Kinetic and thermodynamic description of Cel7A catalyzed hydrolysis of cellulose

Trine Holst Sørensen
Department for Science Systems and Models (NSM)
Roskilde University
Denmark

PhD thesis
August 2015

Cover: The crystal structure of *Hypocrea jecorina* Cel7A (E217Q) in complex with cellohexaose and cellobiose (PDB code CEL7) PyMol.

Summary

Commercial utilization of biomass for the production of second generation bioethanol constitutes a sustainable and clean alternative to fossil fuels. This has recently been emphasized by the European Union, which estimated that 56 % of renewable energy generation in 2020 would be derived from biomass. The main constituent of biomass, cellulose, is a simple chemical polymer whose energy-rich component, glucose, can be fermented into ethanol. The breakdown or hydrolysis of cellulose to glucose is facilitated by a broad class of enzymes referred to as Glycoside hydrolases (GH). GH and especially cellobiohydrolases from the GH7 family (Cel7As) hold a crucial role as it constitutes the major part in commercial enzymatic cocktails. Crystalline cellulose is highly resistant toward degradation and even though the presence of enzymes speed the reaction up considerably, the breakdown of the polymer into its fermentable components still constitutes a challenging step in the process making of second generation bioethanol.

The major part of this dissertation was dedicated to studying the temperature effect on the kinetics of two native Cel7As; the mesophile two-domain *Hypocrea jecorina* Cel7A, consisting of a catalytic domain and a carbohydrate binding module, which are connected by a linker region, and the thermophile single domain *R. emersonii* Cel7A. In addition we studied a truncated version of *H.jecorina* Cel7A and a chimeric protein composed of the linker and CBM from *H.jecorina* Cel7A and the *R. emersonii* enzyme to highlight the role of CBM. We measured the reaction rates of these four enzymes at different substrate loads (crystalline cellulose) and find that the two-domain enzymes are more catalytically efficient at low substrate concentration. Conversely, the single-domain enzymes show a two-fold increase in the overall reaction rate compared with the two-domain enzymes at high substrate loads. The reaction temperature also affected the kinetics of the two-domain and single-domain enzymes; at increasing temperatures the interval covering the substrate loads at which two-domain enzymes are more catalytically efficient, grew bigger. This means that the advantage of possessing a CBM (in terms of hydrolytic activity) is much dependent on the substrate load and reaction temperature. Another result which appears from changing the temperature and measuring the disturbed kinetics is that little or no temperature sensitivity of the hydrolytic rates (for both two-domain and single-domain enzymes) was found at low substrate load, while the hydrolytic rate is much accelerated at saturating substrate loads.

To get a more detailed description of the kinetic parameters, we applying a steady state model which accounts for the characteristic processive mechanism of Cel7As and measured the degree of processivity. We find that the dissociation rate constant ($_pk_{off}$) is more accelerated by temperature than the rate constant for the catalytic cycle ($_pk_{cat}$) and the rate constant governing the enzyme substrate association ($_pk_{on}$). The monitored temperature effect on the kinetic constants ($_pk_{on}$, $_pk_{cat}$ and $_pk_{off}$) also

formed the basis for making a free energy diagram for *H.jecorina* Cel7A hydrolysis of cellulose. The free energy diagram was formed by applying Transition state theory on the kinetic rate constants and by obtaining equilibrium constants for the enzyme-substrate association. The free energy reaction coordinate profile revealed that free activation energy for the enzyme-substrate association is dominated by entropy contributions, while the free energy barrier for dissociation is dominated by enthalpy contributions.

This thesis additionally embraces a study of one of the characteristic tunnel forming loops. For these experiments, point mutations were introduced in a loop covering the cellulose strand at the -4 subsite in *H.jecorina* and *R.emersonii*. More specifically, two asparagine residues were exchanged with alanine residues in positions 194 and 197 of *R.emersonii* Cel7A. Replacement of one or both of the asparagine residues led to improved maximum velocity and lowered affinity for crystalline cellulose. Substituting the N197 corresponding asparagine residue in *H.jecorina*, N200, with 12 different amino acids, additionally led to decreased affinity for cellulose. This site saturation approach in *H.jecorina* suggests that the role of the asparagine in this position is to facilitate 1) a high degree of processivity and 2) high affinity. In continuation thereof, plots of processivity and $_pK_M$ for all the N200 variants suggested a possible link between processivity and affinity. Furthermore, the kinetic data for the N200 variants suggest a relationship between the maximum velocity and low affinity.

Resume

Kommerciel udnyttelse af biomasse til produktion af anden generations bioethanol udgør et bæredygtigt og rent alternativ til fossile brændstoffer. Dette er for nylig blevet understreget af Den Europæiske Union, som anslår, at 56% af vedvarende energi i 2020 vil være afledt af biomasse. Hovedbestanddelen af biomasse, cellulose, er en kemisk simpel polymer, hvis energirige komponent, glukose, kan fermenteres til ethanol. Nedbrydningen eller hydrolysen af cellulose til glukose katalyseres af en bred klasse af enzymer kaldet glycosid hydrolaser (GH) og især cellobiohydrolaser fra GH7 familien (Cel7A) spiller en vigtig rolle, da de udgør den største del i kommercielle enzym cocktails. Krystallinsk cellulose er svært nedbrydeligt og selv om tilstedeværelsen af enzymer fremskynder reaktionen betydeligt, er nedbrydningen af polymeren stadig en af de større udfordringer i fremstillingen af anden generations bioethanol.

Denne afhandling er primært dedikeret til at studere temperatureffekten på kinetikken af to naturligt forekommende Cel7A; *Hypocrea jecorina* Cel7A, som består af to domæner; et katalytisk domæne og et carbohydrate-binding module (CBM) forbundet af en linker region og *Rasamsonia emersonii* Cel7A som kun består af det katalytiske domæne. Derudover undersøgte vi en trunkeret form af *H.jecorina* Cel7A (kun det katalytiske domæne) og et kimær enzyme som var sammensat af det katalytiske domæne fra *R.emersonii* og linker og CBM fra *H.jecorina*. For de fire enzymer målte vi reaktionshastighederne ved forskellige substrat koncentrationer (krystallinsk cellulose) og varierende temperaturer. Vi finder at to-domæne enzymerne har højere reaktionshastigheder ved lav substrate koncentration i forhold til to-domæne enzymerne. Omvendt viser single-domæne enzymer en fordobling i reaktionshastigheden sammenlignet med de to-domæne-enzymer ved høje substrat koncentration. Reaktionstemperaturen påvirket også kinetikken af to-domæner og enkelt-domæne-enzymerne; ved forhøjede temperaturer, øges det substratet koncentrations interval, som definerer de mere katalytisk effektive to-domæne enzymer. Ved at variere temperaturen og måle den ændrede kinetik, finder vi også at temperaturfølsomheden er meget lille for den enzymatiske reaktions hastighed ved lave substrat koncentrationer, hvorimod den hydrolytiske rate ved højere substrat koncentrationer er betydelig accelereret af højere temperaturer. For at få en mere detaljeret beskrivelse af de kinetiske parametre, anvendte vi en steady state model, som tager hensyn til den karakteristiske processive mekanisme for Cel7A. Vi finder at rate konstanten for dissociationen af enzym fra substrat ($_pk_{off}$) er mere accelereret ved øget temperatur end hastighedskonstanterne for katalysen ($_pk_{cat}$) og hastigheds konstanten som bestemmer Cel7A-cellulose associering ($_pk_{on}$). På baggrund af de målte temperatur effekter på de kinetiske rate konstanter ($_pk_{on}$, $_pk_{cat}$ og $_pk_{off}$) dannede vi et fri energi diagram for den *H.jecorina* Cel7A katalyserede hydrolyse af cellulose. Det fri energi diagram blev dannet ved at anvende Transition state teori på de kinetiske

