



Faculteit Bio-ingenieurswetenschappen

Academiejaar 2015 – 2016

Expression and study of the stability of a bacterial
lytic polysaccharide monooxygenase

Confidential - Confidentieel

Matthias Last

Promotor: Prof. dr. Tom Desmet

Tutor: ir. Magali Tanghe

Masterproef voorgedragen tot het behalen van de graad van
Master of Science in de industriële wetenschappen: biochemie



Faculteit Bio-ingenieurswetenschappen

Academiejaar 2015 – 2016

Expression and study of the stability of a bacterial
lytic polysaccharide monooxygenase

Confidential - Confidentieel

Matthias Last

Promotor: Prof. dr. Tom Desmet

Tutor: ir. Magali Tanghe

Masterproef voorgedragen tot het behalen van de graad van
Master of Science in de industriële wetenschappen: biochemie

Acknowledgements

In de loop van het gehele jaar, heb ik veel geleerd over teamwork, tegenslagen aanvaarden en hard werken. Daarom zou ik heel graag iedereen bedanken voor de steun en de aangename sfeer.

Allereerst zou ik graag mijn promotor Prof. Dr. Tom Desmet bedanken om ervoor te zorgen dat ik de kans gekregen heb om ervaring op te doen en mijn thesis te schrijven bij InBio.

Vervolgens zou ik ook heel graag mijn tutor Ir. Magali Tanghe heel hard bedanken. Met dank aan haar hulp, inzet, kritische visie en aanmoedigingen was ik in staat een mooi en volledig resultaat af te geven.

Verder wens ik ook de overige leden van Inbio en vooral Glycodirect te bedanken voor de goede sfeer en steun. Bovendien zou ik graag mijn mede thesisstudenten bedanken. In de eerste plaats Marijn De Munter. Gezien het gelijkaardige onderzoek hebben we elkaar kunnen steunen en soms samen kunnen werken. Net als hem, zou ik dan ook de andere medestudenten: Laurens de Brauwer, Jeff Moens, Ellen Vanderherten en Iris Verhasselt willen bedanken voor hun bijdrage aan de toffe sfeer in het labo.

Ten slotte zou ik graag mijn familie heel hard bedanken, zonder wie ik hier nooit zou staan. Mama, papa en zusje, echt merci voor deze kans en de nooit aflatende steun.

Abstract

In the evolution of the fossil-based economy into a renewable one, the application of cellulose as renewable resource is hindered by its crystalline structure and high recalcitrance towards conventional enzymatic breakdown. Recently, Lytic Polysaccharide MonoOxygenases (LPMOs) were discovered and considered the key element for an efficient enzymatic cellulose breakdown. Oxidative cleavage of the glucose chain disrupts the crystalline structure, thus boosting the activity of the classical cellulases. For a successful application however, the LPMOs require an increased stability to meet the industrial conditions. So, this thesis aims to increase the thermal stability of a specific LPMO, namely *Streptomyces coelicolor* LPMO10C.

First, an expression system was optimized to ensure the required N-terminal end and sufficient enzyme yield and purity. The activity was evaluated on phosphoric acid swollen cellulose (PASC) via HPAEC-PAD analysis and the melting temperatures were obtained via differential scanning fluorimetry (DSF). Next, flexible regions, determined with B-fitter, were targeted in prospect of a higher stability. One set involved the introduction of salt bridges, ionic networks and polar residues at the protein's surface, while another set was based on the introduction of disulphide bonds.

As a result, none of the mutations changed the activity drastically. Furthermore, introducing extra disulphide bonds resulted in 4 mutants with increased melting temperature (+ 3 °C, + 3 °C, + 5 °C, + 9° C) Therefrom one final mutant combination was created and resulted in an increase from 51 °C for the wild type to 63 °C (+ 12 °C) for the mutant.

Key words: Lytic polysaccharide monooxygenase, ScLPMO10C, rational engineering, thermostability, disulphide bonds

Samenvatting

In de evolutie naar een duurzame economie, wordt de toepassing van cellulose voor de bio-ethanol productie bemoeilijkt door zijn kristalliene structuur en zijn enorme weerstand tegenover conventionele enzymatische afbraak. Sinds de recente ontdekking van Lytische Polysaccharide MonoOxygenases (LPMO's) worden ze beschouwd als de ontbrekende factor voor een efficiënte enzymatische cellulose afbraak. Door oxidatieve splitsing van de glucoseketens verbreken ze de kristalliene structuur, waardoor ze de klassieke cellulase activiteit stimuleren. De succesvolle toepassing van LPMOs vereist echter voldoende stabiliteit. Daarom is het doel van deze thesis om de thermische stabiliteit te verhogen van een specifiek LPMO, namelijk *Sterptomyces coelicolor* LPMO10C.

Als eerste werd een expressiesysteem op punt gesteld waarbij de correcte amino-terminus van het eiwit en voldoende eiwitopbrengst en zuiverheid verzekerd werden. De activiteit werd getest op het cellulosesubstraat *Phosphoric Acid Swollen cellulose* (PASC) en geanalyseerd via HPAEC-PAD, terwijl de smeltemperatuur gemeten werd via *differential scanning fluorimetry* (DSF). Vervolgens werden flexibele regio's bepaald met B-fitter en werden ze gebruikt als doelwit voor mutagenese. Eén set betreft de introductie van zout bruggen, ionische netwerken en meer polaire residuen op het eiwitoppervlak, terwijl een andere set gebaseerd was op de introductie van disulfidebruggen.

Als resultaat had geen van de verkregen mutanten een drastische verandering in activiteit. Maar 4 mutanten met extra disulfidebruggen resulteerden in een verhoogde smeltemperatuur (+ 3 °C, + 3 °C, + 5 °C, + 9° C). Hiervan werd 1 finale combinatie gemaakt, die resulteerde in een verhoging van 51°C voor het wild type naar 63 °C voor de mutant.

Sleutelwoorden: Lytisch polysaccharide monooxygenase, ScLPMO10C, rational engineering, thermostabiliteit, disulfidebruggen

List of figures

FIGURE 1: SCHEMATIC REPRESENTATION OF THE COMPOSITION OF LIGNOCELLULOSIC BIOMASS. (ADAPTED FROM (RATANAKHANOKCHAI ET AL., 2013))	3
FIGURE 2: SCHEMATIC REPRESENTATION OF THE STEPWISE ENZYMATIC CATABOLISM OF CELLULOSE BY THE THREE TYPES OF CELLULASES (PROMON, PROGRAM, & SCIENCES, 2015).	6
FIGURE 3: A) REPRESENTATION OF THE CLEFT/GROOVE CONFORMATION IN ENDOGLUCANASE (THIS EXAMPLE IS FROM E2 FROM <i>THERMOBIFIDA FUSCA</i>). B) REPRESENTATION OF THE TUNNEL CONFORMATION IN CELLOBIOHYDROLASE II (THIS EXAMPLE IS FROM <i>TRICHODERMA REESEI</i>). THE PROPOSED CATALYTIC RESIDUES ARE SHADED IN RED. SOURCE:(DIES ET AL., 1995)	6
FIGURE 4: SCHEMATIC REPRESENTATION OF THE DIFFERENT CELLULOSE ACTIVE ENZYMES: CELLOBIOHYDROLASE (CBH) REPRESENTED IN LIGHT BLUE, ENDOGLUCANASE REPRESENTED IN PURPLE AND THE LPMO REPRESENTED IN PINK SOURCE:(FUSHINOBU, 2013)	8
FIGURE 5: VARIETY IN SUBSTRATE SPECIFICITY AND ORIGIN OF THE DIFFERENT LPMO FAMILIES (SPAN & MARLETTA, 2015)	9
FIGURE 6: STRUCTURAL PRESENTATION OF THE DIFFERENT AUXILIARY ACTIVITY FAMILIES AND THEIR RESPECTIVE COPPER BINDING IN THE ACTIVE SITE. A) THE REPRESENTATION OF THE AA9 FAMILY, B) REPRESENTATION OF THE AA10 FAMILY AND C) REPRESENTATION OF THE AA11 FAMILY. ADAPTED FROM:(HEMSWORTH ET AL., 2015)	11
FIGURE 7: THE OXIDATIVE CLEAVAGE REACTION OF LPMOS ON CELLULOSE(HEMSWORTH ET AL., 2015).	12
FIGURE 8: SCHEMATIC REPRESENTATION OF THE GROUPING OF THE DIFFERENT INACTIVATING INTERACTIONS(IYER & ANANTHANARAYAN, 2008)	15
FIGURE 9: (A) THE DETERMINATION OF THE MELTING TEMPERATURE THROUGH THE MELTING CURVE. (B) THE DETERMINATION OF THE TM AND CHANGE IN ENTHALPY THROUGH DIFFERENTIAL SCANNING CALORIMETRY. SOURCE: (POLIZZI ET AL., 2007)	16
FIGURE 10: SCHEMATIC REPRESENTATION OF THE EFFECT OF SPECIFIC MULTIPOINT IMMOBILIZATION ON ENZYME STABILITY (MATEO ET AL., 2007)	18
FIGURE 11: THE SELECTION PLAN OF THE DIFFERENT APPROACHES IN ENZYME ENGINEERING (CHICA, DOUCET, & PELLETIER, 2005)	20
FIGURE 12: THE DIFFERENT STEPS IN THE SANCHIS PCR PRINCIPLE FROM (SANCHIS ET AL., 2008). STARTING WITH THE CREATION OF A MEGA-PRIMER IN THE FIRST STAGE, IT IS FOLLOWED BY THE AMPLIFICATION OF THE ENTIRE PLASMID IN THE SECOND. DEPENDING ON THE POSITION OF THE ANNEALING SITES OF THE STARTING PRIMERS, THE LENGTH OF THE MEGA-PRIMER CAN BE VARIED AS DEPICTED IN THE LEFT COLUMN. BY STARTING WITH MUTATIONS IN BOTH THE FORWARD AND REVERSE PRIMERS, TWO MUTATIONS CAN BE INTRODUCED AT ONCE, AS DEPICTED IN THE RIGHT COLUMN.	24
FIGURE 13: SCHEMATIC REPRESENTATION OF THE PRINCIPLE OF THE HIS TAG BINDING TO THE NICKEL CHELATED NTA-COLUMN(ROWINSKA-ZYREK, DANUTA, & SLAWOMIR, 2013)	31
FIGURE 14: SCHEMATIC REPRESENTATION OF THE PRINCIPLE OF DIFFERENTIAL SCANNING FLUORIMETRY (RODRIGUES, PROSINECKI, MARRUCHO, & REBELO, 2011)	35
FIGURE 15: SCHEMATIC REPRESENTATION OF THE CONSIDERED EXPRESSION SYSTEMS. (A) SCLPMO10C-XA: THE CYTOPLASMIC EXPRESSION OF THE ENZYME, LINKED WITH A XA-FACTOR AT THE N-TERMINAL END. (B)	

SCLPMO10C-SECR THE PERIPLASMIC EXPRESSION OF AN ENZYME WITH A GUARANTEED CORRECT N-TERMINUS THROUGH THE NATURAL CLEAVAGE OF THE SECRETION SIGNAL AT THE N-TERMINAL END UPON SECRETION TO THE PERIPLASM. 37

FIGURE 16: SDS-PAGE ANALYSIS OF THE PURIFICATION FROM THE CYTOPLASMIC AND PERIPLASMIC FRACTION OF LPMO10C-SECR, COMPARING THE PROTEINS COLLECTED DURING THE DIFFERENT PURIFICATION STEPS . (INSOL = INSOLUBLE FRACTION , FT= FLOW-THROUGH, W= WASH STEPS (1 AND 3), EL= ELUTION STEP, EL+CONC= ELUTED AND CONCENTRATED ENZYME AFTER BUFFER EXCHANGE) 39

FIGURE 17: SDS-PAGE ANALYSIS OF THE PURIFICATION FROM THE CYTOPLASMIC FRACTION OF SCLPMO10C-XA FRACTION, COMPARING THE PROTEINS COLLECTED DURING THE DIFFERENT PURIFICATION STEPS . (INSOL = INSOLUBLE FRACTION , FT= FLOW-THROUGH, W= WASH STEPS (1 AND 3), EL= ELUTION STEP, EL+CONC= ELUTED AND CONCENTRATED ENZYME AFTER BUFFER EXCHANGE) 39

FIGURE 18: THE HPAEC-PAD ANALYSIS OF THE XA-CLEAVED ENZYME COMPARED TO THE UNCLEAVED FRACTION OF THE SAME SCLPMO10C-XA (BOTH AFTER 20H INCUBATION WITH PASC). THE LABELS INDICATE THE AMOUNT OF GLUCOSE LINKED SO G2 REPRESENTS CELLOBIOSE AND SO ON. THE ASTERISK (*) INDICATES THE ALDONIC ACID FORM. 40

FIGURE 19: AN OVERLAY OF THE RESULTS OF THE DIFFERENT FRACTIONS OF BOTH STRAINS, WHERE THE UNCLEAVED SCLPMO10C-XA IS REPRESENTED BY THE BROWN, THE CLEAVED SCLPMO10C-XA BY THE PINK, THE PERIPLASM FRACTION OF THE SCLPMO10C-SECR BY THE BLUE AND THE CYTOPLASMIC FRACTION OF THE SCLPMO10C-SECR BY THE BLACK CHROMATOGRAM. 41

FIGURE 20: SDS-PAGE ANALYSIS COMPAIRING THE PROTEIN COLLECTED DURING THE PURIFICATION STEPS: THE FLOW-THROUGH, AND THE DIFFERENT FRACTIONS FROM ELUTION BUFFERS CONTAINING IMIDAZOL CONCENTRATIONS RANGING FROM 20 TO 100 MILLIMOLAIR. 42

FIGURE 21: SCHEMATIC REPRESENTATION OF THE PRINCIPLE BEHIND THE H₂O₂ ASSAY. THE REDUCTION OF THE TYPE-2 COPPER ION OF THE LPMO LEADS TO THE ONE ELECTRON REDUCTION OF THE OXYGEN. THE PRESENCE OF THE RESULTING HYDROGEN PEROXIDE, STARTS THE HRP COUPLED REACTION CONVERTING AMPLEX RED INTO THE FLUORESCENT RESORUFIN. 44

FIGURE 22: DSF ANALYSIS OF THE DIFFERENTLY PURIFIED SAMPLES OF LPMO10C. THE DIFFERENT COLORS INDICATE A DIFFERENCE IN CONCENTRATION IMIDAZOLE IN THE WASHING BUFFER: 40 MM (BLUE CURVE), 20 MM (GREEN CURVE AND RED CURVE). THE SAMPLES OF THE RED CURVE WERE DILUTED TO THE SAME ENZYME CONCENTRATIONS AS THE BLUE CURVES. THE LEFT SIDE IS A REPRESENTATION OF THE MELTING CURVE. THE RIGHT SIDE IS A VISUALISATION OF THE DERIVATIVE. 46

FIGURE 23: DSF ANALYSIS OF THE LPMO10C-XA SAMPLE BEFORE THE CLEAVAGE (NO HIS-1). THE LEFT SITE OF THE FIGURE IS THE VISUALIZATION OF THE MELTING CURVE, AS THE RIGHT SIDE IS THE DERIVATIVE OF THE MELTING CURVE. 47

FIGURE 24: REPRESENTATION OF THE B-FACTORS IN FUNCTION OF THE LOCATION OF THE AMINO ACID. THE 20 AMINO ACIDS WITH THE HIGHEST B-FACTOR ARE REPRESENTED IN YELLOW, THE AMINO ACIDS AT A DISTANCE SHORTER THAN 6 ANGSTROM ARE REPRESENTED IN BLUE AND THE OTHER AMINO ACIDS IN RED. 48

FIGURE 25: SCHEMATICAL REPRESENTATION OF SCLPMO10C, WHERE THE YELLOW LOOPS REPRESENT THE TWO MOST FLEXIBLE REGIONS AND THE BLEU PART ARE THE AMINO-ACIDS CLOSE TO THE COPPER-ION (<6Å). 48

FIGURE 26: REPRESENTATION OF THE ELECTRICALLY CHARGED AMINO ACIDS 49

FIGURE 27: REPRESENTATION OF THE AMINO RESIDUES INVOLVED IN ONE OF THE CONSIDERED MUTATIONS 50

FIGURE 28: SCHEMATIC REPRESENTATION OF THE SELECTED MUTATIONS. A) A FIRST SELECTION WAS MADE BASED ON THE PREDICTION BY DBD AND AN A GRADE SUCCESS RATE BY MODIP. B) A SECOND SET OF DISULPHIDE BONDS WAS SELECTED FOR AMINO ACIDS AMONG THE 20 FLEXIBLE POSITIONS THAT WERE PROPOSED BY DBD AND THE B GRADE LIST OF MODIP. 52

FIGURE 29: SCHEMATIC REPRESENTATION OF THE LOCATIONS OF THE 3 STABILIZING MUTATIONS TO BE COMBINED. THE YELLOW SPHERES REPRESENT THE A143C-P183C MUTATION, THE BLUE SPHERES REPRESENT THE A52C-P61C MUTATION AND THE PINK SPHERES REPRESENT THE S73C-A115C MUTATION. 53

List of tables

TABLE 1: COMPARISON OF APPROACHES FOR ENGINEERING IMPROVED ENZYME STABILITY(CHICA ET AL., 2005)	20
TABLE 2: PROTOCOL FOR THE MEGAPRIMER CREATION FOR A 2-STEP SANCHIS PCR	25
TABLE 3: PROTOCOL FOR THE PLASMID AMPLIFICATION IN THE SECOND STEP OF THE SANCHIS PROTOCOL	25
TABLE 4: OVERVIEW OF THE CONSIDERED RATIONAL MUTATIONS AND THEIR RESPECTIVE PRIMERS (WITH THE MUTATION UNDERLINED IN THE SEQUENCE)	26
TABLE 5: OVERVIEW OF THE CONSIDERED DISULPHIDE BOND INSERTING MUTATIONS AND THEIR RESPECTIVE PRIMERS (WITH THE MUTATION UNDERLINED IN THE SEQUENCE)	26
TABLE 6: PCR PROGRAM FOR THE COLONY PCR	28
TABLE 7: OVERVIEW OF THE MELTING TEMPERATURE OF SAMPLES OF DIFFERENT PURITY AND CONCENTRATION.	46
TABLE 8: OVERVIEW OF THE CREATED RATIONAL MUTANTS. THEIR MUTATION(S), MELTING TEMPERATURE AND WHETHER OR NOT THEY ARE ACTIVE IS INDICATED IN THE TABLE.	50
TABLE 9: OVERVIEW OF THE CONSIDERED DISULPHIDE BRIDGE INTRODUCTIONS: THEIR MUTATIONS, WHETHER OR NOT THEY ARE ACTIVE AND THEIR MELTING TEMPERATURE	52

Table of contents

Acknowledgements	i
Abstract	ii
Samenvatting.....	iii
List of figures	iv
List of tables	vii
Table of contents.....	viii
Introduction.....	x
Thesis.....	1
Literature study	1
1. Cellulose as renewable resource.....	1
2. Cellulases.....	5
3. Lytic polysaccharide monooxygenases (LPMOs).....	8
4. Enzyme stability.....	13
5. Enzyme engineering	20
Materials and methods	23
1. Determination of mutations.....	23
2. Molecular DNA work	23
2.1. Single point mutations	23
2.2. Agarose gel electrophoresis and concentration measurement.....	27
2.3. Production of BL 21 (DE3) electrocompetent <i>E. coli</i> cells	27
2.4. <i>E.coli</i> transformation	28
2.5. Colony PCR.....	28
2.6. Plasmid extraction and sequencing.....	29
3. Enzyme production and purification	29
3.1. Growth.....	29
3.2. Periplasmic isolation of the protein	30
3.3. Cytoplasmic isolation of the protein	31

3.4. Enzyme purification.....	31
3.5. Enzyme concentration.....	32
3.6. SDS-PAGE.....	33
4. Enzyme activity.....	34
5. Enzyme stability with Differential Scanning Fluorimetry (DSF).....	35
Results and discussion.....	36
1. Optimization of expression of ScLPMO10C.....	36
1.1. Determination of an expression system.....	36
1.2. Optimize the purification protocol.....	41
2. Optimization of assay	43
2.1. Activity tests	43
2.2. Measurement of apparent melting temperatures.....	44
3. Increasing Protein stability.....	47
3.1. Importance of the N-terminus	47
3.2. Mapping the flexible regions in ScLPMO10C via B-fitter.....	47
3.3. Rational mutants:	49
3.4. Disulphide bridges:	51
General conclusion	55
Future perspectives.....	57
Appendix.....	i
appendix 1: sequence of the used plasmid encoding the ScLPMO10C	i
Appendix 2: Map of the features of the used plasmid encoding the ScLPMO10C	viii
Sources	ix

Introduction

Due to the strain of the ever increasing energy demands on the limited supply of fossil fuels and the increased awareness of global warming, the interest in renewable resources has strongly increased. Therefore a transition to a more renewable bio based economy is required. The first generation of bio-ethanol, derived from sugar and starch containing crops (e.g. sugar cane, corn, wheat, ...), is currently the most commonly used source of alternative energy. While already commercially available, they are highly contested due to the impact on the prices and availability of the food and feed stocks.

A promising alternative for these food crops is the use of cellulose. As the most abundant biopolymer on earth, cellulose demonstrates a big potential as alternative resource. Lignocellulosic biomass obtained from wood, agricultural residues and municipal waste streams can be applied as carbon source for the creation of second generation bio-ethanol, without interfering with the food and feed market.

Due to the strong and stable β -1,4-glycosidic bonds and the recalcitrant nature of the crystalline structure of cellulose, enzymatic breakdown proved difficult. A recently discovered group of enzymes, however is believed to hold the key to the development of an efficient enzyme mix for the degradation of cellulose. This new group of enzymes, called lytic polysaccharide monoxygenases (LPMOs), however requires sufficient stability in order to retain sufficient activity at process conditions to contribute to the efficiency of the cellulose breakdown.

The goal of the research reported in this thesis, is the expression of a bacterial lytic polysaccharide monoxygenase with the potential for application in industrial processes. This involves the creation of a more thermostable and still active mutant of the cellulose degrading enzyme ScLPMO10C (former ScCelS2). As the expressed recombinant protein needs to be tested for both the activity and the melting temperature, the first requirement in the search of improved mutants is to optimize the expression system. This expression system is thus expected to yield an active enzyme, with the possibility to be purified in sufficient concentrations. In this regard, the enzyme yield, correct N-terminal end and the his-tag are crucial aspects. After the optimization of the expression system, the assays need to be optimized as well to successfully determine differences in the activity and apparent melting temperatures between the enzymes from wild type and mutated genes. The third aspect of the research is the actual mutating, expressing, purifying and measuring of the different strains.

The first part of this thesis consists of a literature study to provide the necessary background information on the discussed topics. The next chapter of this work will report all the materials and methods to have been applied during this research. Subsequently the results of the performed tests are reported and discussed, followed by a conclusion and future perspectives of this research.

Thesis

Literature study

1. Cellulose as renewable resource

1.1. The transition to a biobased economy

Ever since the industrial revolution, the increasing machinery, globalisation, industry and transportation provided the world with an ever increasing energy demand. This rising consumption has put a strain on the natural fossil resource supplies, which become gradually depleted (Naik, Goud, Rout, & Dalai, 2010). This concern along with the increased awareness of the climate change is forcing a change in the global fossil-based economy into a more renewable one. Governments worldwide have been stimulating the transition to other types of energy generation. While different approaches, such as wind turbines, hydroelectric stations and solar power plants, have been implemented over time, fossil fuels still contribute to up to about 87% of the worldwide energy supply (Balat & Balat, 2009; Sandra T. Merino, 2007; Wyman, 1994).

Since cellulose is the most abundant biopolymer on Earth, biomass is an excellent carbon source to use as renewable resource (Y Percival Zhang, Himmel, & Mielenz, 2006). A transition in this direction is currently on-going (Balat & Balat, 2009). The power of this so-called biobased economy lies in the potential of biomass to be applied for not only bio-ethanol generation, but all kinds of non-food products such as electricity, heat, a variety of chemicals, other transportation fuels... (Lin & Tanaka, 2006). The substantial cost for the necessary enzymatic conversions has provided researchers with major challenges for the commercialisation (Demain, 2000). In an attempt to evaluate the added value of bio-fuels (mainly bio-ethanol) when compared to the production cost, researchers have proven that an addition of bio-ethanol decreases the exhaust emission and improves combustion rates by increasing the oxygen content of gasoline (Horn, Vaaje-kolstad, Westereng, & Eijsink, 2012). Even though the energy equivalent of bio-ethanol is only 68% of gasoline, its carbon emission during combustion is over 80% lower and the complete lack of sulphur dioxide exhaustion prevents the formation of acid rain (Mussatto et al., 2010). This leads to the conclusion that the benefits from bio-fuels warrant the efforts made in the search to reduce production costs (Å & Hooijdonk, 2005).

