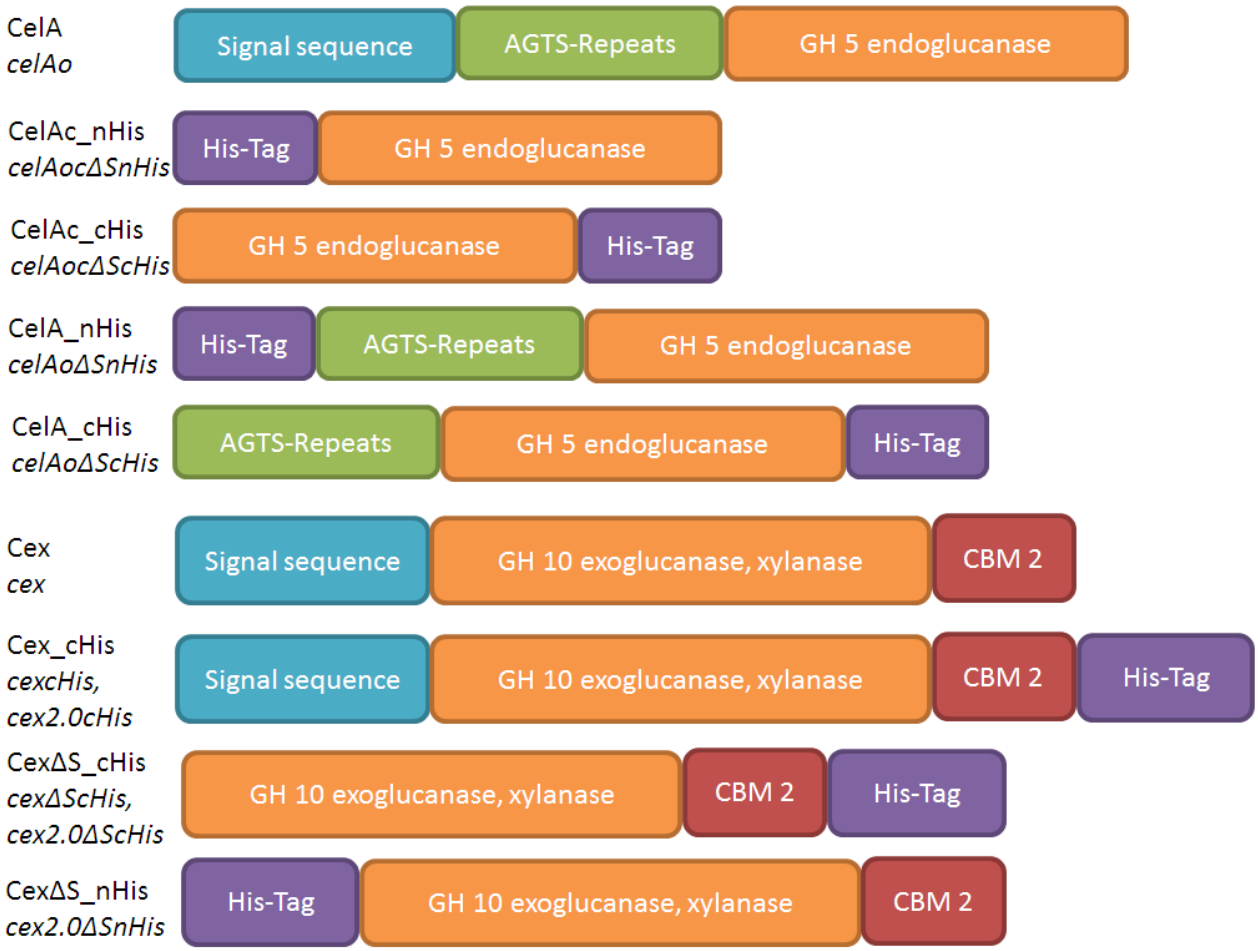

Figure 11: Schemes of the wild type and the truncated CelA and Cex variants with and without N- or C- terminal 6x His Tag.

2.1.4 Standards for SDS-PAGE and Agarose gel

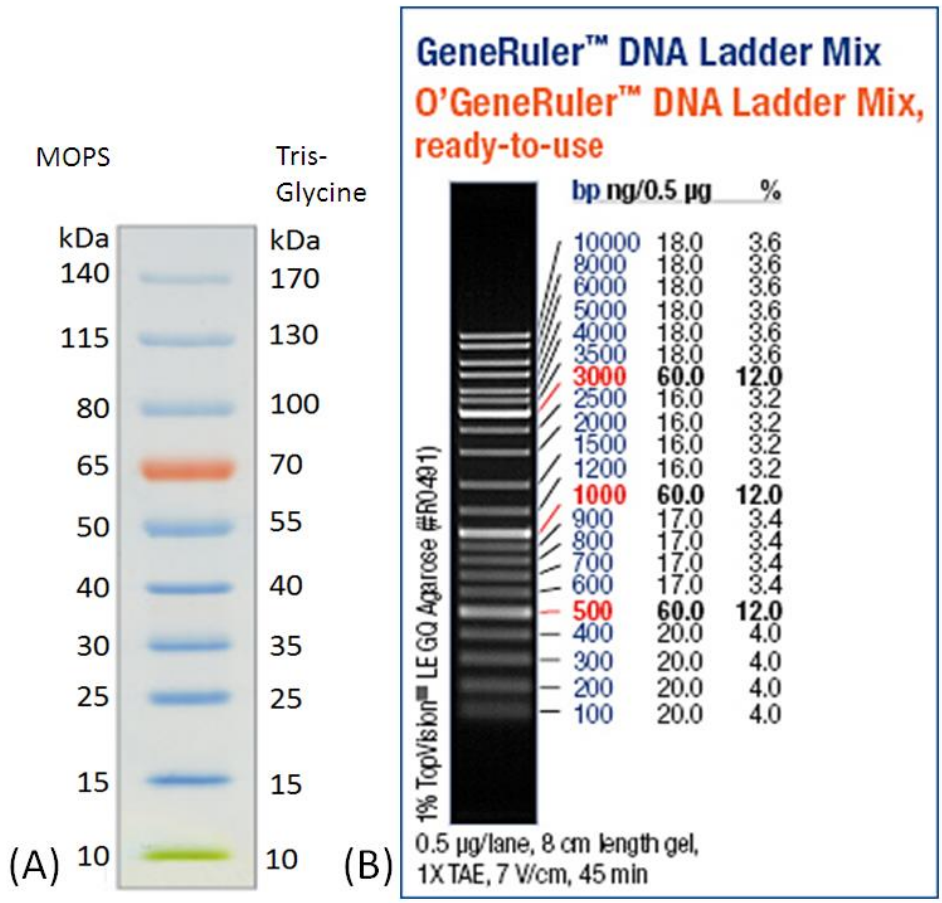

Figure 12: Standards for SDS-PAGE and Agarose gel: (A) PageRuler Prestained Protein Ladder (Fermentas). The size of the protein bands in MOPS and Tris-glycine buffer is given in kDa. (B) GeneRuler™ DNA Ladder Mix (Fermentas).

2.2 General methods

2.2.1 Isolation of Plasmid DNA

For the isolation of plasmid DNA according GeneJET™ Plasmid Miniprep Kit (Fermentas), *E. coli* XL1 yielding pMS470 vector with the corresponding gene, was streaked out on LB-Amp plates and incubated at 37°C overnight. Cells were resuspended in 250 µl Resuspension solution, 250 µl Lysis solution were added and mixed by inverting of the tube. After addition of 350 µl of Neutralization solution, the tube was centrifuged for 10 min with 16,100 x g at room temperature using Eppendorf centrifuge 5415R. The supernatant was loaded onto the GeneJET™ spin column and centrifuged for 1 min with 16,100 x g. The spin column was washed twice with 500 µl Wash solution and centrifuged for 1 min with 16,100 x g. After centrifugation of empty spin column, it was transferred to a new tube and the DNA was eluted in 50 µl ddH_2O after incubation for 2 min.

2.2.2 Polymerase chain reaction

Each PCR reaction contained 25 pmol primer (forward and reverse), 10-15 ng template DNA, 1 U Phusion DNA Polymerase, 10 µmol dNTP`s and 1x GC buffer in a total volume of 50 µl. For amplification of the genes, the following temperature profile was used: 98°C (1 min), 35 cycles: 98°C (10 sec), 55°C (30 sec), 72°C (30 sec); 72°C (10 min), 4°C (hold).

2.2.3 Agarose gel

1 g Agarose (Biozyme LE) was dissolved by heating in 100 ml TAE buffer (40 mM TRIS, 20 mM Acetic acid, 1 mM EDTA). After short cooling, a small amount of Ethidium bromide solution was added and the gel was poured into a tray. For sample preparation DNA loading dye was added to give a final concentration of 1x loading buffer. Preparative gels were developed at 90 Volt, control gels at 120 Volts until a proper separation of DNA was reached. 5 µl GeneRuler DNA Ladder Mix (Fermentas, Thermo Scientific) were loaded as standard.

2.2.4 DNA restriction digest

The vectors and inserts (5 µg vector DNA, or 1µg insert DNA) were digested using 20 U *Nde*I and 10 U *Hind*III, for preparative DNA digest, and 10 U *Nde*I and 5 U *Hind*III for control DNA digest, according to the double digest protocol of Fermentas. The enzyme reaction was incubated overnight in 1x buffer R at 37°C.

2.2.5 DNA purification from Agarose gel slices and PCR purification

Gel slices and PCR products were purified using Promega Wizard SV Gel and PCR Clean Up System. Gel slices, excised from Agarose gels, were dissolved at 65°C in an Eppendorf tube mixed with 1 ml Binding solution per gram gel slice. PCR products were mixed with the equal volume of Binding solution prior loading in SV Minicolumn. After incubation of DNA on Minicolumn for 1 min and centrifugation with 16,100 x g for 1 min using Eppendorf centrifuge 5415R, the Minicolumn was washed twice: the first step was done with 700 µl Wash solution and the second step with 500 µl Wash solution. After the second washing step the tube was centrifuged for 5 min, followed by a centrifugation step of the empty column for 1 min. The Minicolumn was transferred to a new tube, and the DNA was eluted in 30 µl ddH_2O after incubation of 5 min.

2.2.6 Measurement of DNA concentration

The concentration of purified DNA was measured at a wavelength of 260 nm by NanoDrop 2000c (Thermo Scientific).

2.2.7 Ligation of DNA fragments

1.5 U of T4 DNA ligase (Promega) were incubated with 100 ng vector and 3x molar excess of insert in 1x T4 DNA ligase buffer at 8°C overnight. After inactivation at 70°C for 10 min, the whole ligation product was desalted using dialysis plates (Nitrocellulose membrane filter 0.025 µm, Milipore VSW01300).

$$mass\ of\ the\ insert\ [ng] = \frac{mass\ of\ the\ vector\ [ng] \times size\ of\ the\ insert\ [bp] \times molar\ ration\ \frac{vector}{insert}}{size\ if\ the\ vector\ [bp]}$$

Figure 13: Formula for the calculation of the ligation reaction.

2.2.8 Electro-competent cells

20 ml LB media containing Ampicillin 100 µg/ml (LB-Amp) were inoculated with *E. coli* BL21, *E. coli* BL21 (DE3) or *E. coli* XL1, incubated overnight at 37°C, 120 rpm and 80% humidity. Main cultures (each 400 ml LB-Amp) were inoculated to an optical density at 600 nm (OD_{600}) of 0.1 (measurement with Eppendorf BioPhotometer). The cells were grown at 37°C, 120 rpm and 80% humidity until an OD_{600} of 0.7 to 0.8 was reached. The chilled cultures were harvested by centrifugation with 3,220 x g at 4°C for 20 min using Beckman Coulter Avanti J-20XP centrifuge. The cell pellets were washed twice with sterile distilled water (1^{st} step with 250 ml, 2^{nd} with 100 ml), each washing step was followed by centrifugation (3,500 x g, 20 min, 4°C). The cells were washed in 40 ml 10% Glycerol, followed by centrifugation (3,500 x g, 20 min, 4°C). Finally, the cells were resuspended in 1-2 ml 10% glycerol and 90 µl aliquots were pipetted into Eppendorf tubes. After freezing in liquid nitrogen, the cells were stored at - 80°C.

2.2.9 Electrotransformation

The DNA ligation products were transformed into competent *E. coli* XL1 cells and the plasmids containing the target genes, isolated according GeneJET^{TM} Plasmid Miniprep Kit, were transformed into competent *E. coli* BL21 and *E. coli* BL21 (DE3) cells using electroporation. 40 µl of competent cells were mixed with 4 µl ligation mixture or 1 µl plasmid DNA, stored on ice for 10 min and transformed with the electroporator (BioRad Micropulser, 2.5 kV, ≤6 ms). After addition of 1 ml LB media, the cells were regenerated for 45 min at 37°C and 600 rpm in the Thermomixer Eppendorf comfort. 100 µl were plated out on LB-Amp agar plates (LB agar containing 100 µg/ml Ampicillin) and incubated at 37°C overnight.

2.2.10 Cell lysis

Two different cell lysis methods were applied: sonication (Branson Sonifier 250) and chemical lysis with B-PER reagent (43). Cell lysis by sonication was carried out according to the following protocol: harvested cell pellets were resuspended in 30 ml pre-chilled lysis buffer (50 mM NaP_i, 300 mM NaCl, 10 or 15 mM imidazole, pH 8) and lysed twice on ice for 5 min and 80% intensity (duty cycle). Insoluble substances were removed by centrifugation at 4°C and 3,220 x g for 30 min using Eppendorf 5810R centrifuge. The supernatant was collected and stored on ice.

Cell lysis with B-PER reagent was performed according to the recommended protocol from Thermo Scientific (43): 5 ml reagent (B-PER Bacterial Protein Extraction Reagent, in phosphate buffer, Piercenet) was added per 40 ml culture. After resuspension of the cell pellet and incubation at room temperature for 10 min, insoluble substances were removed by centrifugation for 15 min at 16,100 x g and 4°C (Eppendorf Centrifuge 5415R). The supernatant was separated from the pellet and stored on ice.

For the purification of inclusion bodies, the insoluble fraction of cell lysate was resuspended in 5 ml of B-PER reagent. After addition of lysozyme to a final concentration 200 µg/ml, and incubation for 5 min at room temperature, 15 ml 1:10 diluted B-PER reagent were added and insoluble substances were removed by centrifugation (16,100 x g, 15 min, 4°C). The supernatant was separated and the pellet was resuspended again in 1:10 diluted B-PER, incubated and centrifuged.

2.2.11 Gravity flow column chromatography

2.2.11.1 Affinity chromatography

Isolation of CelA variants using His-Tag purification was carried out with a column, containing 1 ml Ni-NTA sepharose (GE Healthcare 17-5318). The column was prepared by washing it with water and equilibration with lysis buffer. The supernatant of B-PER cell lysate was mixed with lysis buffer (ratio 1:3, (50 mM NaP_i, 300 mM NaCl, 10 mM imidazole, pH 8)), loaded onto the column and incubated for 10 min. Unspecific bound proteins were removed by four washing steps, each using 2 ml wash buffer (50 mM NaP_i, 300 mM NaCl, 20 mM imidazole, pH 8). The supernatant of sonication disruption was loaded ed onto the column without incubation, the column was washed with 30 ml washing buffer.

Elution was the same for both cell lysates: it was carried out with 10 column volumes in total. Each of the 10 elution steps was incubated with 1 ml elution buffer (50 mM NaP_i, 300 mM NaCl, 500 mM imidazole, pH 8) for 20 min.

2.2.11.2 Size exclusion chromatography

Eluates were rebuffered using a PD10 column (GE Healthcare, 17-0851-01) according to the gravity protocol (44). The column was prepared by washing it with water and equilibration with 30 ml borate buffer (50 mM citrate, 50 mM NaP_i, 50 mM borate, pH 7) or NaP_i (50 mM NaP_i pH 7). The eluate fractions 1-5 from the His tag purification step were mixed and 2.5 ml were loaded onto the column. The protein fraction was eluted with 3.5 ml buffer. After equilibrating the column again, the remaining 2.5 ml eluate fraction was loaded and eluted.

The purified and rebuffered protein was concentrated using a Vivaspin concentrator 10.000 MWCO PES (Sartorius stedim Biotech Vivaspin 4, VS0403) by centrifugation at 4°C and 3,220 x g (Eppendorf Centrifuge 5810R) until an appropriate protein concentration was reached.

The protein variants were stored at 4°C, at protein concentrations ranging from 1-3 mg/ml.

2.2.12 Measurement of protein concentration

The protein concentration was measured at a wavelength of 280 nm using NanoDrop 2000c (Thermo Scientific) and calculated by considering the extinction coefficient (see Table 4) (45).

Table 4: Properties of CelA variants and Cex_cHis. The values were calculated using Expasy Protparam Tool.

	Number of Amino acids	**Molecular weight [g/mol]**	**Extinction coefficient [M^{-1} cm^{-1}]**	**Abs 0.1% (=1 g/l)**
CelA_cHis	389	40,999.7	79,535	1.940
CelA_nHis	396	41,643.4	79,535	1.910
CelAc_cHis	314	34,654.4	79,535	2.295
CelAc_nHis	321	35,298.1	79,535	2.253
Cex_cHis	489	52,728.8	80,900	1.534

2.2.13 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

Commercial SDS PAGE were developed in 1x MOPS buffer (50 mM MOPS, 50 mM Tris Base, 0.1% SDS, 1 mM EDTA, pH 7.7). Tris-Glycine gels were produced according to the composition in Table 5, and stored at 4°C in 1x running buffer (3 g/l Tris, 14 g/l Glycine, 1 g/l SDS, pH 8.6).

Table 5: Components of SDS PAGE: Volumes of the reagents used for the production of 5 Tris- Glyince gels.

	Stacking gel (4% acrylamide)	**Separation gel (12% acrylamide)**
Stacking gel buffer stock (0.5 M Tris, 0.4% SDS) pH 6.8	2.5 ml	
Separation gel buffer stock (1,5 M Tris, 0.4% SDS) pH 8.8		4 ml
30% Acrylamide (BioRad Laboratories Inc)	1.5 ml	12.8 ml
10% ammonium persulfate (APS)	20 µl	50 µl
Tetramethylethylene diamine (TEMED)	15 µl	10 µl
$_{dd}$H2O	6 ml	5.6 ml
2% bromphenol blue	10 µl	

E. coli cell lysates were separated in soluble and insoluble protein fraction and 0.02 to 0.05 OD_{600} units were loaded onto the gel. Volumes of the samples from protein purification depended on the protein concentration. The samples were mixed with 1.2 µl DTT 500 mM, 3 µl NuPAGE sample buffer and bidistilled water was added to give a final volume of 12 µl. Prior loading onto the gel, the samples were heated to 95°C for 10 min. Gels were developed in the corresponding buffer with 200 V for 10-15 min, then voltage was decreased to 100 or 110 V until the proteins were separated properly. The gels were stained with Coomassie Brilliant Blue solution (100 ml ethanol, 20 ml acetic acid (100%), 80 ml ddH_2O, 0.1 g Brilliant Blue G250), and destained with 10% acetic acid.

2.2.14 Congo red plate assay

A Congo red plate assay was performed to monitor the endocellulase activity while purification. Therefore agar plates containing 0.5% CMC were poured. 5-10 µl of cell lysate or protein purification fraction were pipetted onto the plate, incubated at 37°C for 10 to 20 min and washed off with distilled water. For detection of the enzyme activity the agar plate was stained with 0.2% Congo red solution for 20 min and destained twice with 0.5 M NaCl.

2.2.15 4-Methylumbelliferyl β-D-cellobioside plate assay (MUC plate assay)

For the determination of exocellulase activity during protein purification, agar plates containing 5 mg/L 4-Methylbelliferyl β-D-cellobioside (MUC, Carbosynth, EM08000) were poured. 2-5 µl of cell lysate or protein purification fraction were pipetted onto a plate, incubated at 37°C for 10 to 20 min and the fluorescence was detected at 302 nm.

2.2.16 Reducing sugar assay

2.2.16.1 Enzyme reaction

The enzyme reaction was optimized in the Project Lab Biotechnology (Characterization of *Sorangium cellulosum* CelA, Monika Friedl). In case of a soluble substrate, 600 µl substrate solution were pre-incubated at respective incubation temperature, and the enzyme reaction was started by addition of 2 µl enzyme solution. The enzyme reaction was shaken with 500 rpm in a Thermomixer Eppendorf comfort. All samples were prepared in triplicate.

The degradation of insoluble substrates (filter paper, Avicel) yields to soluble and insoluble reducing sugars. Therefore each enzyme reaction needs to be prepared twice: 16-18 mg punched filter paper (Whatman No.1 filter paper, Machery Nagel MN69, 439009, diameter 5 mm) are weight into an 1.5 ml Eppendorf tube, the exact mass was noted. 240 µl NaP_i buffer (50 mM, pH 7) were added and pre-heated for at least 10 min. For the degradation of Avicel, a 1% solution in NaP_i 50 mM pH 7 was prepared and 240 µl were preheated for each reaction. The enzyme reaction, incubated at 37°C and shaking with 1200 rpm, was started by addition of certain amount of enzyme solution.