rate konstanter og ved eksperimentelt at finde ligevægtskonstanter for enzyme-substrate associering. Profilen for fri energi som funktion af reaktion koordinaterne viste at fri aktiveringsenergi for enzym-substratet association er domineret af et entropi bidrag, mens fri energi barriereen for dissociation er domineret af et enthalpi bidrag.
Denne afhandling omfatter også en undersøgelse af et af de loops som former den karakteristiske katalytiske tunnel i Cel7A. I disse eksperimenter blev punktmutationer introduceret i det loop, der dækker glykosid subsitet (-4) i tunnelen i *H.jecorina* og *R.emersonii*. Mere specifikt blev en asparaginsyre skiftet ud med alanin i position 194 og 197 i B2 loopet af *R.emersonii*. Udskiftningen af en eller begge aminosyrer førte til øget maksimal reaktionshastighed og sænket affinitet for krystallinsk cellulose. Vi udskiftede også en enkelt asparaginsyre i position 200 med 12 forskellige aminosyre i *H.jecorina* Cel7A, svarende til position 197 i *R.emersonii* Cel7A. Kinetiske data for de mange forskellige substitutioner i position 200 antyder, at asparagins rolle i den position er at facilitere høj processivitet og affinitet. I forlængelse heraf, viste et plot af processivitet og $_pK_M$ for alle N200 varianterne, en mulig korrelation mellem processivitet og affinitet. Derudover indikerer disse data også en positiv korrelation mellem den maksimale reaktionshastighed og $_pk_M$.

Preface and Acknowledgements

The work associated with this thesis has been conducted at the Department of Science Systems and Models, Roskilde University and Department of Enzyme Biophysics, Novozymes during the period August 2012 - August 2015. This PhD project was part of the research project RESAB (Rational Engineering of cellulases for improved SAccharification of Biomass) whose main objective is to design variants with higher reaction rates under industrial conditions. RESAB is a collaborative research project between Roskilde University, Denmark, and Novozymes A/S, Denmark, and is partially funded by The Danish Council for Strategic Research.

The PhD project has truly been a very informative and interesting period for which I am eager to spread some credit. For wise, perceptive and broadminded suggestions and advice I would like to thank my supervisors Peter Westh and Kim Borch. I have been encouraged as ever by their contagious enthusiastic and visionary view on the field of cellulases. A special thanks to Nicolaj Cruys-Bagger whose previous work provided the theoretical platform for the study. Equally, I am very grateful Michael Skovbo Windahl for the impressive work associated with engineering and generating a wealth of very interesting variants. I would also like to express my appreciation to Silke Flindt Badino and Kadri Alasepp for good advices and for assisting me in some of the experimental work. Thanks also to other colleagues on the RESAB project Jeppe Kari for creating some of the illustrations used in the thesis and to Johan Olsen and Eva M. Karlsen. Acknowledgments to all my wonderful colleagues at Novozymes, in particular Camilla Hindborg Kristensen for assisting in the work of enzyme purification and Ulla Thyssen Rosenberg for given good advices in the development of kinetic assays. In general I would like to thank all my co-authors for all the innovative inspiring inputs and scientific discussions. Finally, I am very grateful to all those nearest and dearest to me.

List of publications

The thesis is based on the following articles and manuscripts, which will be referred to by their roman numerals **I**, **II**, **III** etc. The articles and manuscripts are appended at the end of the thesis.

I. **Temperature effects on kinetic parameters and substrate affinity of Cel7A cellobiohydrolases**
Trine Holst Sørensen, Nicolaj Cruys-Bagger, Michael Skovbo Windahl, Silke Flindt Badino, Kim Borch & Peter Westh *The journal of biological chemistry 2015 290: 22193-22202*

II. **Free energy diagram for the heterogeneous enzymatic hydrolysis of the glycosidic bonds in cellulose**
Trine Holst Sørensen, Nicolaj Cruys-Bagger, Kim Borch & Peter Westh *The journal of biological chemistry, 2015 290: 22203-22211*

III. **Probing substrate interactions in the active tunnel of a catalytically deficient cellobiohydrolase (Cel7)**
Francieli Colussi, Trine H. Sørensen, Kadri Alasepp, Jeppe Kari, Nicolaj Cruys-Bagger, Michael S. Windahl, Johan P. Olsen, Kim Borch and Peter Westh. *The journal of biological chemistry, 2015 290: 2444-2454.*

IV. **Loop engineering in the thermophile cellobiohydrolasem, Cel7A from Rasamsonia emersonii: Effects on activity and substrate interactions**
Trine Holst Sørensen, Michael S. Windahl, Jeppe Kari, Johan Olsen Kim Borch & Peter Westh *(manuscript in preparation)*

V. **Cellobiohydrolase variants and polynucleotides encoding same. Patent WO2014138672**
Borch, K., Jensen, K., Krogh, K., Mcbrayer, B., Westh, P., Kari, J., Olsen, J. P., Sorensen, T. H., Windahl, M. S., and Xu, H.n , *September 12, 2014*

Other publications

Product inhibition of five *Hypocrea jecorina* cellulases
Leigh Murphy, Christina Bohlin, Martin J. Baumann, Søren N. Olsen, Trine H. Sørensen, Lars Anderson, Kim Borch, Peter Westh
Enzyme microbial technology Enzyme and Microbial Technology, 52, 163-169, 2013

A pyranose dehydrogenase-based biosensor for kinetic analysis of enzymatic hydrolysis of cellulose by cellulases..*Nicolaj Cruys*-Bagger, Silke Flindt Badino, Radina Tokin, Mark Gontsarik, Samin Fathalinejad, Kenneth Jensen, Miguel Duarte Toscano, Trine Holst Sørensen, Kim Borch, Hirosuke Tatsumi, Priit Väljamäe & Peter Westh. *Enzyme and Microbial Technology, 58-59, 68-74, 2014*

In Situ Stability of Substrate-Associated Cellulases Studied by DSC
Kadri Alasepp, Kim Borch, Nicolaj Cruys-Bagger, Silke Badino, Kenneth Jensen, Trine H. Sørensen, Michael S. Windahl, and Peter Westh. *Langmuir, 30, 7134-42. 2014*

Table of Contents

Outline of the dissertation and guide to the reader

The thesis is composed of 3 main Chapters. Three articles and one manuscript are appended.

Chapter 1 provides a general introduction of the theoretical platform for the study. *Chapter 2* principally covers results and discussion for the present investigations. For experimental procedure the reader is referred to the articles, manuscript and patent appended. *Chapter 2* is subdivided into two major sections, where the *first* presents data and discussion of the publications (I-III), the *second* describes the publications (IV-V) The second sections has an extended introduction, which could arguable belong to Chapter 1. However, to facilitate the reading of this thesis, this division was made.
Finally *Chapter 3* provides a conclusion for the results covered in Chapter 2.