1.2. Different generations of bio-ethanol

Inspired by the production of alcoholic beverages, the development of the bio-fuels was initially based on the ethanol fermentation of corn by yeasts. However the potential of micro-organisms doesn't stop at the dissimilation of starch. Since micro-organisms also convert a wider series of polymers into fermentable single saccharide units, other polymers are being considered as potential alternative carbon sources (Balat & Balat, 2009; Demain, 2000).

By general consensus, all molecules currently acknowledged as alternative carbon sources, have been classified in one of three classes. The first class comprises sugar cane and sugar beet. These crops contain sucrose, offering the advantage of fermentation without requiring extensive pre-treatment.

A second class of alternative carbon sources includes starch-based crops such as potatoes, rice and the aforementioned corns. The malting process of corn has been performed for ages and the enzymatic saccharification of potato and rice starch has been well studied. The resulting know-how makes that pre-treatment of this type of carbon source poses few problems in the production process of bio-ethanol (Cardona, 2007). As both this first and second class of biomass require limited pre-treatment, they are considered convenient and they were the first raw materials to be used in the production of bio-fuels (Horn et al., 2012). Therefore, the bio-ethanol produced from these resources are currently referred to as "first generation bio-ethanol". Until today these first generation bio-fuels are still the most commonly produced bio-fuel, with corn-starch (mainly in the USA) and sugar cane (mainly in South-America) as the primary raw materials.

Despite the advantages of these materials, their application for large scale ethanol production has always been highly contested (Cardona, 2007). This controversy finds its origin in the competition the energy production would pose to the food and feed industries. Potatoes, rice, corns, sugar cane and sugar beets all hold major value in the feeding of both people and livestock, so their use on large scale would seriously disarray the food industry (Gray, Zhao, & Emptage, 2012). In order to maintain enough nutrition to feed the increasing world population, the "second generation bio-ethanol" has a higher potential (Å & Hooijdonk, 2005). This type of bio-ethanol differs from the first generation, since it is produced from lignocellulosic biomass. Since this third class of carbon source is the main component in plant cell walls, it can be found in wood, agricultural waste, municipal solid waste, etc. The natural abundance contributes to its value as a renewable resource, as it is not only cheap to obtain, would not undermine different industries but would even contribute to the processing of vegetal/agricultural waste. However despite being constructed by the same glucose subunits as starch, this lignocellulosic biomass is much more difficult to degrade to fermentable glucose units (Mussatto et al., 2010).

1.3. Cellulose

1.3.1. Build

The conversion of lignocellulosic biomass to fermentable units is a difficult venture. Lignocellulose, as found in plant cell wall, consists mainly of cellulose (40% to 50%), combined with hemicellulose (20% to 40%) and lignin molecules (20% to 30%). Figure 1 shows that cellulose is composed of glucose subunits bound by the β -1,4-bond (Ratanakhanokchai, Waeonukul, Sakka, Kosugi, & Mori, 2013). These molecules create hydrogen bonds among themselves, aggregating multiple polymers into microfibrils. As several microfibrils aggregate, lignin and hemicellulose tie those macrofibrils together in a crystalline structure that offers structural stability and strength to plant cell walls (Horn et al., 2012).

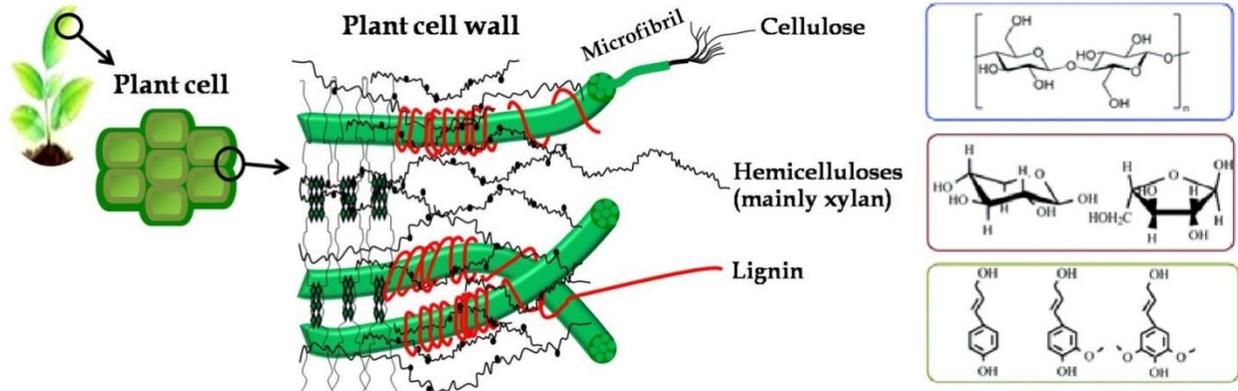


Figure 1: Schematic representation of the composition of lignocellulosic biomass. (Adapted from (Ratanakhanokchai et al., 2013))

1.3.2. Difficulties

Even though starch and cellulose are both built up of D-glucose units, starch is an easily degradable polymer while cellulose is recalcitrant to enzymatic conversion. Different causes can be found for this remarkable difference. First, cellulose polymers are composed of β -1,4- bound glucose units. These bonds have a stability that far exceeds that of the α -1,4- and α -1,6- glucose bonds in starch and form respectively amylose and amylopectin. This results in a much more difficult hydrolysis (Bayer, Lamed, & Himmel, 2007). Secondly, hydrogen bonds are formed between the glucose strings, resulting in aggregation into closely packed microfibrils. The obtained crystalline structure makes it difficult for enzymes to bind the substrate and break the bonds.

As enzymes can only access the outer-ends of the crystalline structures, this leads to a significant prolonging of the conversion and therefore a prolonging of the fermentation process. This will in turn not only result in a higher production cost, but also in the need for more stable enzymes to remain active during the longer conversion time.

Lastly, the presence of the lignin and hemicellulose around the different cellulose fibrils further impedes the conversion of cellulose to glucose. These entwining polymers not only create a physical barrier to further limit the accessibility of the enzymes, but the remnants of their breakdown also show inhibitory effects on the cellulases (Forsberg et al., 2011; Isaksen et al., 2014).

2. Cellulases

As cellulose proves difficult to degrade, nature has provided micro-organisms with an enormous variety of synergistic cellulases. All these enzymes hydrolyse the β -1,4-glycosidic bonds, so that they are classified as glycoside hydrolases (E.C.3.2) (Yi-heng Percival Zhang & Lynd, 2015).

2.1. General build-up

Most cellulose-active enzymes show a modular structure. In order to assure close interaction with the insoluble cellulose biopolymer, a carbohydrate-binding module (CBM) can be present to increase the specific substrate binding capability of the enzyme. After substrate recognition the second important enzyme module, the catalytic domain, will initiate hydrolysis through acid-base catalysis (Sukharnikov, Cantwell, Podar, & Zhulin, 2011).

As for every enzyme, the correct three-dimensional fold of the protein structure is of primordial importance to obtain catalytic substrate activity (Harmsen & Huijgen, 2010; Lynd, Weimer, Zyl, & Pretorius, 2002). This topology of the active site shows three possible characteristic varieties: the pocket, the cleft/groove and the tunnel. All cellulase enzymes are glycoside hydrolases, involved in the conversion of polymers. Only two types of the active sites are commonly found in glycoside hydrolases: the cleft/groove and tunnel conformation. The lack of pocket conformation in glycoside hydrolases is the result of the role of the conformation in the accessibility and therefore the substrate specificity (Demain, 2000; Horn et al., 2012).

2.2. Types and mechanism

Three types of cellulases contribute to the conversion of cellulose. The substrates, products and roles of all three types in the enzymatic breakdown of cellulose are represented in Figure 2.

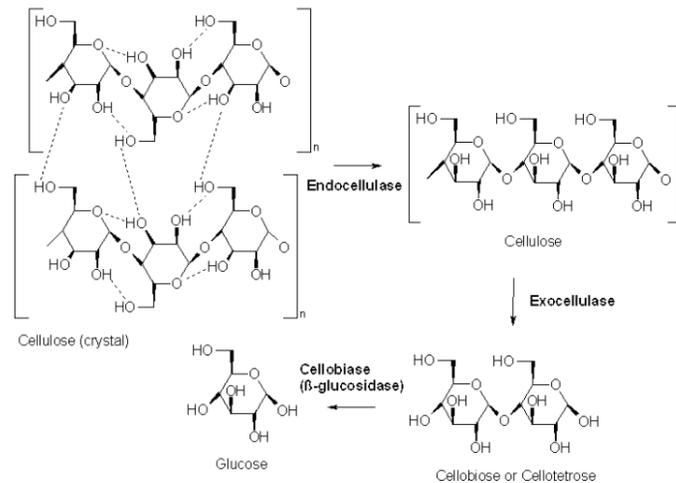


Figure 2: Schematic representation of the stepwise enzymatic catabolism of cellulose by the three types of cellulases (Promon, Program, & Sciences, 2015).

A first group consists of the endo-1,4- β -D-glucanases (EC 3.2.1.4), also known as cellodextrinases. This group of cellulases has a groove shaped active centre (part A on Figure 3), allowing it to bind along the length of the polymer. This type is therefore characterised by a random cleaving of internal glycosidic bonds in the cellulose chain (Dies, Henrissat, Davies, & Henrissat, 1995; Vaaje-Kolstad et al., 2010).

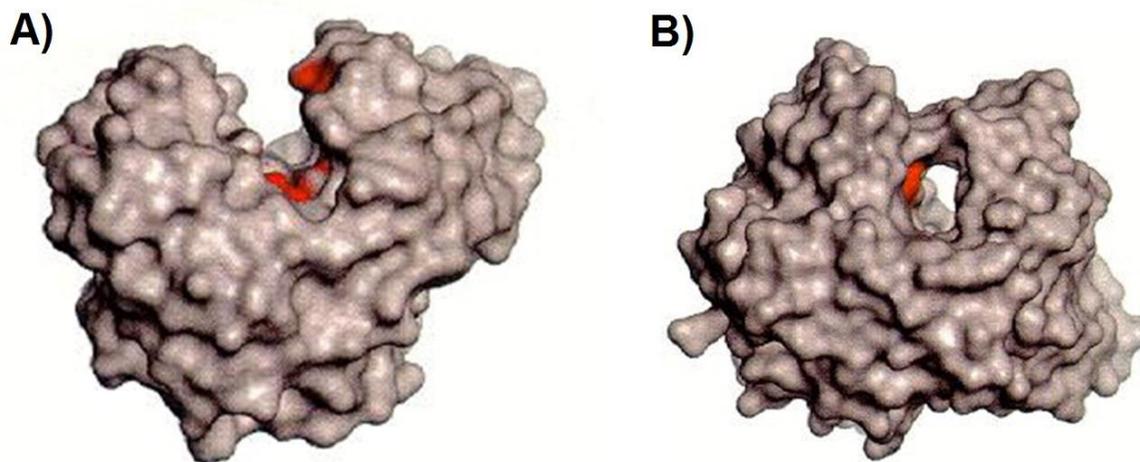


Figure 3: A) Representation of the cleft/groove conformation in endoglucanase (this example is from E2 from *Thermobifida fusca*). B) Representation of the tunnel conformation in cellobiohydrolase II (this example is from *Trichoderma reesei*). The proposed catalytic residues are shaded in red. Source:(Dies et al., 1995)

The second group of cellulases comprises of exo-1,4- β -D-glucanases (EC 3.2.1.176), also called cellobiohydrolases (CBH). The active site of these cellulases shows a tunnel shaped conformation (part B on Figure 3). This conformation only allows the enzyme to bind at the ends of the polymers, because the polymer has to be completely moved through the tunnel. Only when the polymer passes through the tunnel shaped catalytic centre, a cellobiose molecule will be separated from the end of the glucose chain. Depending on the end where the exo-cellulase is able to bind and begin the hydrolysis of the cellulose, two subtypes can be distinguished. Cellobiohydrolases that are able to start the conversion at the reducing ends of the cellulose polymer are named CBHI, while the CBHII can only begin the degradation of cellulose at its non-reducing end (Dies et al., 1995).

As both the endo- as the exo-glucanases are unable to completely break cellulose down to glucose, further conversion of cellobiose to glucose is required. This is where the third group of cellulases: the β -glucosidases (EC 3.2.1.21) come in to play (Horn et al., 2012).

3. Lytic polysaccharide monoxygenases (LPMOs)

3.1. The function of LPMOs in the synergetic mechanism of cellulose breakdown

To use lignocellulosic biomass as carbon source, micro-organisms require a highly efficient enzyme system, comprising of different synergistic enzymes (all of which are displayed in Figure 4). Endoglucanases cut randomly in the polymer, so that more polysaccharide chain-ends become available for degradation by exo-cellulases. Since this results in progressive conversion of cellulose to cellobiose, the enzymatic conversion rate of the β -glucosidases will increase accordingly (Langston et al., 2011). However, this system still remains very inefficient. Since the 1950s, it has been suggested that to achieve a sufficient efficiency on recalcitrant cellulose, the enzyme system would require an additional component.

In 2005, scientists have identified an oxidative enzyme in the chitin breakdown that boosts the degradation process (Forsberg, Røhr, et al., 2014). Chitin degradation is highly comparable to cellulose breakdown and an analogous oxidative enzyme was shortly after identified and presumed to be this missing key element in the enzymatic breakdown (Aachmann, Sørli, Skjåkbræk, Eijsink, & Vaaje-kolstad, 2012). Later on, these enzymes were called Lytic Polysaccharide Monoxygenases. Their flat substrate-binding surface allows the protein to bind the crystalline materials and oxidatively cleave the glucose chain (while the aforementioned cellulases are hydrolases). This disrupts the closely packed recalcitrant structure and provides new accessible chain ends where the exocellulases can work on. This boosts the activity of the classical enzymes. Further research into the nature, classification and origin of this component, led to the discovery of multiple variants in both fungal and bacterial hosts (Aachmann et al., 2012; Forsberg, Mackenzie, et al., 2014; Horn et al., 2012; Isaksen et al., 2014).

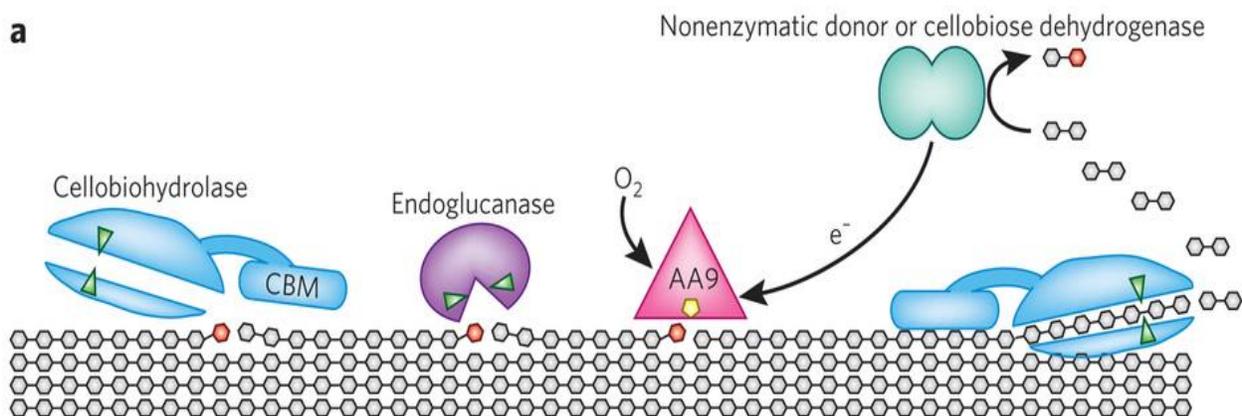


Figure 4: Schematic representation of the different cellulose active enzymes: cellobiohydrolase (CBH) represented in light blue, endoglucanase represented in purple and the LPMO represented in pink Source:(Fushinobu, 2013)

3.2. Classification

Although the discovery of LPMOs is rather recent, four different families have been found and different classifications have been used. The first LPMOs to be identified in 2005 (Forsberg, Røhr, et al., 2014) were firstly classified in the Carbohydrate Binding Module family 33 (CBM33), primarily found in bacteria. Further research has led to the discovery of a second family of LPMOs. These were considered to be a member of the Glycoside Hydrolase family 61 (GH61) because of a very weak endoglucanase activity. These enzymes, mainly produced by fungi, show catalytic activity on cellulose (Aachmann et al., 2012).

As more information about these LPMOs was revealed, it became clear that LPMOs didn't catalyse hydrolysis and a change in nomenclature was considered necessary. The Carbohydrate-Active enZYme (CAZy) database has assigned the different LPMOs into Auxiliary Activity families (Levasseur, Drula, Lombard, Coutinho, & Henrissat, 2013; Span & Marletta, 2015). The families formerly known as GH61 and CBM33 changed to Auxiliary Activity family 9 and 10 respectively (AA9 and AA10). It wasn't until 2013 that scientist discovered a third separate family of LPMOs: the Auxiliary Activity family 11 (AA11) (Hemsworth, Henrissat, Davies, & Walton, 2013). This group of enzymes differentiates from the other families by its activity on chitin. Later, a fourth group of LPMOs was found to reduce starch and named AA13 (Leggio et al., 2015). This group of LPMOs is very remarkable since starch has no crystalline structure and a completely different bond as described above. Both the AA11 and AA13 type are produced by fungi. A schematic representation of the different auxiliary activity classes and their substrate specificity and origin is depicted in Figure 5 (Span & Marletta, 2015).

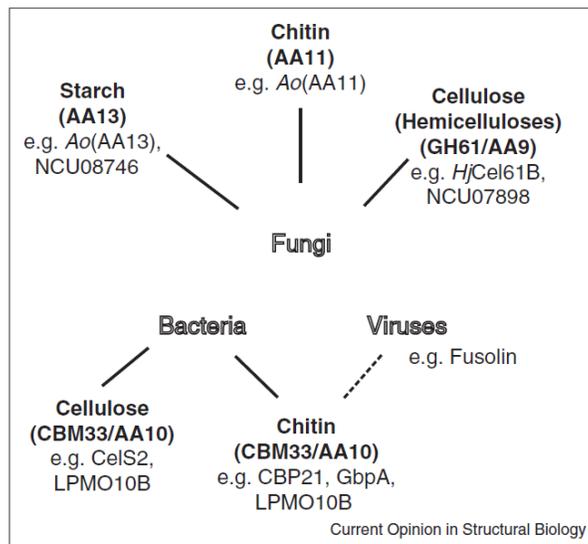


Figure 5: Variety in substrate specificity and origin of the different LPMO families (Span & Marletta, 2015)

3.3. Characteristics

As previously mentioned, LPMOs differ from the conventional cellulases by their different cleavage mechanism. Since LPMOs catalyse a direct oxidation of the glycosidic linkage on the crystalline surface of lignocellulosic biomass, they are considered oxygenases instead of hydrolases.

Although a rather low sequence identity, all four auxiliary active families share a rather similar conformation (visualised in Figure 6). All four families comprise a β -sandwich fold consisting of two anti-parallel oriented β -sheets that are situated on top of each other. One β -sheet consists of three anti-parallel strands, while the other contains four anti-parallel strands. Structural comparison between different LPMOs also revealed several aromatic amino acids at the core are highly conserved. This is speculated to be because of a crucial role in the electron transfer pathway necessary for the reduction of the copper ion (Agger et al., 2014; Hemsworth et al., 2013; Horn et al., 2012). Both β -sheets are connected through a series of loops and turns, formed by α -helices. Upon post-translational modification, the LPMOs are further stabilised with the formation of disulphide bonds (Aachmann et al., 2012; Hemsworth et al., 2013).

Furthermore, LPMOs harbour a divalent metal ion in their active site, that has proven to yield the highest activity in the case of copper (Forsberg, Mackenzie, et al., 2014). According to X-ray crystallographic studies, this copper ion is coordinated in the catalytic centre of the enzymes by a so-called histidine brace (Span & Marletta, 2015). This histidine brace involves 2 histidine residues; wherefrom one of them must be the first amino acid of the protein, since both the side chain and the amino group of the actual N-terminus are involved (Hemsworth et al., 2013). Together with the second and more distal histidine, an equatorial water molecule and one or two water molecules located away from the canonical equatorial positions, this coordinates the copper ion in a T-shaped N_3 arrangement (Hemsworth, Johnston, Davies, & Walton, 2015). The copper ion is further stabilized by other functional groups in amino acid residues, the type of which depending on the family of LPMO. The AA9 family has a strategically placed axial tyrosine coordinating the copper interactions, while the AA10 LPMOs have both phenylalanine and alanine working to stabilize the copper interaction (Hemsworth et al., 2013). The AA11 family makes use of the presence of alanine and tyrosine residues for the chelation of the copper ion in their active centre. In each case, the copper chelation is further stabilised by formation of 2 hydrogen bonds with 2 water molecules. The different types of copper fixation are presented in Figure 6 (Aachmann et al., 2012; Forsberg et al., 2011; Hemsworth et al., 2013, 2015; Horn et al., 2012; Isaksen et al., 2014; Vaaje-Kolstad et al., 2010).

The active site's location in the centre of the flat binding surface provides little information as to the specific interaction with the substrate. Recent findings however have proposed aromatic residues on the surface surrounding the active site, as well as the highly variable loop to be decisive in substrate binding and specificity (X. Li, Beeson, Phillips, Marletta, & Cate, 2012; Vu, Beeson, Phillips, Cate, & Marletta, 2014). Literature reports protuberance of this loop to make up over 50% of the putative substrate binding surface. Once the substrate is correctly bound, oxidative cleavage can be catalysed.

Depending on the LPMO-type, the cleavage can be performed at different carbon positions of the glucose unit. When the oxidation is performed at the C1-position, δ -1,5 lactones are formed that can be hydrated to create aldonic acid. An example is ScLPMO10C. Alternatively, when the C4-position is oxidized 4-ketoaldose are produced or even gemdiol after it was hydrated. An example of this type of enzyme is the NcLPMO9C. Other LPMOs are capable of both oxidations (eg. ScLPMO10B and E7). This third possibility is able to create double oxidized products (Forsberg, Mackenzie, et al., 2014). The rupture in the cellulose polymer induced by these oxidations, renders the crystalline structured macrofibrils more amorphous and accessible for the conventional cellulases (Beeson, Phillips, Cate, & Marletta, 2012).

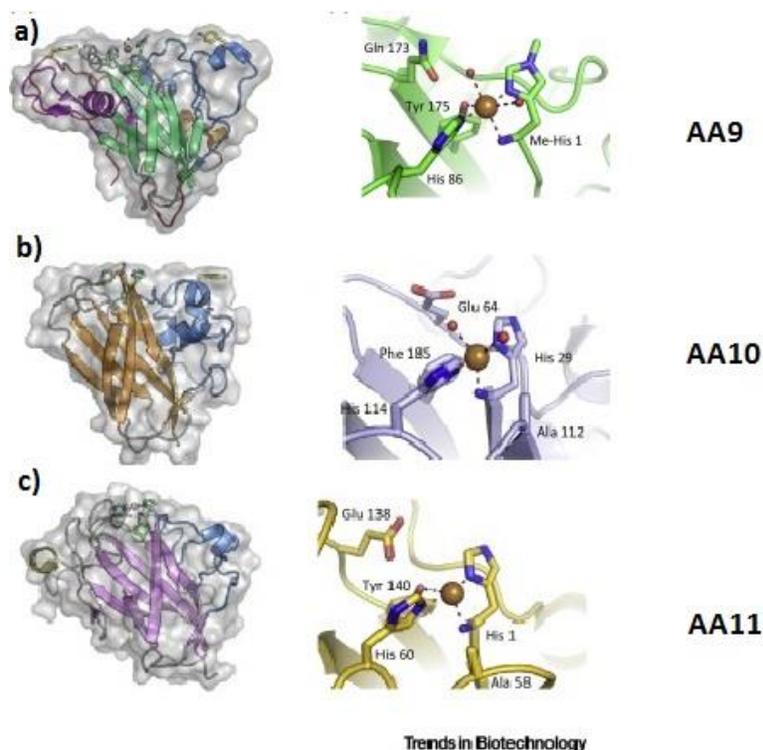


Figure 6: Structural presentation of the different Auxiliary Activity families and their respective copper binding in the active site. A) the representation of the AA9 family, B) representation of the AA10 family and C) representation of the AA11 family. Adapted from:(Hemsworth et al., 2015)

3.4. Reaction mechanism

For each oxidation reaction, molecular oxygen, 2 electrons and 2 protons are required. Therefore an electron donor is necessary to add to the reaction. This can be either an enzymatic or non-enzymatic electron sources. The latter can be redox-active products like ascorbate or gallate (Leggio et al., 2015), but also lignin has been described (Horn et al., 2012). The role of an enzymatic donor, can be accomplished by complexation of the LPMO with cellobiose dehydrogenase (CDH) (Jørgensen & de Campos Giordano, 2015; Tan et al., 2015),

For oxidative cleavage of the cellulose polymer, the electron transfer mechanism of LPMOs requires the presence of Cu^{2+} ion at the histidine brace. This bivalent copper ion, bound in the active centre of the enzyme, will be reduced to Cu^{1+} in the initial step of the electron transfer through the uptake of a first electron. The reduced monovalent copper ion will interact with the molecular oxygen. The reducing effect of the CDH will transfer a second electron to cleave the O-O bond in the intermediary molecule. As this results in the release of a water molecule, a LPMO-Cu(II)-O radical is created (Žifčáková & Baldrian, 2012). This radical is capable of executing an oxidative cleavage of the glycoside bonds of cellulose polymers by coupling with the substrate radical. (Tan et al., 2015). The reaction is represented in Figure 7.