2.2.16.2 Detection reaction

For the detection of reducing sugars, produced in enzyme reaction, 2 assays were applied: p-Hydroxybenzoic acid hydrazide assay (pHBAH-assay) was implemented for determination after incubation with soluble substrate (CMC). A stock solution was produced containing 6.35 g pHBAH (Sigma H9882) and 5 ml HCl conc. per 100 ml. The working solution consists of 20%Vol. stock solution and 12.5%Vol. 4 M NaOH dissolved in ddH_2O. The stock solution is durable for 1 month while working solution needs to be prepared freshly for every measurement. For the detection of reducing sugars after enzyme reaction, 150 µl pre-chilled working solution were mixed with 50 µl enzyme reaction or glucose standard in a microtiter plates (MTP) suitable for a thermocycler (Applied biosystems, GeneAmp PCR system 2700). After heating to 95°C for 5 min and cooling to 4°C, absorbance was measured at 430 nm in the plate reader FLUOstar Omega (150 µl per well). Calibration was done using glucose concentrations from 0.03 to 1.0 mg/ml dissolved in the respective buffer.

For the determination of reducing sugar released from insoluble substrates, Dinitrosalicylic acid assay (DNS assay) was carried out (39). The reagent contained 1% DNS (Sigma, D0550), 1% NaOH, 0.2% Phenol, 0.05% Na-Sulfit and 20% K-Na-Tartrate dissolved in dH_2O. For calibration 120 µl reagent were mixed with 60 µl glucose standard (0.1-0.8 mg/ml), heated in a thermocycler to 95°C for 5 min and cooled to 4°C. 150 µl of the assay samples were transferred into a microtiter plate and the absorbance was measured at 540 nm in a FLUOstar Omega plate reader. After the enzyme reaction, the soluble reducing sugar was determined by mixing 120 µl supernatant (16,100 x g, 1 min) with 240 µl DNS reagent, for the total reducing sugars 480 µl DNS reagent were added to the enzyme reaction (240 µl). The detection reaction was carried out in the Eppendorf comfort thermomixer, heating to 95°C for 10 min and after cooling on ice, 150 µl of the supernatant were pipetted into a microtiter plate and the absorbance was measured by the FLUOstar Omega plate reader at 540 nm.

2.3 Optimization of expression and purification

2.3.1 Plasmid construction

The plasmids pMS470_*Soce_celAocΔSnHis*, pMS470_*Soce_celAoΔSnHis,* pMS470_*Soce_celAocΔScHis*, pMS470_*Soce_celAoΔScHis* and pMS470_*Cf_cexcHis* were constructed in prior work.

In order to optimize the expression of Cex, a variant without signal sequence was cloned using the primers according Table 6. The native *cex* gene was used as template for the PCR amplification. Furthermore codon optimized *cex2.0* was used as a template for the amplification of *cex2.0* and *cex2.0ΔS*. Primers for cloning were ordered with N-terminal and C-terminal 6x His Tag (Table 6).

Table 6: Primers for the amplification of *cex* and *cex2.0* variants. The PCR product was purified by preparative agarose gel electrophoresis, the DNA product with the correct size was cut out of the gel and purified according to the Promega Wizard SV Gel and PCR Clean Up System. The number of the primer numbers is given in Table 3 on page 16. *Δstop* indicates expression without the stop codon, *ΔS* indicates the deletion of the signal sequence.

Gene	Forward primer	Reverse primer	Product size
cex2.0Δstop	175	174	1458 bp
cex2.0ΔSΔstop	173	174	1332 bp
cex2.0ΔS	173	176	1335bp
cexΔSΔstop	168	130	1332bp

The vector pMS470nHis and pMS470cHis were constructed within the CD-laboratory of Genetically Engineered Lactic Acid Bacteria on the Institute of Molecular Biotechnology of the technical University of Graz (IMBT). The purified PCR products *cexΔSΔstop, cex2.0Δstop, cex2.0ΔSΔstop*, and *cex2.0ΔS* as well as the vectors pMS470*nHis* and pMS470*cHis* were digested using 20 U of *Nde*I and 10 U of *Hind*III. The vectors were loaded onto a preparative agarose gel while PCR products were purified according to the Promega Wizard SV Gel and PCR Clean Up System. The purified restriction digest products *cexΔSΔstop, cex2.0Δstop* and *cex2.0ΔSΔstop* were ligated with pMS470cHis, and *cex2.0ΔS* was ligated with pMS470nHis. After transformation in *E. coli* XL1, 4 clones were streaked out for the isolation of plasmid DNA and the verification by restriction analyses. Prior to transformation and expression in *E. coli* BL21 and *E. coli* BL21 (DE3), the recombinant plasmid was confirmed by sequencing (Primer sequences are shown in Table 3 on page 16).

2.3.2 Fermentation

Overnight cultures (5 ml LB-Amp each) were inoculated with the expression clones *E. coli* BL21 or *E. coli* BL21 (DE3) (pMS470_*Cf_cexcHis)*, (pMS470_*Cf_cexΔScHis)*, (pMS470_*Cf_cexnHis)*, (pMS470_*Cf_cex2.0cHis)*, (pMS470_*Cf_cex2.0ΔScHis)*, and (pMS470_*Cf_cex2.0ΔSnHis)*, incubated at 37°C, 80% humidity and shaken with 120 rpm. Main cultures (100 ml or 200 ml LB-Amp in 250 ml and 500 ml flasks) were inoculated to an optical density at 600 nm (OD_{600}) of 0.1 (measurement with Eppendorf BioPhotometer) and incubated at 25 or 37°C, 120 rpm and 80% humidity until an OD_{600} of 0.6 to 0.8 was reached. The expression of the protein variants was induced by addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 100 µM. The main cultures were incubated at 25 or 28°C, 120 rpm and 80% humidity overnight. Additionally main cultures were incubated overnight without induction or induction for 4 h. The cells were harvested by centrifugation of 50 ml culture at 4°C and 3,220 x g for 30 min (Eppendorf Centrifuge 5810R). Cell pellets were stored at -20°C until further use.

Since the overexpression of Cex_cHis resulted in the formation of inclusion bodies, they were dissolved according to the B-PER protocol (see chapter 2.2.10 cell lysis, page 23).

The expression of *celAocΔSnHis, celAoΔSnHis, celAocΔScHis* and *celAoΔScHis* was carried out as described in final purification method (see Appendix 6.1 Fermentation page 70).

2.3.3 Affinity chromatography

CelA variants and Cex2.0_cHis were purified by an automated affinity chromatography, using the ÄKTA purifier system. The protein purification was prepared according to the following protocol: First the column (GE Healthcare HisTrap FF, 17-5355-01) was washed with ethanol (20%) and distilled water and was equilibrated with lysis buffer (50 mM NaP_i, 300 mM NaCl, 10 mM imidazole, pH 8) (44). Then the cell lysates were loaded onto the 5 ml column by pumping with 2 ml/min. All further steps were performed with a flow of 5 ml/min. The unspecific bound proteins were removed by washing the column with 50 mM NaP_i, 300 mM NaCl, 20 mM imidazole, pH 8. For the elution of the proteins, the concentration of imidazole was increased to 500 mM within a slope of 15 CV. To monitor the concentration, the purity and the activity of the proteins, SDS PAGEs and plate activity assays were performed. Using the results of those assays and SDS PAGE analysis in combination with UV absorbance of ÄKTA purifier system, several elution fractions were pooled. For the optimization of the protein purity, a purification of CelA_cHis with 500 mM NaCl in the lysis and elution buffer was performed. Since no improvement was observed in this case, further purifications were performed using the standard salt concentration of 300 mM NaCl in both buffers. The imidazole concentration in lysis buffer was increased to 15 mM to reduce the unspecific bound proteins while loading of the cell lysate.

2.3.4 Desalting

After the His Tag purification using the ÄKTA purifier system, the pooled elution fractions were rebuffered using Vivaspin 20 concentrator. The pooled eluate fractions were pipetted onto the column and centrifuged with 3,220 x g at 4°C to reduce the volume. To get rid of the imidazole, the proteins were washed three times using Sodium phosphate buffer (NaP_i, 50 mM, pH 7) for CelA variants and 0.9% NaCl for Cex_cHis.

Since the proteins precipitated and blocked the membrane, pooled elution fractions were rebuffered by dialysis. Zellutrans membrane (E667, Roth dialysis membranes T3, CARL ROTH GMBH + CO. KG, Karlsruhe) was used for desalting as well as buffer exchange to 50 mM NaP_i pH 7. Therefore the protein variants were rebuffered in 100 x buffer stirring with 180 rpm at 4°C for 1 h. To ensure completed desalting, the buffer was changed 3 times. To verify the protein binding capacity to dialysis column, the protein concentration was determined by SDS PAGE before and after rebuffering.

After rebuffering the proteins were concentrated by Vivaspin 20 or Vivaspin concentrator 4 (10,000 MWCO PES-membrane, Sartorius stedim biotechVS2001) and centrifugation at 4°C and 3,220 x g (Eppendorf Centrifuge 5810R).

2.4 Characterization

2.4.1 Optimal pH and temperature

The optimal pH and temperature was determined using CelA_nHis and CelAc_nHis variants after purification via gravity flow column chromatography. The enzyme reaction was incubated in 1% CMC dissolved in borate buffer (50 mM Citrate, 50 mM NaP_i, 50 mM Borate, pH 7) using 0.5 µg/ml CelAc_nHis and 1.6 µg/ml CelA_nHis. Borate buffer was set to pH 3 to 10 for the determination of the optimal pH, and the optimal temperature was determined by incubation of the enzyme reaction from 10°C to 80°C The enzyme reaction for optimal pH was incubated at 37°C, while the optimal temperature was determined at pH 7, both reactions were incubated for 10 min. The optimal conditions of CelA_nHis were determined during the project lab.

2.4.2 Substrate concentration

In order to find the optimal substrate concentration, the enzyme reaction was incubated using 0.25 to 2.0% CMC dissolved in borate buffer pH 7. The enzyme reaction was incubated with CelAc_nHis (0.5 µg/ml, purified by column affinity chromatography) for 6-30 min.

2.4.3 Buffer conditions

To find the optimal buffer for the enzyme variant CelAc_nHis, the release of reducing sugar was determined in borate buffer, sodium citrate buffer and in sodium phosphate buffer, each 50 mM, pH 7. The enzyme was incubated for 10 min at 37°C in 1% CMC dissolved in the respective buffer. Additionally, the enzyme activity of CelAc_cHis in 1.75% CMC dissolved in 50 mM and 100 mM NaP_i pH 7 was determined. The enzyme concentration was 0.5-0.8 µg/ml.

2.4.4 Specific enzyme activity towards CMC

The specific enzyme activity of all CelA variants was determined using optimal conditions. The enzymes were purified either by ÄKTA or gravity flow column chromatography, and rebuffered in 50 mM NaP_i pH 7. The enzymes were incubated in 1% and 1.75% CMC dissolved in NaP_i, at 37°C and 50°C for several minutes. The specific enzyme activity was calculated according to the formula in Figure 9 on page 15. One unit is defined as the enzyme amount, necessary to release one µmol of reducing sugar, expressed as glucose, in 1 min.

2.4.5 Enzyme stability

The enzyme stability was determined using CelA_cHis and CelAc_cHis. After pre-incubation, the remaining enzyme activity was determined at the following conditions: 1% CMC, dissolved in NaP_i50mM, pH 7 was pre-heated to 37°C, the enzyme reaction was started by addition of CelA_cHis or CelAc_cHis and was incubated for 10 min. Those conditions are referred to standard conditions.

2.4.5.1 Thermal stability

For the determination of thermal stability, the enzymes CelA_cHis 0.2 mg/ml and CelAc_cHis 0.3 mg/ml, dissolved in NaP_i pH 7, were pre-incubated at 10°C to 80°C. The remaining enzyme activity was determined after 30 min to 250 h using standard conditions. The reference value for the calculation of the relative activity is the amount of reducing sugar, detected after incubation of the fresh enzymes.

2.4.5.2 pH stability at 4°C

The pH stability was determined after pre- incubation of the enzymes in borate buffer pH 4-9, at 4°C for 30 min to 220 h. The enzyme concentration during the incubation was 0.2 mg/ml for CelAc_cHis and 0.2-0.3 mg/ml for CelA_cHis. The reference value for the calculation of the relative activity is the amount of reducing sugar, detected after incubation of the fresh enzymes.

2.4.5.3 pH stability at 40°C

The pH stability was determined by pre-incubation of the enzymes, dissolved in borate buffer pH 4-9, at 40°C, for 1 and 2 h. The enzyme concentration during the incubation was 0.2-0.3 mg/ml for CelAc_cHis and 0.3-0.4 mg/ml for CelA_cHis. The reference value for the calculation of the relative activity is the amount of reducing sugar, detected after incubation of the fresh enzymes.

2.4.5.4 Influence of metals ions

The enzymes were pre-incubated with K_2SO_4, $MgSO_4$, $CuSO_4$, $(NH_4)_2SO_4$, $MnSO_4$, $FeSO_4$, $ZnSO_4$, $NiSO_4$, $CaCl_2$ and NaCl, 1 mM and 0.1 mM each for 30 min on ice. The remaining enzyme activity was determined with a substrate solution containing 1 mM or 0.1 mM of the respective metal salt. As a reference value and for the calculation of the relative activity, water was added to the enzymes instead of the metal salt solution to achieve the same final concentration. To verify if metal ions have no influence on the absorbance at 430 nm, the substrate solution containing the respective metal salt concentration was measured as blank for the calculations.

2.4.5.5 Influence of reagents

The enzymes were pre-incubated with EDTA 0.1 mM, DTT 1 mM, DMSO 1 mM, Urea 1 M, Glycerin 1 mM, Ethanol 10% and SDS 0.1%, for 30 min on ice. The remaining enzyme activity was determined in a substrate solution containing the same concentration of the respective reagent. Additionally the enzymes were pre-incubated with SDS 1%, SDS 0.1% and Urea 1 M, and the enzyme reaction was carried out without any addition of the reagent to the substrate solution. The CMC blank was measured containing the same amount of reagent as in the enzyme reaction. The enzyme, incubated with dH_2O instead of any reagent (in the enzyme and substrate solution) was used as reference value for the calculations.

2.4.6 Enzyme inhibition by Cellobiose

Cellobiose was added to CelA_cHis and CelAc_cHis in various concentrations: 20 mM, 2 mM and 0.2 mM to verify whether the disaccharide inhibits the enzyme activity. The enzymes (16 pmol/ml CelA_cHis or 19 pmol/ml CelAc_cHis) were pre-incubated at room temperature for 45 min prior to the determination of the remaining enzyme activity at standard conditions.

2.4.7 Degradation of insoluble substrates

Whatman No.1 filter paper (7% in NaP_i) was degraded using all CelA variants. The enzyme concentration was 5, 8, 12 and 24 µg/ml. The enzyme assays were incubated for 2 h (single determination) or 6 h (measurements in triplicate). The degradation of Avicel (1% in NaP_i) was incubated for 2-6 h using 5-414 µg/ml CelA_cHis and CelAc_cHis.

2.4.8 Evaluation of the synergism of CelA and Cex

2.4.8.1 Degradation of CMC

Synergistic degradations of CMC were carried out using CelAc_nHis (0.3 µg/ml) and several concentrations of Cex_cHis (2-4 µg/ml). The enzyme reactions (single determinations) were incubated at 37°C for 2 to 30 min and 3 h.

2.4.8.2 Degradation of insoluble substrates

The degradation of insoluble substrates was determined using Macherey Nagel MN 619 de filter paper and Avicel. In the case of filter paper, the enzyme reaction (single determinations) was incubated using 4-16 µg/ml Cex_cHis and 0.3-1.2 µg/ml CelAc_nHis for 20 h. The enzyme assay with Avicel as insoluble substrate was incubated for 2 h using 5-12 µg/ml CelA_cHis and CelAc_cHis and 16 µg/ml Cex_cHis.

3 Results

3.1 Optimization of expression and purification

3.1.1 Plasmid construction

The native and the codon optimized *cex* genes were amplified through PCR. Variants with and without signal sequence and stop codon as well as the restriction recognition sites for *Nde*I and *Hind*III were constructed. The PCR products were isolated from a preparative agarose gel and purified. The expression vectors pMS470*nHis* and pMS470*cHis* as well as the PCR products were digested with *Nde*I and *Hind*III. The linearized vectors were separated in a preparative gel and purified according to the Promega Wizard SV Gel and PCR Clean Up System. The *Nde*I and *Hind*III digested gene variants were also purified using the Promega Wizard SV Gel and PCR Clean Up System kit.

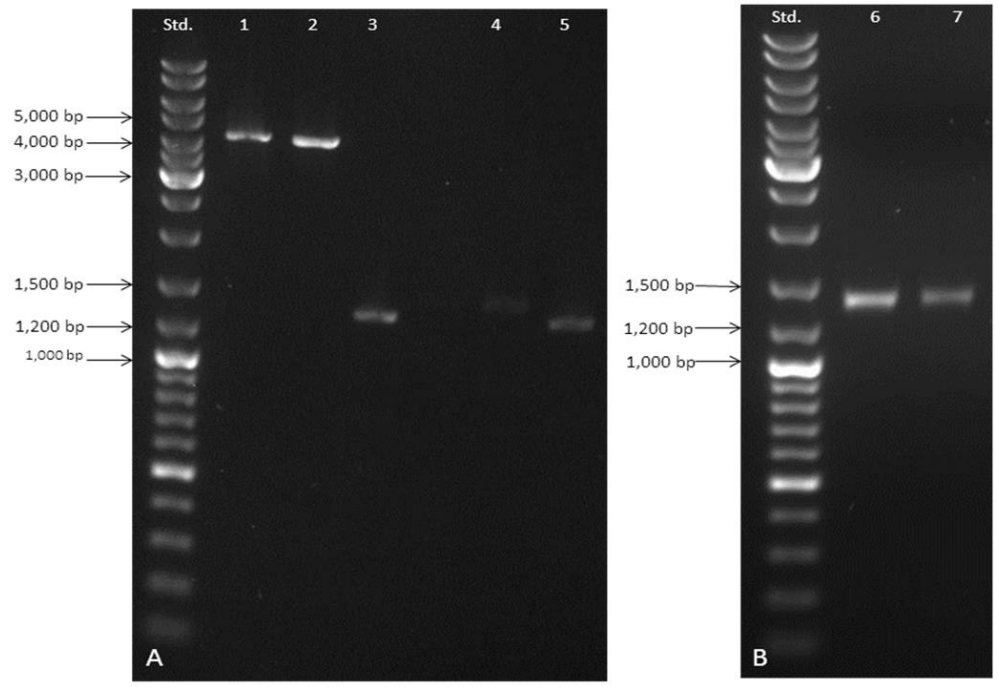

Figure 14: Agarose gel of purified *Nde*I / *Hind*III digested fragments: A: codon optimized *cex2.0* variants, B: native gene sequence; DNA concentration was measured at 280 nm using NanoDrop: 1: pMS470nHis (4040 bp, 42.9 ng/µl), 2: pMS470cHis (4023 bp, 60 ng/µl), 3: *cex2.0ΔS* (1335 bp, 25.2 ng/µl) 4: *cex2.0ΔStop* (1458 bp, 12.2 ng/µl), 5: *cex2.0ΔSΔStop* (1332 bp, 19 ng/µl), 6: *cexΔSΔStop* (1332 bp, 32.7 ng/µl), 7: *cexΔSΔStop* (1332 bp, 17.8 ng/µl).

After confirming the correct size of the DNA fragments by agarose control gel electorphoresis and measuring of the DNA concentration with the NanoDrop 2000c photometer (Figure 14), ligation reactions were prepared. The correct insertion of the genes into the vector backbone was approved twice. First, the isolated plasmids were digested by restriction enzymes and additionally the genes were sequenced at LGC genomics.

3.1.2 Fermentation

The expression of *cexcHis* was optimized using different fermentation conditions. The main cultures were incubated at 28 and or 25°C overnight. As depicted in Figure 15A, only a small amount of the protein was expressed soluble, while large quantities formed inclusion bodies. The decrease of the cultivation temperature to 25°C yielded in no improvement of soluble protein expression (Figure 15B).