List of abbreviations

CBM	*Carbohydrate binding module*
Cel7A	*Cellobiohydrolase, glycoside hydrolase family 7*
Core	*Catalytic domain*
DP	*degree of polymerization*
GEI	*Glycosyl-enzyme intermediate*
GH	*Glucoside hydrolases*
Hj_{CORE}	*The catalytic domain of Hypocrea jecorina Cel7A*
Hj_{CBM} *, HjCel7A*	*Hypocrea jecorina Cel7A*
ITC	*Isothermal titration calorimetry*
PAHBAH:	*4-hydroxybenzoic acid hydrazide*
PCS	*Pretreated corn stower*
p-NPL	*4-nitrophenyl β-D-lactopyranoside*
RC	*Reaction coordinate*
RESAB	*Rational Engeneering of cellulases for improved SAccharification of Biomass*
Re_{CBM}	*Chimeric Cel7A consisting of the catalytic domain from R.emersonii Cel7A connected by a linker region to the CBM from Hypocrea jecorina Cel7A*
Re_{CORE}*, ReCel7A*	*Rasamsonia emersonii Cel7A*
SHF	*Separate hydrolysis and fermentation*
SSF	*Simultaneous sacccharification and fermentation*
TST	*Transition state theory*
wt	*Wild-type*

1 GENERAL INTRODUCTION

1.1 RESAB: Rational Engeneering of cellulases for improved SAccharification of Biomass

The natural setting for cellulase degradation of cellulose barely imitates the conditions for industrial saccharification. As the enzymes are a product of natural evolution, whose apparent function is to adapt the reaction rates to the natural surroundings, improvements in terms of more efficient enzymes for industrial conditions, seems feasible. The main objective of the RESAB[1] project (2012-2016) is to design variants with higher reaction rates under industrial conditions. To meet the main objective an iterative approach is used, as outlined below:

1. *Identification of rate determining step(s)*
2. *Design of variants, with higher hydrolytic activity*
3. *Generation and tests of variants*

The work presented in the thesis is primarily founded on previously conducted work which established the dissociation of cellulases from cellulose to be the rate-determining step. Along with this information, the now considerable collective understanding of cellulase structures offer the crucial structure-function relationship, which is the basis for designing enzymes for industrial application. Within the scope of RESAB, this thesis aims to give a fundamental description of cellulase kinetics under different conditions (temperature and substrate concentration) and to design and test variants with higher dissociation rate.

[1] *RESAB is a collaborative research project between Roskilde University, Denmark, and Novozymes A/S, Denmark, and is partially funded by The Danish Council for Strategic Research.*

1.2 Cellulose

The main constituents of plant based material, also referred to as lignocellulosic biomass, is the high energy polymer cellulose $(C_6H_{10}O_5)_n$, which makes up 30 – 50 % of the plant material (Ververis et al., 2004). This makes cellulose the most abundant organic compound; the annual biosynthesis is estimated to be around 10^{10} - 10^{12} t (Klemm et al., 2005, Hon, 1994). The repeating unit in the polymer cellulose is cellobiose which is two glucose (β-D-glucopyranose) molecules covalently linked between the equatorial OH group of C4 and the C1 carbon atom (β-1,4-glucan) (*Figure 1*). To accommodate the preferred bond angles, every second pyranose ring is rotated 180° in the plane and this rotation is what makes cellobiose and not glucose the repeating unit (*Figure 1*). The cellulose chain consists at one end of a D-glucose unit with an original C4-OH group (the nonreducing end), with the other end terminating with an original C1-OH group, which is in equilibrium with the aldehyde structure (the reducing end). The chain length of cellulose can be expressed in the number of constituent D-glucose molecules (degree of polymerization, DP) and varies (150-15.000) with the origin and treatment of the raw material (Taherzadeh and Karimi, 2008, OSullivan, 1997, Huang et al., 2015).

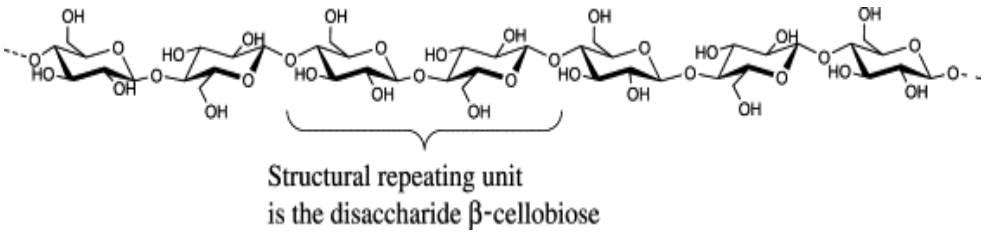


Figure 1. The structure of a single cellulose chain. The β-1,4-linked glucose is the chemical repeating unit, whereas structural repeating unit is cellobiose (Gessler et al., 1994)

Following synthesis by a cellulose synthase complex, a catalyst of the polymerization of the glucan chains, the cellulose chains aggregates to form microfibrils containing the two native allomorphs 1α and 1β (Brett, 2000). More allomorphs exist; however, they are not naturally occurring and are formed during the different physicochemical treatments of cellulose (Atalla, 2008, OSullivan, 1997, Brett, 2000). The microfibrils are remarkably stable and highly resistant toward degradation (Wolfenden and Snider, 2001) as illustrated by the extensive half-life for spontaneous hydrolysis of the glyosidic bonds, which has been reported to be 10^7 years (Wolfenden et al., 1998). This physical and chemical stability of cellulose is primarily brought about by multiple inter- and intramolecular hydrogen bonds and other secondary bonds such as van der Waals

interaction (Nishiyama et al., 2002). The microfibrils are arranged in parallel orientation and the order of this arrangement has traditionally divided cellulose into crystalline (highly ordered) or amorphous (less ordered) regions although several regions fall into an intermediate range between the two extremes, as indicated by the continuous crystallinity index (Park et al., 2010)

1.3 Introduction of Cel7A

The β-1,4-glucosidic bond in cellulose is degraded by a remarkably broad class of enzymes belonging to Glycoside hydrolases (GH). GH enzymes, however, are not restricted to hydrolyze the specific β-1,4-glucosidic bond in cellulose, but are rather defined as enzymes, hydrolysing the glycosidic bond between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety (Cantarel et al., 2009). To date the GHs constitute currently 133 protein families of which we are particularly interested in family number 7 (GH7) (http://www.cazy.org). The GH enzymes are grouped into families based on the amino acid similarity (Cantarel et al., 2009). It should be emphasized the GH families are divided on the basis of amino acid similarity of the catalytic domain and not their carbohydrate binding module (CBM), which will be introduced to the reader in the following section. Enzymes which hydrolyzes the β-1,4-glucosidic bond can also be classified as cellobiohydrolases (CBH), which also suggests that these enzymes hydrolyze their substrate in an *exo*-mode. *Exo*-mode refers to the ability of a glycoside hydrolase to initiate the hydrolysis of the polymeric substrate at the end of a chain. In contrast, enzymes working *endo*-mode can hydrolyze the cellulose chain within the chain. In the classification of cellobiohydrolases there are only two groups; CBHI and CBHII, where the roman numerals indicate an enzymatic attack from either the reducing- or nonreducing end. On the basis of these definitions, we introduce Cel7A as the nomenclature we will use in the remainder of this thesis, which refers to cellobiohydrolases I in the GH7 family.