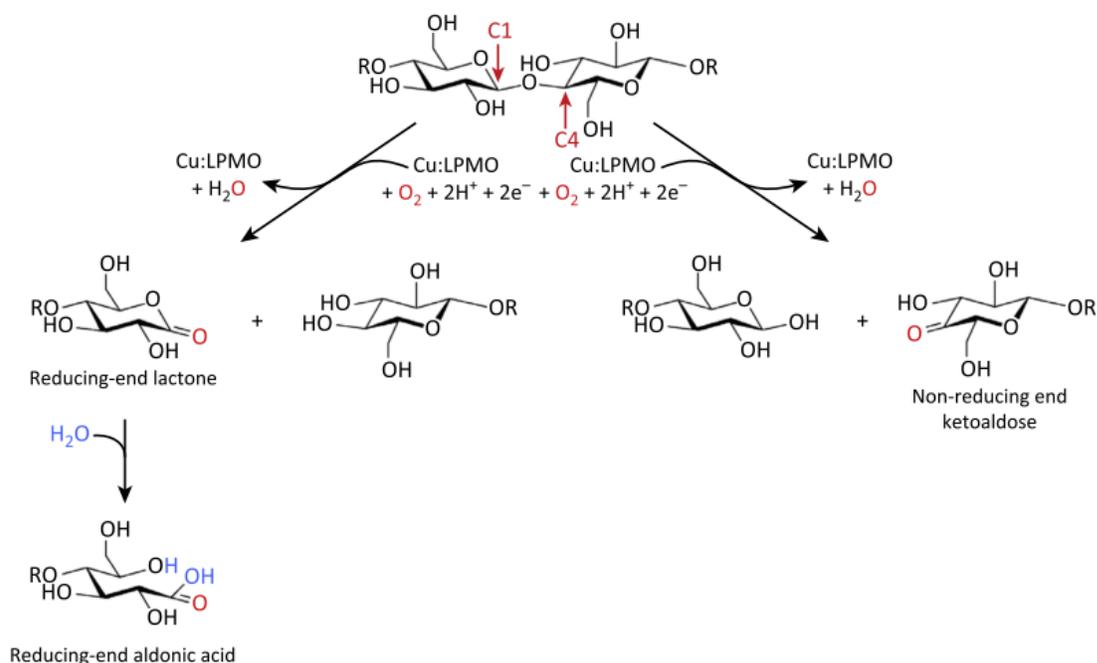


Figure 7: The oxidative cleavage reaction of LPMOs on cellulose (Hemsworth et al., 2015).

4. Enzyme stability

4.1. Relevance

Over the last years, over 4000 enzymes have been studied, while only approximately 200 microbial original types are commercially available and just about 20 applied at truly industrial scale (S. Li, Yang, Yang, Zhu, & Wang, 2012). With the growing understanding of the production, fermentation and recovery processes, the application of enzymes as industrial biocatalysts for more efficient production of biological compounds is expected to increase exponentially. Despite many major advantages, the useful application of enzymes is often counteracted by the lack of stability of the enzyme. While many industrial processes require high temperatures (an increase of temperature from 25°C to 75°C can increase the process rate a 100-fold (Iyer & Ananthanarayan, 2008)) and unfavourable pH ranges to shorten the incubation time, increase mass transfer rate through lower fluid viscosities and reduce risk of contamination, the same harsh conditions could hinder the enzyme activity (Gianfreda & Scarfi, 1991). As such, a higher stability will result in a reduced enzyme turnover and the economic feasibility of the application of an enzyme is thus determined by its stability (Eijsink et al., 2004).

Two different approaches can be taken in defining the stability of enzymes. A first type of stability is often referred to as storage stability or shelf-life: indicating the measure of preserved activity of the enzyme after the time span between manufacturing and use (Fágáin, 2003). A second type of stability is called the operational stability: the persistence of the enzyme activity during the course of the process under the operating conditions required for industrial sized productions (Fágáin, 2003).

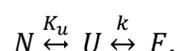
This operational stability of enzymes is of utmost importance. This property will determine whether the enzyme can be applied, while still maintaining the necessary (often denaturing) operating parameters such as increased temperatures, extreme pH conditions, aqueous-organic media,... (Gianfreda & Scarfi, 1991).

4.2. Kinetic vs thermodynamic stability

The ambiguity in literature on protein stability that has accumulated over time, has often caused a lot of confusion due to different terms and meanings. According to the generally accepted Lumry–Eyring model, the inactivation of enzymes is at least a 2-step process. The first step is the reversible inactivation of the enzyme, due to the liability of the interactions responsible for conserving the native structure of the enzyme.

The second step in the Lumry–Eyring model consists of an irreversible inactivation through covalent changes or aggregation to thermodynamically stable protein macromolecules. Based on these two steps, the stability of an enzyme can also be divided into kinetic and thermodynamic stability. The thermodynamic stability of an enzyme is defined as the resistance of the folded enzyme to denaturation, wherein denaturation is defined as the unfolding to a distorted protein. Thus resulting in the formation of a polypeptide where the key amino acid residues are no longer close enough to form the catalytic centre. When the denaturing influence is removed, this unfolding can be reversed and the activity restored (Iyer & Ananthanarayan, 2008). This is the difference with kinetic stability which measures the resistance to irreversible inactivation (Xie et al., 2014).

Both types of stability can be represented in the formula of the Lumry–Eyring model:



Herein: N represents the native state, U the denatured enzymatically inactive state (both represented in Figure 8) and F the final permanently inactivated enzyme. The ratio of N over U is a measure of the thermodynamic stability of the enzyme, whereas the ratio of N over F corresponds to the kinetic or long-term stability (Iyer & Ananthanarayan, 2008; Sanchez-ruiz, 2010).

Each of the stabilities is characterised by its own parameters. For kinetic stability, the time an enzyme retains its activity before undergoing irreversible denaturation is characterised by the half-life ($T_{1/2}$), the overall observed deactivation constant (k), the optimum operating temperature (T_{opt}) and the temperature of half-inactivation (T_{50}). The $T_{1/2}$ is defined as the time, required to decrease the activity of a protein in specific condition to half. The T_{50} on the other hand is the temperature to be maintained during the incubation in order to cut the residual activity by half in a defined period of time (Polizzi, Bommarius, Broering, & Chaparro-Riggers, 2007).

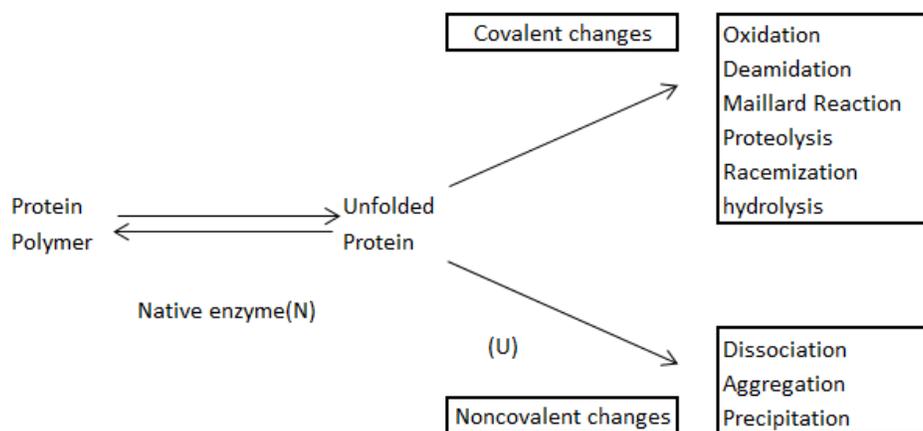


Figure 8: schematic representation of the grouping of the different inactivating interactions(Iyer & Ananthanarayan, 2008)

The most frequent approach in studies on protein stability centres on the determination of the thermodynamic stability, represented by the free energy of unfolding (ΔG_u), an unfolding equilibrium constant (K_u) or the melting temperature of the protein (T_m). The ΔG_u is the difference in Gibbs free energy measured during the transition from folded to unfolded enzyme, while the K_u constant is the ratio of the concentration of the unfolded protein on the concentration of the folded enzyme. The ΔG_u can be directly detected by differential scanning calorimetry, yet the value can be more reliably obtained through interpolation from the difference in enthalpy (Johnson, 2013) determined as visualised in Figure 9(B). The T_m is the temperature at which half of the enzyme has unfolded, which can be determined as visualised in Figure 9(A). These characteristics of the thermodynamic stability can be determined through analysis methods as differential scanning calorimetry, tryptophan fluorescence or changes in tyrosine absorbance (Polizzi et al., 2007).

Research on enzyme stability has shown that enzymes with a thermodynamically unstable native state, may still be capable of retaining their biological activity for physiologically relevant times. In order for the activity to maintain in this situation, the enzyme would require a sufficient difference in free-energy between the native form and the non-functional protein state (Sanchez-ruiz, 2010). Despite these findings, researchers have learned that the difference in ΔG_u of mesophilic and hyper-thermophilic proteins, does not exceed and is often even lower than 100 kJ/mol. As this is only the equivalent of a few non-covalent interactions, composing a consensus strategy on the stabilization of enzymes through specific adaptations for targeted changes in ΔG_u is further impeded(Jaenicke & Gerald, 1998).

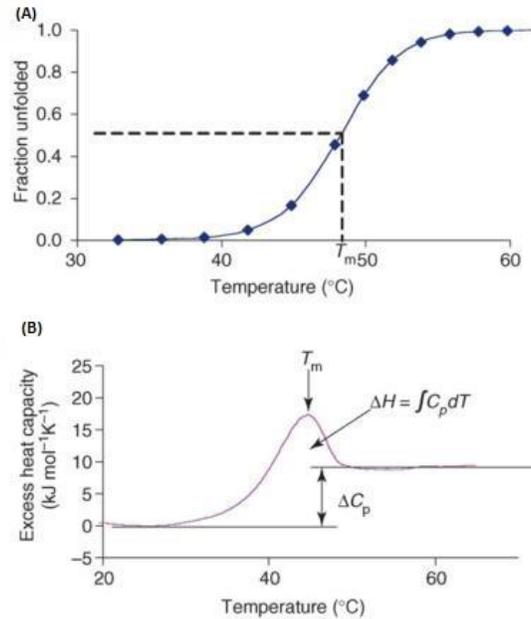


Figure 9: (A) The determination of the melting temperature through the melting curve. (B) The determination of the T_m and change in enthalpy through differential scanning calorimetry. Source: (Polizzi et al., 2007)

4.3. Methods for stability improvement

The active site is generally considered the most fragile part of the enzyme. Due to the importance of the correct folding of this catalytic centre, the main concern in improving the stability of enzymes is the preservation of the active site under more rigid conditions (Xie et al., 2014). Different methods exist such as addition of compounds to stabilise the protein, immobilization of the protein and enzyme engineering. These three strategies will be described in more detail below.

4.3.1. Additives

Across the entire spectrum of fields of enzymology, the oldest and most commonly applied type of stabilization is the use of additives. This addition of substrate, ligand, co-factors (e.g. NADP⁺),... can significantly reduce the production of thermal or ultrasonic-unfolded enzyme. This is the result of the stabilizing effect of the interactions between that compound and the active centre of the enzyme. The addition of salts has also been reported to increase the stability of enzymes, as they contribute to the preferential hydration of the enzyme (Fágáin, 2003).

Both the use of the bridging effects of low concentrations of specific divalent cations and the comprising effect of higher concentrations of non-specific salts through salting out, have positive effects on the enzyme stability. Alternatively, when polyols and sugars are used as additives to aqueous enzyme solutions, their positive effect to the water activity of the medium strengthens the hydrophobic interactions among the enzyme's non-polar amino acid residues (Gianfreda & Scarfi, 1991). This in turn results in a more rigid protein fold and higher resistance to thermal denaturation. Additions thus offers great possibilities in the increase of storage stability, yet they could potentially be incompatible with the required reaction parameters or be economically unviable (Iyer & Ananthanarayan, 2008). For this reason, researchers have been studying different approaches as well.

4.3.2. Immobilization

One of the more recent approaches in the increase of the enzyme stability is the immobilization of enzymes on a porous support material. This immobilization provides a certain structural protection against intermolecular process and external influences (proteolysis, aggregation, inadmissible organic solvents and oxygen) (Mateo, Palomo, Fernandez-lorente, Guisan, & Fernandez-lafuente, 2007). Positive effects from random immobilization is however not guaranteed, as even negative effects from immobilization on stability have been documented (Mateo, Abian, Lafuente, & Guisan, 2000). Several different methods have been tested. Most promising techniques proved to be adsorption, entrapment, membrane confinement and single or multiple covalent bonding (Iyer & Ananthanarayan, 2008). In order to increase conformational stability against thermal influences, the most applied method is the use of multipoint covalent attachment (represented in Figure 10). This approach is based on the stabilizing effect of the fixation of critical positions of the enzyme to a stable surrounding. This is obtained through covalent binding of residues of certain amino acids to reactive groups on spacer arms on the carrier substance. Through locking these amino residues in to place, they retain their relative positions during denaturing conditions, resulting in a higher stability (e.g. 5-fold and 18-fold factor for chymotrypsin and penicillin G acylase respectively as reported by Mateo et al.). Furthermore, this allows the fixation of the different sub-units in multimeric enzymes, preventing inactivation through inter-subunit dissociation or unconstructive intra-subunit crosslinking. When the complex structure of the multimeric enzyme prevents the fixation of every subunit in this multipoint attachment to a plain surface, it can also be complemented with chemical crosslinking with polyfunctional polymers (Mateo et al., 2000).

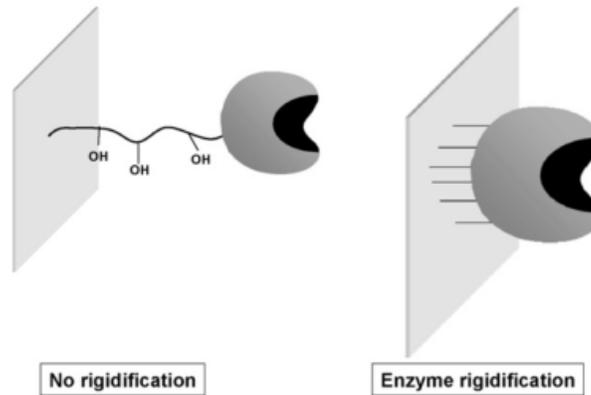


Figure 10: Schematic representation of the effect of specific multipoint immobilization on enzyme stability (Mateo et al., 2007)

The number of enzyme-support bonds can be correlated with the degree of the stabilization provided by the immobilization. Agarose beads, porous glass, zeolite and epoxy resins have been frequently used for fixation of enzymes. Through a successfully executed multi covalent attachment, stabilization factors up to even 30000-folds have been measured. The use of covalent bond enzyme also offers the additional advantages of an increased operational control and enzyme recovery through filtration (Iyer & Ananthanarayan, 2008). Activity studies on the other hand have shown slight decreases after immobilization, with activity recoveries around 60% on average (Mateo et al., 2007). For increasing the thermodynamic stability of enzymes, the second most common technique is the entrapment of enzymes. When compared to the soluble enzyme of Horseradish peroxidase, an increase in activity of 100-fold and an increase in turn-over after 80 minutes of 230-fold was measured after entrapment in a copolymer (Polizzi et al., 2007).

4.3.3.Engineering

Another approach in the stabilization of enzymes consists of genetically enhancing the natural enzyme to a more resistant protein. As immobilization protocols can be rather expensive and time-consuming, random or site-specific mutagenesis can be applied to manufacture a more stable enzyme. In this approach called enzyme engineering, different methods can be applied.

In order to increase enzyme stability, an increase in the transition state energy barrier for denaturation is desired. As this is obtained through formation of numerous non-covalent bonds and interactions, the most common approach is therefore to introduce more hydrophobic interactions, salt bridges, hydrogen bonds and van der Waals forces (Xie et al., 2014).

Through enzyme engineering, the half-life at 60°C of sucrose phosphorylase has been more than doubled through combination of different stabilizing rational mutation and the T_{50} was increased by 3°C (Cerdobbel et al., 2011).

Depending on whether the mutations are random or more thought trough, these methods are called directed evolution, semi-rational design or rational design (Aerts et al., 2013). A more elaborate explanation can be found in the following topic.

5. Enzyme engineering

Three different types of enzyme engineering have developed over the course of the years. The first to be developed in the lab was rational design, followed by directed evolution (which is considered by some as the first technique, as it is based on nature's evolution) and later semi-rational design. The ideal method to be used in certain situations can be evaluated according to the selection plan in Figure 11 while a comparison of the advantages and disadvantages of the different theories is represented in Table 1.

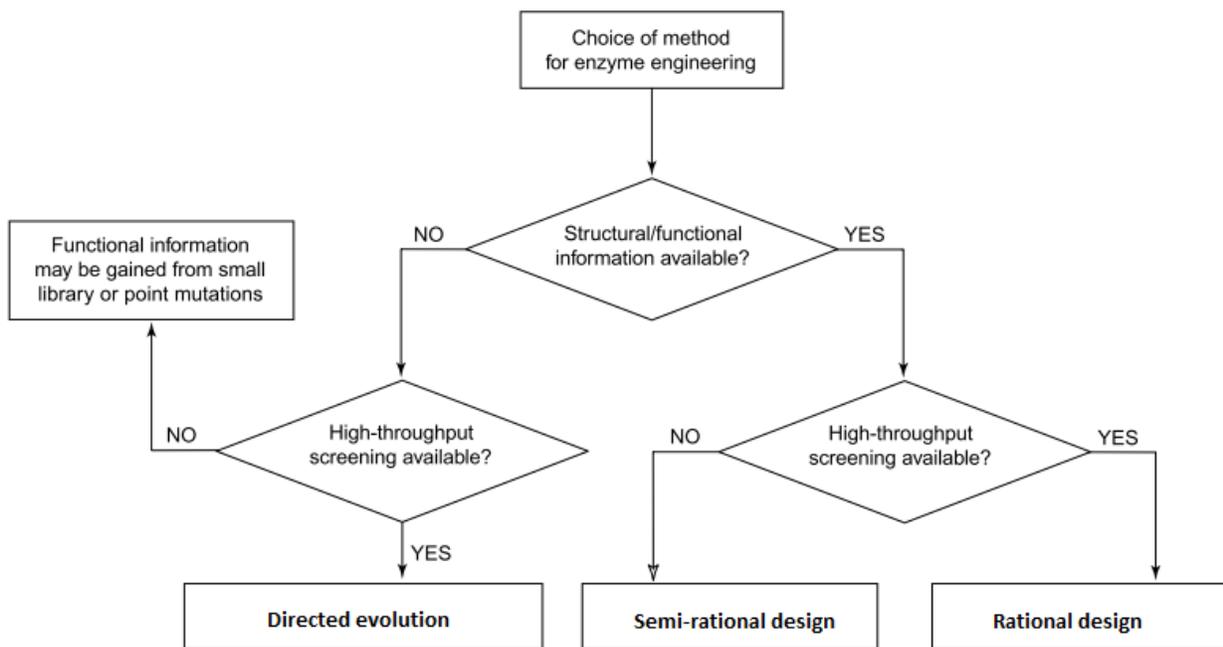


Figure 11: the selection plan of the different approaches in enzyme engineering (Chica, Doucet, & Pelletier, 2005)

Table 1: comparison of approaches for engineering improved enzyme stability(Chica et al., 2005)

	Rational design	Random mutagenesis	Semi-rational design
High-throughput screening or selection method	Not essential	Essential	Advantageous but not essential
Structural and/or functional information	Both essential	Neither essential	Either is sufficient
Sequence space exploration	Low	Moderate, random	Experimental: moderate, targeted Computational: vast, targeted
Probability of obtaining synergistic mutations	Moderate	Low	High

5.1. Types

5.1.1. Directed evolution

Directed evolution is a protein engineering approach involving iterative rounds of selecting beneficial mutations out of substantial libraries obtained through large scale mutagenesis.

The functioning of a protein structure is a characteristic property of the entire polymer, rather than site specific. Ameliorations to enzymes have often proved to be the cumulative effect of improvements by multiple small mutations dispersed across the molecule. Hence in order to improve the enzyme, mutations throughout the entire sequence can be screened on the induction of a desired effect on the enzyme. This approach was introduced in 1994, resulting from the pioneering work of Stemmer, Chen and Arnold, and is called directed evolution as it was inspired by natural evolution (Borchert, Burg, Eijsink, & Ga, 2005; Fágáin, 2003). As the occurrence of successful mutations is of low frequency, this method involves extensive libraries of mutated strains to be tested (Aerts et al., 2013). While the required knowledge of structural or functional information is limited, an efficient and highly specific high through-put screening method is crucial (as can be seen on Figure 11). This is often referred to as “you get what you screen for”.

Several approaches are known for the production of the libraries, such as error prone PCR and UV induced mutagenesis. Another method of creating these mutated strains is through a PCR reaction without addition of primers. The used DNA material is hereby obtained after random fragmentation of a mixture of sequences with one or more point mutations in the original enzyme encoding sequence. This PCR reaction will allow random reassembly, leading to template switching of the different starting point mutations (Fágáin, 2003). In order to maintain feasible sized libraries, “family shuffling” is often applied (Y Percival Zhang et al., 2006). Family shuffling is based on recombining homologues sequences, naturally recurring in related organisms. This approach is based on the assumption that those residues are more likely to successfully stabilize the enzyme, as this has already propagated through natural selection (Cerdobbel et al., 2011). This method has allowed the $t_{1/2}$ of phosphite dehydrogenase at 45°C to be increased over 23000-fold without decreasing the catalytic efficiency (Nannemann, Birmingham, Scism, & Bachmann, 2012).

5.1.2. Rational design

Despite the similarity between directed evolution and natural evolution, the first method of genetically engineering of enzymes to be applied in laboratories was a rational design experiment. Rational design is the approach of enzyme engineering, in which knowledge of both the 3D structure and the mechanistic knowledge of the enzyme is used to determine a handful of specific amino acid exchanges believed to result in a specific property (Polizzi et al., 2007).

Excessive knowledge of both the structure and the function of the different regions is required to successfully determine the specific mutations to be tested (Suplatov, Voevodin, & Švedas, 2015). This approach is thus based on selectively mutating the specific locations believed to determine a specific property while maintaining the other characteristics.

While no decisive rules have been established, certain trends for the rational engineering of thermostability have emerged during thermodynamic assessment experiments on small enzymes able to reversibly unfold. Stabilization can be achieved through increased entropy by insertion of proline residues, while limiting glycine, asparagine and glutamine (Cerdobbel et al., 2011). Other commonly used mutations for potentially stabilizing effects are helix capping and the introduction of extra salt bridges and disulphide bonds (Eijsink et al., 2004). A successful example of enzyme stabilisation was by introducing two disulphide bridges in the tetrameric MDH from *Chloroflexus aurantiacus*. This led to an increase in apparent melting temperature of 15°C at a pH 7,5 (Eijsink et al., 2004).

5.1.3.Semi-rational design

This type of enzyme engineering is based on the saturation mutagenesis of amino acid residues that are assumed to offer the desired improvement, based on sequential comparison to other experiments and sequences. These specific locations have been selected using computational predictive algorithms and what information is known on protein sequence, structure and function. By using saturation mutagenesis, all 20 native amino acids are once mutated in to those specific sites and evaluated for their influence on the specific feature. While still requiring less information than in rational design, this more structural approach provides a higher frequency of successful mutations. Thus this approach requires a smaller mutant library. Further limitation of the libraries can be obtained by taking the degenerate codons in to account (Bommarius, Blum, & Abrahamson, 2011). All positively influencing mutations are afterwards combined for optimal stabilization (Chica et al., 2005).

As the most recent approach to enzyme engineering, semi-rational design thus consists of the combination of the limited need for structural knowledge and the limited requirements concerning the throughput of the screening methods. As sufficient structural and functional information concerning some enzymes is not yet revealed and implementation of high-throughput screening methods often proves difficult, this approach to protein engineering proves more easily applicable (Lutz, 2010)

Materials and methods

1. Determination of mutations

To increase the thermostability, the used approach is based on the fixation of the most flexible amino acids. These locations were detected through analysis of X-ray crystallography data. Analysis was performed through B-fitter, designed by Reetz et al (Reetz, Carballeira, & Vogel, 2006), to rank all amino acids in order of decreasing B-factor values (representing the degree of vibration of that amino acid). To further evaluate these results, the top 20 flexible positions (here this was all positions with B-factors over 15) were visualized in Pymol. In order to reduce the risk of losing the enzyme activity by disrupting the copper binding site, all amino acids at a distance of less than 6 Angström from the Copper atom were considered as unsuitable for mutating. The location of all remaining amino acids with a B-factor over 15, and the surrounding secondary structures were further rationally inspected for the possibility to create stabilizing mutations by use of salt bridges, hydrophobic interactions,...