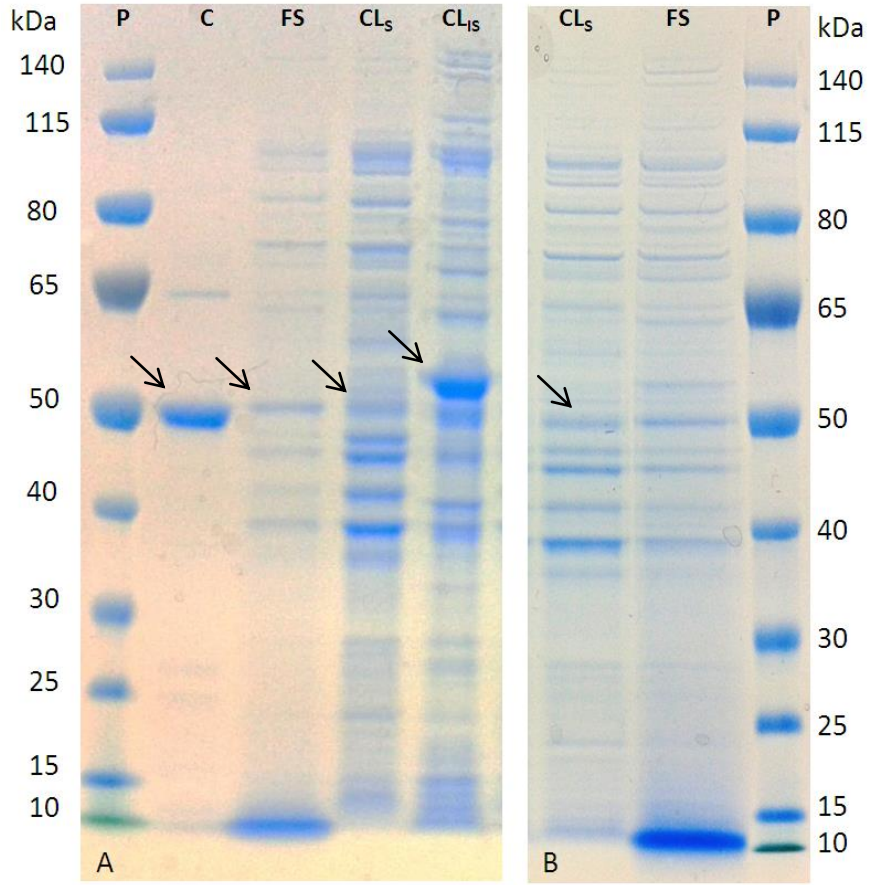

Figure 15: SDS PAGE: Fermentation of *E. coli* BL21 (pMS470_*Cf_cexcHis)*; Incubation at 28°C (A) and 25°C (B) overnight after induction with IPTG. P: PageRuler Prestained Protein Ladder (Fermentas), C: positive control, purified Cex_cHis (1.2 µg, 52.7 kDa), FS: fermentation supernatant, CL_S: cell lysate soluble fraction (0.05 OD units), CL_{IS}: cell lysate insoluble fraction (0.05 OD units),

To find out whether the induction of the strong tac promoter stresses the *E. coli* cells, fermentations with and without induction were performed. The cell lysates of *E. coli* BL21 (pMS470_*cexcHis)* and (pMS470*cHis*) are shown in Figure 16. Higher amount of inclusion bodies were formed in induced cultures, while the total amount of soluble enzyme seems to be nearly equal.

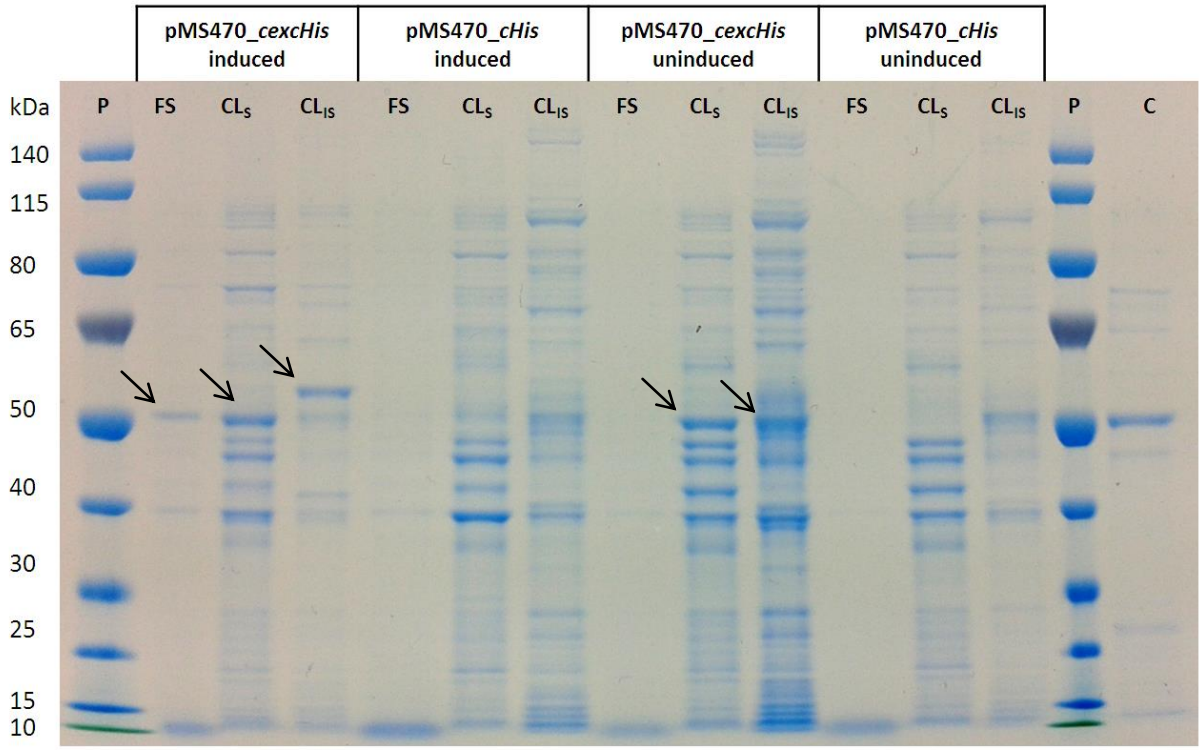

Figure 16: SDS PAGE: Fermentation of *E. coli* BL21 (pMS470_*Cf_cexcHis)* and (pMS470cHis): Incubation at 25°C, with and without IPTG induction. P: PageRuler Prestained Protein Ladder (Fermentas), FS: fermentation supernatant, CL_S: cell lysate soluble fraction (0.02 OD units), CL_{IS}: cell lysate insoluble fraction (0.02 OD units), C: positive control: purified Cex_cHis (0.55 µg, 52.7 kDa).

Thereafter an *E. coli* BL21 (pMS470_*Cf_cexcHis)* culture was induced with IPTG to 100 mM and a sample was taken after 4 h of induction. The remaining culture was incubated overnight. Cex_cHis is already expressed after 4 h, but the amount increased overnight. Cex_cHis was also detected in the fermentation supernatant and the insoluble fraction of the cell lysate after incubation overnight.

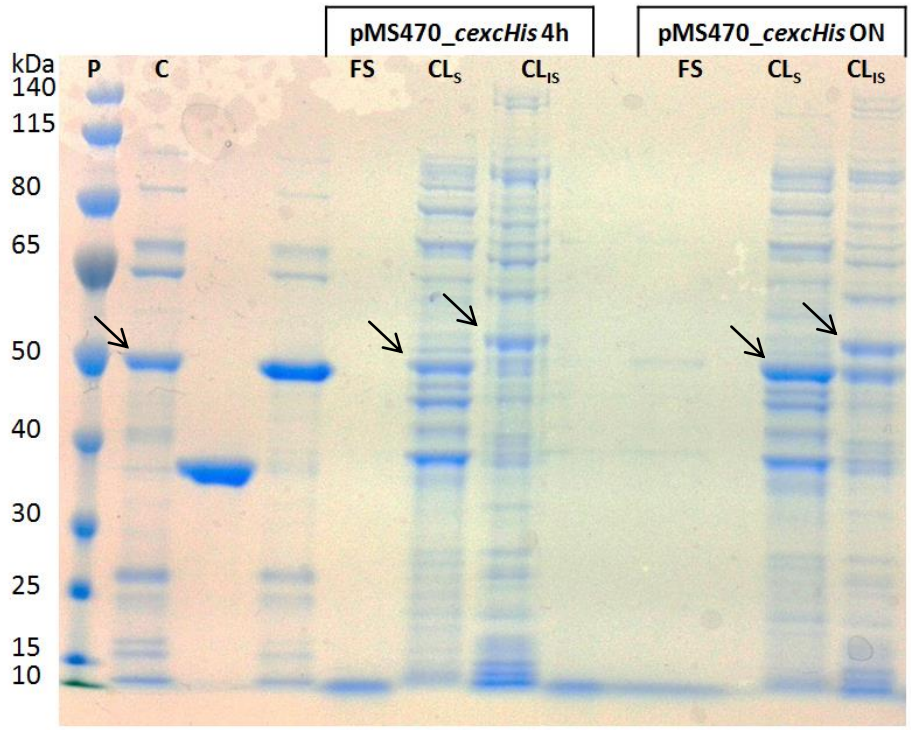

Figure 17: SDS PAGE: Fermentation of *E. coli* BL21 (pMS470_*Cf_cexcHis)*, Incubation at 25°C induction for 4 h and ON. P: PageRuler Prestained Protein Ladder (Fermentas), C: positive control: Cex_cHis purified (3.4 µg/ml, 52.7 kDa), FS: fermentation supernatant, CL_S: cell lysate soluble fraction (0.05 OD units), CL_{IS}: cell lysate insoluble fraction (0.05 OD units)

It was tried to resolve the inclusion bodies according to the B-PER Bacterial Protein Extraction Reagent protocol. As can be seen in Figure 18, Cex_cHis and Cex_nHis were not detected in the soluble fraction of the cell lysate. The inclusion bodies were not solubilized after treatment with lysozyme.

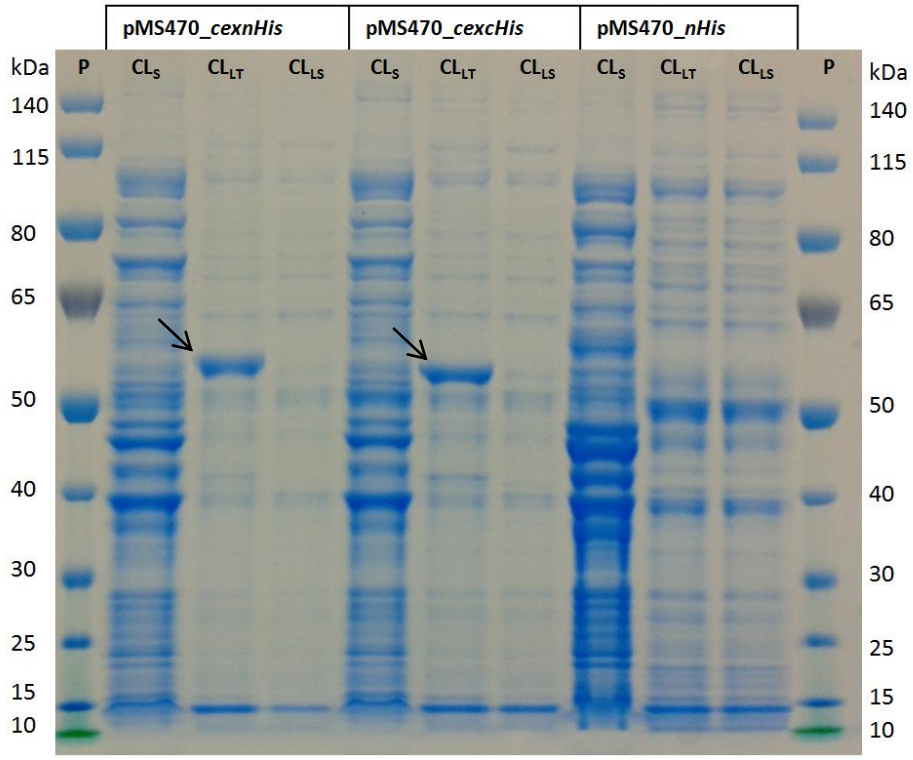

Figure 18: SDS PAGE: Fermentation of *E. coli* BL21 (DE3) (pMS470_*Cf_cexcHis)* and (pMS470_*Cf_cexnHis)*, Incubation at 28°C. P: PageRuler Prestained Protein Ladder (Fermentas), CL_S: cell lysate soluble fraction, CL_{LT}: cell lysate lysozyme total protein fraction, CL_{LS}: cell lysate lysozyme soluble fraction.

For a more effective soluble expression, a variant without signal sequence, *cexΔScHis*, was cloned and expressed mainly intracellular in *E. coli* BL21 cells. No increased expression could be achieved.

Therefore, the *cex* gene was codon optimized and variants were cloned and expressed in *E. coli*.

Three variants of the codon optimized *cex* gene were expressed in *E. coli* BL21 cells: *cex2.0_cHis, cex2.0ΔS_cHis* and *cex2.0ΔS_nHis*. An overview of the variants is shown in Figure 11 on page 18. As can be seen in Figure 19, only *cex2.0_cHis* was mainly soluble expressed. The activity of the enzymes in the cell lysate fractions and the fermentation supernatant were determined using the MUC plate assay. Since only the soluble cell lysate fraction of *cex2.0_cHis* showed activity towards MUC, all further works were continued using this variant.

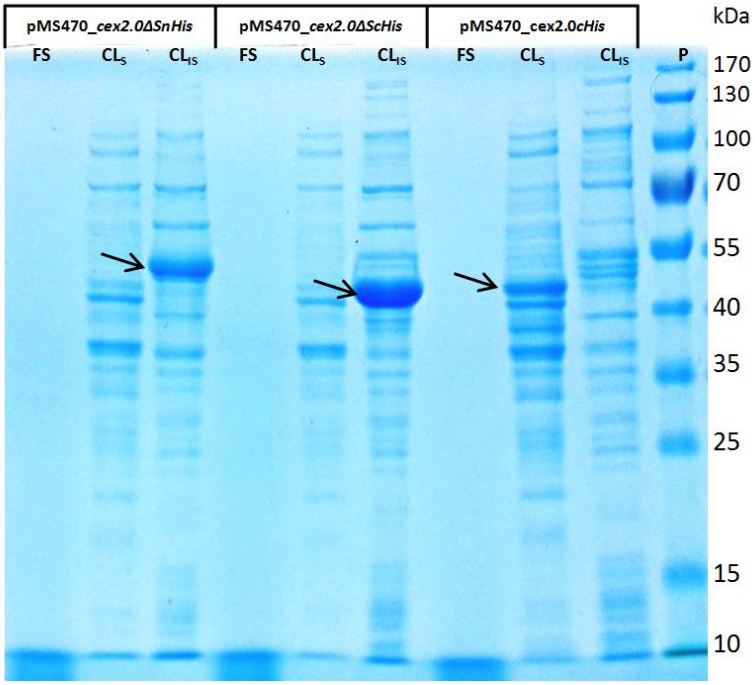

Figure 19: SDS PAGE: Fermentation of *E. coli* BL21 (pMS470_*Cf_cex2.0ΔSnHis), E. coli* BL21 (pMS470_*Cf_cex2.0ΔScHis)* and *E. coli* BL21 (pMS470_*Cf_cex2.0cHis)*: FS: fermentation supernatant (0.05 OD units), CL_S: cell lysate soluble fraction (0.05 OD units), CL_{IS}: cell lysate insoluble fraction (0.05 OD units), P: PageRuler Prestained Protein Ladder (Fermentas).

To compare the expression of all enzymes, a SDS PAGE of the soluble cell lysate fractions was performed after fermentation according to the optimized protocol. The expression of CelAc_cHis and CelAc_nHis was higher than the expression of CelA_cHis, CelA_nHis and Cex_cHis (see Figure 20). Both CelA variants (N- or C- terminal His Tag) with the N-terminal repeat sequence seem to be bigger on the SDS PAGE. This could be due to interaction with the SDS reagent leading to deficient denaturation and a changed resolution on the SDS PAGE.

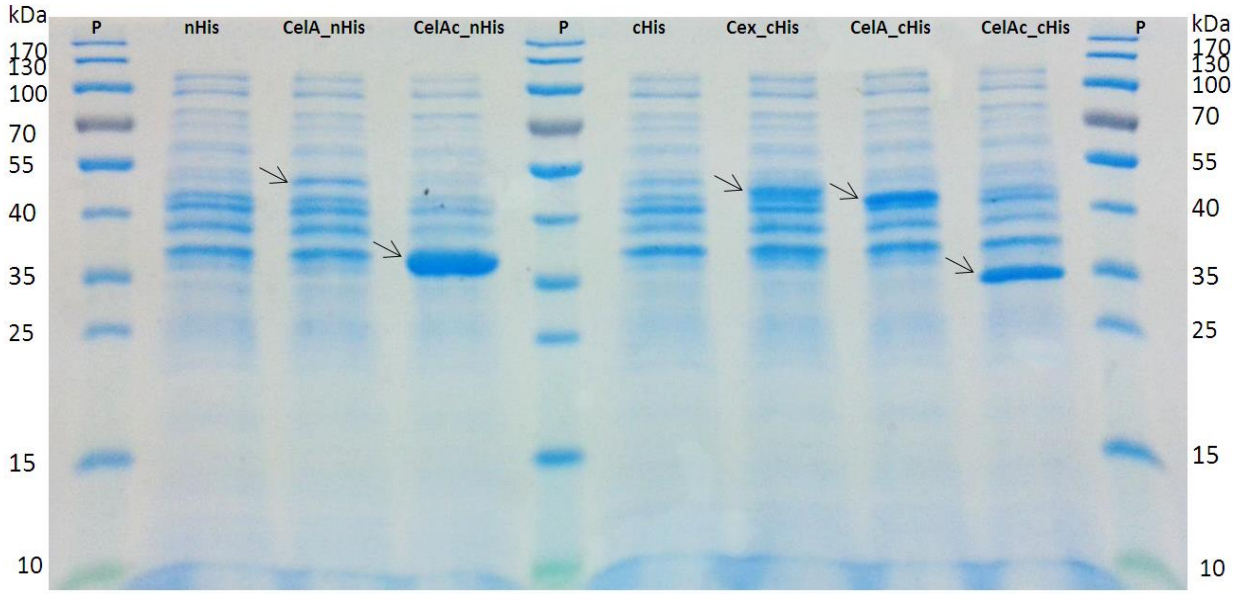

Figure 20: SDS PAGE: *E. coli* BL21 cell lysate soluble fraction (0.05 OD units): P: PageRuler Prestained Protein Ladder (Fermentas), nHis: pMS470*nHis*, CelA_nHis (41.6 kDa): pMS470_*Soce_celAoΔSnHis*, CelAc_nHis (35.3 kDa): pMS470_*Soce_celAocΔSnHis*, cHis: pMS470*cHis*, Cex_cHis (52.7 kDa): pMS470_*Cf_cex2.0cHis*, CelA_cHis (41.0 kDa): pMS470_*Soce_celAoΔScHis*, CelAc_cHis (34.7 kDa): pMS470_*Soce_celAocΔScHis*.

3.1.3 Affinity chromatography

The optimization of the Cex_cHis purification using the ÄKTA purifier and His Tag affinity chromatography is shown in Figure 21. The concentration of the enzyme increases from elution fraction B10 to a maximum at fraction C9. Elution fractions from C12 to D10 are nearly pure. Using MUC activity plate assay, exocellulase activity was detected in the cell lysate and the manual wash fraction while no activity was measureable in the automatic ÄKTA wash fraction. Activity was detected in elution fraction B1 to F3, whereas the highest activity was measured in eluate C1 to D10.

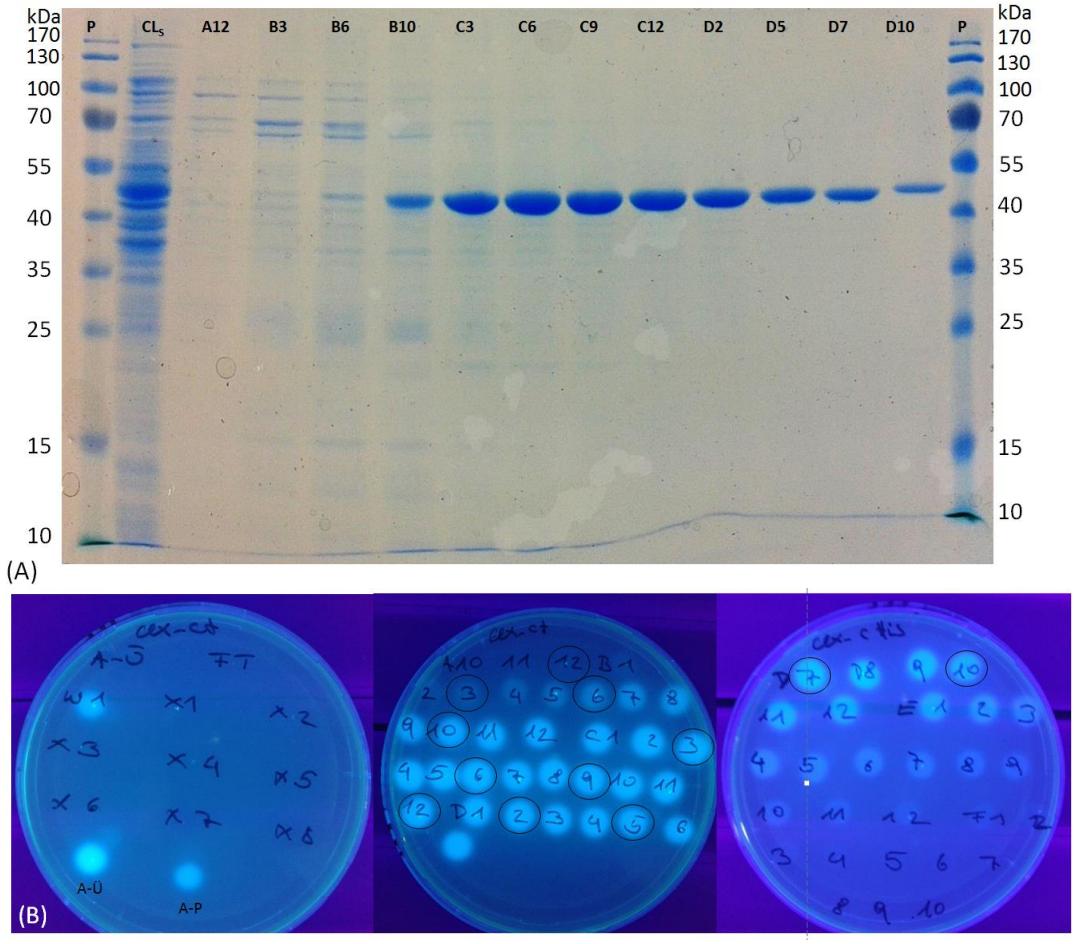

Figure 21: ÄKTA purification of Cex_cHis. (A) SDS PAGE P: PageRuler Prestained Protein Ladder (Fermentas), CL_S: *E. coli* BL21 (pMS470_*Cf_cexcHis)* cell lysate soluble fraction (0.05 OD units), A12-D10: eluate fractions of His Tag purification. (B) MUC plate assay: A-Ü: *E. coli* BL21 (pMS470_*Cf_cexcHis)* cell lysate soluble fraction, A-P: cell lysate insoluble fraction, FT: flow through of column loading, W1: manual wash fraction, X1-8: wash fractions, A10-F10: Elution fractions.