The discovery of the fungus *Trichoderma viride*[2] QM6 and its appertaining cellulases by the US army during World War II (Reese, 1976), initiated the extensive research toward industrial application of these enzymes. Until now, *H. jecorina* remains one of the most important workhorses for production of industrial relevant cellulases. As Cel7A is one of the pillars in commercial cellulase cocktails and accounts for 60% of the cellulolytic proteins in *H.jecorina* (Teeri, 1997), multiple structures have been solved with different stages of ligand complexation

[2] *Trichoderma viride* is also known as *Trichoderma ressei*, which is an anamorph of the fungus *Hypocrea jecorina*. Throughout the thesis, we will consistently use *Hypocrea jecorina*

(Ståhlberg et al., 2001, Divne et al., 1994, Knott et al., 2014). In particular, Cel7A from *H. jecorina* has been well explored and, as most structural data currently is elucidated for *H. jecorina* Cel7A, this enzyme will provide a standard of comparison when describing the structure for the class of Cel7A enzymes. For a review on all reported GH7 crystal structures, consult (Payne et al., 2015).

1.3.1. Structure of Cel7A

The overall structure of Cel7A from *H.jecorina* includes a N-terminal catalytic domain and a C-terminal carbohydrate-binding module (CBM) which are connected by a flexible glycosylated linker (*Figure 2*) (Tomme et al., 1988). Out of this trio, the linker region has attracted the least attention. Amongst different suggested functions of the linker region, it is generally hypothesized to serve as a leash/tether between CBM and the catalytic domain (Boraston et al., 2004). However, lately the linker has been assigned an additional role as assisting in the binding of the enzyme to its substrate during the enzymatic hydrolysis (Payne et al., 2013b).

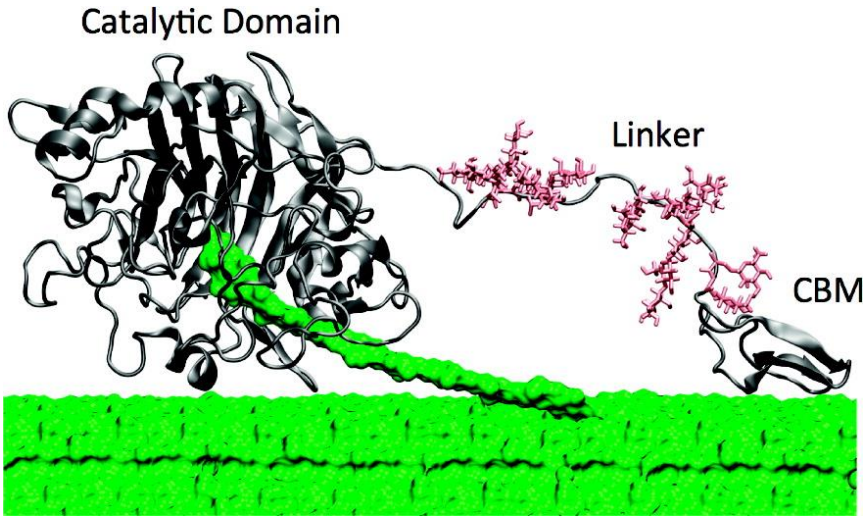


Figure 2. Schematic representation of the two-domain structure of H.jecorina Cel7A on a cellulose surface. The catalytic domain is shown with a single cellulose chain in the active site tunnel. O-glycosylation is shown in pink on the linker (Beckham et al., 2010).

In contrast, the novelty value of the function of CBM dates back to 1988, where the CBM was separated successfully by proteolysis (Tomme et al., 1988, Gilkes et al., 1988). The generally accepted functions of the CBM are to facilitate the enzyme adsorption on the solid substrate surface (Várnai et al., 2013) and to target a specific type or different site of the substrate (Carrard et al., 2000, Boraston et al., 2004). Additionally, some works have suggested the CBM to be more directly involved in the catalytic process (Hall et al., 2011, Wang et al., 2008). The

classification of CBMs based on their amino acid sequence was originally suggested by Tomme et al.(Tomme et al., 1995). However, in keeping with the glycoside hydrolase classification, we will use the nomenclature given by the cazy data base (cazy.org). The CBM belonging to *H.jecorina* Cel7A is classified in the CBM1 family. This family covers peptides of approx. 40 residues, which bind to the cellulose surface via three aromatic residues (*Figure 3*). The NMR solved 3D structure of the CBM1 from *H.jecorina* Cel7A (Kraulis et al., 1989), is a highly-suitable representation of for the CBM1 family and reveals 3 tyrosine residues (Y5, Y31 and Y32) positioned on the flat surface and three anti-parallel β-sheets. Along with the three tyrosine residues, two amino acids asparagnine (N29) and glutamine (Q7) have been assigned an important function in the mechanism of Cel7A hydrolysis of cellulose (Beckham et al., 2010) (*Figure 3*).

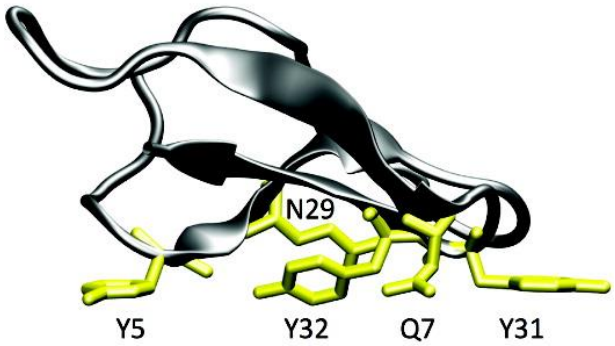


Figure 3. Schematic representation of a wedge shaped carbohydrate-binding module from H.jecorina Cel7A with Y5, Q7, N29, Y31, and Y32 positioned on the flat surface (Beckham et al., 2010).

To date most cellulases characterized have the CBM-linker-Core structure although only the minority of cellulases identified do comprise CBMs (Várnai et al., 2013). A possible reason for overlooking the single domain cellulases is that early studies found a 90 % decrease in activity upon removal of the linker and CBM (Tomme et al., 1988). However, as we will show below, such dramatic reductions in activity for single domain enzymes (both truncated and native cellulases) are not general. The native single domain Cel7A we will investigate in this thesis is the *R.emersonii Cel7A* and the structure of this enzyme has also been solved (again with difference ligands) (Grassick et al., 2004).

One third of the structure of the catalytic domain of *H.jecorina* is arranged as a β-sandwich, which is made of two opposing antiparallel β-sheets, while the rest of the protein consists entirely of loops connecting the β-sheets (Divne et al., 1994). These loops connecting the β-sheets form an enclosed substrate-binding tunnel, where the catalytic site also resides (Divne et al., 1994) and this overall structural description of *H.jecorina* Cel7A is in accordance with other family 7

cellobiohydrolases. The tunnel contains several sub-sites for binding of glucopyranose moieties (-7 to +2), where the sub-site -7 is located at the entrance of the tunnel and the sub-sites +1, and +2 constitute the expulsion site. The catalytic residues Glu212 and Glu217 are positioned on opposite site of the scissile glycosidic bond at subsites -1/+1 (Divne et al., 1998). The GH catalyzed hydrolysis results in either a retention or inversion of the anomeric carbon and this particular stereochemistry of the anomeric carbon classifies the hydrolytic mechanism. The retaining mechanism, which Cel7A employs to hydrolysis the glycosidic bond, is a two-step double displacement mechanism with a nucleophile (Glu212) responsible for the initial attack on the glycosidic bond and a proton donor (Glu217) accountable for the following step (Ståhlberg et al., 1996). In addition to Glu212 and Glu217, Asp214 is assigned a catalytic residue. The role of the latter is to hydrogen bond to the nucleophile and thereby promote catalysis (Ståhlberg et al., 1996). The equivalence to proton donor Glu217 and the nucleophile Glu212 in the single domain *R.emersonii* Cel7A is Glu214 Glu209, respectively (Grassick et al., 2004).