As a second approach, the insertion of disulphide bonds is generally considered to offer great potential in the stabilization of proteins. In order to determine the mutations to allow possible insertions of disulphide bonds, 2 tools were applied: disulphide by design (DbD) and MOdelling of Disulfide bridges In Proteins (MODIP)(Dani, Ramakrishnan, & Varadarajan, 2003; Sowdhamini, R., Srinivasn, N., Shoichet, B., Santi, Daniel, V., Ramakrishnan, C. & Balaram, 1989). The combining of the results from these tools, should increase the chance of successful mutations. All proposed mutations on the A list from the MODIP tool to reoccur on in the results of the DbD tool were selected. Furthermore, all proposed mutations from the B list from the MODIP results that were proposed by the DbD tool, were listed as flexible positions by B-fitter and further than 6 Angström from the Copper binding site were selected as well.

2. Molecular DNA work

2.1. Single point mutations

The protocol of Sanchis et al was used for site directed mutagenesis strategies to substitute one single codon at a time (Sanchis et al., 2008). This method is visualised in Figure 12. During this process, the first step of the PCR protocol used a cycle of denaturation, annealing and elongation steps to create a mega-primer carrying the desired mutation. This was followed by a second step where the created mega-primer was used to amplify the entire plasmid.

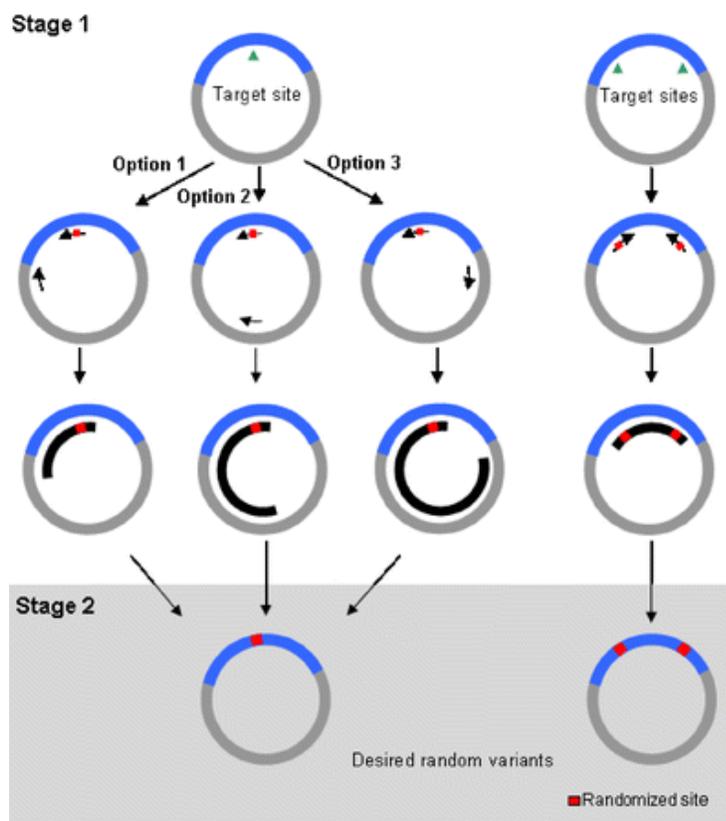


Figure 12: The different steps in the Sanchis PCR principle from (Sanchis et al., 2008). Starting with the creation of a mega-primer in the first stage, it is followed by the amplification of the entire plasmid in the second. Depending on the position of the annealing sites of the starting primers, the length of the mega-primer can be varied as depicted in the left column. By starting with mutations in both the forward and reverse primers, two mutations can be introduced at once, as depicted in the right column.

The initial protocol was a long PCR reaction, that combined the creation of the mega-primer in 5 cycles (stage 1 in Figure 12) and the amplification of the entire plasmid in 20 cycles (stage 2 in Figure 12). The efficiency of this protocol appeared to be very low in our case, therefore, the entire protocol was optimized to a protocol with two separate PCR reactions. In the first reaction the megaprimer was created in a mixture of 10 μL of the 5X Primestar GXL buffer, 0.2 mM dNTPs, 1 μL of a 1/1000 diluted template plasmid DNA, 0.5 μM of the primers, 1 μL of the Primestar GXL DNA polymerase (Takara Bio Inc) and adjusted to a volume of 50 μL with milliQ water. In this thesis the reverse primer, oExtra347_pCX34_Bla_rev (sequence: 5'-CATCCGTCAGGATGGCCTTC-3' able to anneal about 400 bp downstream the end of the protein) was used for all mutants, while the forward primers are listed in Table 4 and Table 5. This mixture was subjected to an initial 30 seconds denaturation, followed by 30 cycles of a 10 second denaturation (98 $^{\circ}\text{C}$), a 15 second annealing and 1 minute per kilobase elongation (here it amounts to 1 minute 30 at 68-72 $^{\circ}\text{C}$). Finally, these 30 cycles were followed by an additional 2 minute elongation step and kept at 16 $^{\circ}\text{C}$ for the remainder of the time spent in the PCR-block. The entire protocol is represented in Table 2.

Table 2: Protocol for the megaprimer creation for a 2-step sanchis PCR

Cycles	Time	Temperature (°C)
1x	30 sec	98°C
30x	10 sec	98°C
	15 sec	55 or 60°C
	1 min 30 sec	68°C
1x	2 min	68°C
	∞	16°C

The second step of the protocol consisted of the actual amplification of the plasmid. The PCR mixture in this reaction consisted of 5 µL of the 10x PfuUltra HF reaction buffer, 0.2 mM dNTPs, ±15 ng template, 2 µL of the previously mentioned megaprimer PCR mixture, 1 µL Pfu Ultra DNA Polymerase AD (Agilent Technologies) and milliQ water to adjust to 50 µL. The protocol for the second step (represented in Table 3), started with a 30 second denaturation (95 °C), 30 cycles of: a 30 second denaturisation (95 °C), 1minute per kilobase annealing and elongation (72 °C). The protocol then ended with a final 5 minute elongation step at 72 °C and the samples were kept at 16 °C for the remainder of time they stayed at the PCR-block.

Table 3: Protocol for the plasmid amplification in the second step of the sanchis protocol

cycles	Time	Temperature
1x	30s	95°C
30x	30s	95°C
	5min 30sec	72°C
1x	5min	72°C
	pause	16°C

The PCR reaction mixture, still containing the starting template, was incubated with 2 µL DNPI at 37 °C for 1 h to degrade the methylated starting plasmid. The reaction mixture was purified with the Immuprep PCRpure kit by Analytic Jena following the manufacturer's instructions.

Table 4: Overview of the considered rational mutations and their respective primers (with the mutation underlined in the sequence)

Mutation	Primer name	Sequence (5' -> 3')
D186E	oExtra348_ScCels2_D186E_fwd	CAGTCCGGGTACAG <u>GA</u> AGGTGGTCACTATTATTGG
K53R	oExtra349_ScCels2_K53R_fwd	TGTCAGCTGGATGCAC <u>GT</u> ACCGGTACAGGTGC
L50D	oExtra364_ScCels2_L50D_fwd	CCTATCTGTGTCA <u>G</u> GATGATGCAAAAACCGG
A58S	oExtra365_ScCels2_A58S_fwd	AAACCGGTACAGGT <u>AGC</u> CTGGATCCGACC
A58K	oExtra366_ScCels2_A58K_fwd	ACCGGTACAGGT <u>AAA</u> CTGGATCCGACC
T62K	oExtra367_ScCels2_T62K_fwd	GCACTGGATCCG <u>AAA</u> AATCCGGCATGTC
T173D	oExtra417_ScCels2_T173D_fwd	GGAAGTATTCA <u>G</u> GATGTTACCAATCCGCC
T173R	oExtra418_ScCels2_T173R_fwd	GGAAGTATTCA <u>GCG</u> GTTACCAATCCGCCTC
A195S	oExtra419_ScCels2_A195S_fwd	TATTATTGGGATCTG <u>AGC</u> CTGCCGAGCGGTC
T175S	oExtra420_ScCels2_T175S_fwd	CTGATTCAAGACCGTT <u>AGC</u> AATCCGCCTCAGC
L170T	oExtra421_ScCels2_L170T_fwd	GGGATGATCTGGAA <u>ACC</u> ATTCAAGACCGTTACC
A94S	oExtra422_ScCels2_A94S_fwd	CCGGTGGTCTGGT <u>AGC</u> GGTTATGTTCC

Table 5: Overview of the considered disulphide bond inserting mutations and their respective primers (with the mutation underlined in the sequence)

Mutation	Primer name	Sequence (5' -> 3')
A143C	oExtra388_CelS2_A143C_fwd	GCAATTGGGCAT <u>TG</u> TATCCGGGTGATTTTCG
A52C	oExtra390_CelS2_A52C_fwd	CTGTGTCAGCTGGAT <u>TGT</u> AAAACCGGTACAGG
P61C	oExtra391_CelS2_P61C_fwd	ACAGGTGCACTGGAT <u>TGT</u> ACCAATCCGGCATG
T129C	oExtra394_CelS2_T129C_fwd	CCTCGTACCCATCTG <u>TGT</u> AGCGGTGCCACCATTCC
Q72C	oExtra396_CelS2_Q72C_fwd	CAGCCCTGGAT <u>TGT</u> AGCGGTGCAAC
S73C	oExtra398_CelS2_S73C_fwd	CTGGATCAGT <u>TGT</u> GGTGCAACCGCAC
T45C	oExtra400_CelS2_T45C_fwd	GCCTGGTAGCCGTT <u>TGT</u> TATCTGTGTCAGC
A105C	oExtra402_CelS2_A105C_fwd	CACCTGTGTAGT <u>TGT</u> GGTGATCGTAGC
P183C	oExtra389_CelS2_P183C_fwd	CCTCAGCAGGGCAGT <u>TGT</u> GGTACAGATGGTGG
L59C	oExtra392_CelS2_L59C_fwd	ACCGGTACAGGTGCAT <u>TGT</u> GATCCGACCAATCC
C48A	oExtra393_CelS2_C48A_fwd	AGCCGTACCTATCTG <u>GCG</u> CAGCTGGATGCAAAAACC
G230C	oExtra395_CelS2_g230C_fwd	GTGGTGTGGATGGT <u>TGT</u> AATGGTGAAGTTACCGG
A115C	oExtra397_CelS2_A115C_fwd	CCGTATGATTTTAGCT <u>TGT</u> TATAATGCAGCACGTAGCG
D112C	oExtra399_CelS2_D112C_fwd	CGATCGTAGCCCGTATT <u>TGT</u> TTTAGCGCATATAATGC
W81C	oExtra401_CelS2_W81C_fwd	CCGCACTGTATAAAT <u>TGT</u> TTTGCCGTTCTGGATAGC
F220C	oExtra403_CelS2_F220C_fwd	GCCAAGAAAACCTTCT <u>TGT</u> AGCTGCAGTGATGTGG
M39C	oExtra439_CelS2_M39C_fwd	ATGGTGTGCAT <u>TGC</u> ATGCCTGGTAGC
P41C	oExtra440_CelS2_P41C_fwd	TTGCAATGATG <u>TGC</u> GGTAGCCGTACCTATC
R92C+Y96C	oExtra441_CelS2_R92C_Y96C_fwd	AATGCCGGTGGT <u>TGT</u> GGTGCAGGT <u>TGC</u> GTTCCGGATGG
R125C	oExtra442_CelS2_R125C_fwd	AGCGATTGGCCT <u>TGT</u> ACCCATCTGACC
S104C	oExtra443_CelS2_S104C_fwd	CACCTGTGTT <u>TGC</u> GCCGGTGATCGTAG
V37C	oExtra444_CelS2_V37C_fwd	GAAGCACATGGT <u>TGT</u> GCAATGATGCCTGG
E137C	oExtra445_CelS2_E137C_fwd	CACCATTCCGGTT <u>TGC</u> TATAGCAATTGGG
T126C	oExtra446_CelS2_T126C_fwd	ATTGGCCTCGT <u>TGC</u> CATCTGACCAGCGG
D224C	oExtra448_CelS2_D124C_fwd	TAGCTGCAGT <u>TGT</u> GTGGTGTGGTGGTGG
S223C	oExtra449_CelS2_S223C_fwd	CTTTAGCTGCT <u>TGT</u> GATGTGGTGTGGTGGTGG
F113C	oExtra450_CelS2_F113C_fwd	GCCCGTATGAT <u>TGT</u> AGCGCATATAATGCAGC
S221C	oExtra447_CelS2_S221C_fwd	AACTCTTTT <u>TGT</u> TGCAGTGATGTGGTGG

2.2. Agarose gel electrophoresis and concentration measurement

The 1 % agarose gels (in 1X TAE buffer, ThermoFisher) were used to analyse DNA-fragments. Three μL 2-log DNA ladder (NEB) or 2 μL of the sample, mixed with 3 μL loading dye were separated on gel. When used to control the colony PCR with OneTaq Quick-Load 2x Master Mix (NEB), 5 μL of the PCR mixture was loaded without the need of adding extra loading dye. The electrophoresis was performed at a potential of 120 V for 25 minutes. Afterwards the gels were stained with ethidium bromide for half an hour, briefly rinsed and visualized using UV-light in a GelDoc system (Bio-Rad). Alternatively, the concentration of the fragments could be measured on the Nanodrop 2000 device (ThermoScientific).

2.3. Production of BL 21 (DE3) electrocompetent *E. coli* cells

Competent cells were created with the density gradient wash method developed by Warren et al. (Warren, 2011). A preculture of 5 mL Luria Bertani (1 % trypton, 0.5 % yeast extract and 0.5 % NaCl) was inoculated with BL21 (DE3) *E. coli* cells from the stock. The overnight grown preculture was then diluted to 1 % (v/v) in 350 mL sterile SOB medium, consisting of 2 % trypton, 0.5 % yeast extract, 0.05 % NaCl, 0.250 mM KCl and 5 mL of a 10 mM MgCl_2 solution per liter. The culture was incubated in a 2 L Erlenmeyer at 37 °C while shaking at 200 rpm until an optical density of 0.600 (at a wave length of 600 nm). Once the desired cell density was obtained, 300 mL of medium was divided over 6 falcon tubes and incubated on ice for 30 minutes. After incubation, the cells were pelleted through centrifugation (5000 g, 4 °C for 15 minutes) and the supernatant was discarded. The cells were resuspended in ice cold milliQ water and 10 mL GM (20 % (w/v) glycerol and 1,5 % (w/v) mannitol in water) was carefully pipetted underneath, to create a two-phase system. Hereafter, the falcon tubes were centrifuged at 2000 g and 4 °C for 15 minutes. The supernatant was then removed and the cells were resuspended in 500 μL GM, divided in 50 μL aliquots and flashfrozen. The cells were frozen at -80 °C until use.

2.4. *E.coli* transformation

To transform the desired construct into *E. coli* cells, 2 μL of the plasmid DNA was added to and mixed with 20 μL of thawed competent cells, while kept on ice. This mixture was incubated for 1 minute and transferred to chilled electroporation cuvettes. These cells were electroporated, with a Bio-Rad gene pulser and Bio-Rad pulse controller with the settings on: 2.0 kV, 200 Ω and 25 μF . Afterwards, 480 μL LB medium was added to the cell mixture, the total volume was transferred to a sterile 1.7 mL tube and incubated for 1 hour at 37 °C while shaking at 200 rpm. After incubation, 100 μL of the mixture was plated on a LB-agar plate containing 100 $\mu\text{g}/\text{mL}$ ampicillin for selection. The remaining medium was centrifuged for 1 minute at 2000 g, the pellet resuspended in 100 μL of the medium and plated on a second plate. Both plates were incubated overnight at 37 °C.

2.5. Colony PCR

In order to control the correct insertion of a plasmid into the cells, a selection of 4 colonies were transferred to a LB-agar backup plate and subjected to a colony PCR. Each reaction mixture of this colony PCR consisted of 11.5 μL milliQ water, 12.5 μL of the One Taq Master Mix, 0.5 μL of a forward (oExtra81-pCX34_fwd1 with sequence 5'-GCTCTAGAAGAAGCTTG-3') and reverse primer (oExtra189-pCXp_rev1 with sequence 5'-GTCTTGCGTCTTCGCCAGAC-3') and some cells of a single colony. The PCR protocol (represented in Table 6) started with an initial denaturation step of 7 minutes at 94 °C. This step was followed with 30 cycles of: a 20 seconds denaturation at 94 °C, a 20 seconds annealing at 52 °C and an elongation step of 1 minute at 68 °C. The final step of the protocol was an elongation of 7 minutes at 68 °C. The obtained PCR mixture was then analyzed through agarose gel electrophoresis as mentioned above.

Table 6: PCR program for the colony PCR

	cycles	Time	Temperature
initial denaturation	1	7 min	94°C
denaturation	30	20s	94°C
Annealing		20s	52°C
Elongation		1min 30sec	68°C
Final Elongation	1	7 min	68°C

2.6. Plasmid extraction and sequencing

After checking the correct insertion by colony PCR, 2 colonies per construct were grown overnight at 37 °C, while shaking at 200 rpm. The plasmids were extracted using the InnuPREP Plasmid Mini Kit (Analytic Jena). The elution however was performed by using 30 µL of milliQ water instead of the 50-100 µL elution buffer. The remainder of the protocol was executed as stated in the manual of the kit. The obtained plasmids were sent to be sequenced by LGC Genomics (Germany) or Macrogen (The Netherlands).

3. Enzyme production and purification

3.1. Growth

The prepared strains were grown at 30 °C, 200 rpm for 18-20 hours in 250 mL Terrific Borth medium containing 100 µg/mL ampicillin. Terrific broth (TB) contains 1.2 % tryptone, 2.4 % yeast extract, 0.4 % glycerol, 0.017 M KH_2PO_4 and 0.072 M K_2HPO_4 . The culture was inoculated with 1 % (v/v) of an overnight grown preculture.

In the optimization phase, growth at 20 °C and 37 °C were also evaluated and the commonly used LB medium was also evaluated instead of TB. The protein was isolated from 2 different fractions: the periplasm or the cytoplasm. Both protocols are described below. In order to determine the ideal conditions for growth, both Luria Bertani broth (LB) and Terrific Broth (TB) were considered. This LB medium consists of 1 % trypton, 0.5 % yeast extract, 0.5 % NaCl and 100 µg/mL ampicillin. Both these growth media were inoculated with the wild type strain and grown once at 30 °C in 250 mL culture volume using a 1 % inoculum of an overnight grown preculture. One culture of both the LB and TB were grown at 20 °C and another of each at 30 °C. After incubation the multiplied cells were harvested and the enzyme extracted.

3.2. Periplasmic isolation of the protein

This periplasmic isolation takes advantage of the difference in osmotic pressure between the periplasm and the buffers used for the isolation of the periplasmic fraction and the osmotic shock fraction. Those buffers consisted of 20 % (w/v) sucrose, 100 mM Tris-HCl (pH8) and 1 mM EDTA for the periplasmic fraction buffer and 5 mM MgCl₂ for the isolation of the osmotic shock buffer. The cells were pelleted by centrifugation at 3200 g for 10 minutes and the pellet was resuspended in 12.5 mL of the periplasmic fraction buffer (1/20 of the culture volume) and incubated on ice for 30 minutes.

This cell suspension was centrifuged at 4500 g for a duration of 20 minutes and the supernatant was stored as the periplasmic fraction. The pellet was then resuspended in the second buffer: in 12.5 mL (1/20 of the volume of the grown culture) of the osmotic shock buffer. After 20 minutes of incubation on ice, the suspension was centrifuged at 4500 g for 20 minutes. These 2 fractions together form the periplasmic extract to be used. The pellet can either be used for a subsequent cytoplasmic extraction or can be discarded.

3.3. Cytoplasmic isolation of the protein

The principle of cytoplasmic extraction is based on the disruptive effect of the soundwaves on the cellular membrane, to release the cytoplasm. The cell pellet of a culture (either way after periplasmic extraction or not) was frozen at -20 °C. After thawing the pellet, the cells were resuspended in 5 mL lysis buffer (1/50 of the culture volume). Lysis buffer consists of PBS at pH 7.4, 1 mg/mL lysozyme, 10 µL/mL PMSF and 10 mM imidazole. The cells were resuspended and kept on ice during 30 minutes. After this incubation the cells were sonicated (output 3, duty cycle 50) for 2.5 minutes to disrupt the cell walls. The cycle was repeated three times. To remove the cell debris, the suspension was centrifuged at 8000 rpm for 1 h (or when transferred in 1.7 mL Eppendorf tubes, they were centrifuged at 14000 rpm for 15 to 20 minutes) to remove the cell debris. The supernatant was saved for further purification.

3.4. Enzyme purification

The nickel affinity chromatography type of protein purification is based on the ability of histidine residues to interact with nickel. When the used vector codes for 6 or more consecutive histidine residues at the end of a protein, this part is called a his-tag. This is used to selectively purify the his-tag linked enzyme out of the periplasmic or cytoplasmic fraction, as this tag will adhere to nickel-chelated NTA-columns (Figure 13). Due to the necessity of the free N-terminal histidine in LPMOs, the 6 consecutive histidines were added to the C-terminus.

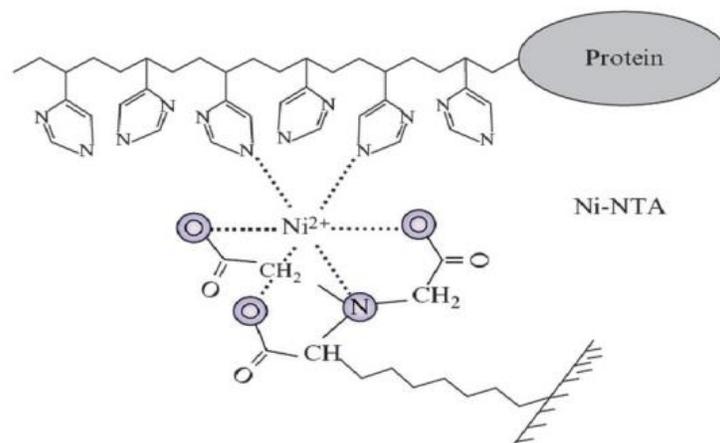


Figure 13: schematic representation of the principle of the his tag binding to the nickel chelated NTA-column (Rowinska-Zyrek, Danuta, & Slawomir, 2013)

For these purifications, 1.5 mL of the nickel NTA-agarose slurry (MC-lab) was added to 10 mL columns. This was washed with 6 mL of sterile distilled water and consecutively equilibrated with 3 volumes of native binding buffer. This buffer was composed of 10 mM imidazole in a PBS buffer (50 mM sodium phosphate and 300 mM sodium chloride at pH 7.4). The sample was bound to the resin by adding 8 mL each time and incubating at 4 °C for 30 to 60 minutes while rotating at a slow rate, where after the resin is settled by gravity and the suspension is aspirated. This step is repeated till the complete sample volume was loaded on the column. Next, the resin was washed three times with 8 mL of the native washing buffer containing 20 mM imidazole in PBS. By washing the column, competition of the imidazole with the histidine residues caused the unwanted histidine rich proteins to be released while the his-tag firmly binds the desired protein. The pure his-tag bound enzyme was then eluted with 10 mL of the elution buffer containing 100 mM imidazole. To remove the excessive amount of imidazole and switch to a more adequate buffer for storage, a buffer exchange was executed. This was achieved through repeated centrifugation of the eluate in spin columns (10000 MWCO PES, Vivaspin, Sartorius) with 50 mM sodium acetate buffer at pH 5.

For optimalization purposes, the concentration of imidazole in the washing buffer and elution buffer was varied. Furthermore, another type of elution was tested, based on the change of protonation at different pH. The elution was hereby induced with 4 different buffers with decreasing pH. All 4 buffers were 0.1 M citrate buffers with a pH of respectively 6, 5, 4 and 3. This method however was discarded, as it was more extensive and proved no more efficient.

3.5. Enzyme concentration

Upon concentration and buffer exchange, the concentration of the obtained enzyme solution was determined by measuring the absorbance. The Lambert-beer law states that the absorbance (A) can be determined by the extinction coefficient of the enzyme (ϵ), the concentration (c) and the distance through the sample (l):

$$A = \epsilon cl.$$

The absorbance was measured using Nanodrop 2000 device (Thermo Scientific), where both the molecular mass and extinction coefficient of the ScLPMO10C enzyme have been entered in the software. The extinction coefficient of the ScLPMO10C enzyme has been determined to be $75775 \text{ M}^{-1}\text{cm}^{-1}$ and the molecular mass was 35,39 kDa (obtained via Expasy tool ProtParam: <http://web.expasy.org/protparam/>).

3.6. SDS-PAGE

Sodium dodecyl sulphate – Polyacrylamide gel electrophoresis or shortly stated SDS-PAGE, was used to check the presence of our desired protein as well as the purity of the samples.