Similar results were obtained in purification of CelA_cHis (Figure 22). The enzyme was detected in fraction A11 – D7, while the elution fractions to B2 contain higher amounts of impurities. CelA_cHis activity was detected in nearly every fraction. Considering the size of the halos, the enzyme activity and enzyme amount seemed to be lower in the wash fractions compared to the elution fractions A8 to D6.

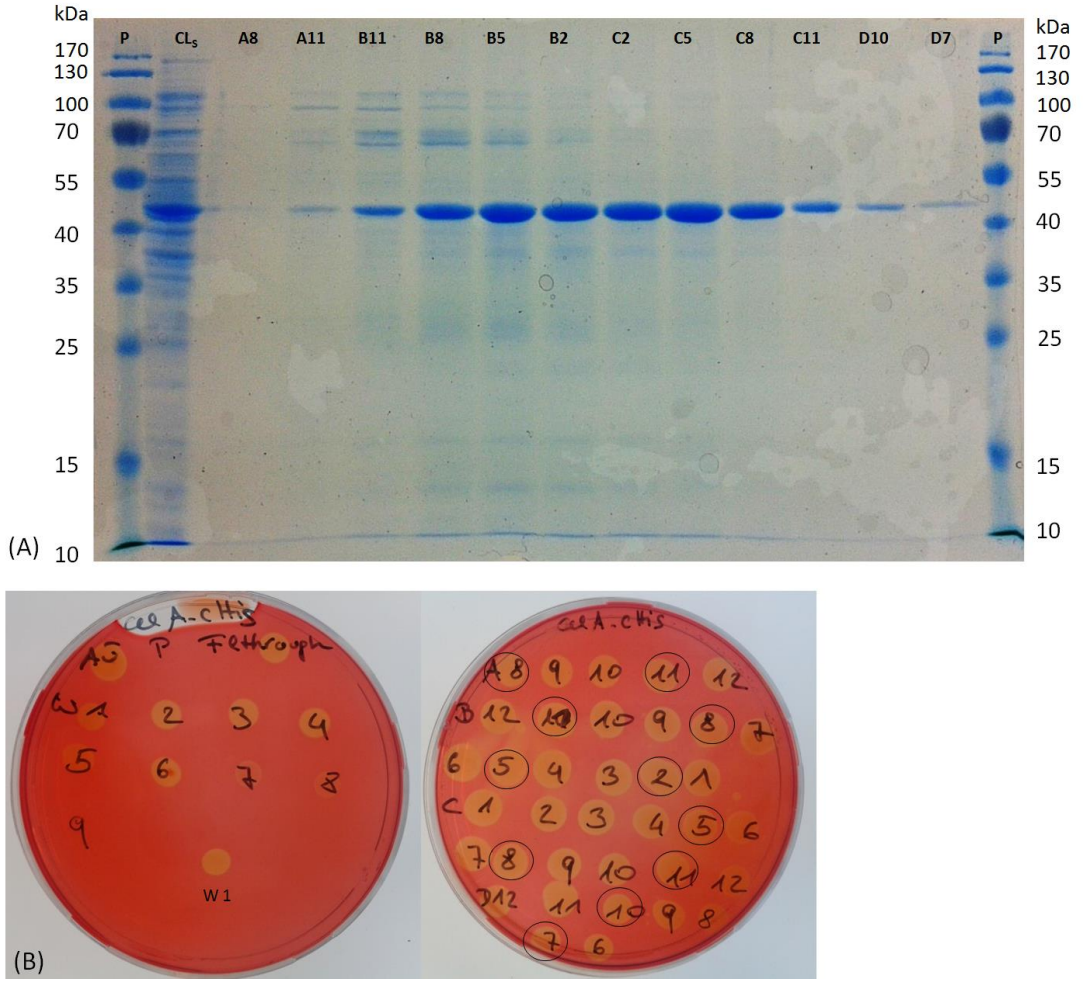

Figure 22: ÄKTA purification of CelA_cHis. (A) SDS-PAGE: P: PageRuler Prestained Protein Ladder (Fermentas), CL_S: *E. coli* BL21 (pMS470_*Soce_celAoΔScHis)* cell lysate soluble fraction (0.05 OD units), A8-D7: eluate fractions of His Tag purification. (B) Congo red plate assay. A-Ü: *E. coli* BL21 (pMS470_*Soce_celAoΔScHis)* cell lysate soluble fraction, P: cell lysate insoluble fraction, Flthrough: flow through of column loading, W1: manual wash fraction, W1-8: wash fractions, A8-D6: Elution fractions.

To reduce the amount of impurities in the elution fractions, CelA_cHis was purified using higher NaCl concentration in lysis and elution buffer. Since the results were comparable to those using lower salt concentrations, the salt concentration was reduced back to 300 mM, therefore the imidazole concentration in the lysis buffer was increased to 15 mM.

All CelA variants as well as Cex_cHis were purified using the optimized protocol (See Appendix 6.3.2 page 71). The eluates were pooled in impure and pure fractions (see Figure 23) prior to buffer exchange. The total protein concentration corresponds to the expression (compare Figure 20 page 42). N- and C-terminal 6x His-Tag CelAc variants have a higher expression compared to CelA variants and Cex_cHis and the total amount of purified protein corresponds to the expression.

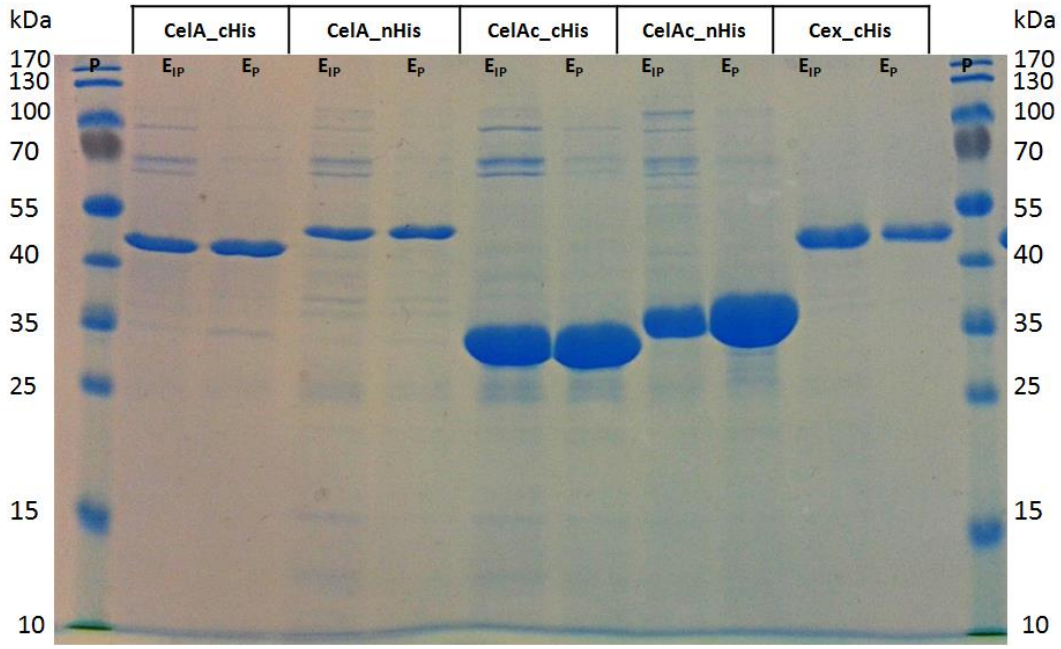

Figure 23: SDS PAGE: Affinity chromatography purified protein eluates: P: PageRuler Prestained Protein Ladder (Fermentas), E_{IP}: Pooled impure eluate fractions, E_P: pooled nearly pure eluate fractions. The imidazole concentrations of the pooled fractions are given in Appendix 6.3.2. automatic affinity chromatography.

3.1.4 Desalting

To get rid of the imidazole, the purified proteins were rebuffered using Vivaspin concentrator. Since the proteins precipited, the membrane was blocked. Therefore dialysis was chosen as a gentle method for rebuffering. Since the dialysis membrane consists of regenerated cellulose, the binding capacity of the enzymes was tested but no significant decrease of protein concentration could be determined.

3.2 Characterization

3.2.1 Optimal pH and temperature

In order to find the optimal conditions of CelA variants, the optimal pH and temperature was determined using CelA_nHis and CelAc_nHis. The optimal pH for both enzyme variants is pH 7, while the enzymes were inactive in pH 3, 4 and 10. The optimal temperature was 50°C and the enzymes were inactivated at 70°C.

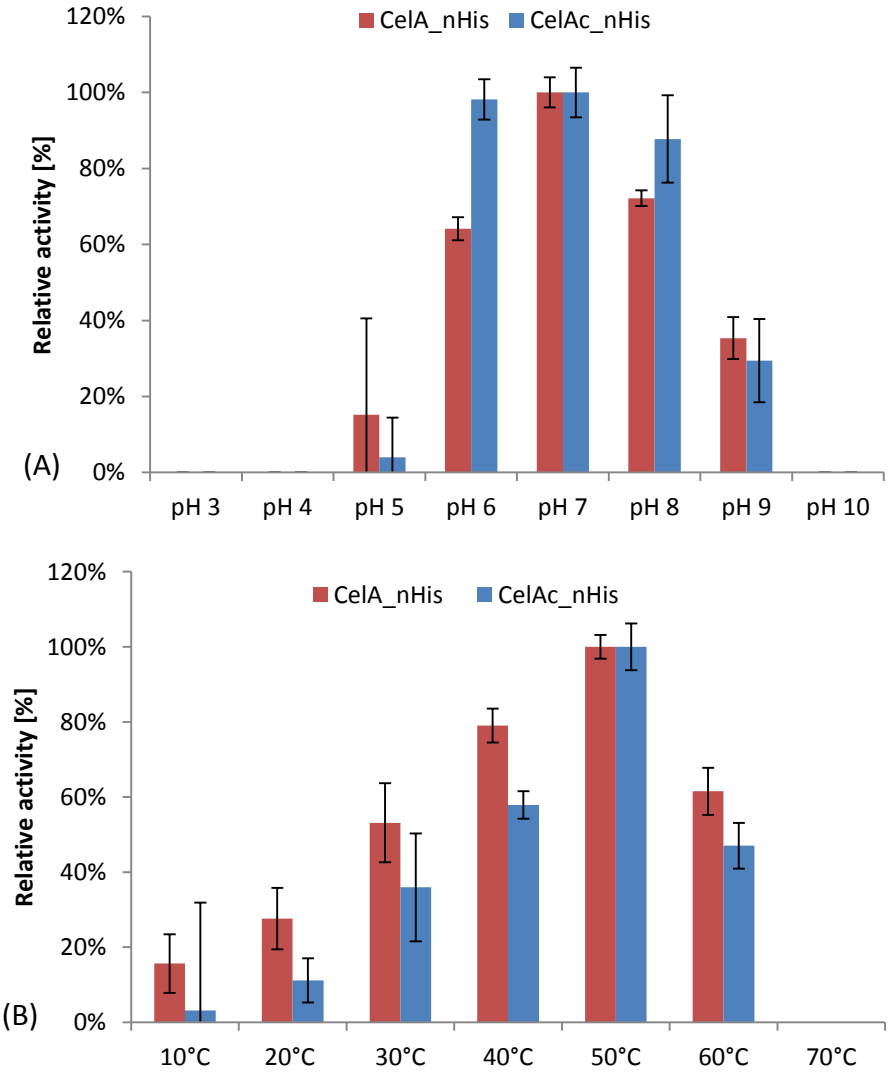

Figure 24: Optimal pH (A) and temperature (B) of CelA_nHis and CelAc_nHis. The release of reducing sugar was determined after incubation at respective conditions. The results are expressed relative to the maximal activity. The error bars represent standard deviation of triplicate experiments.

3.2.2 Substrate concentration

The optimal substrate concentration for CelAc_nHis was determined in borate buffer. The maximal reducing sugar concentration was detected after incubation with 1.75% CMC (see Figure 25).

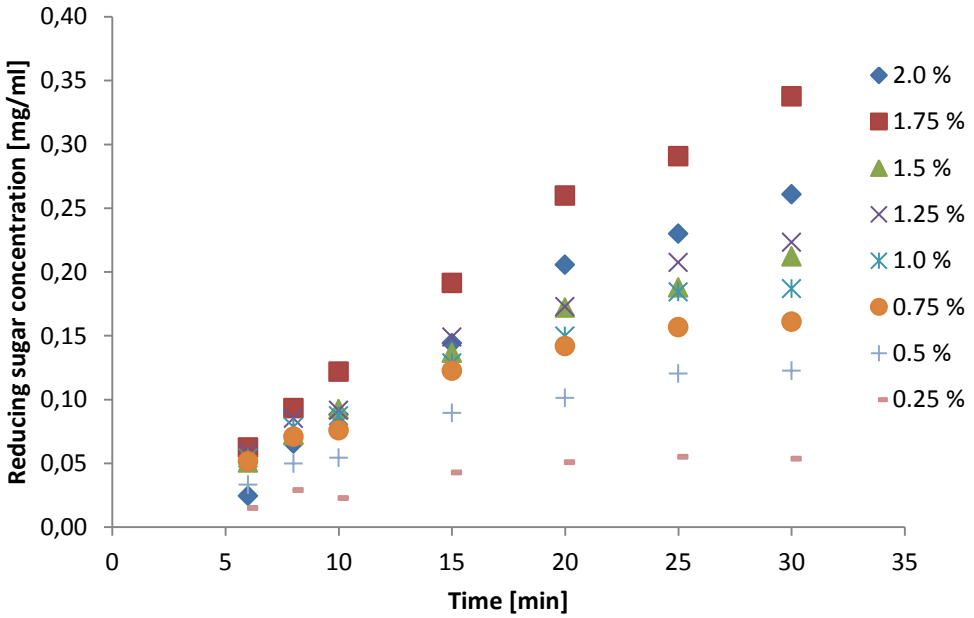

Figure 25: Optimal substrate concentration. CelAc_nHis was incubated in various CMC concentrations at standard conditions and the reducing sugar concentration was determined. The samples were measured in single evaluation.

3.2.3 Buffer condition

To check if the salts of borate buffer influence the enzyme activity, the release of reducing sugar was determined using different buffer systems. The highest enzyme activity of CelAc_nHis was detected in NaP_i Figure 26A. The doubling of NaP_i concentration decreased the activity of CelAc_cHis (see Figure B).

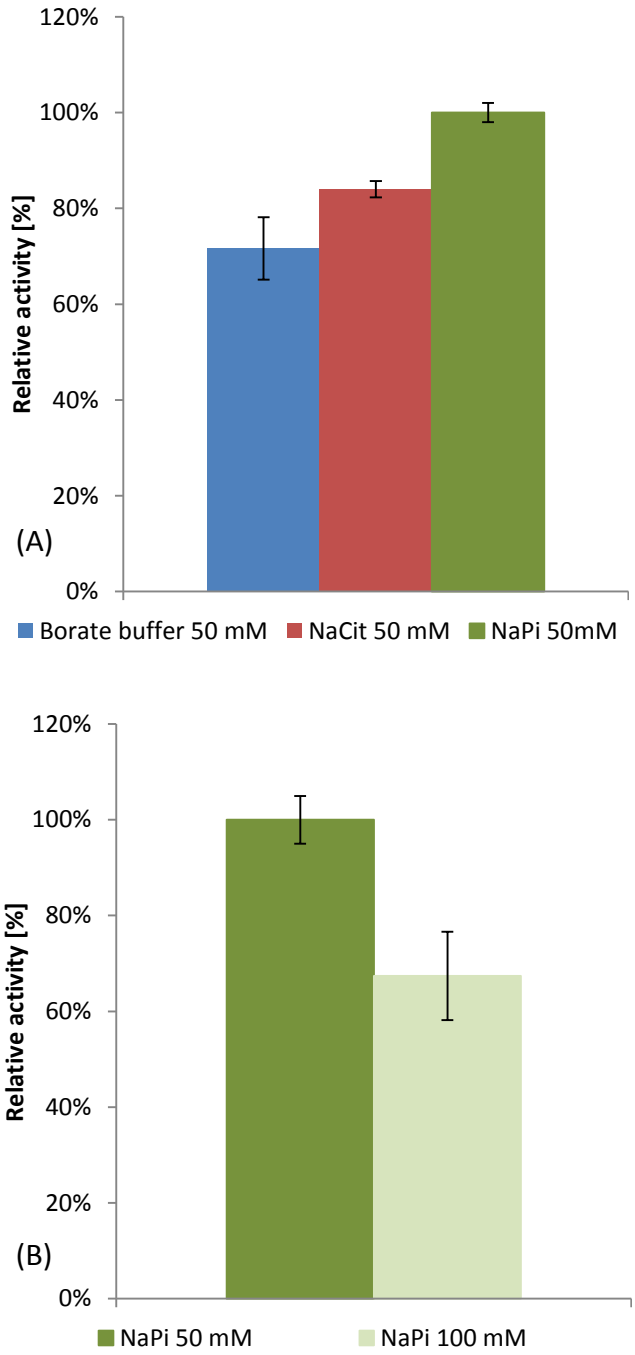

Figure 26: Influence of salts on the enzyme activity of CelAc_nHis and CelAc_cHis. (A) 1% CMC dissolved in borate, Na-citrate and NaP_i buffer, each 50 mM at pH 7 (CelAc_nHis 0.5-0.6 µg/ml), (B) 1.75% CMC in NaP_i 50mM and NaP_i 100 mM each pH 7 (CelAc_cHis 0.8µg/ml). The enzyme activity was determined using standard conditions, and the results are expressed relative to the activity of NaP_i 50 mM. The error bars represent standard deviation of triplicate experiments.

3.2.4 Specific enzyme activity towards CMC

The enzyme activity of all CelA variants was determined under several conditions. The specific activity is higher at 50°C compared to 37°C, and the incubation with higher substrate concentrations increased the activity too (see Table 7).

Table 7: Specific enzyme activity measured after ÄKTA purification and dialysis in U/mg.

	50°C 1.75% CMC	50°C 1% CMC	37°C 1.75% CMC	37°C 1% CMC
CelAc_cHis	296 ± 7	217 ± 5	239 ± 0.3	190 ± 9
CelAc_nHis	245 ± 15	177 ± 10	186 ± 3	149 ± 9
CelA_cHis	179 ± 10	140 ± 12	149 ± 5	138 ± 7
CelA_nHis	185 ± 7	125 ± 7	108 ± 8	95 ± 9

The specific enzyme activity of the truncated CelAc is higher compared to the native CelA. The molar specific enzyme activity was calculated by considering the molecular weight of the enzymes (see Table 4, page 25). The molecular weight of CelA is approx. 15% higher compared to CelAc, resulting in higher relative activity after calculation using the molar specific enzyme activity (see Table 8).

Table 8: Comparison of the specific enzyme activities of the CelA variants. The results are expressed relative to the activity of CelAc_cHis, calculated either from the specific activity U/mg or U/mmol.

	Relative activity U/mg	Relative activity U/mmol
CelAc_cHis	100%	100%
CelAc_nHis	83%	84%
CelA_cHis	60%	72%
CelA_nHis	63%	75%

The activity of the enzymes is higher after purification by gravity flow column chromatography and PD10 rebuffering compared to ÄKTA purification and dialysis (Figure 27). To verify if incomplete rebuffering influences the enzyme activity, the enzymes, purified by ÄKTA and dialysis, were additionally rebuffered using PD10 columns. Those enzymes had higher activities, but the high specific enzyme activities after gravity purification could not be achieved.

Table 9: Specific enzyme activity measured after gravity flow column purification and PD10 rebuffering in U/mg.

	50°C 1.75% CMC	50°C 1% CMC	37°C 1.75% CMC	37°C 1% CMC
CelAc_cHis	423 ± 8	331 ± 11	282 ± 6	233± 1
CelA_cHis	231 ± 4	187± 3	139 ± 9	127 ± 1

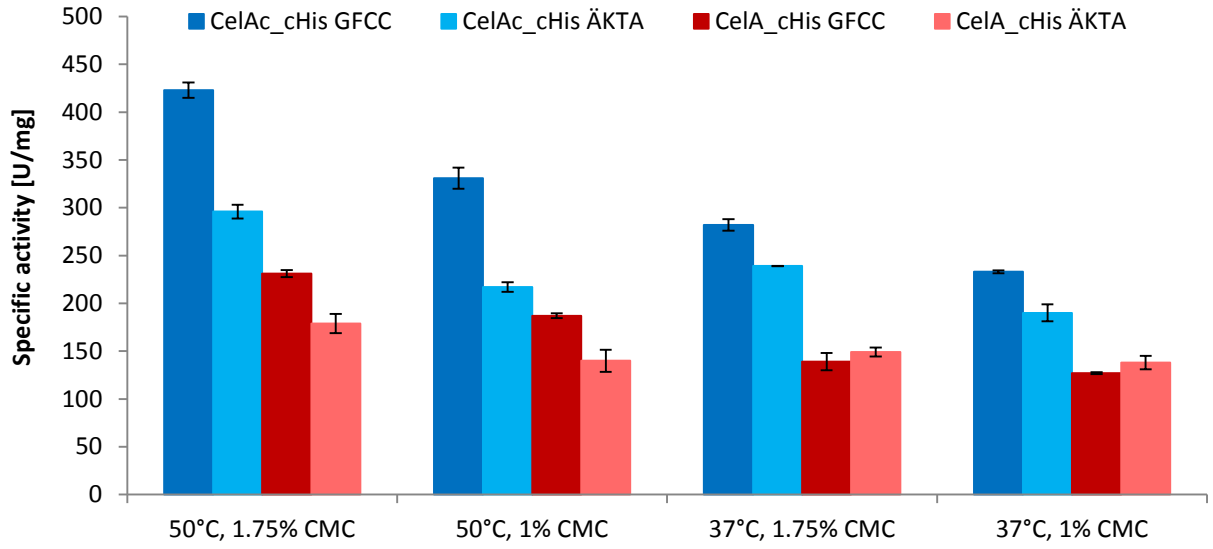

Figure 27: Specific activities of CelA variants. The enzymes were purified either by gravity flow column chromatography (GFCC) or ÄKTA chromatography (ÄKTA) and incubated with CMC in NaP_i 50 mM pH 7.The results are expressed as specific activity [U/mg].The error bars represent standard deviation of triplicate experiments.