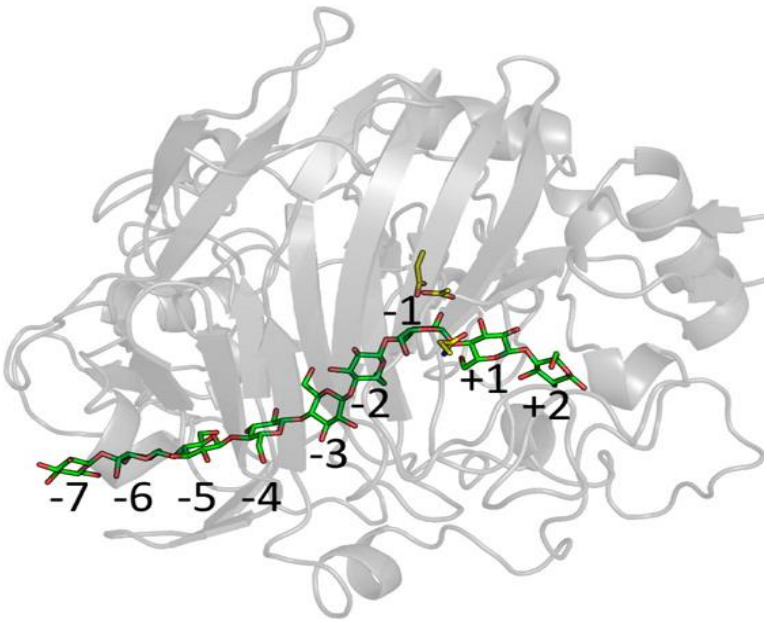


Figure 4. Overall structure of the Michaelis complex. The sub-sites for binding of glucopyranose moieties are indicated: the sub-site -7 is located at the entrance of the tunnel and the + sub-sites constitute the expulsion site. Side chains are drawn for the catalytic residues Glu212 and Glu217 (yellow) which are positioned on the opposite sites of the cleavable glycosidic bond at subsites -1/+1. (Knott et al., 2014)

Although the formation of the glycosyl-enzyme complex has been acknowledged for a considerable number of years, the first picture of the complex based on experimentally determined structure has recently been published (*Figure 4*) (Knott et al., 2014). Moreover, they suggested a structure of the Michaelis-Menten

complex, which appears after the substrate had filled the whole tunnel prior to catalysis of the cellulase chain into the active site tunnel but prior to catalytic event in subsite +1 and -1 (Knott et al., 2014). A more thorough description of the mechanism is given in section 1.6.

1.4 Cellulase kinetics

The enzymatic hydrolysis of cellulose occurs in the liquid-solid interface (it is a heterogeneous reaction), where cellulose makes up the solid sorbent, while cellulase and water constitute the diffusive elements. One of the fundamental differences between heterogeneous hydrolysis and classical enzyme kinetic is the number of reaction steps. Some important reaction steps in the mechanistic action of *H.jecorina* Cel7A (Bansal et al., 2009a) are listed below and illustrated in *Figure 5*. In addition to the heterogeneous nature, Cel7A hydrolyses cellulose in a processive manner (i.e. completes several catalytic cycles before dissociating from the cellulose strand) which further adds to the complexity of reaction path (see point 4).

1) *Adsorption* of Cel7A from the aqueous phase to the solid cellulose surface
2) *Location* of an accessible free end on a cellulose chain or an appropriate attack-site for endo-mode initiation
3) *Formation of enzyme–substrate complex* by decrystalizing a single chain from the cellulose surface and threading the chain into the tunnel of the Core domain.
4) *Inner catalytic cycle* which includes i) hydrolysis of the β-glycosidic bond ii) expulsion of the formed product, cellobiose iii) sliding of the cellulose chain forward in the tunnel to form a new Michaelis complex. Steps i to iii are repeated which gives rise to the processive movement of the enzyme along the cellulose chain.
5) *Unbinding* of the cellulose chain from the active site but still adsorbed to the cellulose surface via either the CBM and/or the Core domain
6) *Desorption* of the enzyme from the cellulose surface into the aqueous phase or repetition of step 2

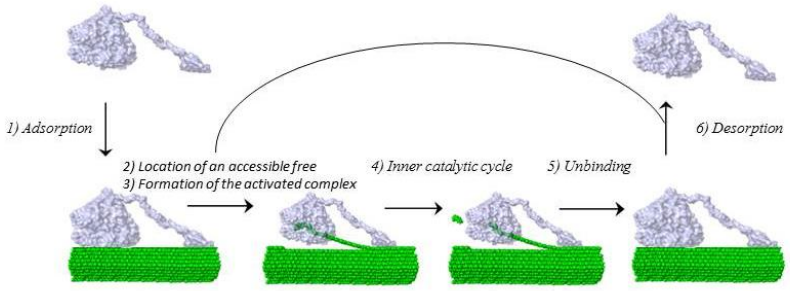


Figure 5: Schematic illustration of the mechanism by which processive cellobiohydrolases are hypothesized to hydrolyze cellulose. Step 1) Adsorption 2-3) location and formation of enzyme-substrate complex 4) Catalytic cycle 5) dethreading and 6) desorption or repetition of step 2-5.

Cel7A hydrolysis of insoluble cellulose is characterized by nonlinear kinetics and is one of the more distinctive features of Cel7A degradation of cellulose. The explanations for the nonlinearity have been assigned as either enzyme (deactivation, inhibition, jamming, clogging, irreversible adsorption, processivity) (Yang et al., 2006) or substrate (insoluble nature, crystallinity, surface area, porosity, degree of polymerization) (Zhang et al., 1999, Nidetzky and Steiner, 1993, Bubner et al., 2012) origin and the explanations most likely fall into an ambiguous range between these. Many of these factors such as product inhibition, substrate depletion and enzyme jamming/crowding however, fail to explain the fast declining rate which even appears in the transient (pre-steady state) regime of the hydrolysis (Kipper et al., 2005, Jalak and Valjamae, 2010, Murphy et al., 2010a, Murphy et al., 2010b). Thus the non-linear kinetics obviously adds more to the complex picture.

1.4.1. Modelling

To gain an understanding of the many reaction steps and to account for the nonlinearity, various models have been proposed (for reviews see (Bansal et al., 2009b, Sousa et al., 2011). One of the principal goals of modelling cellulase kinetics has been to identify rate-limiting step(s). Elucidating the rate-limiting step is not only crucial for explaining the fundamental mechanism of cellulase, but moreover it is critical for optimization of cellulases in terms of increasing the rate for conversion of biomass into fermentable sugars. One approach to elucidate the rate-limiting step(s) is to quantify the rate for individual steps in an enzyme reaction pathway by pre-steady-state kinetic measurements (Johnson, 1992). However, if both rapid and slow steps in the reaction path are to be elucidated, experiments must cover the pre-stady state (transient) time regime and this requires a suitable fast experimental technique and a realistic model that can be used to derive the kinetic parameters. Both of these are challenging tasks. However, recent work has shown promise in both developing sensitive techniques and in elucidating an useable model (Praestgaard et al., 2011, Cruys-Bagger et al., 2012b).