A 12 % resolving and 5 % concentrating gel was used for SDS-PAGE with self-made gels. The resolving gel consists of 12 % Acrylamide/Bis-acrylamide, 30 % solution (Sigma), 0.375 M Tris-HCl pH 8.8, 0.1 % SDS, 0.05 % ammonium persulfate and 0.05 % TEMED. The mixture was carefully mixed and transferred between the thin glass plate and spacer plate from Biorad. Isopropanol was added on top, to create an even upper edge on the resolving gel. After half an hour, the resolving gel was solidified and the isopropanol could be removed. The glass plates were carefully dried and the stacking gel could be loaded on top. This gel consisted of 5.1 % Acrylamide/Bis-acrylamide, 30 % solution (Sigma), 0.125 M Tris-HCl pH 6.8, 0.10 % SDS, 0.10 % APS and 5 μ L TEMED. This mixture was carefully mixed and added on top of the resolving gel and a comb was added to form 15 wells. After half an hour, the comb was removed from the gel and it was stored at 4 °C, wrapped in a moist paper towel until use.

To perform SDS-PAGE, the gels were placed in the gel electrophoresis unit (Mini- Protean Tetra Cell, Biorad) and submerged in running buffer (0.303 % Tris base, 1.44 % glycine and 0.1 % SDS). Before samples could be loaded, sample preparation required the production of Laemmli buffer (10 % β -Mercaptoethanol, in 62.50 mM Tris-HCl pH 6.8, 26.3 % glycerol, 2 % (w/v) SDS and 0.01 % (w/v) bromophenol blue. Ten μ L of the sample buffer has been added to 20 μ L of the sample and was boiled at 95 °C for 3-5 minutes. For a sample, 12 μ L of this mixture or 4 μ L of ladder (Pageruler Prestained protein ladder, ThermoScientific) was then pipetted into the wells and the electrophoresis was performed for 1 hour at 200 V.

The gel was removed from the spacer plate and was stained with the QC Colloidal coomassie stain form Biorad following the manufacturer's instructions.

4. Enzyme activity

To measure the activity, 150 μL of the obtained purified enzyme solution was diluted in 450 μL of a 1 mM sodium acetate buffer at pH 5 and mixed with 200 μL 10 mM ascorbic acid in buffer and 250 μL of an approximately 2 % Phosphoric acid swollen cellulose (PASC) solution. As negative control, a similar solution was prepared, where the 100 μL enzyme solution was replaced with an additional 100 μL of the 1 mM sodium acetate buffer at pH 5. These solutions were incubated at 37 °C at a shaking rate of 14000 rpm for 20 hours. The activity was terminated by heating the solution at 95 °C for 15 minutes, followed by a 15 minute incubation time on ice. Afterwards the denatured enzyme and the insoluble PASC were removed by centrifugation at maximum speed for 20 minutes. The remaining sample was analyzed through HPAEC-PAD following the protocol from Forsberg et al (Forsberg et al., 2011).

5. Enzyme stability with Differential Scanning Fluorimetry (DSF)

The stability of the mutants was evaluated by measuring the melting temperature (T_m) via differential scanning fluorimetry. This method is based upon the increasing accessibility of the fluorescent dye to the hydrophobic regions of the protein during the denaturation and a decrease upon aggregation of the denatured protein. This thus results in an increasing fluorescent signal upon denaturation, followed by the decreasing fluorescent signal upon aggregation, as visualized in Figure 14. This assay consisted of 20 μL sample and 10 μL 1/400 diluted Sypro Orange (Protein gel stain, Sigma-Aldrich). All these combinations were tested in triplicate. The measurement was performed in a qPCR cycler (Bio-rad CFX Connect) by excitation at 470/20 nm and emission at 570/10 nm while increasing the temperature from 25 $^{\circ}\text{C}$ to 95 $^{\circ}\text{C}$ with 1 $^{\circ}\text{C}/\text{min}$ increment.

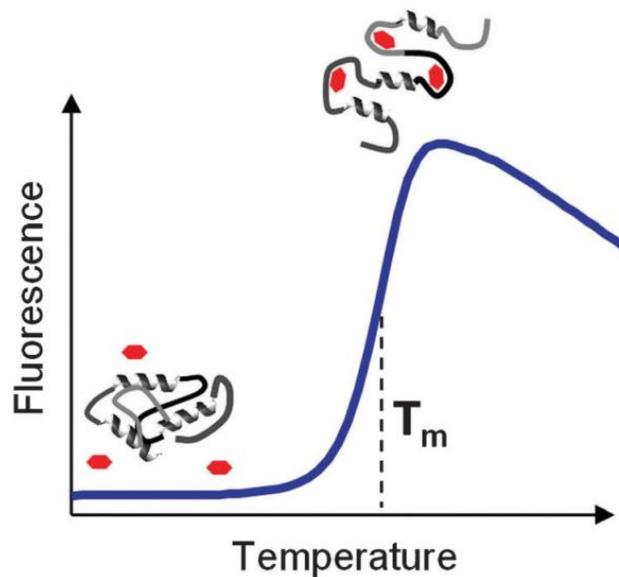


Figure 14: schematic representation of the principle of differential scanning fluorimetry (Rodrigues, Prosinecki, Marrucho, & Rebelo, 2011)

Results and discussion

1. Optimization of expression of ScLPMO10C

1.1. Determination of an expression system

The first objective in this thesis work is to express the enzyme ScLPMO10C in a thought-out way to obtain an active protein. This is a substantial challenge since the amino terminus is fixed and must be a histidine residue to result in an active protein, a feature that is quite rare. The reason is that this histidine and the actual amino-terminus are involved in binding a copper ion in the active site (see literature study section 3.3). The presence of this his-1 will be checked by activity measurement followed by HPAEC-PAD analysis. Furthermore, a descend protein yield and purity is required for testing activity and melting temperatures. The purity will be evaluated by SDS-PAGE and the apparent melting temperature will be checked by differential scanning fluorimetry (DSF).

Previous tests have narrowed the possible options to the use of 2 distinct expression systems in *E. coli* as schematically visualized in Figure 15.

In the first expression system, called LPMO10C-Xa, the protein is expressed in the cytoplasm like most proteins are expressed in *E. coli*. To ensure the correct N-terminus, a protease-restrictionsite (Xa-factor) precedes the N-terminal histidine so that this small part can be cleaved off manually after expression to obtain an active protein. The protein is expressed in *E. coli* Origami 2 (DE3) cells and the gene is encoded on a pET22b vector, harboring a T7-promotor for inducible expression with IPTG (strain was created in former work by Ir B. Danneels).

The second expression system, called LPMO10C-secr, makes use of a natural cleavage system to ensure the his-1, namely the use of its natural secretion signal that transports the protein to the periplasm (and later out of the cell). This protein is expressed in *E. coli* BL21 (DE3) cells and the gene is encoded on a pCXp34 vector (Aerts & Desmet, 2011), resulting in a constitutive expression system.

Both the plasmids used for the ScLPMO10C-Xa and the ScLPMO10C-secr also encode a his-tag, linked to the carboxyl terminus to allow purification using Ni-NTA resin.

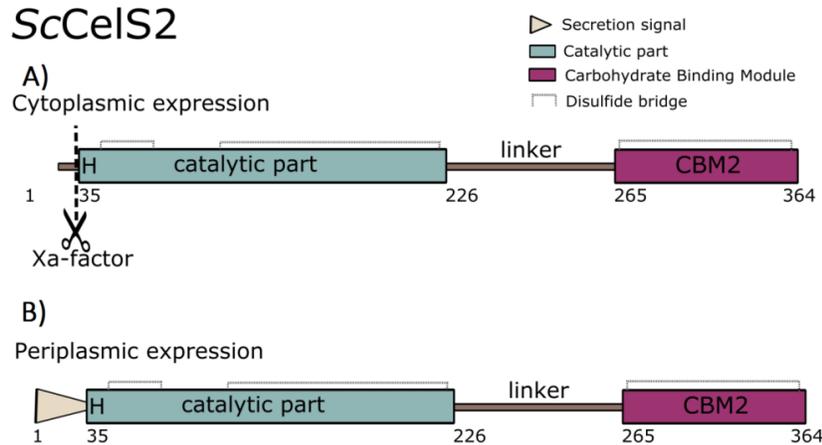


Figure 15: Schematic representation of the considered expression systems. (A) *ScLpmo10C-Xa*: The cytoplasmic expression of the enzyme, linked with a Xa-factor at the N-terminal end. (B) *ScLpmo10C-secr*: The periplasmic expression of an enzyme with a guaranteed correct N-terminus through the natural cleavage of the secretion signal at the N-terminal end upon secretion to the periplasm.

Firstly, the Xa-cleavable construct will be compared to the use of the secretion signal and the preferred fraction to isolate the protein from. Therefore, 3 enzyme preparations can be compared: (1) *ScLPMO10C-Xa* as isolated from the cytoplasm, (2) *ScLPMO10C-secr* isolated from the cytoplasm and (3) *ScLPMO10C-secr* isolated from the periplasm. The applicability of the second option depends on whether or not it will show activity, since the secretion signal might still be attached.

As first parameter, the expression level and purity during and after his-tag purification needs evaluation. The first construct, *ScLPMO10C-secr*, contains a secretion signal to send the protein out of the cell. The protein passes through the cytoplasm in any case and therefore, a protein isolation from the cytoplasm and periplasmic space was done. Both fractions were his-tag purified and their SDS-PAGE analysis is shown in Figure 16. As clearly visible, the enzyme was both expressed in the cytoplasm and the periplasm. However the concentration of the *ScLPMO10C-secr* enzyme present in the cytoplasm slightly exceeds the concentration in the periplasmic fraction.

The *ScLPMO10C-Xa* strain was only expressed in the cytoplasm (represented in Figure 17). This SDS analysis clearly shows an abundance in unwanted proteins in the insoluble fraction and flow through, that were not entirely removed during the purification. Since none of the produced *ScLPMO10C-Xa* was transported out of the cell this led to a significant enzyme build-up, resulting in a higher concentration in the obtained cytoplasmic fraction than in the case of the *ScLPMO10C-secr*. As a result, the elution and especially the concentrated sample of the elution showed a high amount of the *ScLPMO10C* as well as the presence of several smaller proteins.

Furthermore, this gel (Figure 17) and the SDS-analysis of the ScLPMO10C-secr (Figure 16) also visualize the difference in the amount of proteins present at the start of the purification of a cytoplasmic and a periplasmic fraction. The periplasm contained much less unwanted proteins than the cytoplasm, what makes it easier to obtain a pure protein.

The size of the protein in the LPMO10C-secr strain as observed from SDS-PAGE analysis implies the presence of the enzyme with an attached secretion signal in both the periplasmic and cytoplasmic fraction. While the enzyme has a molecular weight of 35,4 kDa, the position of the bands seems to be located rather at the height of 40 kDa, than 35 kDa, comparing to the reference ladder. However, only the cytoplasmic fraction was expected to contain a heavier protein because the secretion signal might still be attached to the protein, resulting in a 38 kDa protein.

Since the activity requires a free N-terminal histidine, only the secretion signal free fraction should be able to show enzyme activity. As such, the first impression of the gel would suggest the presence of only inactive enzymes. However, MS-analysis on a sample of LPMO10C-secr from the cytoplasmic and periplasmic space, equaled the mass of LPMO10C without secretion signal. This would be expected from the periplasmic fraction, yet proves to be the case for the cytoplasmic fraction as well. Despite the first impression of the SDS-gels, this would thus suggest the presence of active protein in both the cytoplasm and periplasm.

Furthermore, the obtained mass from the MS-analysis was 6 Da lower than the theoretical molecular weight. A result that can be explained by the loss of 6 H-atoms due to presence of 3 disulphide bonds. These were considered unlikely, as the prokaryotic host was not expected to be able to create an oxidative environment in the cytoplasm, required for the formation of disulphide bonds. Only the periplasm is known to provide an oxidative environment and the required chaperones. Further examination of the literature however learns the possibility of the formation of intracellular disulphide bonds in bacterial hosts has already been reported (Mallick, Boutz, Eisenberg, & Yeates, 2002). The same source suspects the intracellular formation of disulphide bonds to result from potential differences in redox environment or redox mechanisms, making the formation of these disulphide bonds energetically favorable.

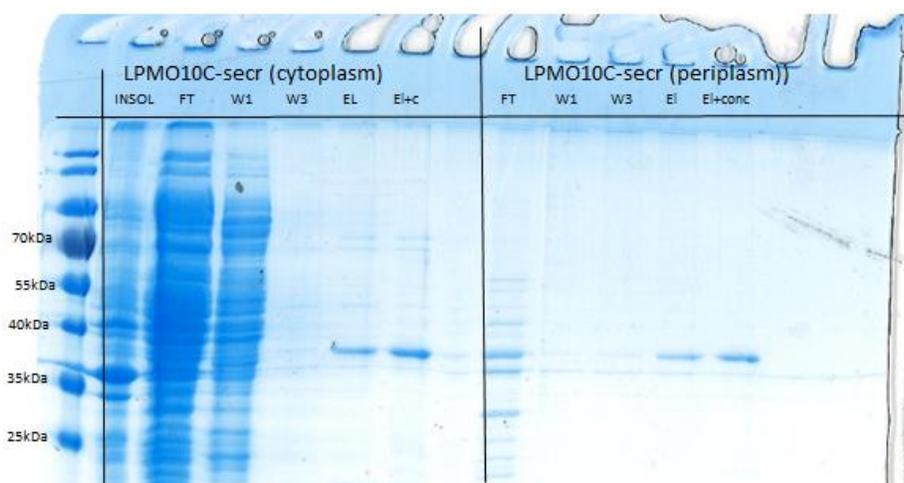


Figure 16: SDS-PAGE analysis of the purification from the cytoplasmic and periplasmic fraction of LPMO10C-secr, comparing the proteins collected during the different purification steps . (Insol = insoluble fraction , FT= flow-through, W= wash steps (1 and 3), EL= elution step, EL+conc= eluted and concentrated enzyme after buffer exchange)

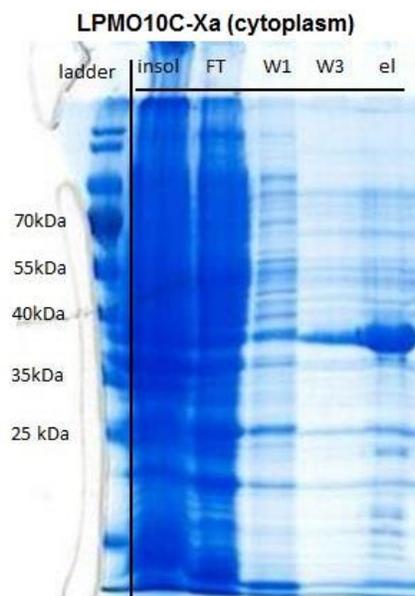


Figure 17: SDS-PAGE analysis of the purification from the cytoplasmic fraction of ScLPMO10C-Xa fraction, comparing the proteins collected during the different purification steps . (Insol = insoluble fraction , FT= flow-through, W= wash steps (1 and 3), EL= elution step, EL+conc= eluted and concentrated enzyme after buffer exchange)

As second parameter, the activity on the substrate Phosphoric Acid Swollen Cellulose (PASC) was measured after incubation at 37 °C and analysis of the HPAEC-PAD profiles. For an active C1 oxidizing LPMO as ScLPMO10C, a peak pattern with native and C1-oxidized cello-oligosaccharides, as described by Forsberg et al (Forsberg et al., 2011)) is expected.

When applying this on the ScLPMO10C-Xa, the inevitability of the his-1 was proven, as Figure 18 visualizes the differences in the chromatogram between the cleaved and uncleaved ScLPMO10C-Xa strains. While the peaks from the polymers of 2 to 5 beta-1,4 bound glucose units (indicated G2 till G5) can still be recognized, the rest of the typical pattern of aldonic acids (indicated with *) is absent.

Since these molecules are to be created by the C1 oxidation, their absence confirms the presumption that the uncleaved ScLPMO10C-Xa strain shows no activity or for sure a seriously distorted activity when incubated with PASC. Furthermore, these peaks are much smaller than the ones for the Xa-cleaved variant. The Xa-cleaved protein did show the expected pattern with native and oxidized products, these results thus confirm the activity of the ScLPMO10C-Xa on the condition that the Xa-factor has correctly been cleaved.

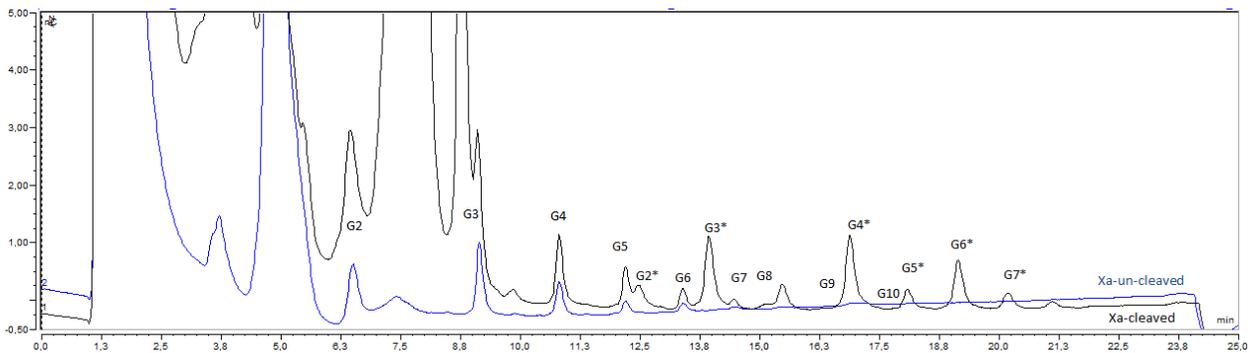


Figure 18: The HPAEC-PAD analysis of the Xa-cleaved enzyme compared to the uncleaved fraction of the same ScLPMO10C-Xa (both after 20h incubation with PASC). The labels indicate the amount of glucoses linked so G2 represents cellobiose and so on. The asterisk (*) indicates the aldonic acid form.

Furthermore, the results from the expression analysis on SDS-PAGE and accompanying MS-data confirmed the correct N-terminus for LPMO10C-secr both in cytoplasm and periplasm. The activity test indeed showed a positive result for both protein fractions. Thus proving the secretion signal was correctly removed and the resulting enzyme is active and contains the required his-1. An overlay of the chromatograms of the different enzyme fractions from both the expression approaches was made to evaluate the options. This overview is represented in Figure 19.

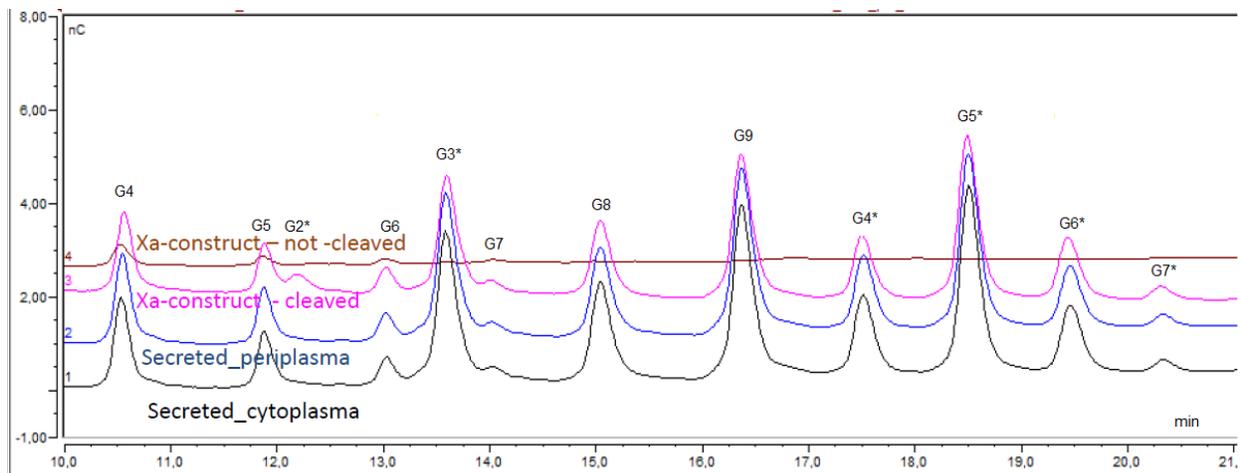


Figure 19: An overlay of the results of the different fractions of both strains, where the uncleaved ScLPMO10C-Xa is represented by the brown, the cleaved ScLPMO10C-Xa by the pink, the periplasm fraction of the ScLPMO10C-secr by the blue and the cytoplasmic fraction of the ScLPMO10C-secr by the black chromatogram.

As a general conclusion for the activity test with the HPAEC-PAD, it is clear that both the ScLPMO10C-secr and the ScLPMO10C-Xa strain produced an active ScLPMO10C. Even though a much higher protein yield was reached for ScLPMO10C-Xa, the increased process duration and extra cost due to the additional cleavage and purification step made the Xa-linked strain less interesting to work with. Despite activity for both the fractions and a slightly higher expression yield for the cytoplasmic fraction, the periplasmic fraction proved to be the preferred option. This choice was mainly based on practical considerations, namely the easier downstream processing, as the abundance of unwanted proteins in the cytoplasmic fraction seriously hindered the purification by clogging the filter columns. Since the obtained concentrations still sufficed for the planned tests, the slightly lower yields could not outweigh the benefits from selecting to work with the periplasmic fraction.

1.2. Optimize the purification protocol

All previously mentioned purified enzyme fractions were obtained through Immobilized Metal Affinity Chromatography (IMAC), based on the interaction between chelated nickel ions, the his-tag bound to the enzymes and imidazole from the buffers. In order to optimize the downstream processing of the enzymes, a second elution method based on the different protonation of the his-tag in different pH conditions has been considered. In order to evaluate this method, purification of a 250 mL culture of the wild type enzyme was performed with elution through a pH gradient ranging from pH 6 to 3. The elution through pH-gradient proved to be more work intensive, yet yielded no benefits in regard to obtaining a purer sample or higher concentration.

Since the pH elution offered no advantages, this approach was abandoned and further tests were performed to determine the optimal imidazole concentrations in the wash and elution buffers for the imidazole purification method. In order to determine the ideal imidazole concentration in the wash buffer and elution buffer, a new purification of a 250 mL culture was performed. During this purification the enzyme was eluted through an increasing imidazole gradient: starting at 10 mM imidazole in the equilibration buffer and stepwise increased with 10 mM up to a concentration of 60 mM, after which the concentration was increased to 100 mM and finally 250 mM for elution. All samples were analyzed via SDS-PAGE (visualized in Figure 20). This gel clearly shows a 20 mM wash already removed most unwanted proteins but the protein of interest also slightly eluted already (the band observed between 35 and 40 kDa). The 30 mM fraction still contained a small unwanted protein (about 10 kDa, circled on Figure 20). Since this protein is no longer visible on the 40mM fraction, this imidazole concentration in the wash-buffer would be advisable for higher purification. This preference is based on the observation that washing with a 40 mM instead of 20 mM would result in a higher purity, yet this would also increase the loss of enzyme during the purification. As for the imidazole concentration in the elution buffer, the gel clearly shows most of the enzyme had already eluted at a concentration of 60 mM (as indicated on Figure 20 by an arrow). However a significant amount of enzyme could still be detected in the 100mM fraction, while the 250 mM fraction was completely void of the enzyme.

To conclude, the optimal purification conditions that will be used in all further experiments uses 10 mM imidazole for equilibration and at least 20 mM imidazole for wash and 100 mM imidazole for elution, considering most of the protein is already eluted from 60 mM imidazole concentration.

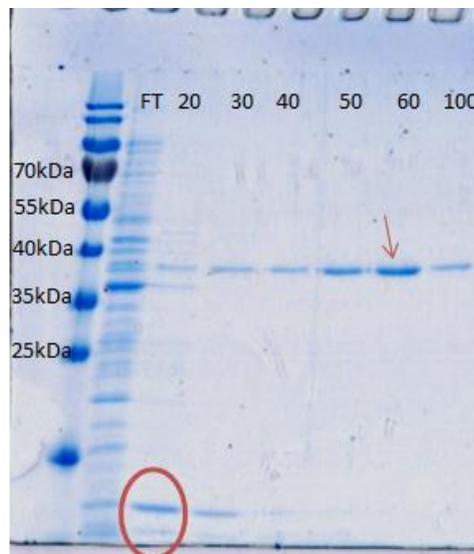


Figure 20: SDS-PAGE analysis comparing the protein collected during the purification steps: the flow-through, and the different fractions from elution buffers containing imidazole concentrations ranging from 20 to 100 milliMolair.

2. Optimization of assay

Upon finishing the optimization of the expression and purification of the enzyme, the next concern for the adequate evaluation of any created mutants is to find a suitable measurement method to identify the positive mutants.