3.2.5 Enzyme stability

3.2.5.1 Thermal stability

The thermal stability of CelA_cHis and CelAc_cHis was determined in a range of 5 to 70°C. The enzymes were inactivated at temperatures higher than 60°C, while 30% activity was measurable after pre-incubation at 50°C for 30 min. The remaining activity after pre-incubation for 100 h at 40°C was 65-70%, while pre-incubation at 5-30°C decreased the enzyme activity to 80% (Figure 28). The relative activity after pre-incubation for 250 h ranged from 30-40% (pre-incubation at 40°C) to 60-70% (pre-incubation at 5°C).

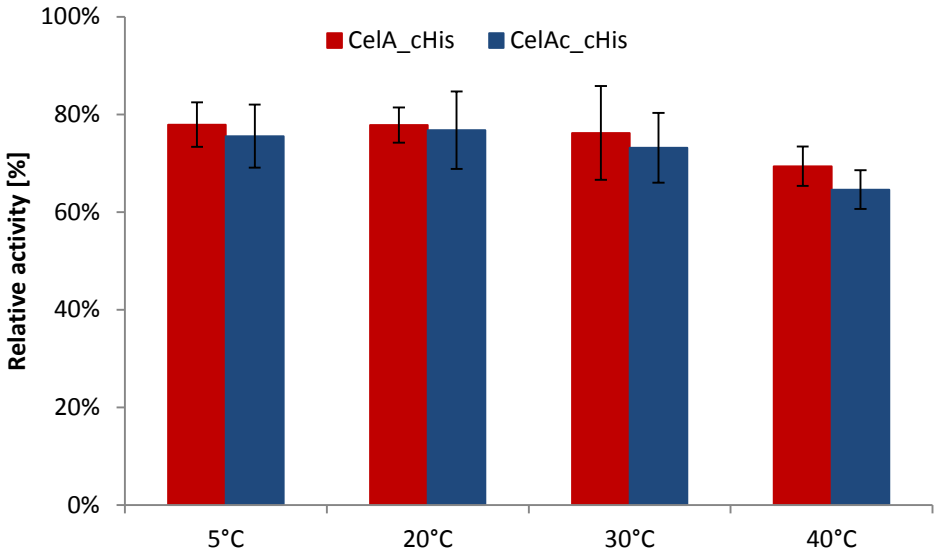

Figure 28: Thermal stability of CelA_cHis (red) and CelAc_cHis (blue). The release of reducing sugar was determined using standard conditions after pre-incubation at respective temperature for 100 h. The results are expressed relative to the standard activity of CelA_cHis or CelAc_cHis. The error bars represent standard deviation of triplicate experiments.

3.2.5.2 pH stability at 4°C

The pH stability was determined after pre-incubation of CelA_cHis and CelAc_cHis in borate buffer pH 4-9. The remaining activity of the enzymes is shown in Figure 29A and B. While the enzymes are stable from pH 5 to 9, they are inactivated in pH 4 after 120 h. The pre-incubation of CelA_cHis at pH 5 slightly reduced the enzyme activity, while CelAc_cHis is not affected.

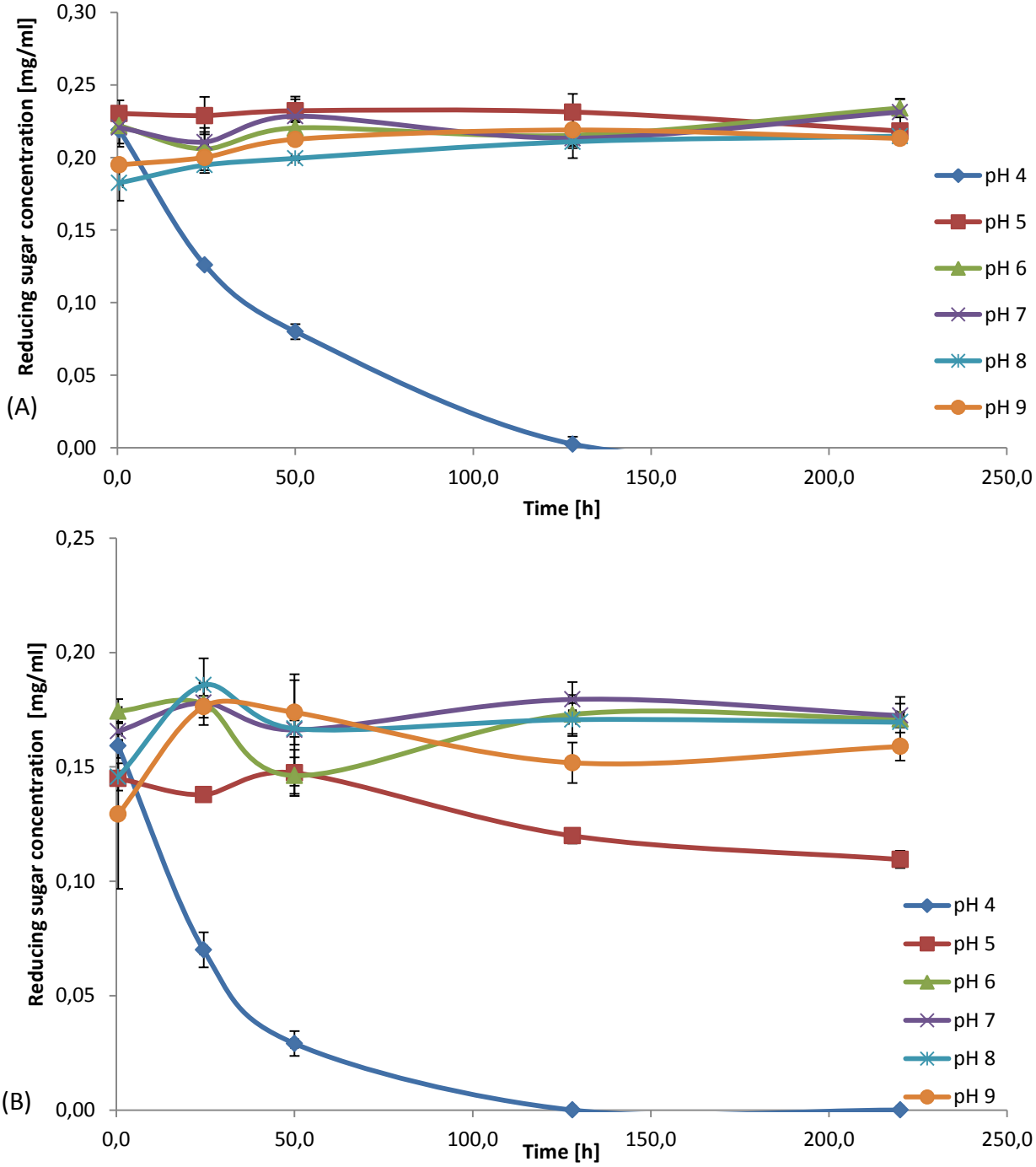

Figure 29: pH stability of CelAc_cHis (A) and CelA_cHis (B). The release of reducing sugar was determined using standard conditions after pre-incubation. The error bars represent standard deviation of triplicate experiments.

3.2.5.3 pH stability at 40°C

Due to the high pH stability after pre-incubation at 4°C, the pH stability of the enzymes was measured after pre-incubation at 40°C. The enzymes were inactivated after incubation at pH 4. The relative activity after 1 h reached about 60% for pre-incubation at pH 5 and pH 9 (see Figure 30). 80-90% remaining enzyme activity was detected after pre-incubation at pH 5-8 for 1 h. The influence after 2 h was stronger: the remaining enzyme activity at pH 5 and pH 9 was about 20%; incubation on pH 5-8 decreased the activity to 60-80%.

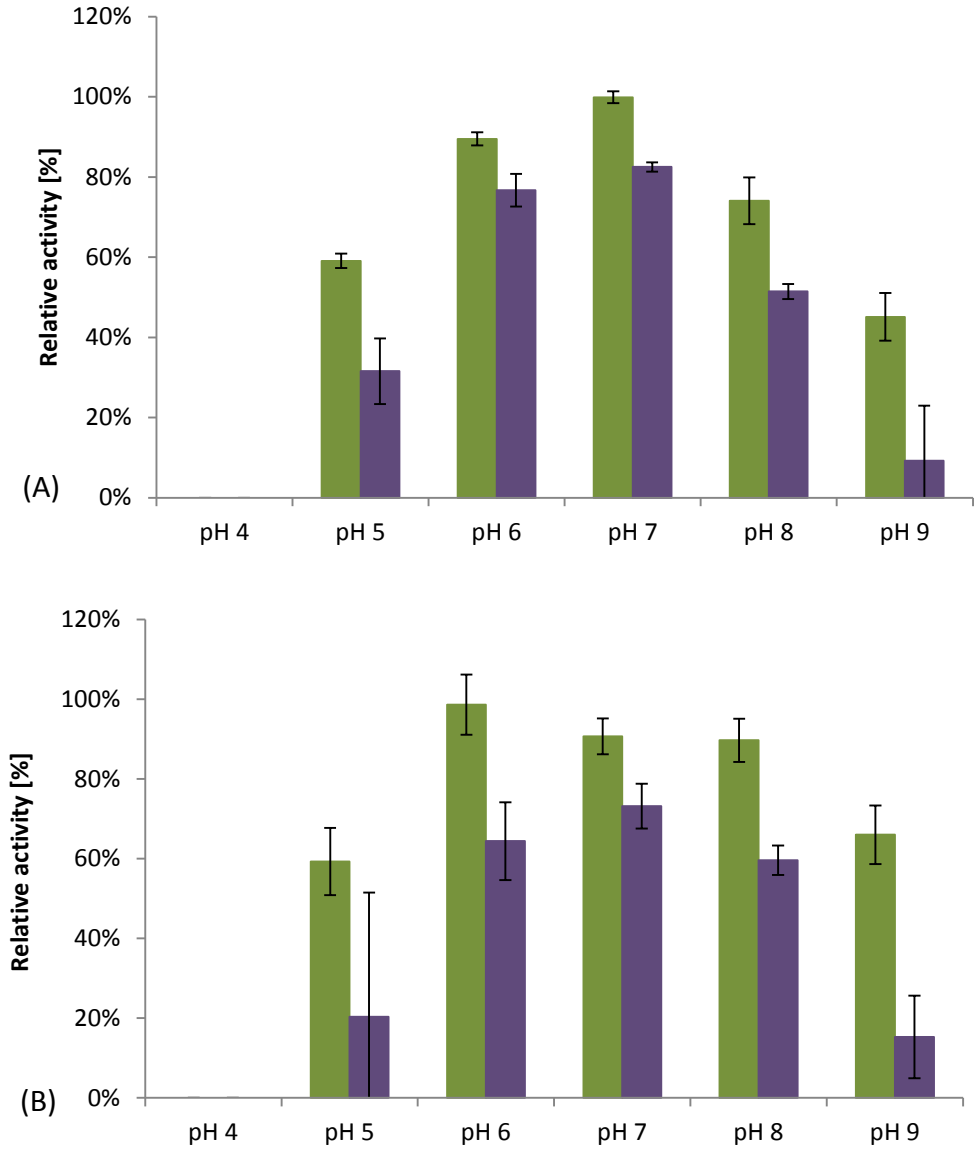

Figure 30: pH stability of CelAc_cHis (A) and CelA_cHis (B) after 1 h (green) and 2 h (violet) of incubation at 40°C. The release of reducing sugar was determined after pre-incubation using standard conditions. The results are expressed relative to the maximal activity. The error bars represent standard deviation of triplicate experiments.

3.2.5.4 Influence of metals

To check whether the enzyme activity is influenced by metal ions, several metal salts were pre-incubated with the enzymes and the enzyme activity was determined in presence of the metal ions. As shown in Table 10 no significant influence could be determined. The calculated relative activities range from 60 to 140%. Considering the influence of $CuSO_4$, $MnSO_4$ and $FeSO_4$ in the detection reaction, the influence is even less significant.

Table 10: Effect of metal salts on the enzyme activity of CelA_cHis and CelAc_cHis. The enzymes were pre incubated with respective metal salt and the enzyme activity was measured using standard conditions plus addition of metal salt to the substrate solution. The results are expressed relative to the activity in absence of metal salts. * Indicates an influence of the metal salt in the detection reaction. The standard deviation was calculated of triplicate experiments.

	Relative activity of CelAc_cHis [%]		**Relative activity of CelA_cHis [%]**	
	1 mM	**0,1 mM**	**1 mM**	**0,1 mM**
K_2SO_4	100 ± 3	101 ± 8	105 ± 4	89 ± 8
$MgSO_4$	117 ± 2	105 ±6	117 ± 8	80 ± 7
$CuSO_4$*	143 ± 8	105 ± 9	103 ± 7	59 ± 12
$(NH_4)_2SO_4$	126 ± 6	100 ± 2	92 ± 2	82 ± 0.4
$MnSO_4$*	111 ± 5	81 ± 12	100 ± 8	65 ± 6
$FeSO_4$*	118 ± 1	106 ± 6	128 ± 7	89 ± 1
$ZnSO_4$	97 ± 8	97 ± 3	102 ± 8	73 ± 10
$NiSO_4$	118 ± 9	92 ± 4	135 ± 0.3	72 ± 5
$CaCl_2$	108 ± 1	95 ± 4	108 ± 10	74 ± 3
NaCl	110 ± 5	107 ± 2	121 ± 8	77 ± 8

3.2.5.5 Influence of reagents

No significant change of the enzyme activity could be detected after incubation with EDTA, DTT and DMSO. Urea and Glycerin reduced the activity towards CMC, but the enzymes are still active, while SDS completely inactivated both enzymes. Ethanol seemed to slightly improve the enzyme activity. After pre-incubation of the enzymes in SDS and Urea, the enzyme activity was determined without supplemental addition of the reagent to the substrate solution. Those conditions enhanced the enzyme activity except CelA_cHis which was not affected by Urea.

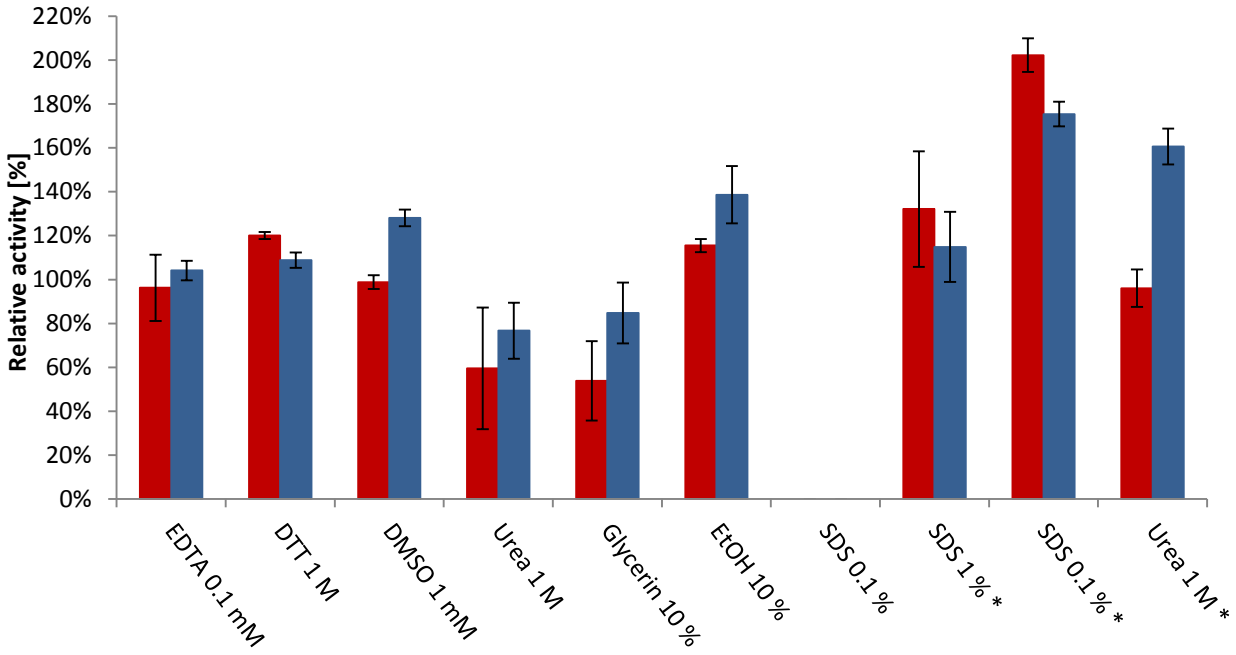

Figure 31: Effect of reagents on the enzyme activities of CelA_cHis (red) and CelAc_cHis (blue). The enzymes were pre-incubated with respective reagent and the enzyme activity was measured using standard conditions plus addition of the reagent to the substrate solution. * Indicates pre-incubation of the enzymes with the reagent and assay conditions without addition of reagent to the substrate solution. The results are expressed relative to the activity in absence of the reagent. The error bars represent standard deviation of triplicate experiments.

3.2.6 Enzyme inhibition by Cellobiose

The enzyme inhibition of CelA_cHis and CelAc_cHis was determined after pre-incubation with several Cellobiose concentrations. The remaining enzyme activity was measured and no significant influence of Cellobiose towards the enzyme was determined.

3.2.7 Degradation of insoluble substrates

The degradation of Whatman No.1 filter paper was determined after 2 h of incubation at standard conditions. The enzyme activity of the CelAc variants is higher compared to the CelA variants (Figure 32 A and B). The total reducing sugar concentration was higher after incubation with higher enzyme amounts.

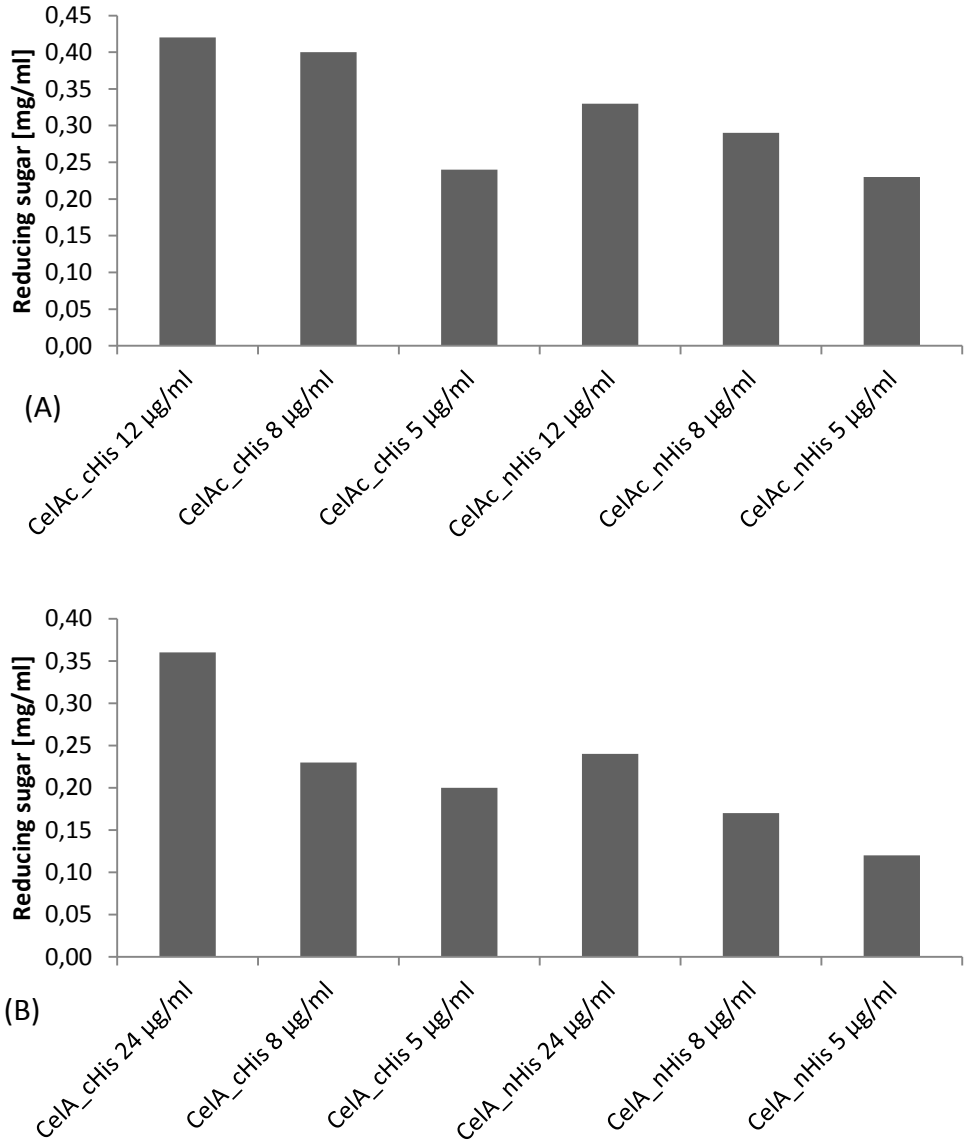

Figure 32: Degradation of Whatman No. 1 filter paper using CelAc (A) and CelA (B) variants. The amount of total reducing sugar was determined after incubation of filter paper for 2 h at standard conditions. The samples were measured in single evaluation.