Praestgaard et al. suggested a kinetic model based on a simplified reaction scheme (scheme 1):

$$E + C_m \xrightarrow{{}_p k_{on}} EC_m \xrightarrow{{}_p k_{cat}} EC_{m-1} + C \xrightarrow{{}_p k_{cat}} EC_{m-2} + C \cdots \xrightarrow{{}_p k_{cat}} EC_{m-n}$$

$$\downarrow {}_p k_{off} \qquad \downarrow {}_p k_{off} \qquad \downarrow {}_p k_{off} \qquad \downarrow {}_p k_{off}$$

$$E + C_m \qquad E + C_{m-1} \qquad E + C_{m-2} \qquad E + C_{m-n}$$

Sheme 1: Simplified Reaction Scheme for the Hydrolysis of a Cellulose Strand by a Processive enzyme. The enzyme, E, attacks the end of a strand, C_m, with m cellobiose units and moves processively along the strand. Cellobiose, C, is released and the strand is sequentially shortened to C_{m-1}, C_{m-2}, and so forth. The reaction is governed by four parameters: rate constants for association (i.e., formation of an activated complex), catalysis, and dissociation and a processivity number, n. (Cruys-Bagger et al., 2013b).

Reaction scheme 1 is described by 3 rate constants ${}_pk_{on}$, ${}_pk_{cat}$, ${}_pk_{off}$ and a processivity number (n). The subscript 'p' in front of a symbol indicates that it pertains to the processive reaction scheme. The enzyme (E) first combines with a substrate strand consisting of m cellobiose units (C_m) to form a Michaelis complex (EC_m). This association is governed by the rate constant ${}_pk_{on}$. When the activated complex is formed (EC_m) it may either dissociate (${}_pk_{off}$) or go through a catalytic cycle (${}_pk_{cat}$) to produce a product (cellobiose, C) and a new activated complex with a slightly shorter substrate strand (EC_{m-1}). The parameter, n, specifies the average number of steps following one association (the processivity number). *Figure 6* shows an illustration of this model which matches scheme 1.

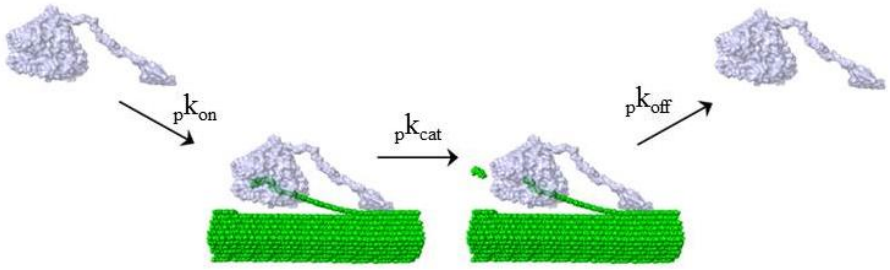


Figure 6: Schematic illustration of the simplified processive model showing ${}_pk_{on}$, ${}_pk_{cat}$, ${}_pk_{off}$. The figure was adapted and modified from (Cruys-Bagger et al., 2013b) .

Kinetic equations can be written on the basis of the reaction scheme 1 and solved with respect to the time-dependent cellobiose concentration (Praestgaard et al., 2011). Although progress has been made to obtain experimental data in the transient regime (Cruys-Bagger et al., 2014, Cruys-Bagger et al., 2012b), it still remains challenging and time consuming to obtain data in this timeframe and the mathematical analysis is not straightforward (Cruys-Bagger et al., 2013a). This led

us to employ a steady state model which also offers physically well-defined kinetic parameters and is based on the processive reactions scheme (scheme 1) (Cruys-Bagger et al., 2013a). With regard to the choice of analyzing steady state data as opposed to pre-steady state data, we note that these measurements only give information on the rate-limiting step as the fast steps in the reaction path has almost no influence on the overall rate at steady state.

1.4.2. Steady state model

Given that the cellulose conversion is minor, the conventional rate equation for scheme (1) can be solved analytically (Praestgaard et al., 2011). In the special case of quasi steady-state, which occurs when the concentrations of the complexes EC_m, EC_{m-1}, EC_{m-2} ... are almost independent of time, the steady state rate of cellobiose production ${}_pV_{ss} = d[C]/dt$ can be expressed ;

$$ {}_pV_{ss} = \frac{S_0 E_0 \, {}_pk_{cat}\left(1-\left(\frac{{}_pk_{cat}}{{}_pk_{cat}+{}_pk_{off}}\right)^n\right)}{\frac{{}_pk_{off}}{{}_pk_{on}}+S_0} \qquad \text{(eq. 1)} $$

where E_0 is the total concentration of enzyme (in µM) and S_0 is the of substrate concentration (g/l). To simplify eq. (1) and illustrate its relationship to the usual Michaelis Menten equation, we define a maximal processive rate, ${}_pV_{\max}$, and a processive analog of the Michaelis´ constant, ${}_pK_M$

$$ {}_pV_{\max} \equiv E_0 \, {}_pk_{cat}\left(1-\left(\frac{{}_pk_{cat}}{{}_pk_{cat}+{}_pk_{off}}\right)^n\right) \qquad \text{(eq. 2)} $$

$$ {}_pK_M \equiv \frac{{}_pk_{off}}{{}_pk_{on}} \qquad \text{(eq. 3)} $$

Inserting eq. (2) and (3) into eq. (1) yields the usual hyperbolic form

$$ {}_pV_{ss} = \frac{{}_pV_{\max} S_0}{{}_pK_M + S_0} \qquad \text{(eq. 4)} $$

The general strategy for the work in this thesis is to fit eq. (4) to experimental data for ${}_pV_{ss}(S_0)$ at different temperatures and/or for different enzymes and to extract the resulting parameters, ${}_pV_{max}$ and ${}_pK_M$. In addition to the parameters ${}_pV_{max}$ and ${}_pK_M$, the steady state theory allows a more detailed kinetic analysis by approximations and introduction of independent experimental data. With the approximation that ${}_pk_{cat} >> {}_pk_{off}$ (see the next section) at high substrate loads, the maximal specific rate is governed by (Cruys-Bagger et al., 2013a)

$$\frac{{}_pV_{\max}}{E_0} \approx n_p k_{off} \qquad \text{(eq. 5)}$$

Provided that we know the parameter (n), we can find ${}_pk_{off}$. The processivity number (n) can be can be derived from the relative amounts of products (e.g. glucose, cellobiose and cellotriose) which can be estimated from independent experimental data (Horn et al., 2012). The association constant ${}_pk_{on}$ can be related to the specificity constant, ${}_p\eta$;

$${}_p\eta = \frac{\left(\frac{{}_pV_{\max}}{E_0}\right)}{{}_pK_M} \approx n_p k_{on} \qquad \text{(eq. 6)}$$

1.4.3. Rate-limiting step

Compared with one of the most efficient enzymes, carbonic anhydrase, which has a rate constant k_{cat} around 1 x 10^6 s^{-1} (Kernohan, 1964), the reported values for k_{cat} for Cel7A are very low and fall in the range of 2.2 – 7.2 s^{-1} (Payne et al., 2015), but actually the overall rate of Cel7A hydrolysis of crystalline cellulose is even one or two orders of magnitude lower than this (Jalak and Valjamae, 2010, Kurasin and Valjamae, 2011, Cruys-Bagger et al., 2013a). The slow rate of Cel7A has even led to speculation as to whether the hydrolytic degradation of cellulose would be the primary function and if the genes encoding it would be under evolutionary pressure at all to increase the catalytic efficiency (Konstantinidis et al., 1993). However, the general catalytic power of Cel7A is still remarkable when seen in relation to the uncatalyzed hydrolysis of the primary bond in cellulose, which has a half-life of 10^7 years (Wolfenden et al., 1998).