2.1. Activity tests

The first aspect of the enzymes to be verified is the activity. As the goal of this thesis consists of improving the thermostability for better use of this LPMO in industrial sized reactions, the conservation of its activity after mutagenesis is of utmost importance. In order to determine the effect of the mutation on the activity, two assays were proposed based on literature. The first test includes a colorimetric assay, based on the production of hydrogen peroxide (Kittl, Kracher, Burgstaller, Haltrich, & Ludwig, 2012). The assay was firstly described in 2012 and includes an unprofitable side reaction with reducing activity converting O_2 to H_2O_2 when in contact to an electron donor. Indeed, this reaction does not imply the natural substrate cellulose. The mechanism of this reaction, represented in Figure 21, is based on the linear relation between the concentration of the H_2O_2 production and the concentration of the LPMO in the reaction mix. Literature reports that the use of a horseradish peroxidase coupled reaction on Amplex red for the visualization of the produced hydrogen peroxide, leads to a fast and robust assay for the determination of LPMO activity. When applying this to the his-tag purified ScLPMO10C, it could not be implemented in our work. Standard curves with H_2O_2 could be measured appropriately and also another LPMO, that was ultra-filtrated but not his-tag purified, could be analyzed without any problems. For a still unknown reason, any and all tests with different concentrations of the wild type ScLPMO10C, resulted in a decreased fluorescent signal when compared to the blank samples. The effect of PBS buffer and of imidazole were evaluated but did not show to be the cause. Due to an unknown inhibiting phenomenon, this approach for the evaluation of the enzyme was abandoned.

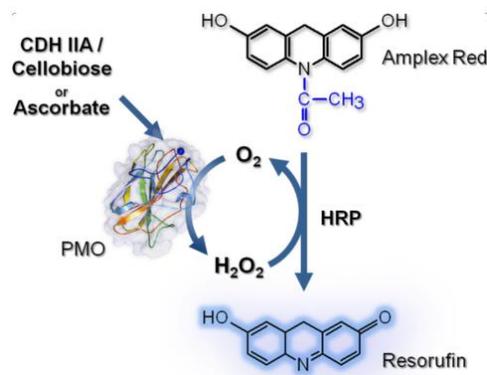


Figure 21: schematic representation of the principle behind the H₂O₂ assay. the reduction of the type-2 copper ion of the LPMO leads to the one electron reduction of the oxygen. The presence of the Resulting hydrogen peroxide, starts the HRP coupled reaction converting amplex red into the fluorescent Resorufin.

The second method for assessing the activity of LPMOs proves more work intensive, yet measures the actual activity on a cellulose based substrate and is therefore a more realistic test. This approach thus not only measures the reducing capabilities but also confirms the conservation of the substrate specificity and affinity. This activity test is based on the HPAEC-PAD analysis and identification of the reaction products upon incubation of a cellulose based substrate with the enzyme to be tested and ascorbic acid as electron donor. As explained in the discussion of Figure 18, when selecting the appropriate expression system, the obtained chromatogram proved high resemblance to the chromatogram reported by Forsberg et al. in 2011 for activity of LPMO10C (Forsberg et al., 2011).

2.2. Measurement of apparent melting temperatures

While the detection of the degradation products of cellulose by HPAEC-PAD provided sufficient data for the evaluation of the enzyme activity, a second enzyme characteristic needed to be analyzed in the screening for successful thermostable mutants. To determine the melting temperature, a differential scanning fluorimetry assay (also called a thermal shift assay) was tested and optimized. Using a fluorescent dye (Sypro Orange) and an increasing temperature, the intensity of fluorescence corresponds with the quantity of denatured protein present at that temperature. From the increasing fluorescence, the apparent melting temperature can be determined as the inflection point of the curve. Upon calculating the first derivative of the measured fluorescence units, the minimum of the negative derivative corresponds to the apparent melting temperature.

To assess the variability of the assay and the effect of a difference in purity and enzyme concentration on the thermal shift assay, a set of wild type enzyme cultures was grown and purified. Three of these cultures were purified using a wash buffer containing 40 mM imidazole, while two more were purified using a wash buffer containing 20 mM. The former resulting in a higher enzyme concentration but as a downside also a lower purity, while the first resulted in a lower enzyme concentration, but a higher purity as explained in part 1.2 of the result section: the optimization of the purification protocol. In order to gain insight into the effect of the concentration on the melting temperatures, the same samples from the purification with the 20 mM wash buffer were diluted to the same concentration as the samples washed with 40 mM imidazole (seemed about 3 times dilution). The apparent melting temperature of the resulting enzymes, separately purified for each Erlenmeyer, were determined in triplicate (see Figure 22 and Table 7). An apparent melting temperature of 51 °C was obtained for all samples that were washed with a concentration of 40 mM imidazole. All samples from the purification with the 20 mM wash buffer on the other hand resulted in a melting temperature of 50 °C. This 1 °C difference in the melting temperature, thus suggests a difference in purity or concentration influences the measurement. The diluted samples however, did result in an apparent melting temperature of 51 °C.

Since this melting temperature is an exact match to the melting temperature of the 40 mM imidazole washed samples, it shows a 1°C difference to the sample it was diluted from. This suggests a small effect on the apparent melting temperature depending on the concentration of the enzyme. Seeing as the 20 mM washed samples resulted in a much higher concentration (and as a result in a higher signal on the DSF) and the purity proved to suffice, further purifications will be performed using a 20 mM imidazole wash.

This set of data (represented in Table 7) thus learns that in order to apply this test to the determination of an improvement in melting temperature, the reference wild type needs to be produced at the same time, purified with the same imidazole concentrations and all samples will be diluted to the same enzyme concentration prior to the thermal shift assay. This because we learned here that enzyme concentration and purity can have a small influence. When taking these requirements in to account, this methodology shows an extremely low variability and thus a suitable method to distinguish positive mutants.

Table 7: overview of the melting temperature of samples of different purity and concentration.

	Imidazole concentration of the washing step	Concentration (mg/mL)	T _m (°C)			mean	stdev
Erlenmeyer 1	40 mM	0,089	51	51	51	51	0
Erlenmeyer 2	40 mM	0,073	51	51	51	51	0
Erlenmeyer 3	40 mM	0,086	51	51	51	51	0
Erlenmeyer 4	20 mM	0,298	50	50	50	50	0
Erlenmeyer 5	20 mM	0,263	50	50	50	50	0
1/3 dilution from Erlenmeyer 5	20 mM	0,088	51	51	51	51	0

---20 mM washed sample ---40 mM washed sample --- 1/3 dilution of the 20 mM washed sample

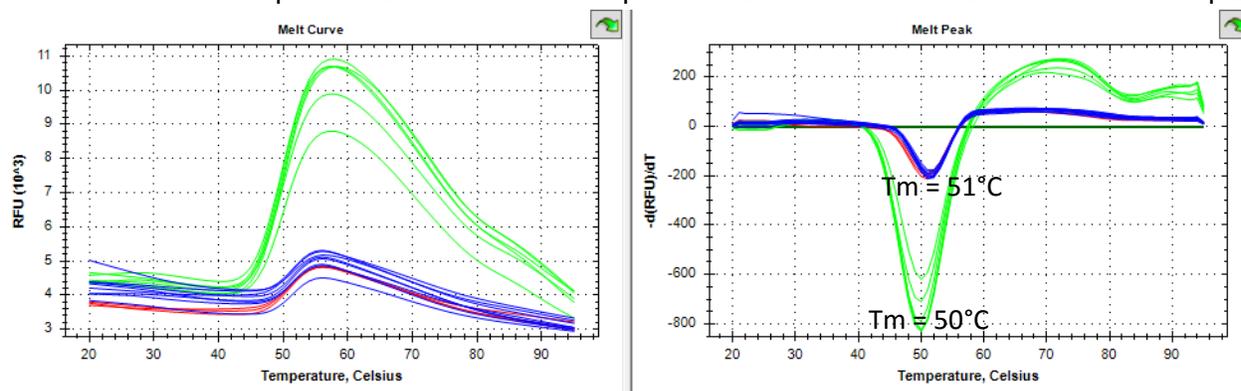


Figure 22: DSF analysis of the differently purified samples of LPMO10C. The different colors indicate a difference in concentration imidazole in the washing buffer: 40 mM (blue curve), 20 mM (green curve and red curve). The samples of the red curve were diluted to the same enzyme concentrations as the blue curves. The left side is a representation of the melting curve. The right side is a visualisation of the derivative.

3. Increasing Protein stability

3.1. Importance of the N-terminus

While literature provides abundant information on the importance of the correct N-terminal end for the activity of LPMOs, little is reported on the importance of that N-terminus for the enzyme's thermostability. In order to assess the potential influence of the free histidine on position 1, the melting temperature of the previously mentioned LPMO10C-Xa (as explained in part 1.1 of the results section: the determination of an expression system) was determined in triplicate before the cleavage of the Xa-factor on the N-terminal end (Figure 15). The obtained melting temperature of 43°C was significantly lower than the 50°C obtained with a correct N-terminus (Figure 23). This decrease in melting temperature indicates a lower thermostability. As a possible hypothesis, the interaction between the free histidine on position 1 and the copper binding site could not only be crucial to the activity but could also contribute to the conservation of the overall fold of the enzyme.

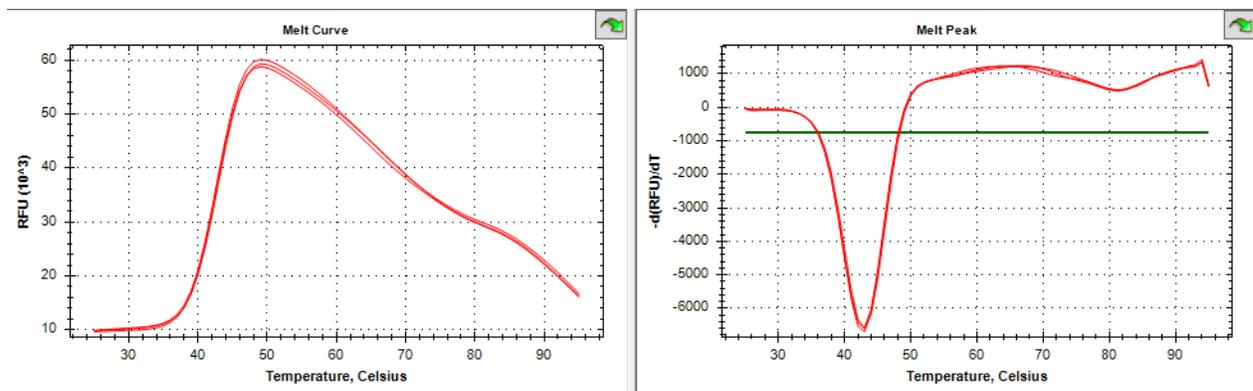


Figure 23: DSF ANALYSIS OF THE LPMO10C-Xa SAMPLE before the cleavage (no his-1). The left site of the figure is the visualization of the meltingcurve, as the right side is the derivative of the melting curve.

3.2. Mapping the flexible regions in ScLPMO10C via B-fitter

Rational engineering of thermostability in enzymes requires knowledge on the structural and functional properties of that enzyme, in order to determine interesting locations for mutagenesis. In order to stabilize the most vulnerable positions in the enzyme, the B-fitter method as described by Reetz et al was used to find the most flexible positions (Reetz et al., 2006). Using measurements of atomic displacement parameters obtained from X-ray data, the method assigns every amino acid with a B-factor to reflect a measure of localized instability inside a protein.

This value is in accordance to the smearing of atomic electron densities in comparison to their equilibrium positions due to thermal motion and positional disorder (Reetz et al., 2006). Thus the locations with the highest value correspond with the amino acids to move most, these are reasoned to be the most sensitive to start the denaturation (Aerts et al., 2013; Wijma, Floor, & Janssen, 2013; Xie et al., 2014). When considering the 20 amino acids with the highest B-factor, several clusters can be observed (as visualised in yellow on Figure 24 and Figure 25) to correspond with flexible loops in the protein structure. Loops are prone to be more flexible so this is no surprise.

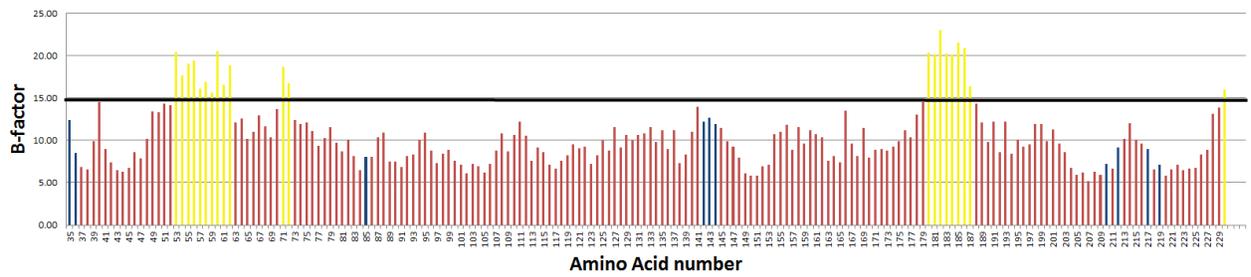


Figure 24: representation of the B-factors in function of the location of the amino acid. The 20 amino acids with the highest B-factor are represented in yellow, the amino acids at a distance shorter than 6 angstrom are represented in blue and the other amino acids in red.

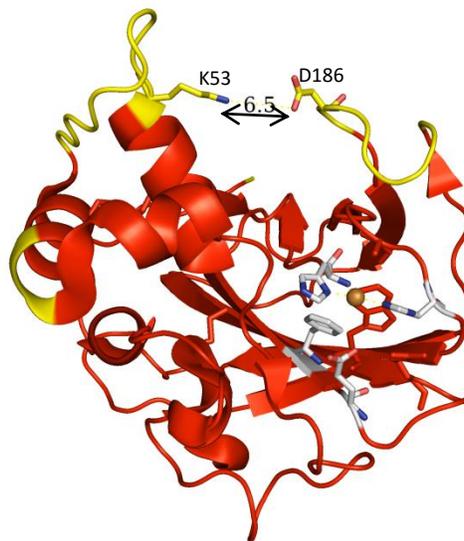


Figure 25: schematical representation of SclPMO10C, where The yellow loops represent the two most flexible regions and the bleu part are the amino-acids close to the copper-ion (<6Å).

3.3. Rational mutants:

The combination of the information on the copper binding site (from literature) and the locations of the most flexible positions as determined through the B-fitter method, provides the basis to design a more stable enzyme variant in a rational way.

As visualized in Figure 25, the most flexible regions are clustered in two loops at the enzyme surface. A first strategy to stabilize ScLPMO10C is to fixate both loops by facilitating the formation of a salt bridge between the two closest residues. As indicated on Figure 25, these residues are at a distance of 6.5 Angström in their crystalline structure. While these residues are oppositely charged, this distance seems a little too large to indeed form an interaction. An attempt to facilitate the formation of a salt bridge, was based on the decrease of this distance by using an amino acid with a larger side chain, while maintaining the opposite charges (the possible amino acids are represented in Figure 26). In order to achieve this, the aspartic acid (D) on position 186 was replaced by a Glutamate (E) and the lysine (K) on position 53 was switched with an Arginine residue (R), abbreviated to D186E and E53K respectively. Both these mutations were introduced individually and combined. As reported in Table 8 (inside the yellow rectangle), only the mutation K53R seems to yield a 1° C improvement when compared to the wild type ScLPMO10C. However, this is not a significant difference. When combining this mutation with D186E, the beneficial effect disappeared. As possible hypothesis for the explanation, it can be reasoned that the combined mutations resulted in a less favourable distance than the one single mutation. The results can also be explained if the initial slight increase was not even the result of the actual formation of a salt bridge in the first place, rather than the result of interaction with other amino residues.

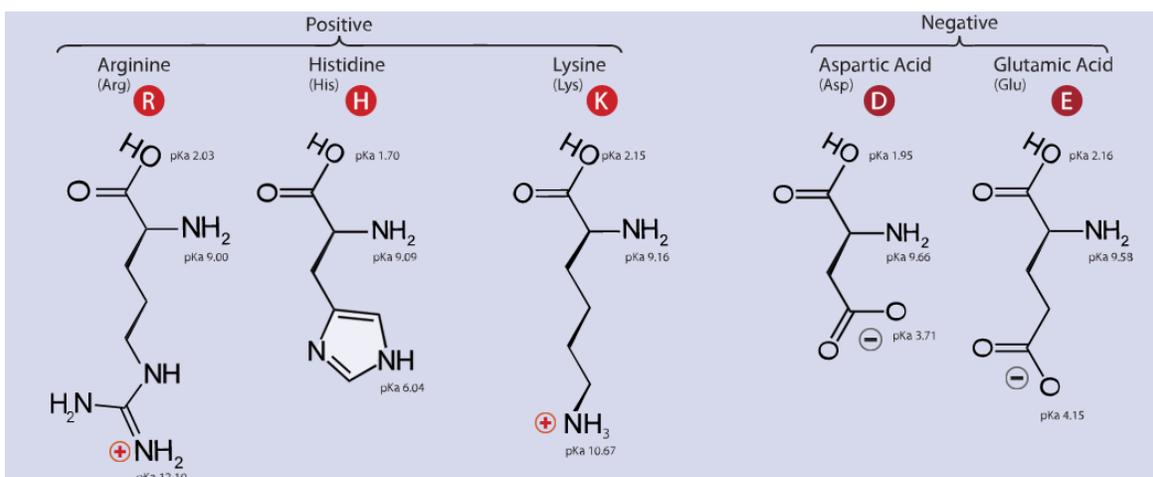


Figure 26: representation of the electrically charged amino acids

Other rational mutants were considered by visual inspection of the protein's surface. The aim was to introduce interactions like ionic networks and avoid apolar groups at the surface (preferably at flexible positions).

A complete list of the designed mutants can be found in Table 8, with their respective melting temperature and whether their activity on PASC was maintained. As example, the Leucine at position 50 might make some hydrophobic interactions with Y46, yet by introducing a aspartic acid instead of the leucine (mutant L50D), a salt bridge with K53 or Q49 might result in a more stable protein (the concerning amino acids are visualized in Figure 27).

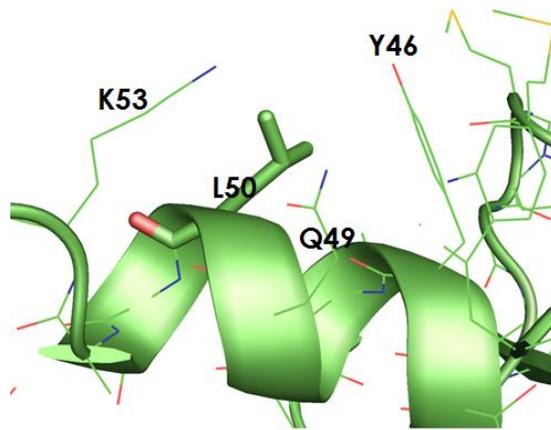


Figure 27: representation of the amino residues involved in one of the considered mutations

Table 8: Overview of the created rational mutants. Their mutation(s), melting temperature and whether or not they are active is indicated in the table.

Nr	Mutation	T _m (°C)	Active?
	WT	51 ± 1	V
1	<u>K53R</u>	52 ± 1	V
2	<u>D186E</u>	51 ± 1	V
3	<u>K53R + D186E</u>	51 ± 1	V
4	<u>A58S</u>	50 ± 1	V
5	<u>L50D</u>	45 ± 1	V
6	<u>T62K</u>	49 ± 1	V
7	<u>A58K</u>	51 ± 1	V
8	<u>T173D</u>	48 ± 1	V
9	<u>T173R</u>	47 ± 1	V
10	<u>A195S</u>	50 ± 1	V
11	<u>T175S</u>	51 ± 1	V
12	<u>L170T</u>	45 ± 1	V

Similar to the results of the mutations designed to fixate the loops, no significant increase in apparent melting temperature could be detected for the other rational mutants. Four showed no significant change in apparent melting temperature and five mutants, such as the previously explained L50D mutation, even resulted in a significant decrease in apparent melting temperature of up to 6 °C.

Despite the rational selection of the mutations, the desired interactions either do not correctly occur or disrupt other interactions with a stronger impact on the enzyme stability.

3.4. Disulphide bridges:

In order to stabilize the wild type ScLPMO10C, one approach is the stabilization of the flexible regions of the protein through the introduction of disulphide bridges. While the formation of the new disulphide bonds might not be achieved in the cytoplasmic fraction, the previous choice to work with the periplasmic fraction should support this strategy.

Disulphide bonds are generally considered to be of great importance. Both in earlier research in the rational engineering of thermostable proteins (Matsumura, Signor, & Matthews, 1989) and in the examination of enzymes in thermophilic bacteria (Mallick et al., 2002), the importance of disulphide bonds has been well established. As these interactions are strong covalent bonds, the additional energy required to break them in reductive environments could add to the stability of the enzyme. The basis of the determination of these mutants are two computational tools, namely “MOdelling of Disulphide bridges In Proteins” (MODIP) (Dani et al., 2003; Sowdhamini, R., Srinivasn, N., Shoichet, B., Santi, Daniel, V., Ramakrishnan, C. & Balaram, 1989) and “disulphide by design” (DbD) (Craig & Dombkowski, 2013). While both tools are designed to determine the amino acids at positions in the protein where the geometry theoretically meets the stereochemical criteria for the formation of the disulphide bonds when it's replaced by a cysteine, the differences in algorithms result in different lists of possibilities. While MODIP results are divided in a list A till D based on their likeliness to succeed, the DbD tool rendered but one list. From all these possible mutations, 17 were selected as they were predicted by both tools. More specifically, the selection consists of all mutations predicted by the DbD tool that were either on the A list from the MODIP tool (part A on Figure 28) or, if they were on the B list they should at least include one residue from the top 20 most flexible residues as obtained from the B-fitter tool. Furthermore, all involved residues must be at a distance greater than 6Å from the copper binding site in order not to disrupt the copper active center (part B on Figure 28). One exception was made as a mutant was proposed by DbD, that did not fit the requirements for MODIP.

Another interesting case was made to gauge the effect of a redirection of a native disulphide bond. In the wild type, C66 was linked to C48. Following the predicted linkage, C66 will be linked to P61C. Therefore, the native cysteine at position 48 was mutated to an alanine (C48A). In total, a list of 17 possible mutants was to be considered (represented in Table 9).

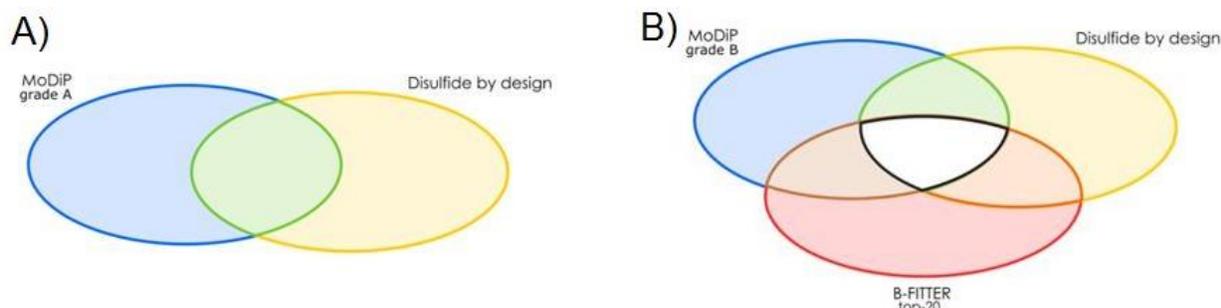


Figure 28: schematic representation of the selected mutations. A) A first selection was made based on the prediction by DbD and an A grade success rate by MoDiP. B) A second set of disulphide bonds was selected for amino acids among the 20 flexible positions that were proposed by DbD and the B grade list of MoDiP.

Table 9: overview of the considered disulphide bridge introductions: their mutations, whether or not they are active and their melting temperature

Nr	Mutation 1	Mutation 2	MODIP-grade	Disulfide by Design	Top20 B-FIT	>6A from Cu	Tm (°C)	Active
1	A143C	P183C	A	X	X	X	54 ± 1	V
2	A52C	P61C	A	X	X	X	56 ± 1	V
3	A52C	L59C	B	X	X	X	45 ± 1	V
4	P61C	C66 (native!)	B	X	X	X	42 ± 1	V
5	T129C	G230C	-	X	X	X	45 ± 1	V
6	Q72C	A115C	B	X	X	X	49 ± 1	V
7	S73C	D112C	A	X	-	X	54 ± 1	V
8	S73C	A115C	A	X	-	X	60 ± 1	V
9	T45C	W81C	A	X	-	X	44 ± 1	V
10	A105C	F220C	A	X	-	X	46 ± 1	V
11	M39C	E137C	A	X	-	X	47 ± 1	V
12	P41C	T126C	A	X	-	X	Stuck in molecular work	
13	R92C	Y96C	A	X	-	X	Stuck in molecular work	
14	R125C	D224C	A	X	-	X	40 ± 1	V
15	P41C	S223C	A	X	-	X	Stuck in molecular work	
16	S104C	F113C	A	X	-	X	Stuck in molecular work	
17	V37C	S221C	A	X	-	X	44 ± 1	V

Among all seventeen mutants, four show a significant increase in melting temperature. While the disulphide bridges between the cysteines at positions A143C-P183C and S73C-D112C lead to an increase of 3 °C, the disulphide introduced between the positions A52C and P61C leads to an increase of 5 °C and the additional disulphide bond between the positions S73C and A115C yields an even bigger increase of 9 °C. Since four of the mutants are stuck in the molecular phase, only 13 of the mutants were successfully expressed and purified in this thesis.