The enzyme assay with Whatman No. 1 filter paper was incubated with CelAc_cHis and CelA_cHis for 6 h. The total and soluble reducing sugar concentration is shown in Table 10. Again, the amount of total and soluble reducing sugar is higher after incubation with CelAc_cHis compared to CelA_cHis.

Table 11: Enzyme reaction with Whatman No. 1 filter paper and CelA_cHis and CelAc_cHis. The amount of total and soluble reducing sugar was determined after incubation of filter paper for 6 h. The standard deviation was calculated of triplicate experiments.

Enzyme	Enzyme concentration [µg/ml]	total RS [mg/ml]	soluble RS [mg/ml]
CelAc_cHis	12	0.77 ± 0.04	0.24 ± 0.02
CelAc_cHis	8	0.60 ± 0.09	0.18 ± 0.02
CelAc_cHis	5	0.52 ± 0.08	0.13 ± 0.02
CelA_cHis	12	0.54 ± 0.10	0.12 ± 0.02
CelA_cHis	8	0.41 ± 0.07	0.12 ± 0.01
CelA_cHis	5	0.29 ± 0.06	0.09 ± 0.00

The reaction with Avicel was incubated with CelA_cHis and CelAc_cHis. 0.3 to 0.4 mg/ml total reducing sugar were detected after incubation for 6 h at 37°C using 40-199 µg/ml CelA_cHis or 83-414 µg/ml CelAc_cHis. No reducing sugar was detected after incubation with less enzyme or shorter incubation times.

3.2.8 Evaluation of the synergism of CelA and Cex

3.2.8.1 Degradation of CMC

The synergistic effect of Cex_cHis and CelAc_nHis towards CMC was determined. Cex_cHis showed no enzyme activity towards CMC. However, Cex_cHis enhances the activity of CelAc_nHis. Increasing the concentration of exocellulase in the enzyme reactions, led to higher release of reducing sugar (see Figure 33).

The increase of reducing sugar after 30 min is shown in Table 12.

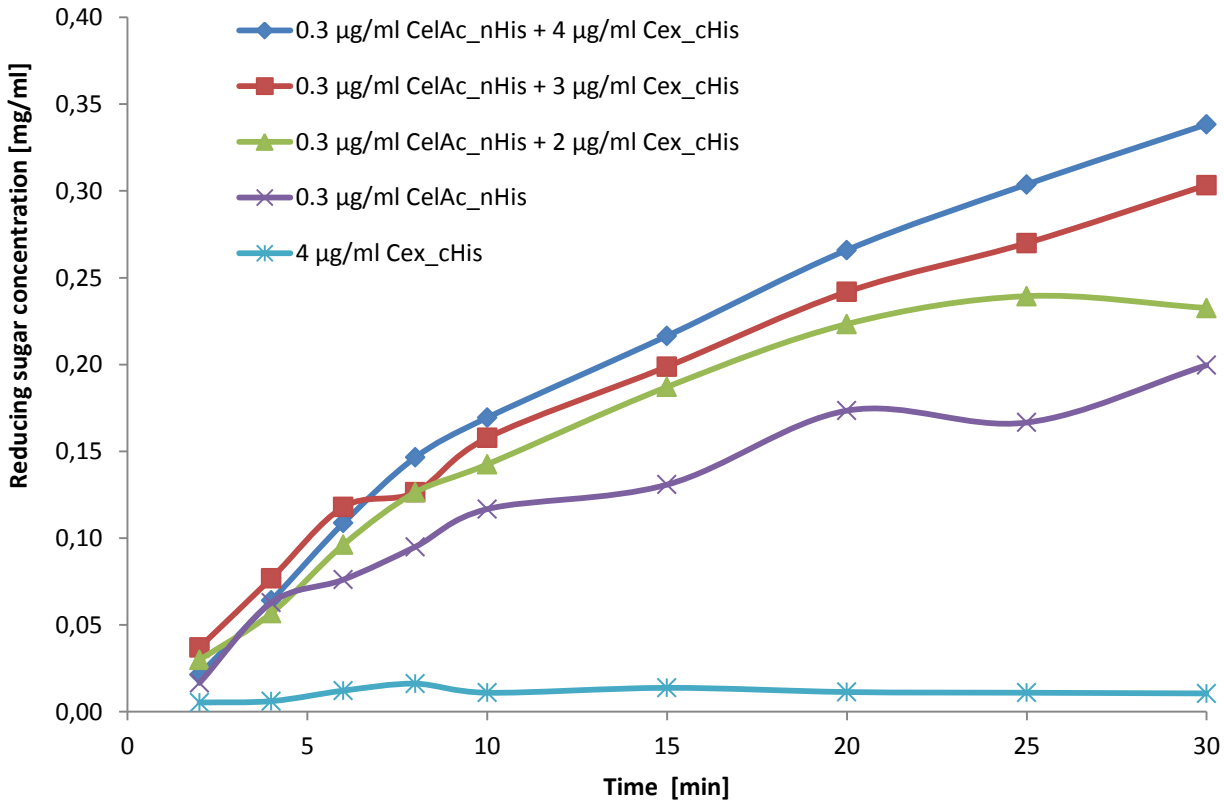

Figure 33: Synergistic effect of CelAc_nHis and Cex_cHis towards CMC. The enzyme reactions were incubated at standard conditions and the reducing sugar concentration was determined. The samples were measured in single evaluation.

Table 12: Synergistic effect of CelAc_nHis and Cex_cHis towards CMC. The relative activities of reducing sugar release were calculated after incubation at standard conditions for 30 min.

0.3 µg/ml CelAc_nHis	+ 2 µg/ml Cex_cHis	+ 3 µg/ml Cex_cHis	+ 4 µg/ml Cex_cHis
100%	117%	152%	169%

3.2.8.2 Degradation of insoluble substrates

The filter paper Macherey Nagel MN 619 de was degraded by CelAc_nHis, Cex_cHis and combinations of both enzymes. The amount of total and soluble reducing sugar is shown in Figure 34. The highest RS concentration was measured after incubation with Cex_cHis (16 µg/ml) and CelAc_nHis (1.2 µg/ml). Adding the RS of the controls CelA_nHis and Cex_cHis incubated solely, nearly the same concentration was reached as in combinatorial incubation.

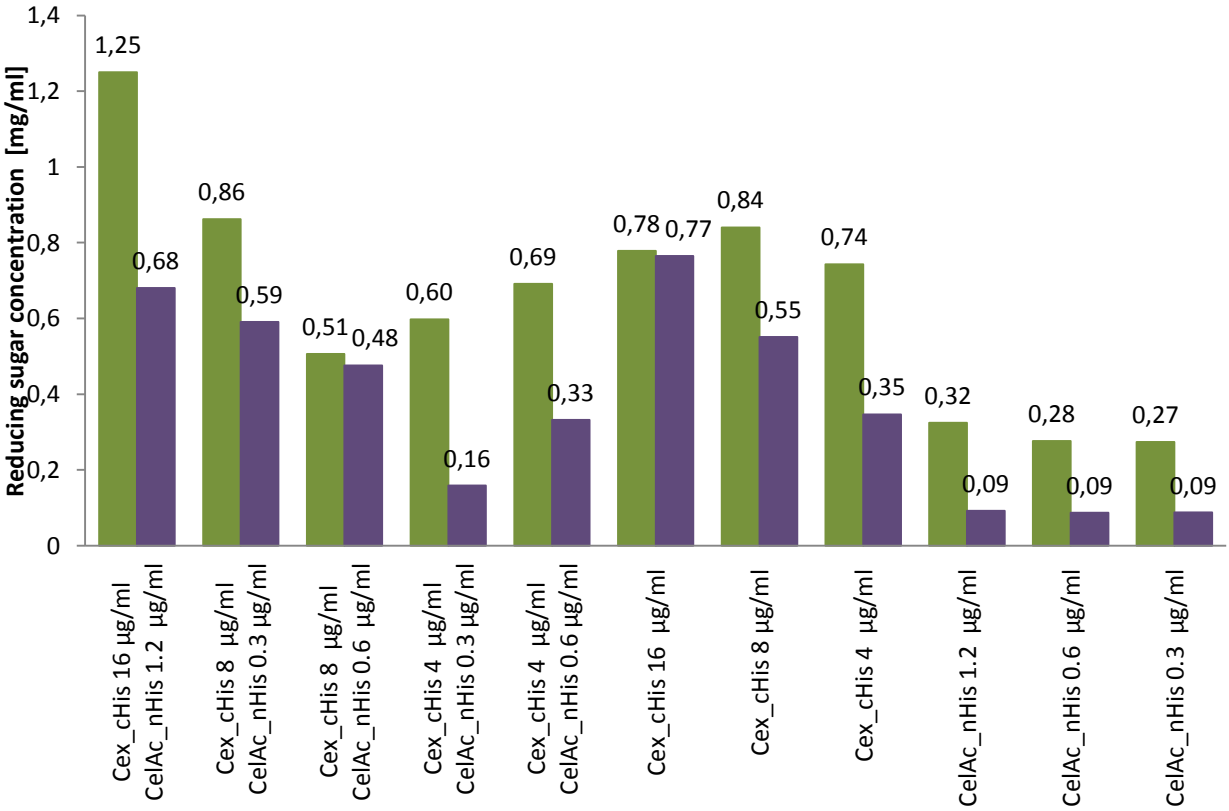

Figure 34: Degradation of and Macherey Nagel MN 619 de filter paper using CelAc_nHis and Cex_cHis. The amount of total (green) and soluble (violet) reducing sugar was determined after incubation of filter paper for 20 h. The samples were measured in single evaluation.

No release of reducing sugar was measured after incubation of Avicel for 2 h.

4 Discussion

Due to the broad availability of cellulose, it is of great interest to convert it for further industrial applications. Cellulose degrading enzymes should correspond to the requirements, given by a technical process. Therefore novel enzymes need to be characterized regarding their optimal conditions, enzyme activities and substrate specificities. In this work, the new GH5 endocellulase CelA from *Sorangium cellulosum* was characterized as well as the synergism with *Cellulomonas fimi* exocellulase Cex was determined.

4.1 Optimization of expression and purification

Caused by the high GC content of *celA* and *cex*, both genes were codon optimized for the expression in *Lactobacillus plantarum* (37; 28). Less inclusion bodies were formed through the expression of the optimized gene sequences. The growth at 25°C instead of 37°C and 28°C lead to higher expression. The higher metabolic load could cause a decrease of the growth rate (46). Incubation at reduced temperatures and the optimized gene sequences reduce the cell stress and lead to increased expression of soluble protein.

As a result of the higher expressed core enzyme CelAc compared to CelA (see Figure 19), an influence of the N-terminal repeat structure on the expression can be suggested. The Glycine, Serine, Threonine and Alanine rich sequence (80 AA), encoded by similar codons, could lead to a lack of the appropriate tRNA and therefore to a limitation in protein expression.

After reaching a moderate overexpression of the protein variants, the affinity chromatography was optimized to increase the yield of purified protein and reduce impurities. Since the gravity flow column chromatography is time consuming, the protein variants were isolated using automatic ÄKTA purifier system. High yields of nearly pure proteins could be reached of all CelA variants and Cex_cHis (see Figure 23). The protein yield of the CelA variants is, analogous to the expression, lower than the yield of the CelAc variants.

4.2 Characterization

The optimal temperature of CelA variants is 50°C. This thermophilic temperature was often found in bacterial and fungal endocellulases (47; 48; 49). The optimal pH is 7, whereas the proteins were active from pH 5 to pH 9. The thermal and pH stability was determined by pre-incubation of both enzymes at 4-70°C. Even though the optimal temperature was 50°C, inactivation occurred after 1 h. Both enzyme variants showed a high stability at temperatures from 5-40°C (see Figure 28, page 51). In terms of the application in silage process, the thermal stability is beneficial, since the processes take place at moderate temperatures.

CelA and CelAc were highly stable in a broad pH range and incubation at 4°C. The proteins were inactivated at pH 4 after 120 h and the enzyme activity of CelA seemed to be slightly reduced at pH 5. The influence of the pH was higher at 40°C (Figure 30, page 53). This could be due to higher Brownian motion of the molecules what could cause a faster inhibition. Both enzyme variants were stable at pH 6-8 whereas the reduction of enzyme activity at pH 5 and 9 was higher. Incubation at pH 4 led to inactivation of CelA and CelAc. Many thermostable endocellulases are reported, some show a broad pH range and stability (50; 51; 49) and other enzymes have a narrow pH range (52). The endocellulase should be expressed in *Lactobacillus plantarum*, to improve the silage process. Therefore the enzyme must be active at low pH and moderate temperature conditions (pH 4, 20°C). To reach a higher enzyme activity under those conditions, the enzyme needs to be engineered towards a higher activity and pH stability.

Metal ions had no significant influence on the enzyme activity. EDTA, DMSO and DTT showed no influence, while SDS, added to the enzyme reaction, completely inactivated the enzymes. Pre-incubation with SDS enhanced the enzyme activity of both variants, while Urea increased only the activity of CelAc. The increased enzyme activity after pre-incubation with SDS, or Ethanol in the enzyme reaction could be due to improved substrate accessibility in the active site through opening of the cleft. The thermophile endocellulase from *Chaetomium thermophilum* is inactivated by metals and reagents (52), while other enzymes are less sensitive (49; 48; 50; 53). Hyun Woo Zin et al. characterized a salt-tolerant carboxymethyl cellulase from *Artemia salina*, which gets strongly activated by $CoCl_2$, $CuCl_2$ and $MnCl_2$ (54). The interaction of metal-salts or reagents with the enzyme could cause changes in electrostatic binding, and thus affect the enzyme activity (52).

Béra-Maillet et al. reported an inhibitory effect of cationic metals towards GH 5 endocellulases. Cations like Zn^{2+} or Cu^{2+} could oxidize cysteins leading to denaturation of the protein (55). Reducing agents, like DTT or β-Mercaptoethanol, increase the activity of SH-class enzymes by keeping the SH group in a reduced state (49). Since DTT did not affect the enzyme activity of CelA variants, the enzyme does not belong to the SH enzyme class.

CelAc as well as CelA were not inhibited by Cellobiose. Since the remaining enzyme activity was determined by a reducing sugar assay, Cellobiose was absent while the enzyme reaction. It is possible that Cellobiose bound to the active site while pre-incubation and dissociated in the enzyme reaction, reactivating the enzyme. For an exact inhibition assay, further trials using different detection methods need to be done.

The highest enzyme activity was determined in 1.75% CMC; higher substrate concentrations slightly reduced the enzyme activity. A reason could be the solubility of CMC in buffer, leading to solid particles, and a lower activity compared to the soluble substrate.

The specific enzyme activity towards CMC was determined of all CelA variants and the activity of the core enzyme CelAc is higher compared to CelA with the N-terminal repeat structure. Thus, the repeat sequence has no catalytic function. As reported by Hou et al., the cellulolytic enzymes of *Sorangium cellulosum* are arranged at the cell surface to a cellulosome-like complex (56). The N-terminal repeat sequence of CelA could be responsible for the attachment to the cell surface or for interaction with other cellulases.

The specific enzyme activities of CelAc and CelA were 420 U/mg and 180 U/mg (see Table 7 and 9, page 49-50). The GH5 endoglucanase Cel5A from the hyperthermophilic bacteria *Thermotoga maritime* has a specific activity of 39 U/mg towards CMC, determined at optimal conditions (80°C, pH 4.8) (57). The endocellulase from the thermophilic fungus *Chaetomium thermophilum* CT2 had a specific activity of 28 U/mg towards CMC (52). As reported by Kyriacou et al. (58), the commercial cellulase mix Celluclast was separated into its compounds endoglucanase I-III (EG I-III) and cellobiohydrolase II (CBH II). The enzyme activity of the purified enzymes was determined towards CMC and 30 U/mg were measured for EG I, whereas EG II and EG III showed higher activities, 90 and 100 U/mg respectively. Comparing those results, the activity of the isolated CelA variants is high at optimal conditions.

The specific enzyme activity towards CMC was determined after isolation via gravity flow column chromatography or automatic affinity chromatography and rebuffering using size exclusion columns or dialysis. All CelA variants were more active after gravity flow column chromatography and size exclusion chromatography. Incomplete rebuffering via dialyses was not the only reason for the reduced enzyme activity. An influence of the automatic affinity chromatography is suggested. The automatic purification takes place at a maximal pressure of 3 bar. Compression and partly defolding of the protein could be a further reason for the reduced activity.

Using CMC as soluble substrate, the degree of substitution has to be considered. Therefore the degradation rate for CMC (DS = 0.7) should be limited to 2% (2). The hydrolysis rate of CMC decreases with time (see Figure 25, page 47). The decrease is faster after incubation with lower CMC concentrations, considering this is caused by limitations of hydrolyzable β-glycosidic bonds. High specific enzyme activities were measured at maximal hydrolyzation rate of 3%.

The synergistic degradation of CMC was also examined. The addition of the endoxylanase and exocellulase Cex to the CelAc enzyme reaction led to an increase of reducing sugar release. After incubation for 3 h 6% of the substrate was hydrolyzed, whereas 4% were hydrolyzed after exclusive incubation of CelAc.

The degradation of insoluble substrates was determined with Whatman No. 1 filter paper and Avicel. The CelAc variants were more active towards Whatman No. 1 filter paper than the natural variants CelA. Hence, the N-terminal sequence has no affinity to bind cellulose. CBM are important for the degradation of crystalline substrate since the enzymes are bound to the substrate and the processivity is enhanced. Thongekkaew et al. (24) reported an increased activity towards Avicel and Sigmacell (both microcrystalline cellulose) after 1 or 2 CBM I from *Trichoderma reesei* were tagged on the C-terminal end of *Cryptococcus* sp. S-2 Carboxymethyl cellulase. Due to the lack of CBM in CelA variants Avicel, a highly crystalline model substrate, was not degraded properly.

The synergism of CelAc and Cex was also tested towards the filter paper Macherey Nagel MN 619 de. The exocellulase and xylanase Cex was more active compared to CelAc. This could be due to the C-terminal CBM 2, or minor contents of Xylan in the filter paper. The enzymes did not enhance their activity compared to the incubation of the isolated enzymes.

The synergistic degradation of insoluble substrates using Cex or alternative exocellulases needs to be further optimized regarding the enzyme enzyme and enzyme substrate ratio. Due to the low repeatability of enzyme assays with filter paper as substrate, phosphor acid swollen cellulose (PASC) could be used as alternative model substrate. However, this substrate consists mostly of amorphous cellulose and is not representative for natural, crystalline cellulose.

Concluding, the endocellulase CelA variants were characterized towards the specific enzyme activity, optimal conditions, enzyme stabilities and the synergism with the exocellulase Cex. For the application in silage processes, the enzyme needs to be engineered to optimize the activity and stability at low pH and moderate temperature conditions.

5 Literature

1. **Dewey D. Y. RYU, Mary Mandels.** Cellulases: biosynthesis and applications. *Enzyme Microb. Technol. .* 1980.

2. **Y.-H- Pericival Zhang, Michael E. Himmel, Jonathan R. Mielenz.** Outlook for cellulase improvement: Screening and selection strategies. *Biotechnology Advances.* 24, 2006.

3. **Bhat, M.K.** Cellulase and related enzymes in biotechnology. *Biotechnology Advances.* 18, 2000.

4. **Nitin Trivedi, Vishal Gupta, C.R.K. Reddy et al.** Enzymatic hydrolysis and production of bioethanol from common macrophytic green alga Ulva fasciata Delile. *Bioresource Technology.* 150, 2013.

5. **N. Sarkar, S. K. Ghosh, S. Bannerjee et al.** Bioethanol production from agricultural wastes: An overview. *Renewable Energy.* 37, 2012.

6. **M.G. Tabka, I. Herpoël-Gimbert, F. Monod et al.** Enzymatic saccharification of wheat straw for bioethanol production by a combined cellulase xylanase and feruloyl esterase treatment. *Enzyme and Microbial Technology.* 38, 2006.

7. **Z.H. Sun, S.M. Liu, G.O. Tayo et al.** Effects of cellulase or lactic acid bacteria on silage fermentation and in vitro gas production of several morphological fractions of maize stover. *Animal Feed Sciene and Technology.* 152, 2009.

8. **Z.G. Weinberg, Y. Chen.** Effects of storage period on the composition of whole crop wheat and corn silages. *Animal Feed Science and Technology.* 185, 2013.

9. **Z.G. Weinberg, R.E. Muck.** New trends and opportunities in the development and use of inoculants for silage. *FEMS Microbiology Reviews.* 19, 1996.