The overall rate of an enzymatic reaction is governed by the slowest step, also known as the rate-limiting or rate-determining step. Therefore, in order to increase

the enzymatic degradation of the cellulose by for instance site-directed mutagenesis (see next section), the rate determining step has first to be identified. The endeavors to determine the rate-limiting step have led to divergent interpretations ranging from association or complexation of the cellulose chain into the active site (Shang et al., 2013, Fox et al., 2012), glycosylation step (formation of the glycosyl-enzyme complex) (Knott et al., 2014) to dissociation (Jalak and Väljamäe, 2010, Cruys-Bagger et al., 2012a, Cruys-Bagger et al., 2013b). A recent review (Payne et al., 2015) states that the rate limiting step for GH7 CBH hydrolysis of cellulose in the presence of synergistic enzymes is likely to be the processive velocity (i.e., steps of hydrolysis, product expulsion and the processive motion), while the rate limiting step in the absence of synergistic enzymes is likely to be substrate dissociation. Specifically, they suggest that the slowest step for processive velocity is the glycosylation step (or the first step) in the hydrolysis (Knott et al., 2014) (see section 1.6). This means that the glycosylation step (first step) should be a target for designing faster variants. However, the two transition states in the two-step mechanism present a challenge for rational design that targets the catalytically involved residues. For example, if lowering of the energy of the transition state for the first step (the glycosylation step) by, for instance, stabilizing the glycosyl-enzyme complex is successful, the rate of the following step (the deglycosylation step) will be slowed down as the glycosyl-enzyme complex might become too stable. In cases of two-step mechanism and appertaining transition states, the cellulase has to juggle the energy level of glycosyl-enzyme complex so that it is stable enough to form but not too stable to decay. The explanation of the dissociation as being the rate limiting factor has been rationalized with the difference between the measured processivity and the theoretical processivity, where the term theoretical processivity signifies the ideal processivity on immaculate substrate and is given by $n_{theo} \approx \frac{{}_{p}k_{cat}}{{}_{p}k_{off}}$ (Cruys-Bagger et al., 2013a). Specifically, n_{theo} has been found to be around 100-400, while the experimental processivity range between 14-23 (Cruys-Bagger et al., 2013b). This means that the processivity is stopped before the cellulose would dissociate from the strand, a behavior which has been rationalized with the occurrence of steric obstacles or imperfections of the substrate (Jalak and Väljamäe, 2010, Praestgaard et al., 2011, Eriksson et al., 2002). In the case of obstacles on the cellulose strand, the processive movement of Cel7As would be stalled, which would result in a non-productive population of enzyme. Consequently, Cel7A has to be dethreaded (captured in the composite dissociate rate constant) to re-enter the possible active population. Considering that the dissociation rate constant must overcome all the forces responsible for binding, the energy barrier for this event is likely to be significant. In addition to the obstacle-explanation, the explanation of occurring

traffic jams proposed by Igarashi et al, could also give rise to the same rationale (Igarashi et al., 2011).

With the emerging data suggesting a slow dissociating constant (Cruys-Bagger et al., 2013b, Kurasin and Väljamäe, 2011, Jalak and Väljamäe, 2010) the work presented in the thesis adopts the premise that the dissociation of cellulase from cellulose is the rate determining step. At low substrate concentration however, the complexation rate becomes important (Cruys-Bagger et al., 2013b, Cruys-Bagger et al., 2013a) which implies that the rate-limiting step depends on the substrate concentration. The shift in rate-limiting step has been coupled to the characteristic burst, which occurred at high substrate concentrations but vanished at low substrate loads (Cruys-Bagger et al., 2013b)

Another interesting point made regarding the rate limiting step was that a different rate limiting step for a cellobiohydrolases from the GH6 family applies for diverse cellulose substrate types (Zhang et al., 2000b). In line with this finding, a similar conclusion was reached for another class of processive enzymes, the chitinases (Zakariassen et al., 2010). Another factor which could potentially change the rate-limiting step is changes in temperature, as every kinetic constant has its own temperature dependence. A nice demonstration of a temperature dependent rate-limiting step is given for GH3 β-glycosidase hydrolysis of *p*-nitrophenyl-β-D-glucoside; at low temperature the rate-limiting step is the degradation of the enzyme-substrate intermediate (i.e. analogues to the dissociation limitations discussed above) whereas the slowest step at higher temperature is the formation of the enzyme-substrate intermediate (Weber and Fink, 1980). The change in rate-limiting step was also here elegantly coupled to the occurrence of a burst phase (which only occurs if dissociation is rate-limiting) at low temperature, though not at higher temperatures. In the light of this it is noteworthy that the majority of kinetic constants reported for GH7, are obtained in the range of 20 °C – 30 °C (Igarashi et al., 2011, Igarashi et al., 2012, Kurasin and Väljamäe, 2011, Jalak and Väljamäe, 2010, Cruys-Bagger et al., 2013b, Gruno et al., 2004).

1.5 Temperature

The simplest way to increase the reaction rate of an enzymatic reaction is to increase the reaction's temperature. In conjugation with this general statement, a high process temperature for saccharification and fermentation is also advantageous for restricting contamination risks, reducing viscosity and reducing cooling cost, as pretreatment normally occurs at rather high temperatures (Chokhawala et al., 2015). Altogether this makes the relation temperature and cellulosic enzymes a central topic in the development of industrially relevant enzymes.

1.5.1 Temperature and the saccharification and fermentation processes.

Enzymatic hydrolysis and fermentation in industrial settings can be carried out by two separate processes referred to as **s**eparate **h**ydrolysis and **f**ermentation (SHF) or by one simultaneous process, termed **s**imultaneous **s**acccharification and **f**ermentation process (SSF) (Larsen et al., 2008, Jorgensen et al., 2007). One significant discrepancy between the saccharification and fermentation is the optimal temperature for the two processes; ethanol fermenting yeasts such as *Saccharomyces cerevisiae* have an optimal temperature of 30–35 °C for fermentation while the optimum temperature for cellulosic enzymes is 45–50 °C (Mutturi and Liden, 2013). For SSF this discrepancy in optimal temperature makes the selection of process temperature a critical factor, as it needs to be a compromise between the two, although the process temperature can be tailored for the two independently processes in SHF. In the case of SSF process selection or the development of more thermostable yeast strain, it also seems to be a good solution in order to increase the ethanol yield (de Souza et al., 2012, Caspeta et al., 2014). Another approach which aims to increase the ethanol yield is to shift/change the temperature during the course of the process. One theoretical (Chen and Wang, 2010) and some experimental studies (Huang and Chen, 1988, de Souza et al., 2012, Kang et al., 2012) have shown that changing the temperature during a SSF process results in higher ethanol yield compared with the corresponding isothermal processes.