The 4 positive results on 13 tested strains thus means a success ratio of 31% of the evaluated mutations that were predicted by the combination of the 2 tools. This percentage is rather high as the angles and distances between 2 cysteines are needed to exactly meet geometric requirements. It is however even more remarkable to point out that despite the specific geometric requirements, the same amino residue on position S73C was able to create disulphide bonds with two different cysteines on positions D112 and A115.

The introduced cysteine residue on position A52 on the other hand, shows a 5 °C increase in apparent melting temperature when forming a disulphide bridge with a cysteine on position P61, yet a 6 °C decrease when forming a disulphide bridge with position L59C. This decrease, together with the 9 °C decrease in apparent melting temperature of the redirected native disulphide bridge, would however support the theoretical narrow requirements for angles and distances.

To further increase the thermostability of the improved mutants, combinations can be made. While four mutants yielded positive results, two involve the same cysteine (on position S73). Apart from these two mutants all four mutants can be combined to double and even triple mutants for a possibility of an additive increase in melting temperature. A schematic representation of a triple mutant has been depicted in Figure 29. The first mutants to be created with an increased melting temperature were the mutants A143C-P183C, S73C-D112C and S73C-A115C. While the other combinations were not ready in the time course of this thesis, the double mutant with additional disulphide bridges between positions A143C-P183C and S73C-A115C shows a melting temperature of 12 ± 1 °C higher than the wild type. This melting temperature is exactly the sum of the 9 °C and 3 °C increase in apparent melting temperature of these mutations individually.

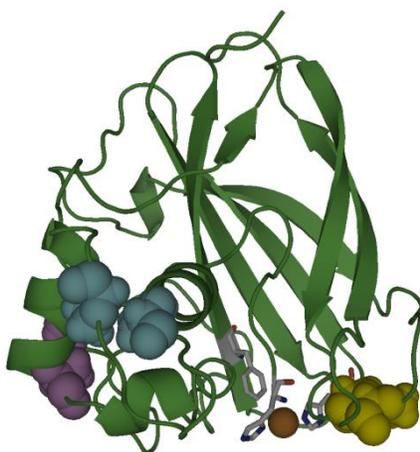


Figure 29: Schematic representation of the locations of the 3 stabilizing mutations to be combined. The yellow spheres represent the A143C-P183C mutation, the blue spheres represent the A52C-P61C mutation and the pink spheres represent the S73C-A115C mutation.

As previously stated, the strategy applied in the determination of the mutations to be considered resulted in a 31% success rate. However, all successful mutations were listed as A-grade according to the MODIP tool. This suggests the list of possible mutants could have been reduced. When only regarding the mutants proposed by DbD and listed with A grade by MODIP, this would have led to 4 successful mutations out of 9 or a success rate of 44.44%.

While these increased melting temperatures can be considered a great success, studying the available literature suggests that relatively large increases in T_m can be expected from disulphide introduction if there is a stabilizing effect. The introduction of disulphide bridges in Phage T4 lysosome for example, resulted in mutants with a change of melting temperature of -0.5, -2.4, 4.8, 6.4 and 11.0°C (Matsumurat, Beckett, Levitr, & Matthewstl, 1989). Several other projects on the stabilization of subtilisin have reported increases of melting temperatures by the formation of disulphide bonds as well, like an increase of 4.5 °C (Takagi, Takahashi, & Momose, 1990). Applying the same “disulphide by design” algorithm for the determination of possible mutants in lipase (as was employed in this work), an improvement in T_m of 7°C was also reported by Yu and coworkers (Yu, Tan, Xiao, & Xu, 2012). However both our results and these examples from the literature still can't bear the comparison with the increase in melting temperature of around 40°C, resulting from a single disulphide bond in two separate Arc repressors (Robinson & Sauer, 2000).

When evaluating the apparent melting temperatures, multiple significant decreases were measured. Most of these mutants exhibit a decrease in melting temperature up to about 6°C, yet some decreases exceed up to 11°C. While most articles focus on the positive results, literature also reveals that negative effects on the melting temperature due to added disulphide bridges have been reported in other studies as well. These results lead to believe that the formation of those disulphide bonds can result in the disruption of other stabilizing effects, or put a strain on the enzyme conformation when introduced in more rigid regions of the protein (Dombkowski, Zakia, & Craig, 2014; Matsumurat et al., 1989). This last reference, as well as (Takagi et al., 1990) also report the introduced disulphide bond may result in an increased melting temperature, but under reducing conditions will result in an unchanged or even decreased melting temperature.

Considering all these theories and positive and negative literature results, our improvement is a real success. Moreover, the T_m has been increased from 51 °C to 63 °C. Since cellulose degradation with cellulose mixtures is usually done at temperatures starting at 50 °C, our improved enzyme is already a better candidate than the wild type to keep up with the classical cellulases. However, more tests are still to be done, such as testing how long the enzyme will retain its activity at elevated temperatures. This will give a good indication for real use of the enzyme.

General conclusion

The first aspect of this thesis was the determination of a suitable expression system to assure the activity of the ScLPMO10C. The most crucial feature of the expression, is the necessity of the free histidine at the amino terminus involved in the binding of a copper ion in the active site. Both considered approaches, use of a secretion signal or a protease cleavage site, proved successful at producing sufficient amounts of active ScLPMO10C. However, the preferred approach was to insert a secretion signal. By a natural cleavage mechanism, his-1 is ensured, while transporting the protein to the periplasm (and later out of the cell). In this case, HPAEC-PAD analysis of the breakdown products when incubated with PASC and MS-analysis of a sample from the cytoplasmic and periplasmic space, confirmed the correct N-terminus and activity in both the cytoplasm and periplasm. While the other approach achieved similar results, the requirement of additional cleavage and purification steps proved more work efficient. Furthermore, the use of the secretion signal allowed the expression in the periplasm. This fraction proved to contain lower quantities of unwanted protein, resulting in a shorter and easier down-stream processing while maintaining sufficient concentrations of purified active ScLPMO10C.

To further optimize the down-stream processing of the enzymes, tests were performed to determine the optimal purification protocol for his-tag purification. Elution protocols based on a pH gradient and different concentrations of imidazole were considered. These test clearly showed the optimal purification is imidazole based, with concentrations of 10 mM in the equilibration buffer, 20 mM in the wash buffer and 100 mM in the elution buffer. Even higher purity can be achieved through a 40 mM Imidazole wash buffer, yet this was not required for the assays used in this thesis and will lead to a much lower enzyme concentration and only a slightly higher purity.

In order to determine the effect of the mutation on the activity, two assays were proposed based on literature. The first test included the measurement of an unprofitable side reaction, yet could not be used due to an unknown inhibiting phenomenon. The activity was thus tested by HPAEC-PAD analysis and identification of the reaction products, upon incubation of a cellulose based substrate (PASC) with the enzyme to be tested and ascorbic acid as an electron donor.

To evaluate the thermostability, the apparent melting temperature was determined by differential scanning fluorimetry assay (DSF, or also called a thermal shift assay). Samples were analyzed in triplicate and from different purifications during optimization of the assay. They showed an extremely low variability, yet comparison of different samples proved to require a similar

concentration and purity. The melting temperature of the wild type enzyme was determined $51 \pm 1^\circ\text{C}$.

In order to assess the potential influence of the free histidine on position 1, the apparent melting temperature of ScLPMO10C with an N-terminal linked Xa-restriction factor was determined before and after the cleavage of that Xa-factor. The obtained apparent melting temperature of 43°C in the absence of the free his-1 was significantly lower than the 51°C obtained with a correct N-terminus, once more proving the importance of the histidine on position 1. This not only for the activity, but apparently also for the correct fold of the protein.

After determining a good expression system and several analysis systems, 2 strategies for mutagenesis were evaluated: (1) rational engineering by simply inspecting the structure of the protein and (2) the introduction of additional disulphide bonds predicted by a combination of the “disulfide by design” and “MOdelling of Disulphide bridges In Proteins” tools.

Though the initial set of mutants were rationally selected to stabilize the protein surface or fixate 2 flexible loops, the determined apparent melting temperatures showed no significant increase for any of these mutants. Several significant changes were recorded, yet these mutants showed a decrease in apparent melting temperature of 3 to 6°C .

The second approach to stabilize the ScLPMO10C was thus considered based on the introduction of additional disulphide bonds. Twelve ScLPMO10C variants were obtained with an added disulphide bond, while one mutant was obtained by redirecting a native disulphide bond. While the redirected disulphide bond resulted in a decrease of 9°C in the apparent melting temperature, 4 of the other mutations resulted in an increase of 3°C (twice), 5°C and even 9°C . The 4 positive results on 13 tested strains thus means a success ratio of 31% of the evaluated mutations that were predicted by the combination of the 2 tools. Since the formation of disulphide bonds requires very specific angles and distances between two cysteines, it is remarkable to note two of the positive mutants (one of a 3°C and one of a 9°C increase) involve the same cysteine. Finally, 2 disulphide bonds were also combined, resulting in an additive effect, namely an increase in melting temperature of 12°C .

To conclude, the optimization of the expression system, the purification protocol and assays, allowed the successful creation and evaluation of PASC active variants of LPMO10Cs. Among these mutants, 4 proved to result in an increased apparent melting temperature. Combining a couple of these introduced disulphide bonds, even resulted in an increase in apparent melting temperature from 51°C for the wild type to 63°C for the mutated strain of ScLPMO10C. As the degradation of cellulose involves several enzymes with optimal temperature ranges starting at 50°C , this increase can be considered a success as it results in a higher chance of successful implementation of the ScLPMO10C for industrial applications.

Future perspectives

While several positive results have been obtained, not all proposed mutations were completed in time for this thesis. It stands to reason, the completion of the list could yield more stabilizing mutations.

Since the combination of two stabilizing mutations showed an additive effect on the apparent melting temperature, a next step in the further stabilization involves the creation and evaluation of all the other possible combinations of successful mutations.

In order to fully evaluate the impact of the mutations on the applicability in industrial processes, the actual stability should be tested at the actual process conditions to see if the ScLPMO10C can remain active long enough to contribute to the cellulose breakdown and boost the efficiency of the conventional cellulases mixtures.

Other approaches in the stabilization, such as Immobilization by adsorption, entrapment, membrane confinement and single or multiple covalent bonding could be considered as well. Such fixation could allow the further rigidifying of the enzyme structure for a higher stability.

Appendix

appendix 1: sequence of the used plasmid encoding the ScLPMO10C

```

                                                                 BstXI
                                                                 -----+-----
1  caccatcacc atcaccatta aggtcgacca tatgggagag ctcccaacgc gttggatgca
   gtggtagtgg tagtggtaat tccagctggt ataccctctc gagggttgcg caacctacgt
   >>.....His-tag.....>>
     h h h h h h h -
   >.....Cels2'.....>>
     h h h h h h h -

   NspI
   -----+
   SphI
   -----+

61  ggcattgcaag cttggctggt ttggcggatg agagaagatt ttcagcctga tacagattaa
    ccgtacgttc gaaccgaaa aaccgcctac tctcttctaa aagtccgact atgtctaatt

    < oExtra189_pCXhPX_rev1
    GTCTTGCG TCTTCGCCAG AC
121  atcagaacgc agaagcggtc tgataaaaca gaatttgcct ggcggcagta gcgcggtggt
    tagtcttgcg tcttcgccag actattttgt cttaaaccgga ccgccgtcat cgcgccacca

181  cccacctgac cccatgccga actcagaagt gaaacgccgt agcgccgatg gtagtgtggg
    ggggtggactg gggtagggct tgagtcttca ctttgccgca tcgcccgtac catcacacc

241  gtctcccat  gcgagagtag ggaactgcc  ggcattcaat aaaacgaaag gctcagtcga
    cagaggggta cgctctcatc cttgacgggt ccgtagttta ttttgctttc cgagtcagct

301  aagactgggc  ctttcgtttt atctgttgtt tgtcggtgaa cgctctcctg agtaggaaa
    ttctgaccgg gaaagcaaaa tagacaacaa acagccactt gcgagaggac tcactctggt

361  atccgccggg  agcggatttg aacgttgcca agcaacggcc cggaggggtg cgggcaggac
    taggcggccc tcgcctaaac ttgcaacgct tcggtgccgg gcctcccacc gccgtcctg

                                                                 < oExtra347_pCX34_Bla_rev
                                                                 CTTCCG GTAGGACTGC CTAC
421  gcccgccata aactgccagg catcaaatta agcagaaggc catcctgacg gatggccttt
    cgggcgggat ttgacgggtc gtagtttaat tcgtcttccg gtaggactgc ctaccggaaa

481  ttgcgtttct acaaactctt tttgtttatt tttctaaata cattcaaata tgtatccgct
    aacgcaaaga tgtttgagaa aaacaataaa aaagatttat gtaagtttat acataggcga

541  catgagacaa taaccctgat aaatgcttca ataatttga aaaaggaaga gtatgagtat
    gtactctggt attgggacta tttacgaagt tattataact ttttccttct catactcata
                                                                 >>bla'.>
                                                                 m s

601  tcaacatttc cgtgtcgccc ttattccctt ttttgccgca ttttgccctc ctgtttttgc
    agttgtaaag gcacagcggg aataagggaa aaaacgccgt aaaacggaag gacaaaaacg
    >.....bla'.....>
    i q h f r v a l i p f f a a f c l p v f
```

661 tcacccagaa acgctggtga aagtaaaaga tgctgaagat cagttgggtg cacgagtggg
agtgggtcct tgcgaccact ttcattttct acgacttcta gtcaaccac gtgctcacc
>.....bla'.....>
a h p e t l v k v k d a e d q l g a r v

721 ttacatcgaa ctggatctca acagcggtaa gatccttgag agttttcgcc ccgaagaacg
aatgtagctt gacctagagt tgtcgccatt ctaggaactc tcaaaagcgg ggcttcttgc
>.....bla'.....>
g y i e l d l n s g k i l e s f r p e e

781 ttttccaatg atgagcactt ttaaagttct gctatgtggc gcggtattat cccgtgttga
aaaaggttac tactcgtgaa aatttcaaga cgatacaccg cgccataata gggcacaact
>.....bla'.....>
r f p m m s t f k v l l c g a v l s r v

ScaI

---+---

841 cgccgggcaa gagcaactcg gtcgccgcat aactattct cagaatgact tggttgagta
gcggcccgtt ctcggtgagc cagcggcgta tgtgataaga gtcttactga accaactcat
>.....bla'.....>
d a g q e q l g r r i h y s q n d l v e

901 ctcaccagtc acagaaaagc atcttacgga tggcatgaca gtaagagaat tatgcagtgc
gagtggctag tgtcttttcg tagaatgctt accgtactgt cattctctta atacgtcacg
>.....bla'.....>
y s p v t e k h l t d g m t v r e l c s

961 tgccataacc atgagtgata acactgcggc caacttactt ctgacaacga tcggaggacc
acggatttgg tactcactat tgtgacgccg gttgaatgaa gactgttgct agcctcctgg
>.....bla'.....>
a a i t m s d n t a a n l l l t t i g g

1021 gaaggagcta accgcttttt tgcacaacat gggggatcat gtaactcgcc ttgatcgttg
cttctctgat tggcgaaaaa acgtgttgta ccccctagta cattgagcgg aactagcaac
>.....bla'.....>
p k e l t a f l h n m g d h v t r l d r

1081 ggaaccggag ctgaatgaag ccataccaaa cgacgagcgt gacaccacga tgcctacagc
ccttggcctc gacttacttc ggtatggtt gctgctcgca ctgtggtgct acggatgctg
>.....bla'.....>
w e p e l n e a i p n d e r d t t m p t

1141 aatggcaaca acgttgcgca aactattaac tggcgaacta cttactctag cttcccggca
ttaccgttgt tgcaacgcgt ttgataattg accgcttgat gaatgagatc gaagggcctg
>.....bla'.....>
a m a t t l r k l l t g e l l t l a s r

AseI

---+---

1201 acaattaata gactggatgg aggcggataa agttgcagga ccacttctgc gctcggccct
tgtaattat ctgacctacc tccgcctatt tcaacgtcct ggtgaagacg cgagccggga
>.....bla'.....>
q q l i d w m e a d k v a g p l l r s a

BglI

-----+-----

BpmI

-----+-----

1261 tccggctggc tggtttattg ctgataaatc tggagccggt gagcgtgggt ctgcggtat
aggccgaccg accaaataac gactatttag acctcggcca ctgcaccca gagcgccata
>.....bla'.....>
l p a g w f i a d k s g a g e r g s r g

AhdI

-----+-----

1321 cattgcagca ctggggccag atggtaagcc ctcccgtatc gtagttatct acacgacggg
gtaacgtcgt gaccccggtc taccattcgg gagggcatag catcaataga tgtgctgccc
>.....bla'.....>
i i a a l g p d g k p s r i v v i y t t

1381 gagtcaggca actatggatg aacgaaatag acagatcgct gagatagggtg cctcactgat
ctcagtcctg tgatacctac ttgctttatc tgtctagcga ctctatccac ggagtgacta
>.....bla'.....>
g s q a t m d e r n r q i a e i g a s l

1441 taagcattgg taactgtcag accaagttaa ctcatatata ctttagattg atttaaaact
attcgtaacc attgacagtc tggttcaaat gagtatatat gaaatctaac taaattttga
>.....bla'....>>
i k h w -

1501 tcatttttaa tttaaaagga tctaggtgaa gatccttttt gataatctca tgaccaaact
agtaaaaatt aaattttcct agatccactt ctaggaaaa ctattagagt actggtttta

1561 cccttaacgt gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc
gggaattgca ctcaaaaagca aggtgactcg cagtctgggg catcttttct agtttcctag

1621 ttcttgagat cctttttttc tgcgcgtaat ctgctgcttg caaacaacaaa aaccaccgct
aagaactcta ggaaaaaaag acgcgcatta gacgacgaac gtttgttttt ttggtggcga

1681 accagcgggtg gtttgtttgc cggatcaaga gctaccaact ctttttccga aggtaactgg
tggtcgccac caaacaacag gcctagttct cgatggttga gaaaaaggct tccattgacc

1741 cttcagcaga gcgagatac caaatactgt ctttctagtg tagccgtagt taggccacca
gaagtcgtct cgcgtctatg gtttatgaca ggaagatcac atcggcatca atccggtggt

AlwNI

-----+-----

1801 cttcaagaac tctgtagcac cgcctacata cctcgcctctg ctaatcctgt taccagtggc
gaagttcttg agacatcgtg gcggatgtat ggagcgagac gattaggaca atggtcaccg

1861 tgctgccagt ggcgataagt cgtgtcttac cgggttgagc tcaagacgat agttaccgga
acgacggtca ccgctattca gcacagaatg gcccaacctg agttctgcta tcaatggcct

1921 taaggcgcag cggtcgggct gaacgggggg ttcgtgcaca cagcccagct tggagcgaac
attccgcgct gccagcccga cttgcccccc aagcacgtgt gtcgggctga acctcgcctg

1981 gacctacacc gaactgagat acctacagcg tgagctatga gaaagcgcca cgcttcccga
ctggatgtgg cttgactcta tggatgtcgc actcgatact ctttcgcggt gcgaagggt

2041 agggagaaaag gcgacaggt atccggttag cggcagggtc ggaacaggag agcgcacgag
tccctctttc cgctgtcca taggccattc gccgtcccag cttgtctctc tcgctgtctc

2101 ggagcttcca gggggaaacg cctggtatct ttatagtcct gtcgggtttc gccacctctg
cctcgaagggt ccccctttgc ggaccataga aatatcagga cagcccaaag cgggtggagac

2161 acttgagcgt cgatTTTTgt gatgctcgtc aggggggCGG agcctatgga aaaacGCCag
tgaactcgca gctaaaaaca ctacgagcag tccccCGCC tcgatacct ttttgcggtc

NspI

-----+

2221 caacgcggcc tttttacggt tcttggcctt ttgctggcct tttgctcaca tgttctttcc
gttgcgcccg aaaaatgcca aggaccggaa aacgaccgga aaacgagtgt acaagaaagg

2281 tgcgttatcc cctgattctg tggataaccg tattaccgcc tttgagttag ctgataccgc
acgcaatagg ggactaagac acctattggc ataatggcgg aactcactc gactatggcg

SapI

2341 tcgccgcagc cgaacgaccg agcgcagcga gtcagttagc gaggaagcgg aagagcgcct
agcggcgctc gcttgcctggc tcgcgctcgt cagtcactcg ctctctcgcc ttctcgcgga

BstAPI

-----+---

2401 gatgcggtat tttctcctta cgcactctgt cggatattca caccgcatat ggtgcactct
ctacgccata aaagaggaat gcgtagacac gccataaagt gtggcgata ccacgtgaga

Bst1107I

---+---

2461 cagtacaatc tgctctgatg ccgcatagtt aagccagtat aactcgcct atcgctacgt
gtcatgttag acgagactac ggcgatcaaa ttcggtcata tgtgaggcga tagcgatgca

2521 gactgggtca tggctgcgcc ccgacaccgg ccaacaccgg ctgacgcgcc ctgacgggct
ctgaccagat accgacgcgg ggctgtgggc ggttgtgggc gactgcgcgg gactgcccga

BsmBI

2581 tgtctgctcc cggcatccgc ttacagacaa gctgtgaccg tctccgggag agctcgatat
acagacgagg gccgtaggcg aatgtctggt cgacactggc agaggccctc tcgagctata

SmaI

----+---

AvaI

--+-----

XmaI

--+-----

NspI

-----+

2641 cccgggCGGC cgtttgagga caaaaggTTC ttgacatgTT ctaatgtTTA tgctataaTT
gggcccCGCC gcaaactcCT gttttccaAG aactgtacAA gattacaAAT acgatataAA

NotI

---+-----

NcoI

---+-----

StyI

---+-----

XbaI

---+-----

BamHI

---+-----

EcoRI

---+-----

oExtra081_pCXhP34_fwd_1 >

GCC GCTCTAGAAG AAGCTTG

2701 tttcggatc catggcggcc gctctagaag aagcttggga tccgctgacc tcgaattcgg
aaagccatag gtaccgcccg cgagatcttc ttcgaaccct aggcagctgg agcttaagcc

2761 aggaacaaaa gatggttcgt cgtaccgcgc tgctgaccct ggcagcagtt ctggcaaccc
tcctttgttt ctaccaagca gcatgggcag acgactggga ccgctcgtcaa gaccgttggg

>>.....CelS2'.....>

```

                m v r r t r l l t l a a v l a t
                >>.....Signal peptide Cels2'.....>
                m v r r t r l l t l a a v l a t

2821 tgctgggtag cctgggtggt accctgctgc tgggtcaggg tcgtgccgaa gcacatggg
acgacccatc ggaccacaa tgggacgacg acccagtccc agcacggctt cgtgtaccac
>.....Cels2'.....>
l l g s l g v t l l l g q g r a e a h g
>.....Signal peptide Cels2'.....>
l l g s l g v t l l l g q g r a e a

2881 ttgcaatgat gcctggtagc cgtacctatc tgtgtcagct ggatgcaaaa accggtacag
aacgttacta cggaccatcg gcatggatag acacagtcga cctacgtttt tggccatgac
>.....Cels2'.....>
v a m m p g s r t y l c q l d a k t g t

                BamHI          NspI
                -+-----          -----+

2941 gtgcactgga tccgaccaat ccggcatgtc aggcagcctt ggatcagagc ggtgcaaccg
cacgtgacct aggctgggta ggccgtacag tccgtcggga cctagtctcg ccacgttggc
>.....Cels2'.....>
g a l d p t n p a c q a a l d q s g a t

                                                BspMI
                                                -----
                                                BsgI
                                                -----

3001 cactgtataa ttggtttgcc gttctggata gcaatgccgg tggtcgtggt gcaggttatg
gtgacatatt aaccaaaccg caagacctat cgttacggcc accagacca cgtccaatac
>.....Cels2'.....>
a l y n w f a v l d s n a g g r g a g y

                AccIII          BstAPI
                -+-----          -----+

3061 ttccggatgg caccctgtgt agtgccgggtg atcgtagccc gtatgatttt agcgcataata
aaggcctacc gtgggacaca tcacggccac tagcatcggg cataactaaaa tcgcgtatat
>.....Cels2'.....>
v p d g t l c s a g d r s p y d f s a y

3121 atgcagcacg tagcgattgg cctcgtaccc atctgaccag cggtgccacc attccggttg
tacgtcgtgc atcgtaacc ggagcatggg tagactggtc gccacgggtg taaggccaac
>.....Cels2'.....>
n a a r s d w p r t h l t s g a t i p v

                MunI
                -+-----

3181 aatatagcaa ttgggcagca catccgggtg attttcgtgt ttatctgacc aaaccgggtt
ttatatcgtt aaccgcgtcg ttaggccac taaaagcaca aatagactgg tttggcccaa
>.....Cels2'.....>
e y s n w a a h p g d f r v y l t k p g

                BstXI
                -----+-----

3241 ggagcccgac cagcgaactg ggttgggatg atctggaact gattcagacc gttaccaatc
cctcgggctg gtcgcttgac ccaaccctac tagacctga ctaagtctgg caatggttag
>.....Cels2'.....>
w s p t s e l g w d d l e l i q t v t n