10. **S. H. Mohd-Setapar, N. Abd-Talib, R. Aziz.** Review on Crucial Parameters of Silage Quality. *APCBEE Procedia.* 3, 2012.

11. **Z.G. Weinberg, G. Ashbell, Yaira Hen et al.** The effect of cellulase and hemicellulase plus pectinase on the aerobic stability and fibre analysis of peas and wheat silages. *Animal Feed Science and Technology.* 55, 1995.

12. **Serge Pérez, Daniel Samain.** Structure and Engineering of Cellulose. *Advances in Carbohydrate Chemistry and Biochemistry.* 64, 2010.

13. **Peters, Jana.** Strukturanalysen an Cellulose und Cellulosederivaten aus ionischen Lösungen. 2004.

14. **Chun You, Xiao-Zhou Zhang, Y.-H. Percival Zhang.** Mini-scaffolding enhanced mini-cellulosome hydrolysis on low-accessibility cellulose (Avicel) more than on high-accessibility amorphous cellulose. *Biochemical Engineering Journal.* 63, 2012.

15. **Andersen, Natalija.** Enzymatic Hydroylsis of Cellulose. 2007.

16. *Sigma Aldrich.* [Online] [Cited: 01 02, 2013.] http://www.sigmaaldrich.com.

17. **Hehuan Liao, Xiao-Zhou Zhang, Joseph A. Rollin et al.** A minimal set of bacterial cellulases for consolidated bioprocessing of lignocellulose. *Biotechnology Journal.* 6, 2011.

18. **Alexander N. Glazer, Hiroshi Nikaido.** *Microbial Biotechnology. Fundamentals of applied Microbiology.* s.l. : W.H. Freeman and Company, 1998.

19. **E. Vlasenko, M. Schülein, J. Cherry et al.** Substrate specificity of family 5, 6, 7, 9, 12 and 45 endoglucanases. *Bioresource Technology.* 101, 2010.

20. **T. Scheper, G.T. Tsao, et al.** *Advances in Biochemical Engineering / Biotechnology 65: Recent Progress in Bioconversion of Lignocellulosics.* s.l. : Springer, 1999.

21. Cazypedia. [Online] [Cited: 12 15, 2013.] http://www.cazypedia.org.

22. **Edward A. Bayer, Linda J. W. Shimon, Yuval Shoham.** Cellulosomes-Strucutre and Ultrastructure. *Journal of Structural Biology.* 124, 1998.

23. *Cazy.* [Online] [Cited: 11 14, 2013.] www.cazy.org.

24. **Jantaporn Thongekkaew, Hiroko Ikeda, Kazuo Masaki et al.** Fusion of cellulose binding domain from Trichoderma reesei CBHI to Cryptococcus sp. S-2 cellulase enhances its binding affinity and its cellulolytic activity to insoluble cellulosic substrates. *Enzyme and Microbial Technology.* 52, 2013.

25. **Marcos Henrique Luciano Silveira, Martinho Rau, Elba Pinto da Silva Bonet al.** A simple and fast method for the determination of endo- and exo-cellulase activity in cellulase preparations using filter paper. *Enzyme and Microbial Technology.* 51, 2012.

26. **John D. McCarter, Stehen G. Withers.** Mechanisms of enzymatic glycoside hydrolysis. *Current Opinion in Structural Biology.* 4, 1994.

27. **Rolf Müller, Klaus Gerth.** Development of simple media which allow investigations into the global regulation of chivosazol biosynthesis with Sorangium cellulosum S ce56. *Journal of Biotechnology.* 121, 2006.

28. **Susanne Schneiker, Olena Perlova, Olaf Kaiser et al.** Complete genome sequence of the myxobacterium Sorangium cellulosum. *Nature Biotechnology.* 25, 2007.

29. **Matthias Keck, Nicolas Gisch, Hermann Moll et al.** Unusual Outer Membrane Lipid Composition of the Gram-negative, Lipopolysaccharide-lacking Myxobacterium Sorangium cellulosum Soce56. *Journal of biological Chemistry.* 286, 2011.

30. **Pradella.** Characterisation, genome size and genetic manipulation of the myxobacterium Sorangium cellulosum So ce56. *Arch Microbiol.* 2002, 178, pp. 484-492.

31. SignalP 4.1 Server. [Online] [Cited: 01 02, 2013.] http://www.cbs.dtu.dk/services/SignalP/.

32. NCBI BLAST. [Online] [Cited: 11 21, 2013.] http://blast.ncbi.nlm.nih.gov/Blast.cgi.

33. psipred. [Online] [Cited: 11 21, 2013.] http://bioinf.cs.ucl.ac.uk/psipred/.

34. M4T Server. [Online] ver. 3.0. [Cited: 11 21, 2013.] http://manaslu.aecom.yu.edu/M4T/results/DIR_ETyC3e/out.html.

35. RCSB Protein Data Bank. [Online] [Cited: 11 21, 2013.] http://www.rcsb.org/pdb/home/home.do.

36. **Christopher M. Bianchetti, Phillip Brumm, Robert W. Smith et al.** Structure, Dynamics, and Specificity of EndoglucanaseD fgrom Clostridium cellulovorans. *Journal of Molecular Biology.* 425, 2013.

37. **P. Chaudhary, N.N. Kumar, D.N. Deobagkar.** The glucanases of Cellulomonas. *Biotechnology Advances.* 15, 1997.

38. **Lever, M.** Colorimetric and Fluorometric Carbohydrate Determiantion withp-Hydroxybenzoic acid Hydrazide. *Biochemical medicine.* 7, 1973.

39. **Miller, Gail Lorenz.** Use of Dinitrosalicyclic acid Reagent for Determinatin of Reducing sugar. *Analytical chemistry.* 31, 1959.

40. Carl Roth. [Online] [Cited: 02 05, 2014.] www.carlroth.com.

41. life technologies. [Online] [Cited: 11 14, 2013.] www.lifetechnologies.com.

42. IDT Integrated DNA Technology. [Online] [Cited: 11 17, 2013.] http://eu.idtdna.com/analyzer/applications/oligoanalyzer/default.aspx.

43. Thermo Scientific. [Online] [Cited: 10 11, 2013.] http://www.piercenet.com.

44. GE Healthcare. [Online] [Cited: 10 14, 2013.] http://www.gelifesciences.com.

45. Expasy ProtParam. [Online] [Cited: 11 21, 2013.] http://www.expasy.org/proteomics.

46. **L.A. Palomares, S. Estrada-Mondaca, O.T. Ramírez.** Production of Recombinant Proteins Challenges and Solutions. *Methods in Molecular Biology.* 267, 2004.

47. **Sangrila Sadhu, Tushar Kanti Maiti.** Cellulase Production by Bacteria: A Review. *British Microbiology Research Journal.* 3, 2013.

48. **Jatinder Kaur, Bhupinder S. Chadha, Badhan A. Kumar et al.** Purification and characterization of two endoglucanases from Melanocarpus sp. MTCC3922. *Bioresource Technology.* 98, 2007.

49. **Kyoung-Mi Lee, Marimuthu Jeya, Ah-Reum Joo et al.** Purification and characterization of a thermostable endo-ß-1,4-glucanase from a novel strain of Penicillium purpurogenum. *Enzyme and Microbial Technology.* 46, 2010.

50. **Dongyang Liu, Ruifu Zhang, Xingming Yang et al.** Expression, purification and characterization of two thermostable endoglucanases cloned from a lignocellulosic

decomposing fungi Aspergillus fumigatus Z5 isolated from compost. *Protein Expression and Purification.* 79, 2011.

51. **Crispen Mawadza, Rajni Hatti-Kaul, Remigio Zvauya et al.** Purification and characterization of cellulases produced by two Bacillus strains. *Journal of Biotechnology.* 83, 2000.

52. **Duo-Chuan Li, Mei Lu, Ya-Ling Li, Jing Lu.** Purification and characterization of an endocellulase from the thermophilic fungus Chaetomium thermophilum CT2. *Enzyme and Microbial Technology.* 33, 2003.

53. **Jing Liu, Wenshui Xia.** Purification and characterization of a bifunctional enzyme with chitosanase and cellulase activity from commercial cellulase. *Biochemical Engineering Journal.* 30, 2006.

54. **Hyun Woo Zin, Kwang-Hyun Park, Tae Jin Choi.** Purifiction and characterization of a carboxymethyl cellulase from Artemia salina. . *Biochemical and Biophysical Research Comminucations.* 443, 2014.

55. **C. Béra-Maillet, L. Arthaud, P.Abad, M. Rosso.** Biochemical characterization of MI-ENG1, a family 5 endoglucanase secreted by the root-knot nematode Meloidogyne incognita. *Eur. J. Biochem. .* 267, 2000.

56. **Pei-bin Hou, Yue-zhong Li, Bin-hui Wu et al.** Cellulolytic complex exists in cellulolytic myxobacterium Sorangium. *Enzyme and microbial Technology.* 38, 2006.

57. **Jose H. Pereira, Zhiwei Chen, Ryan P. McAndrew et al.** Biochemical characterization and crystal structure of endoglucanase Cel5A from the hyperthermophilic Thermotoga maritima. *Journal of Structural Biology.* 172, 2010.

58. **Andreas Kyriacou, C. Roger MacKenzie, Ronald J. Neufeld.** Detection and characterization of the specific and nonspecific endoglucanases of Trichoderma reesei: evidence demonstrating endoglucanase activity by cellobiohydrolase II. *Enzyme ans Microbial Technology.* 9, 1987.

59. **Guodong Liua, Yuqi Qina, Yibo Hua et al.** An endo-1,4-ß-glucanase PdCel5C from cellulolytic fungus Penicillium decumbens with distinctive domain composition and hydrolysis product profile. *Enzyme and Microbial Technology.* 52, 2013.

6 Appendix

6.1 Fermentation

- Overnight culture:
  - 5 ml LB-Amp are inoculated with a glycerol stock or a single colony
  - The cultures are incubated at 37°C, 120 rpm ON
- Main culture:
  - 200 ml LB-Amp in a 500 ml shake flask are inoculated to an OD_{600} 0.05-0.1
  - The cultures are incubated at 25°C, 100-120 rpm until an OD_{600} of 0.6-0.8 is reached
  - The expression is induced with IPTG to a concentration of 100 µM
  - The cultures are incubated at 25°C, 100-120 rpm ON
- Harvest:
  - 50 ml culture are poured into a falcon tube and centrifuged (3,200 x g, 4°C, 30 min)
  - The pellet is stored at -20°C

6.2 Cell lysis using Sonication

- The pellets of 100-200 ml culture are resuspended in 30 ml lysis buffer (50 mM NaP_i, 300 mM NaCl, 10-15 mM imidazole, pH 8)
- The cells are lysated twice through sonication in the Branson Sonifier for 5 min on ice (80% intensity)
- Insoluble substances are removed by centrifugation (3,220 x g for 30 min at 4°C)

6.3 Protein purification

6.3.1 Gravity flow column chromatography

- Affinity chromatography
  - Column with 1 ml Ni-NTA Sepharose (GE Healthcare 17-5218)
  - Equilibration of the column
    - The column is washed with
      - 10 ml 20% Ethanol
      - 10 ml ddH_2O
      - 10 ml lysis buffer
  - The soluble fraction of the cell lysate is loaded onto the column
  - Washing
    - Unspecific bound protein are washed with 50 ml wash buffer (50 mM NaP_i, 300 mM NaCl, 20 mM imidazole, pH 8)
  - Elution
    - The protein is eluted with 1 ml elution buffer (50 mM NaP_i, 300 mM NaCl, 500 mM imidazole, pH 8), the elution is repeated 10 times
    - each step is incubated for 20 min
    - Fraction 1-5 or 8 are pooled for desalting

- Size exclusion chromatography
    - PD10 column (GE Healthcare 17-0851-01)
    - Equilibration of the column
        - The column is washed with
            - 10 ml 20% Ethanol
            - 25 ml ddH_2O
            - 25 ml NaP_i (50 mM, pH 7) for CelA variants or borate buffer (50 mM Citrate, 50 mM NaP_i, 50 mM Borate, pH 7), or
            - 25 ml 0.9% NaCl pH 7 for Cex_cHis
    - Loading
        - 2.5 ml pooled affinity chromatography fraction are loaded
    - Elution
        - The protein is eluted in 3.5 ml NaP_i or 0.9% NaCl
    - The column is washed with 25 ml buffer prior the next protein fraction is loaded

6.3.2 Automatic affinity chromatography with ÄKTA purifier

- Column (GE Healthcare HisTrap FF, 17-5355-01)
    - Equilibration of the column (5 ml/min)
        - The column is washed with
            - 20 ml 20% Ethanol
            - 20 ml ddH_2O
            - 20 ml lysis buffer (50 mM NaP_i, 300 mM NaCl, 10 mM imidazole, pH 8)
    - The cell lysate is loaded onto the column using the external pumping system
        - The column is assembled to the external pump
        - The soluble cell lysate fraction is loaded (2 ml/min)
    - Washing
        - The external pumping system is washed with 10 ml lysis buffer
    - Assemble the column into the ÄKTA purifier system
    - Washing (5 ml/min)
        - Unspecific bound protein are removed with 50 ml wash buffer (lysis buffer and elution buffer are mixed to give a final concentration of 20 mM imidazole)
    - Elution (5 ml/min)
        - The protein is eluted in 75 ml in total, and 1ml fractions are collected in a deep well plate
        - The imidazole concentration is increased to 500 mM within a slope of 15 CV
        - The fractions are pooled according to the imidazole concentration in Table 13

Table 13: Imidazole concentrations at the start and end of the pooled fractions

	Start pure fraction	End Peak/Fraction
CelA_cHis	114 mM I	240 mM I
CelA_nHis	116 mM I	240 mM I
CelAc_cHis	140 mM I	245 mM I
CelAc_nHis	265 mM I	377 mM I
Cex_cHis	222 mM I	375 mM I

- Desalting using dialysis
  - Zellutrans membrane (Roth E667)
  - The membrane is equilibrated in NaP_i or 0.9% NaCl for 10 min
  - The membrane is filled with the pooled enzyme fraction
  - The membrane is stirred for 1 h in 100 x buffer at 4°C and this step is repeated twice with fresh buffer (3 buffer exchanges in total)
- Concentration
  - Vivaspin 4 or Vivaspin 20 concentrator (Sartorius stedim biotech VS2001)
    - The desalted protein is load onto the membrane
    - And centrifuged at 4°C and 3,220 x g
- The proteins are stored at 4°C

- Measurements of the protein concentration
  - Nano Drop A280 nm

Table 14: Extinction coefficients

	Abs 0.1% (=1 g/l)
CelA_cHis	1.940
CelA_nHis	1.910
CelAc_cHis	2.295
CelAc_nHis	2.253
Cex_cHis	1.534

6.4 pHBAH assay protocol

6.4.1 Enzyme reaction

- Substrate solution: 1% or 1.75% CMC in Sodium phosphate buffer (50 mM NaP_i pH 7)
    - CMC is dissolved by autoclaving and, hot stirring to ensure a proper dissolution
- 600 µl substrate solution are pre-heated in a Thermomixer
- The enzyme reaction is started through addition of 2 µl enzyme dilution
- The assay is incubated for several minutes and stirred with 500 rpm
- Specific activity: 37°C, pH 7, 1,75% CMC
    - CelAc: 0.6 - 0.4 µg/ml
    - CelA: 0.9 - 0.6 µg/ml
- Each reaction is measured in triplicate

6.4.2 Detection reaction

- Stock-solution: 2.5 g pHBAH (Sigma H9882), 2 ml HCl (37%), 40 ml ddH_2O
    - Durable for approx. 1 month at 4°C
- Working solution: 4 ml pHBAH stock, 2.5ml 4 M NaOH => 20 ml ddH_2O
    - Final concentration: 0.078 mmol/ml pHBAH, 0.0966 mmol/ml HCl, 0.5 mmol/ml NaOH
- 150 µl WS (pre-chilled on ice in MTP, suitable for PCR Thermocycler) are mixed with 50 µl Enzyme reaction or glucose standard
    - The MTP is sealed properly with a cover
    - The detection reaction is carried out in a Thermocycler at 95°C for 5 min, followed by cooling to 4°C
- Measurement:
    - 150µl are transferred into a MTP suitable for the measurement in a Plate reader
    - Plate reader FLUOstar Omega, 430 nm
- Calibration: Glucose standards (0.03-1 mg/ml) in NaP_i

6.5 DNS assay protocol

6.5.1 Enzyme reaction

- Substrate solution:
  - 16-18 mg punched Whatman No.1 filter paper in Sodium phosphate buffer (50 mM NaP_i pH 7)
  - Or 1% Avicel in NaP_i
- 240 µl substrate + enzyme
- Incubation for several hours, at 37°C stirring with 1200 rpm
  - CelA variants: 5-12 µg/ml
  - Cex_cHis: 1-116 µg/ml
- Each reaction is measured in triplicate (soluble and total reducing sugar)

6.5.2 Detection reaction

- DNS reagent (storage at 4°C):
  - 1% DNS
  - 1% NaOH
  - 0.2% Phenol
  - 0.02% Na-Sulfit
  - 20% K-Na-Tartrate
- Calibration:
  - 120 µl reagent (pre-chilled on ice in MTP, suitable for PCR Thermocycler) are mixed with 60 µl glucose standard (0.03-1 mg/ml) in NaP_i
  - The MTP is sealed properly with a cover
  - The detection reaction is carried out in a Thermocycler at 95°C for 5 min, followed by cooling to 4°C
- Soluble reducing sugar:
  - 120µl of the enzyme reaction supernatant (13,200 g, 1 min) are mixed with 240 µl reagent
  - The detection reaction is carried out in a Thermomixer at 95°C for 10 min, followed by cooling on ice
- Total reducing sugar:
  - 240 µl enzyme reaction are mixed with 480 µl reagent
  - The detection reaction is carried out in a Thermomixer at 95°C for 10 min, followed by cooling on ice
- Measurement:
  - 150 µl are transferred into a MTP suitable for the measurement in a Plate reader
  - Plate reader FLUOstar Omega, 540 nm

6.6 Gene- and amino acid sequences

6.6.1 CelA

MKRVLLRGVLGRSKWSFGLAVVALQATVLIGCGDAGTDDGVGGAGGGGGATSASATTSSASTTSAGATTSS
ASTTSASATTSSASTTSSASTTSATTSASSTTSASTGGVPTDGTPVERHGRLRVMNGNIVGEHGSPVQLKGMS
LFWSQWSNYYNGNVVNSLADNWESTVVRAAMGIEGEDGYLQDAGAQKAKAKTIADAAIAKGIYVILDWH
DHNAHQHLDLAKSYFREVAQAYKNTPNVIFEVFNEPLNTNTWPAVKSYAEAVISEIRGQGANNLVIVGSPN
WSQDVDIAADNPLSDQNVAYTLHFYANTHKASLRDKAQKAINKKLALFVTEWGTCSADGNGQLNLGESQT
WLDFLDSHNISWANWSLGDKAEACSALRPNANQMGNWNDNDLTESGKWVKAKIAE

celA

TTCGCCTCACGGTCTCGCTGGAGACCACCGGCCGCAGTTGTCACCTGCGCGCACCGATACTGTATTTTGC
CATACAAGCGGCGTGGATTTCATCCCTTGGTCAGAGGCACGCGGAGGGCTTGCCATGGCGCTGAACTG
CCCTTCTAATGCCGCGCCATCCGGATACTTCTCTTGGAGGACATGATGAAAAGAGTTTTGCTTCGCGGTG
TGCTTGGCAGGTCGAAGTGGTCATTCGGCCTTGCAGTGGTTGCGTTGCAAGCCACCGTGCTGATCGGCT
GCGGTGACGCGGGGACGGACGACGGGGTGGGCGGCGCGGGCGGCGGCGGCGGGGCGACCTCCGCG
AGCGCGACGACCTCGAGCGCCTCGACGACCTCCGCGGGCGCGACGACCTCGAGCGCCTCGACGACCTC
CGCGAGCGCGACGACCTCGAGCGCCTCGACGACCTCGAGCGCCTCGACGACCTCGGCGACGACATCGG
CCTCGTCCACGACGAGCGCGAGCACCGGGGGGGTGCCCACGGACGGCACGCCGGTCGAGCGTCATGG
CCGGCTCCGGGTCATGAACGGCAACATCGTCGGCGAGCACGGGTCCCCGGTCCAGCTCAAGGGCATGA
GCCTGTTCTGGAGCCAGTGGAGCAACTACTACAACGGCAATGTCGTCAACTCGCTCGCGGACAACTGGG
AGTCGACCGTCGTCCGTGCCGCCATGGGCATCGAGGGCGAAGACGGGTACCTCCAGGACGCGGGCGCC
CAGAAGGCCAAGGCGAAGACCATCGCCGACGCGGCGATCGCGAAGGGCATCTACGTCATCCTCGACTG
GCACGATCACAACGCGCACCAGCACCTGGACCTGGCGAAGAGCTACTTCCGCGAGGTGGCGCAGGCCT
ACAAGAACACGCCCAACGTCATCTTCGAGGTCTTCAACGAGCCCCTCAACACGAACACCTGGCCCGCGG
TGAAGTCCTACGCCGAGGCGGTGATCTCCGAGATCCGCGGCCAGGGCGCGAACAACCTGGTCATCGTC
GGCTCGCCGAACTGGTCGCAGGACGTCGACATCGCCGCGGACAATCCCCTCAGCGATCAGAACGTGGC
GTACACCCTCCACTTCTACGCCAACACGCACAAGGCGTCGCTGCGCGACAAGGCCCAGAAAGCCATCAA
CAAGAAGCTGGCGCTGTTCGTGACCGAGTGGGGGACCTGCTCCGCCGACGGCAACGGCCAGCTCAACC
TCGGCGAGTCGCAGACCTGGCTCGACTTCCTGGACTCGCACAACATCAGCTGGGCCAACTGGTCGCTGG
GCGACAAGGCCGAGGCGTGCTCCGCGCTCAGGCCCAACGCAAACCAGATGGGCAACTGGAACGACAAC
GACCTCACCGAGTCGGGCAAGTGGGTGAAGGCGAAGATCGCAGAGTGATCCCCGTGCGGGTCACCGG
CCGCTCGCGCCGGTGACCCGGCCCCGACCGAGGGCGCCGCTCACCCGTGCGCTCGTTTGCTCTCTACGG
ACAGGGGCCGCCCGGCACAGAAGGGGCATTGATCGTCCGAGCACCAGTCGTCGTCTTCGAGGCCGAGC
GCCCGGAGCTGCTCCAGCTCCCACGCGCG