```

```

      BbvCI
      ---+-----
      Bpu10I
      ---+-----
3301  cgcctcagca gggcagtcg ggtacagatg gtggtcacta ttattgggat ctggcactgc
      gcggagtcgt cccgtcaggc ccatgtctac caccagtgat aataacccta gaccgtgacg
      >.....Cels2'.....>
      p p q q g s p g t d g g h y y w d l a l

3361  cgagcggtcg tagcggatg gcaactgatc ttatgcagtg ggttcgtagc gatagccaag
      gctcgccagc atcgccacta cgtgactaga aatacgtcac ccaagcatcg ctatcggttc
      >.....Cels2'.....>
      p s g r s g d a l i f m q w v r s d s q

      PstI
      -----+
3421  aaaacttctt tagctgcagt gatgtggtgt ttgatgggtg taatggtgaa gttaccggta
      ttttgaagaa atcgacgtca ctacaccaca aactaccacc attaccactt caatggccat
      >.....Cels2'.....>
      e n f f s c s d v v f d g g n g e v t g

3481  ttctgtgtag cggtagcacc cctgatcctg atccgacccc gacaccgacc gatccgacaa
      aagcaccatc gccatcgtgg ggactaggac taggctgggg ctgtggctgg ctaggctggt
      >.....Cels2'.....>
      i r g s g s t p d p d p t p t p t d p t

3541  cccctccgac ccataccggt agctgtatgg cagtttatag cgttgaaaat agctggtcag
      ggggaggctg ggtatggcca tcgacatacc gtcaaatatc gcaactttta tcgaccagtc
      >.....Cels2'.....>
      t p p t h t g s c m a v y s v e n s w s

3601  gtggttttca gggtagcgtg gaagttatga atcatggcac cgaaccgctg aatggttggg
      caccaaaagt cccatcgcac cttcaatact tagtaccgtg gcttggcgac ttaccaacc
      >.....Cels2'.....>
      g g f q g s v e v m n h g t e p l n g w

      PasI
      ---+-----
3661  cagttcagtg gcagcctggt ggtggtacaa ccctgggtgg tgtgtggaat ggtagcctga
      gtcaagtcac cgtcggacca ccaccatggt gggaccacc acacacctta ccacggact
      >.....Cels2'.....>
      a v q w q p g g g t t l g g v w n g s l

      BclI
      -+-----
3721  ccagtggtag tgatggcacc gttaccgttc gtaatgttga tcataatcgt gttgttccgc
      ggtcaccatc actaccgtgg caatggcaag cattacaact agtattagca caacaaggcg
      >.....Cels2'.....>
      t s g s d g t v t v r n v d h n r v v p

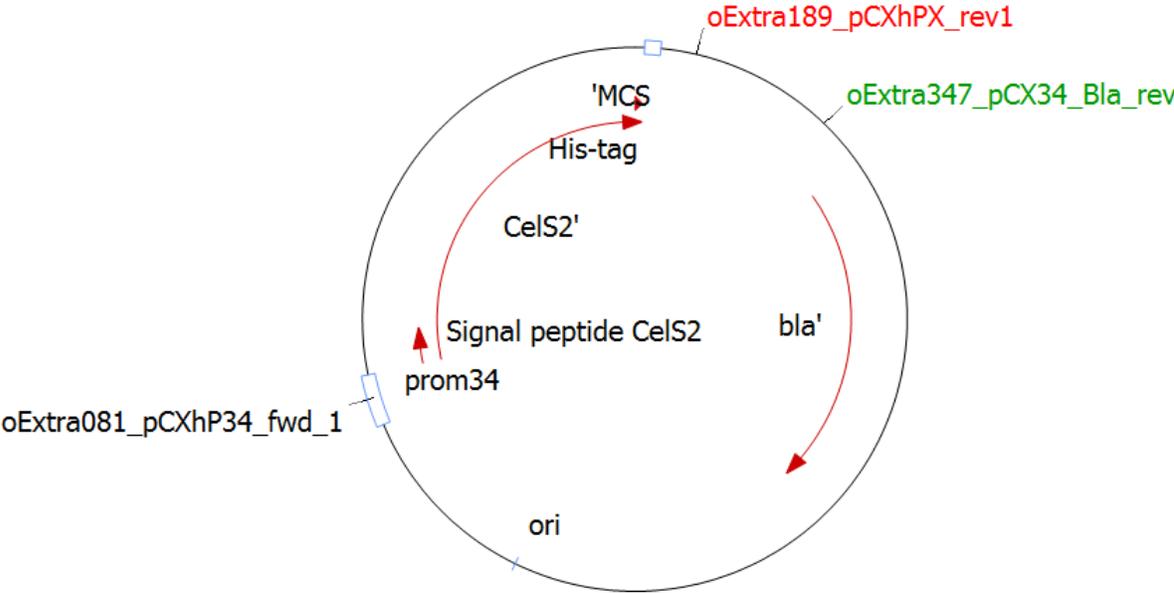
3781  ctgatggtag cgttacctt ggttttaccg caaccagcac cggtaatgat tttccgggtg
      gactaccatc gcaatggaaa ccaaaatggc gttggtcgtg gccattacta aaaggccacc
      >.....Cels2'.....>
      p d g s v t f g f t a t s t g n d f p v

3841  atagcattgg ttgtgttgca ccg

```

```
tatcgtaacc aacacaacgt ggc  
>.....CelS2'.....>  
d s i g c v a p
```

Appendix 2: Map of the features of the used plasmid encoding the ScLPMO10C



Sources

- Ã, C. N. H., & Hooijdonk, G. Van. (2005). Ethanol from lignocellulosic biomass : techno-economic performance in short- , middle- and long-term, *28*, 384–410. <http://doi.org/10.1016/j.biombioe.2004.09.002>
- Aachmann, F. L., Sørli, M., Skjåk-bræk, G., Eijsink, V. G. H., & Vaaje-kolstad, G. (2012). NMR structure of a lytic polysaccharide monoxygenase provides insight into copper binding , protein dynamics , and substrate interactions, *109*(46). <http://doi.org/10.1073/pnas.1208822109>
- Aerts, D., & Desmet, T. (2011). A constitutive expression system for high- throughput screening, (1), 10–19. <http://doi.org/10.1002/elsc.201000065>
- Aerts, D., Verhaeghe, T., Joosten, H.-J., Vriend, G., Soetaert, W., & Desmet, T. (2013). Consensus engineering of sucrose phosphorylase: The outcome reflects the sequence input. *Biotechnology and Bioengineering*, *110*(10), 2563–2572. <http://doi.org/10.1002/bit.24940>
- Agger, J. W., Isaksen, T., Várnai, A., Vidal-melgosa, S., Willats, W. G. T., & Ludwig, R. (2014). Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation, *111*(17). <http://doi.org/10.1073/pnas.1323629111>
- Balat, M., & Balat, H. (2009). Recent trends in global production and utilization of bio-ethanol fuel. *Applied Energy*, *86*(11), 2273–2282. <http://doi.org/10.1016/j.apenergy.2009.03.015>
- Bayer, E. A., Lamed, R., & Himmel, M. E. (2007). The potential of cellulases and cellulosomes for cellulosic waste management, (Figure 1). <http://doi.org/10.1016/j.copbio.2007.04.004>
- Beeson, W. T., Phillips, C. M., Cate, J. H. D., & Marletta, M. A. (2012). Oxidative Cleavage of Cellulose by Fungal Copper-Dependent Polysaccharide Monoxygenases, 2011–2013.
- Bommarius, A. S., Blum, J. K., & Abrahamson, M. J. (2011). Status of protein engineering for biocatalysts : how to design an industrially useful biocatalyst. *Current Opinion in Chemical Biology*, *15*(2), 194–200. <http://doi.org/10.1016/j.cbpa.2010.11.011>
- Borchert, T. V., Burg, B. Van Den, Eijsink, V. G. H., & Ga, S. (2005). Directed evolution of enzyme stability, *22*, 21–30. <http://doi.org/10.1016/j.bioeng.2004.12.003>
- Cardona, C. A. (2007). Fuel ethanol production : Process design trends and integration opportunities, *98*, 2415–2457. <http://doi.org/10.1016/j.biortech.2007.01.002>
- Cerdobbel, A., De Winter, K., Aerts, D., Kuipers, R., Joosten, H.-J., Soetaert, W., & Desmet, T. (2011). Increasing the thermostability of sucrose phosphorylase by a combination of sequence- and structure-based mutagenesis. *Protein Engineering Design and Selection*, *24*(11), 829–834. <http://doi.org/10.1093/protein/gzr042>
- Chica, R. A., Doucet, N., & Pelletier, J. N. (2005). Semi-rational approaches to engineering enzyme activity : combining the benefits of directed evolution and rational design, 378–384.

<http://doi.org/10.1016/j.copbio.2005.06.004>

Craig, D. B., & Dombkowski, A. A. (2013). Disulfide by Design 2.0 : a web-based tool for disulfide engineering in proteins, 0–6.

Dani, V. S., Ramakrishnan, C., & Varadarajan, R. (2003). MODIP revisited : re-evaluation and refinement of an automated procedure for modeling of disulfide bonds in proteins, *16*(3), 187–193.
<http://doi.org/10.1093/proeng/gzg024>

Demain, A. L. (2000). Small bugs , big business : The economic power of the microbe, *18*, 499–514.

Dies, G., Henrissat, B., Davies, G., & Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases.

Dombkowski, A. A., Zakia, K., & Craig, D. B. (2014). Protein disulfide engineering. *FEBS Letters*, *588*(2), 206–212. <http://doi.org/10.1016/j.febslet.2013.11.024>

Eijsink, V. G. H., Bjørk, A., Sigrid, G., Sirev, R., Synstad, B., Burg, B. Van Den, & Vriend, G. (2004). Rational engineering of enzyme stability, *113*, 105–120. <http://doi.org/10.1016/j.jbiotec.2004.03.026>

Fágáin, C. Ó. (2003). Enzyme stabilization — recent experimental progress, *33*, 137–149.
[http://doi.org/10.1016/S0141-0229\(03\)00160-1](http://doi.org/10.1016/S0141-0229(03)00160-1)

Forsberg, Z., Mackenzie, A. K., Sørli, M., Røhr, Å. K., Helland, R., Arvai, A. S., ... Eijsink, V. G. H. (2014). Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases, *111*(23), 8446–8451. <http://doi.org/10.1073/pnas.1402771111>

Forsberg, Z., Røhr, Å. K., Mekasha, S., Kristo, K., Eijsink, V. G. H., Vaaje-kolstad, G., & Sørli, M. (2014). Comparative Study of Two Chitin-Active and Two Cellulose-Active AA10-Type Lytic Polysaccharide Monooxygenases.

Forsberg, Z., Vaaje-kolstad, G., Westereng, B., Bunæs, A. C., Stenstrøm, Y., Mackenzie, A., ... Eijsink, V. G. H. (2011). Cleavage of cellulose by a CBM33 protein, *20*, 1479–1483.
<http://doi.org/10.1002/pro.689>

Gianfreda, L., & Scarfi, M. R. (1991). Enzyme Stabilization : State of the Art, (MARCH 1991).
<http://doi.org/10.1007/BF00234161>

Gray, K. A., Zhao, L., & Emptage, M. (2012). Bioethanol. <http://doi.org/10.1016/j.cbpa.2006.02.035>

Harmsen, P., & Huijgen, W. (2010). Literature Review of Physical and Chemical Pretreatment Processes for Lignocellulosic Literature Review of Physical and Chemical Pretreatment Processes for Lignocellulosic Biomass, (November 2015).

Hemsworth, G. R., Henrissat, B., Davies, G. J., & Walton, P. H. (2013). Discovery and characterization of a new family of lytic polysaccharide monooxygenases. *Nature Chemical Biology*, *10*(2), 122–126.
<http://doi.org/10.1038/nchembio.1417>

Hemsworth, G. R., Johnston, E. M., Davies, G. J., & Walton, P. H. (2015). Lytic Polysaccharide

- Monooxygenases in Biomass Conversion. *Trends in Biotechnology*, xx, 1–15.
<http://doi.org/10.1016/j.tibtech.2015.09.006>
- Horn, S. J., Vaaje-kolstad, G., Westereng, B., & Eijsink, V. G. H. (2012). Novel enzymes for the degradation of cellulose.
- Isaksen, T., Westereng, B., Aachmann, F. L., Agger, J. W., Kracher, D., Kittl, R., ... Horn, S. J. (2014). A C4-oxidizing Lytic Polysaccharide Monooxygenase Cleaving Both Cellulose and Cello-oligosaccharides. *Journal of Biological Chemistry*, 289(5), 2632–2642. <http://doi.org/10.1074/jbc.M113.530196>
- Iyer, P. V., & Ananthanarayan, L. (2008). Enzyme stability and stabilization—Aqueous and non-aqueous environment. *Process Biochemistry*, 43(10), 1019–1032.
<http://doi.org/10.1016/j.procbio.2008.06.004>
- Jaenicke, R., & Gerald, B. (1998). The stability of proteins in extreme environments. *Current Opinion in Structural Biology*, 738–748.
- Johnson, C. M. (2013). Differential scanning calorimetry as a tool for protein folding and stability. *Archives of Biochemistry and Biophysics*, 531(1-2), 100–109.
<http://doi.org/10.1016/j.abb.2012.09.008>
- Jørgensen, H., & de Campos Giordano, R. (2015). Lignocellulose pretreatment technologies affect the level of enzymatic cellulose oxidation by LPMO. *Green Chemistry*, (February 2016).
<http://doi.org/10.1039/C4GC02179G>
- Kittl, R., Kracher, D., Burgstaller, D., Haltrich, D., & Ludwig, R. (2012). Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnology for Biofuels*, 5(1), 79. <http://doi.org/10.1186/1754-6834-5-79>
- Langston, J. A., Shaghasi, T., Abbate, E., Xu, F., Vlasenko, E., & Sweeney, M. D. (2011). Oxidoreductive Cellulose Depolymerization by the Enzymes Cellobiose Dehydrogenase and Glycoside Hydrolase 61 [†], 77(19), 7007–7015. <http://doi.org/10.1128/AEM.05815-11>
- Leggio, L. Lo, Simmons, T. J., Poulsen, J. N., Frandsen, K. E. H., Hemsworth, G. R., Stringer, M. A., ... Walton, P. H. (2015). Structure and boosting activity of a starch-degrading lytic polysaccharide monooxygenase. *Nature Communications*, 6, 1–9. <http://doi.org/10.1038/ncomms6961>
- Levasseur, A., Drula, E., Lombard, V., Coutinho, P. M., & Henrissat, B. (2013). Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnology for Biofuels*, 6(1), 41. <http://doi.org/10.1186/1754-6834-6-41>
- Li, S., Yang, X., Yang, S., Zhu, M., & Wang, X. (2012). Technology Prospecting on Enzymes : Application , Marketing and Engineering, (September).
- Li, X., Beeson, W. T., Phillips, C. M., Marletta, M. A., & Cate, J. H. D. (2012). Structural Basis for Substrate Targeting and Catalysis by Fungal Polysaccharide Monooxygenases. *Structure*, 20(6), 1051–1061.
<http://doi.org/10.1016/j.str.2012.04.002>

- Lin, M. Y., & Tanaka, S. (2006). Ethanol fermentation from biomass resources : current state and prospects, 627–642. <http://doi.org/10.1007/s00253-005-0229-x>
- Lutz, S. (2010). Beyond directed evolution — semi-rational protein engineering and design. *Current Opinion in Biotechnology*, 21(6), 734–743. <http://doi.org/10.1016/j.copbio.2010.08.011>
- Lynd, L. R., Weimer, P. J., Zyl, W. H. Van, & Pretorius, I. S. (2002). Microbial Cellulose Utilization : Fundamentals and Biotechnology, 66(3), 506–577. <http://doi.org/10.1128/MMBR.66.3.506>
- Mallick, P., Boutz, D. R., Eisenberg, D., & Yeates, T. O. (2002). Genomic evidence that the intracellular proteins of archaeal microbes contain disulfide bonds, 99(15), 9679–9684.
- Mateo, C., Abian, O., Lafuente, R. F., & Guisan, J. M. (2000). Increase in conformational stability of enzymes immobilized on epoxy-activated supports by favoring additional multipoint covalent attachment, 26, 509–515.
- Mateo, C., Palomo, J. M., Fernandez-lorente, G., Guisan, J. M., & Fernandez-lafuente, R. (2007). Improvement of enzyme activity , stability and selectivity via immobilization techniques, 40, 1451–1463. <http://doi.org/10.1016/j.enzmictec.2007.01.018>
- Matsumura, M., Signor, G., & Matthews, W. B. (1989). Substantial increase of protein stability by multiple disulphide bonds. *Nature*.
- Matsumurat, M., Bechtelt, W. J., Levitr, M., & Matthewstl, B. W. (1989). Stabilization of phage T4 lysozyme by engineered disulfide bonds, 86(September), 6562–6566.
- Mussatto, S. I., Dragone, G., Guimarães, P. M. R., Paulo, J., Silva, A., Carneiro, L. M., ... Teixeira, J. A. (2010). Technological trends , global market , and challenges of bio-ethanol production. *Biotechnology Advances*, 28(6), 817–830. <http://doi.org/10.1016/j.biotechadv.2010.07.001>
- Naik, S. N., Goud, V. V, Rout, P. K., & Dalai, A. K. (2010). Production of first and second generation biofuels : A comprehensive review, 14, 578–597. <http://doi.org/10.1016/j.rser.2009.10.003>
- Nannemann, D., Birmingham, W., Scism, R., & Bachmann, B. (2012). Assessing directed evolution methods for the generation of biosynthetic enzymes with potential in drug biosynthesis. *Future Med Chem*, 3(7), 809–819. <http://doi.org/10.4155/fmc.11.48>.Assessing
- Polizzi, K. M., Bommarius, A. S., Broering, J. M., & Chaparro-Riggers, J. F. (2007). Stability of biocatalysts. *Current Opinion in Chemical Biology*, 11, 220–225. <http://doi.org/10.1016/j.cbpa.2007.01.685>
- Promon, S. K., Program, B., & Sciences, N. (2015). Studies on Isolation of Yeasts from Natural Sources for Bioethanol production from Vegetable Peels and the Role of Cellulose Degrading Bacteria (*Bacillus subtilis*) on Ethanol Production, (August).
- Ratanakhanokchai, K., Waeonukul, R., Sakka, K., Kosugi, A., & Mori, Y. (2013). Strain B-6 Multienzyme Complex : A Novel System for Biomass Utilization.
- Reetz, M. T., Carballeira, J. D., & Vogel, A. (2006). Iterative saturation mutagenesis on the basis of b

- factors as a strategy for increasing protein thermostability. *Angewandte Chemie - International Edition*, 45(46), 7745–7751. <http://doi.org/10.1002/anie.200602795>
- Robinson, C. R., & Sauer, R. T. (2000). Striking Stabilization of Arc Repressor by an Engineered Disulfide Bond, 12494–12502.
- Rodrigues, J., Prosinecki, V., Marrucho, I., & Rebelo, L. P. N. (2011). Protein stability in an ionic liquid milieu : on the use of differential scanning fluorimetry, (Cd), 13614–13616. <http://doi.org/10.1039/c1cp21187k>
- Rowinska-Zyrek, M., Danuta, W., & Slawomir, P. (2013). His-rich sequences – is plagiarism from nature a good idea? †, 58–70. <http://doi.org/10.1039/c2nj40558j>
- Sanchez-ruiz, J. M. (2010). Biophysical Chemistry Protein kinetic stability. *Biophysical Chemistry*, 148(1-3), 1–15. <http://doi.org/10.1016/j.bpc.2010.02.004>
- Sanchis, J., Fernández, L., Carballeira, J. D., Drone, J., Gumulya, Y., Höbenreich, H., ... Reetz, M. T. (2008). Improved PCR method for the creation of saturation mutagenesis libraries in directed evolution: application to difficult-to-amplify templates. *Applied Microbiology and Biotechnology*, 81(2), 387–397. <http://doi.org/10.1007/s00253-008-1678-9>
- Sandra T. Merino, J. C. (2007). *Progress and challenges in enzyme development for biomass utilization*.
- Sowdhamini, R., Srinivasn, N., Shoichet, B., Santi, Daniel, V., Ramakrishnan, C. & Balaram, P. (1989). Stereochemical modeling of disulfide bridges. Criteria for introduction into proteins by site-directed mutagenesis. *Protein Engineering*.
- Span, E. A., & Marletta, M. A. (2015). The framework of polysaccharide monooxygenase structure and chemistry. *Sciencs Direct Current Opinion in Structural Biology*, 35, 93–99. <http://doi.org/10.1016/j.sbi.2015.10.002>
- Sukharnikov, L. O., Cantwell, B. J., Podar, M., & Zhulin, I. B. (2011). Cellulases : ambiguous nonhomologous enzymes in a genomic perspective, 1–7. <http://doi.org/10.1016/j.tibtech.2011.04.008>
- Suplatov, D., Voevodin, V., & Švedas, V. (2015). Robust enzyme design: Bioinformatic tools for improved protein stability. *Biotechnology Journal*, 10(3), 344–355. <http://doi.org/10.1002/biot.201400150>
- Takagi, H., Takahashi, T., & Momose, H. (1990). Enhancement of the Thermostability of Subtilisin E by Introduction a Disulfide Bond Engineered on the Basis of Structural Comparison with a Thermophilic Serine Protease. *Journal of Biological Chemistry*, 265(12), 6674–6676.
- Tan, T., Kracher, D., Gandini, R., Sygmund, C., Kittl, R., Haltrich, D., ... Ha, B. M. (2015). Structural basis for cellobiose dehydrogenase action during oxidative cellulose degradation, (May). <http://doi.org/10.1038/ncomms8542>
- Vaaje-Kolstad, G., Westereng, B., Horn, S. J., Liu, Z., Zhai, H., Sorlie, M., & Eijsink, V. G. H. (2010). An

- Oxidative Enzyme Boosting the Enzymatic Conversion of Recalcitrant Polysaccharides. *Science*, 330(6001), 219–222. <http://doi.org/10.1126/science.1192231>
- Vu, V. V., Beeson, W. T., Phillips, C. M., Cate, J. H. D., & Marletta, M. A. (2014). Determinants of Regioselective Hydroxylation in the Fungal Polysaccharide Monooxygenases, 2–5.
- Warren, D. J. (2011). Preparation of highly efficient electrocompetent *Escherichia coli* using glycerol / mannitol density step centrifugation. *Analytical Biochemistry*, 413(2), 206–207. <http://doi.org/10.1016/j.ab.2011.02.036>
- Wijma, H. J., Floor, R. J., & Janssen, D. B. (2013). Structure- and sequence-analysis inspired engineering of proteins for enhanced thermostability. *Current Opinion in Structural Biology*, 23(4), 588–594. <http://doi.org/10.1016/j.sbi.2013.04.008>
- Wyman, C. E. (1994). ETHANOL FROM LIGNOCELLULOSIC BIOMASS : TECHNOLOGY , ECONOMICS , AND OPPORTUNITIES, 50.
- Xie, Y., An, J., Yang, G., Wu, G., Zhang, Y., Cui, L., & Feng, Y. (2014). Enhanced Enzyme Kinetic Stability by Increasing Rigidity within the Active Site * □, 289(11), 7994–8006. <http://doi.org/10.1074/jbc.M113.536045>
- Yu, X., Tan, N., Xiao, R., & Xu, Y. (2012). Engineering a Disulfide Bond in the Lid Hinge Region of *Rhizopus chinensis* Lipase : Increased Thermostability and Altered Acyl Chain Length Specificity, 7(10), 1–7. <http://doi.org/10.1371/journal.pone.0046388>
- Zhang, Y. P., Himmel, M. E., & Mielenz, J. R. (2006). Outlook for cellulase improvement : Screening and selection strategies, 24, 452–481. <http://doi.org/10.1016/j.biotechadv.2006.03.003>
- Zhang, Y. P., & Lynd, L. R. (2015). Toward an aggregated understanding of enzymatic hydrolysis of cellulose : Noncomplexed cellulase systems, (November). <http://doi.org/10.1002/bit.20282>
- Žifčáková, L., & Baldrian, P. (2012). Fungal polysaccharide monooxygenases: new players in the decomposition of cellulose. *Fungal Ecology*, 5(5), 481–489. <http://doi.org/10.1016/j.funeco.2012.05.001>