6.6.2 CelA_nHis

MGSSHHHHHHSSGLVPRGSHMAGTDDGVGGAGGGGGATSASATTSSASTTSAGATTSSASTTSASATTSS
ASTTSSASTTSATTSASSTTSASTGGVPTDGTPVERHGRLRVMNGNIVGEHGSPVQLKGMSLFWSQWSNYY
NGNVVNSLADNWESTVVRAAMGIEGEDGYLQDAGAQKAKAKTIADAAIAKGIYVILDWHDHNAHQHLDL
AKSYFREVAQAYKNTPNVIFEVFNEPLNTNTWPAVKSYAEAVISEIRGQGANNLVIVGSPNWSQDVDIAADN
PLSDQNVAYTLHFYANTHKASLRDKAQKAINKKLALFVTEWGTCSADGNGQLNLGESQTWLDFLDSHNISW
ANWSLGDKAEACSALRPNANQMGNWNDNDLTESGKWVKAKIAE

celAoΔS_nHis

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTGGT
ACTGATGATGGTGTTGGTGGTGCAGGTGGTGGTGGTGGTGCCACTAGTGCCTCAGCGACTACGAGTTC
AGCTAGTACAACCTCAGCTGGTGCAACTACGAGTTCAGCAAGTACAACCTCAGCTAGTGCTACTACGAG
TTCAGCGTCAACAACCAGTTCAGCTAGTACTACGTCAGCAACAACCAGTGCCAGTTCAACTACGAGTGCA
TCAACTGGTGGTGTTCCAACTGATGGTACGCCAGTTGAACGTCATGGTCGTTTACGGGTTATGAATGGT
AATATTGTTGGTGAACACGGTAGTCCAGTTCAATTGAAGGGCATGTCATTATTTTGGAGTCAATGGTCAA
ACTACTACAACGGTAACGTTGTTAACAGTTTGGCTGATAACTGGGAATCAACGGTTGTTCGGGCTGCAA
TGGGTATTGAAGGTGAAGATGGTTATTTACAAGATGCTGGTGCACAAAAAGCCAAGGCGAAAACTATT
GCGGATGCCGCGATTGCTAAGGGTATTTACGTTATTTTGGATTGGCATGATCACAATGCACATCAACACT
TAGATTTGGCCAAGAGTTATTTTCGTGAAGTTGCTCAAGCATACAAGAACACTCCAAACGTTATTTTTGA
AGTTTTTAATGAACCATTGAATACAAATACCTGGCCAGCCGTTAAAAGTTATGCCGAAGCGGTTATTTCA
GAAATTCGGGGTCAAGGTGCTAACAACTTGGTTATTGTTGGTAGTCCAAATTGGTCACAAGATGTTGAT
ATTGCTGCTGATAATCCATTGAGTGATCAAAACGTTGCATACACTTTGCATTTTTATGCAAATACGCACAA
GGCCTCATTGCGTGATAAGGCACAAAAGGCCATTAACAAGAAGTTGGCGTTGTTTGTTACAGAATGGGG
TACCTGTAGTGCCGATGGTAATGGTCAATTAAATTTGGGTGAATCACAAACTTGGTTGGATTTCTTAGAT
AGTCATAACATTTCATGGGCTAACTGGAGTTTGGGTGATAAGGCTGAAGCATGTTCAGCGTTACGGCCA
AATGCTAATCAAATGGGTAACTGGAACGATAACGATTTGACGGAAAGTGGCAAGTGGGTTAAGGCGAA
AATTGCTGAATAA

6.6.3 CelAc_nHis

MGSSHHHHHHSSGLVPRGSHMAGTDGTPVERHGRLRVMNGNIVGEHGSPVQLKGMSLFWSQWSNYYN
GNVVNSLADNWESTVVRAAMGIEGEDGYLQDAGAQKAKAKTIADAAIAKGIYVILDWHDHNAHQHLDLA
KSYFREVAQAYKNTPNVIFEVFNEPLNTNTWPAVKSYAEAVISEIRGQGANNLVIVGSPNWSQDVDIAADNP
LSDQNVAYTLHFYANTHKASLRDKAQKAINKKLALFVTEWGTCSADGNGQLNLGESQTWLDFLDSHNISWA
NWSLGDKAEACSALRPNANQMGNWNDNDLTESGKWVKAKIAE

celAocΔS_nHis

ATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTGGT
ACTGATGGTACGCCAGTTGAACGTCATGGTCGTTTACGGGTTATGAATGGTAATATTGTTGGTGAACAC
GGTAGTCCAGTTCAATTGAAGGGCATGTCATTATTTTGGAGTCAATGGTCAAACTACTACAACGGTAAC
GTTGTTAACAGTTTGGCTGATAACTGGGAATCAACGGTTGTTCGGGCTGCAATGGGTATTGAAGGTGAA
GATGGTTATTTACAAGATGCTGGTGCACAAAAAGCCAAGGCGAAAACTATTGCGGATGCCGCGATTGCT
AAGGGTATTTACGTTATTTTGGATTGGCATGATCACAATGCACATCAACACTTAGATTTGGCCAAGAGTT
ATTTTCGTGAAGTTGCTCAAGCATACAAGAACACTCCAAACGTTATTTTTGAAGTTTTTAATGAACCATTG
AATACAAATACCTGGCCAGCCGTTAAAAGTTATGCCGAAGCGGTTATTTCAGAAATTCGGGGTCAAGGT
GCTAACAACTTGGTTATTGTTGGTAGTCCAAATTGGTCACAAGATGTTGATATTGCTGCTGATAATCCAT
TGAGTGATCAAAACGTTGCATACACTTTGCATTTTTATGCAAATACGCACAAGGCCTCATTGCGTGATAA
GGCACAAAAGGCCATTAACAAGAAGTTGGCGTTGTTTGTTACAGAATGGGGTACCTGTAGTGCCGATG
GTAATGGTCAATTAAATTTGGGTGAATCACAAACTTGGTTGGATTTCTTAGATAGTCATAACATTTCATG
GGCTAACTGGAGTTTGGGTGATAAGGCTGAAGCATGTTCAGCGTTACGGCCAAATGCTAATCAAATGG
GTAACTGGAACGATAACGATTTGACGGAAAGTGGCAAGTGGGTTAAGGCGAAAATTGCTGAATAA

6.6.4 CelA_cHis

MAGTDDGVGGAGGGGGATSASATTSSASTTSAGATTSSASTTSASATTSSASTTSSASTTSATTSASSTTSAS
TGGVPTDGTPVERHGRLRVMNGNIVGEHGSPVQLKGMSLFWSQWSNYYNGNVVNSLADNWESTVVRA
AMGIEGEDGYLQDAGAQKAKAKTIADAAIAKGIYVILDWHDHNAHQHLDLAKSYFREVAQAYKNTPNVIFE
VFNEPLNTNTWPAVKSYAEAVISEIRGQGANNLVIVGSPNWSQDVDIAADNPLSDQNVAYTLHFYANTHKA
SLRDKAQKAINKKLALFVTEWGTCSADGNGQLNLGESQTWLDFLDSHNISWANWSLGDKAEACSALRPNA
NQMGNWNDNDLTESGKWVKAKIAEKLAAALEHHHHHH

celAoΔS_cHis

ATGGCTGGTACTGATGATGGTGTTGGTGGTGCAGGTGGTGGTGGTGGTGCCACTAGTGCCTCAGCGAC
TACGAGTTCAGCTAGTACAACCTCAGCTGGTGCAACTACGAGTTCAGCAAGTACAACCTCAGCTAGTGCT
ACTACGAGTTCAGCGTCAACAACCAGTTCAGCTAGTACTACGTCAGCAACAACCAGTGCCAGTTCAACTA
CGAGTGCATCAACTGGTGGTGTTCCAACTGATGGTACGCCAGTTGAACGTCATGGTCGTTTACGGGTTA
TGAATGGTAATATTGTTGGTGAACACGGTAGTCCAGTTCAATTGAAGGGCATGTCATTATTTTGGAGTCA
ATGGTCAAACTACTACAACGGTAACGTTGTTAACAGTTTGGCTGATAACTGGGAATCAACGGTTGTTCG
GGCTGCAATGGGTATTGAAGGTGAAGATGGTTATTTACAAGATGCTGGTGCACAAAAAGCCAAGGCGA
AAACTATTGCGGATGCCGCGATTGCTAAGGGTATTTACGTTATTTTGGATTGGCATGATCACAATGCACA
TCAACACTTAGATTTGGCCAAGAGTTATTTTCGTGAAGTTGCTCAAGCATACAAGAACACTCCAAACGTT
ATTTTTGAAGTTTTTAATGAACCATTGAATACAAATACCTGGCCAGCCGTTAAAAGTTATGCCGAAGCGG
TTATTTCAGAAATTCGGGGTCAAGGTGCTAACAACTTGGTTATTGTTGGTAGTCCAAATTGGTCACAAGA
TGTTGATATTGCTGCTGATAATCCATTGAGTGATCAAAACGTTGCATACACTTTGCATTTTTATGCAAATA
CGCACAAGGCCTCATTGCGTGATAAGGCACAAAAGGCCATTAACAAGAAGTTGGCGTTGTTTGTTACAG
AATGGGGTACCTGTAGTGCCGATGGTAATGGTCAATTAAATTTGGGTGAATCACAAACTTGGTTGGATT
TCTTAGATAGTCATAACATTTCATGGGCTAACTGGAGTTTGGGTGATAAGGCTGAAGCATGTTCAGCGTT
ACGGCCAAATGCTAATCAAATGGGTAACTGGAACGATAACGATTTGACGGAAAGTGGCAAGTGGGTTA
AGGCGAAAATTGCTGAAAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCACTGA

6.6.5 CelAc_cHis

MAGTDGTPVERHGRLRVMNGNIVGEHGSPVQLKGMSLFWSQWSNYYNGNVVNSLADNWESTVVRAA
MGIEGEDGYLQDAGAQKAKAKTIADAAIAKGIYVILDWHDHNAHQHLDLAKSYFREVAQAYKNTPNVIFEV
FNEPLNTNTWPAVKSYAEAVISEIRGQGANNLVIVGSPNWSQDVDIAADNPLSDQNVAYTLHFYANTHKAS
LRDKAQKAINKKLALFVTEWGTCSADGNGQLNLGESQTWLDFLDSHNISWANWSLGDKAEACSALRPNAN
QMGNWNDNDLTESGKWVKAKIAEKLAAALEHHHHHH#

celAocΔS_cHis

ATGGCTGGTACTGATGGTACGCCAGTTGAACGTCATGGTCGTTTACGGGTTATGAATGGTAATATTGTT
GGTGAACACGGTAGTCCAGTTCAATTGAAGGGCATGTCATTATTTTGGAGTCAATGGTCAAACTACTAC
AACGGTAACGTTGTTAACAGTTTGGCTGATAACTGGGAATCAACGGTTGTTCGGGCTGCAATGGGTATT
GAAGGTGAAGATGGTTATTTACAAGATGCTGGTGCACAAAAAGCCAAGGCGAAAACTATTGCGGATGC
CGCGATTGCTAAGGGTATTTACGTTATTTTGGATTGGCATGATCACAATGCACATCAACACTTAGATTTG
GCCAAGAGTTATTTTCGTGAAGTTGCTCAAGCATACAAGAACACTCCAAACGTTATTTTTGAAGTTTTTAA
TGAACCATTGAATACAAATACCTGGCCAGCCGTTAAAAGTTATGCCGAAGCGGTTATTTCAGAAATTCG
GGGTCAAGGTGCTAACAACTTGGTTATTGTTGGTAGTCCAAATTGGTCACAAGATGTTGATATTGCTGCT
GATAATCCATTGAGTGATCAAAACGTTGCATACACTTTGCATTTTTATGCAAATACGCACAAGGCCTCATT
GCGTGATAAGGCACAAAAGGCCATTAACAAGAAGTTGGCGTTGTTTGTTACAGAATGGGGTACCTGTA
GTGCCGATGGTAATGGTCAATTAAATTTGGGTGAATCACAAACTTGGTTGGATTTCTTAGATAGTCATAA
CATTTCATGGGCTAACTGGAGTTTGGGTGATAAGGCTGAAGCATGTTCAGCGTTACGGCCAAATGCTAA
TCAAATGGGTAACTGGAACGATAACGATTTGACGGAAAGTGGCAAGTGGGTTAAGGCGAAAATTGCTG
AAAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCACTGA

6.6.6 Cex_cHis

MPRTTPAPGHPARGARTALRTTLAAAAATLVVGATVVLPAQAATTLKEAADGAGRDFGFALDPNRLSEAQY
KAIADSEFNLVVAENAMKWDATEPSQNSFSFGAGDRVASYAADTGKELYGHTLVWHSQLPDWAKNLNGS
AFESAMVNHVTKVADHFEGKVASWDVVNEAFADGGGRRQDSAFQQKLGNGYIETAFRAARAADPTAKLCI
NDYNVEGINAKSNSLYDLVKDFKARGVPLDCVGFQSHLIVGQVPGDFRQNLQRFADLGVDVRITELDIRMRT
PSDATKLATQAADYKKVVQACMQVTRCQGVTVWGITDKYSWVPDVFPGEGAALVWDASYAKKPAYAAV
MEAFGASPTPTPTTPTPTPTTPTPTPTSGPAGCQVLWGVNQWNTGFTANVTVKNTSSAPVDGWTLTFSFP
SGQQVTQAWSSTVTQSGSAVTVRNAPWNGSIPAGGTAQFGFNGSHTGTNAAPTAFSLNGTPCTVGKLAA
ALEHHHHHH

cex_cHis

ATGCCTAGGACCACGCCCGCACCCGGCCACCCGGCCCGCGGCGCCCGCACCGCTCTGCGCACGACGCTC
GCCGCCGCGGCGGCGACGCTCGTCGTCGGCGCCACGGTCGTGCTGCCCGCCCAGGCCGCGACCACGCT
CAAGGAGGCCGCCGACGGCGCCGGCCGGGACTTCGGCTTCGCGCTCGACCCCAACCGGCTCTCGGAGG
CGCAGTACAAGGCGATCGCCGACAGCGAGTTCAACCTCGTCGTCGCCGAGAACGCGATGAAGTGGGAC
GCCACCGAGCCCTCGCAGAACAGCTTCTCCTTCGGCGCGGGCGACCGCGTCGCGAGCTACGCCGCCGAC
ACCGGCAAGGAGCTGTACGGCCACACGCTCGTCTGGCACTCGCAGCTGCCCGACTGGGCGAAGAACCT
CAACGGCTCCGCGTTCGAGAGCGCGATGGTCAACCACGTGACGAAGGTCGCCGACCACTTCGAGGGCA
AGGTCGCGTCGTGGGACGTCGTCAACGAGGCGTTCGCCGACGGCGGCGGCCGCCGGCAGGACTCGGC
GTTCCAGCAGAAGCTCGGCAACGGCTACATCGAGACCGCGTTCCGGGCGGCACGTGCGGCGGACCCGA
CCGCCAAGCTGTGCATCAACGACTACAACGTCGAGGGCATCAACGCGAAGAGCAACTCGCTCTACGACC
TCGTCAAGGACTTCAAGGCGCGCGGCGTCCCGCTCGACTGCGTCGGGTTCCAGTCGCACCTCATCGTCG
GCCAGGTGCCGGGCGACTTCCGGCAGAACCTGCAGCGGTTCGCGGACCTGGGCGTGGACGTGCGCATC
ACCGAGCTCGACATCCGCATGCGGACGCCCTCCGACGCGACCAAGCTCGCGACCCAGGCGGCCGACTA
CAAGAAGGTCGTGCAGGCCTGCATGCAGGTGACCCGCTGCCAGGGCGTGACCGTCTGGGGCATCACCG
ACAAGTACTCGTGGGTGCCGGACGTCTTCCCGGGCGAGGGGGCCGCGCTGGTGTGGGACGCGAGCTA
CGCCAAGAAGCCGGCCTACGCCGCCGTGATGGAGGCCTTCGGCGCGAGCCCGACGCCGACGCCCACCA
CGCCGACCCCGACGCCCACGACGCCGACGCCGACCCCGACGTCCGGTCCGGCCGGGTGCCAGGTGCTG
TGGGGCGTCAACCAGTGGAACACCGGCTTCACCGCGAACGTCACCGTGAAGAACACGTCCTCCGCTCCG
GTCGACGGCTGGACGCTCACGTTCAGCTTCCCGTCCGGCCAGCAGGTCACCCAGGCGTGGAGCTCGAC
GGTCACGCAGTCCGGCTCGGCCGTGACGGTCCGCAACGCCCCGTGGAACGGCTCGATCCCGGCGGGCG

GCACCGCGCAGTTCGGCTTCAACGGCTCGCACACGGGCACCAACGCCGCGCCGACGGCGTTCTCGCTCA
ACGGCACGCCCTGCACGGTCGGCAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCACTGA

cex2.0_cHis

ATGGCCACCACCCTGAAAGAAGCGGCAGATGGTGCTGGACGTGATTTTGGGTTCGCCTTAGACCCGAAC
CGACTAAGCGAAGCTCAGTACAAGGCGATTGCAGACTCTGAATTCAATTTAGTTGTTGCGGAAAACGCT
ATGAAGTGGGATGCTACTGAACCTTCTCAAAATAGTTTCTCGTTTGGCGCTGGCGATCGTGTTGCGAGTT
ATGCTGCGGATACTGGTAAAGAGTTGTATGGTCATACGTTGGTGTGGCACAGTCAATTGCCGGACTGGG
CGAAAAACCTAAACGGTTCTGCCTTTGAGTCGGCGATGGTCAATCACGTAACCAAAGTTGCTGACCATTT
TGAAGGTAAAGTAGCATCATGGGATGTGGTCAATGAGGCGTTCGCTGATGGCGGTGGGCGACGGCAG
GATAGTGCATTTCAACAGAAGTTGGGTAACGGCTACATTGAAACGGCATTTCGCGCAGCGCGTGCCGCT
GATCCTACCGCCAAGCTATGTATCAACGACTATAACGTGGAAGGGATTAATGCTAAATCGAATAGTTTGT
ATGATTTAGTCAAAGACTTTAAAGCGCGTGGTGTTCCGCTGGATTGTGTGGGTTTTCAAAGCCACTTAAT
TGTTGGTCAGGTTCCGGGTGATTTTCGTCAGAATTTACAACGGTTCGCCGATTTGGGTGTGGATGTCCG
GATTACCGAATTGGATATCCGCATGCGCACGCCAAGTGACGCAACAAAGCTGGCCACTCAAGCCGCTGA
CTACAAAAAGGTTGTCCAAGCATGTATGCAAGTCACGCGGTGTCAAGGGGTTACAGTCTGGGGAATCAC
AGATAAGTATTCATGGGTGCCAGATGTGTTCCCAGGCGAAGGTGCAGCCTTGGTGTGGGATGCGTCGT
ACGCCAAGAAACCAGCCTATGCCGCGGTCATGGAAGCATTTGGTGCTTCTCCGACCCCGACACCAACGA
CACCCACCCCAACTCCGACTACGCCGACTCCAACGCCAACCTCGGGACCTGCTGGCTGCCAGGTACTGTG
GGGTGTTAATCAATGGAATACGGGATTTACTGCGAACGTTACAGTAAAGAATACGAGTAGCGCGCCTGT
TGACGGCTGGACGTTAACTTTCTCATTTCCAAGTGGGCAACAAGTTACTCAGGCGTGGTCAAGTACTGTC
ACCCAGTCAGGATCAGCTGTCACGGTGCGTAATGCCCCATGGAATGGAAGTATTCCAGCCGGTGGGAC
GGCACAATTTGGGTTTAATGGGAGCCATACAGGTACGAATGCAGCCCCCACGGCATTCAGCTTAAACGG
AACACCCTGCACCGTCGGCAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCACTGA

6.7 Overview of the culture collection

Table 15: Assignment of the used *E. coli* cells to the IMBT culture collection number

IMBT culture collection	Culture	Plasmid
7174	*E. coli* BL21	pMS470_celAoΔS_cHis
7175	*E. coli* BL21	pMS470_celAoΔS_nHis
7176	*E. coli* BL21	pMS470_celAoCΔS_cHis
7177	*E. coli* BL21	pMS470_celAoCΔS_nHis
7178	*E. coli* BL21	pMS470_cex_cHis
7179	*E. coli* BL21	pMS470_cex2.0_cHis