1.5.2. Temperature and enzymes

The overall rate for enzyme-catalyzed reactions plotted against temperature, has a characteristic curve which resembles a bell. The explanation for the characteristic occurrence of a bell-shaped curve was first offered by Tamann (Laidler and Peterman, 1979) and states that two independent processes account for the shape; the catalyzed reaction and the thermal inactivation of the enzyme. At low temperatures, the catalyzed reactions are accelerated as temperature increases, up to a point where this effect is balanced out by thermal inactivation/denaturation of the enzyme. The optimum temperature is recognized as the peak/maximum of the curve and has been used as a standard value for different wild type enzymes or engineering variants with improved thermal stability (Heinzelman et al., 2009, Voutilainen et al., 2008, Voutilainen et al., 2010, Wu and Arnold, 2013, Teter et al., 2005, Day et al., 2004). In this work we will focus on the ascending branch of this bell curve which has also been the focus of various earlier studies (*Table 1*). Collectively these studies find values for the activation energy (Ea) based on Arrhenius plots (*i.e.* natural logarithm of the parameter plotted against the reciprocal of the absolute temperature) (Arrhenius, 1889). Although the Arrhenius

equation strictly applies to single rate constants (Laidler and Peterman, 1979), it has been shown to also apply to overwhelmingly complicated processes, which must certainly involve several rate constants (Laidler and Peterman, 1979). This is also the case for the studies given in *Table 1*, as none of them aims to find the activation energy corresponding to a single elementary step. Actually, most of the studies use commercial cellulase mixtures (*Table 1*), which make the activation energy for the already complicated reaction of cellulases even more diffuse. In essence, this means that reported activation energies provide a practical measure of how the rate of the process depends on temperature, and the value of Ea is without direct meaning with regard to the energy change associated with a certain structural change.

Another characteristic of these studies is that their primary aim is not to elucidate the activation energies, but to report them in relation to models (Drissen et al., 2007, Brown et al., 2010) or different methods (Hu et al., 2009, He et al., 2000). However, for some the primary goal is to investigate the effect of temperature on the enzymatic hydrolysis (Zhang et al., 2012, Ye and Berson, 2014). By employing a model which takes cellulase inactivation into account, Ye at al. found that the increase in reaction temperature resulted in an increase in reaction rate constant, which was similar to the increase of their rate of enzyme inactivation. Accordingly, they concluded that prolonged hydrolysis time at high temperature is not an efficient way to improve cellulose conversion (Ye and Berson, 2014) because it is directly linked to thermal degradation of the enzyme.

The publication of Garsoux et al. differs from the other listed papers in *Table 1* as it reports the temperature effect on mono-component enzymes. Additionally, it provides a different view as the aim of this paper is to elucidate cold-adaptive feature of a psychrophilic GH5 enzyme (Garsoux et al., 2004) rather than investigating different aspects, which relate to the commercial degradation of cellulose. This field of natural adaption of enzymes to temperature is also most relevant when discussing temperature effects on enzymes. The article of Garsoux et al. is a suitable example for work in this field (German et al., 2012, Lonhienne et al., 2000, Stone et al., 2012) and exemplifies some general trends evident in the temperature adaption strategy highlighted by the comparison of the psycrophilic CelB with a mesophile Cel5. Although changes in temperature significantly change reaction rates, CelB and Cel5 displayed the same hydrolytic activity at 4 °C and 35 °C respectively (Garsoux et al., 2004). Another trend which is exemplified by the CelB and Cel5 relation is a lower Gibbs free energy of activation ($\Delta G^{\ddagger}$)[3] for the psycrophilic enzyme (Garsoux et al., 2004, Lonhienne et al., 2000). Although all

[3] The notation refers to the value of interest at the transition state. This will be further explained in the next section.

these selected papers have contributed much to a better understanding of the cellulases-temperature relationship on soluble substrates, we are not aware of any earlier studies which have investigated the temperature dependence of mono-component cellulases on insoluble substrate and analyzed this relation with respect to a relevant model.

Enzymes	Aim	Hydrolysis conditions	Temperature	Quantified by	Ea	ref
Crude cellulase powder (Shanghai Bio Life Science & Technology Co., Ltd. of China)	Investigating the effect of temperature on the enzymatic hydrolysis by fractal-like kinetics model.	1.72 g crude cellulase powder and 12.50 g pretreated rice straw (10 FPU/g substrate) Continous assay (72 h)	37 °C -50 °C (7 points)	DNS (3,5-dinitrosalicylic acid) method	Ea 21.1 kJ/mol	(Zhang et al., 2012)
Spezyme CP	Investigating the effect of temperature on the enzymatic hydrolysis by a mechanistic model which considers first order inactivation of adsorbed enzyme.	20 g/l Sigmacell 0.6 ml Spezyme/100 ml (50 FPU (filter paper unit)/mL cellulases activity) Continous (72 h)	20 °C, 30 °C and 50 °C	YSI 2700 Select Biochemistry Analyzer.	Ea 16.3 kcal/mol	(Ye and Berson, 2014)
T. reesei cellulase (Genencor 300P) and β-glucosidase (Novozyme 188).	Testing 6 different models and using the Arrhenius plot to validate 6 different models	AFEX treated wheat straw Enzyme and substarte concentartion varied (see article)	35 °C,42 °C and 49 °C	YSI Model 27 glucose analyser	Ea 47.6 kJ/mol.	(Brown et al., 2010)
Celluclast cellulase	Introduction of a Quartz crystal microbalance technique. This method was used to investigate the enzymes kinetics at different temperatures.	0.1-0.5 mg/ml Thin films of cellulose on polyvinylamine precoated quartz crystal sensors continuous	25 °C 32 °C and 40 °C	Microgravimetry with a Quartz Crystal Microbalance	Ea 37 kJ/mol	(Hu et al., 2009)
Pseudoalteromonas haloplanktis	Highlights the cold-adaptive features of a of the psychrophilic enzyme CelB from *Pseudoalteromonas haloplanktis*	*0.2 mM - 30 mM* *p*-nitrophenyl-*β*-D-cellobioside	25 °C 32 °C 40 °C	Absorbance at 405 nm	CelG 48.5 Cel5: Ea 68.1 kJ/mol)	(Garsoux et al., 2004)
Trichoderma viride, Cellulase Shanghai Lizhu Dongfeng Biological Technology Limited Corporation.	Introduction of an acoustic wave (BAW) cellulase sensing technique. This method was used to investigate the enzyme's kinetics at different temperatures.	Continous 2 mg/ml cellulase 2.5 to 20 mg ml_1 carboxymethylcellulose	30 °C. and 35 °C. (two points)	bulk acoustic wave (BAW) cellulase sensing technique,	51.99+/-1.26 kJ/ mol	(He et al., 2000)
Endo-b-1,4-glucanase (EC 3.2.1.4) from *Aspergillus niger*	To test the effect of carboxyl group modification on kinetic properties.	1.5% (w/v) carboxymethylcellulose-Na salt (CMC) 15 min (endpoint)	10 °C – 99 °C (various points)	DNS (3,5-dinitrosalicylic acid) method	53 kJ/mol (wt) 44 – 46 kJ/mol (modified enzyme)	(Shibafuji et al., 2014)
Cellubrix and β-glucosidase Novozymes 188 (Novozymes Corp., Denmark)	Presentation of a mechanistic model for cellulose degradation by cellulases. Temperature is a parameter in the model.	12 g/l Avicel 17.6 FPU/g cellulose (enzyme) Continuous assay 1 h	25 °C, 30 °C 35°C 40 °C 50 °C	oxidase/peroxidase enzyme kit	29.8 kJ/mol	(Drissen et al., 2007)

Table 1. Literature survey on reported activation energies (Ea).

1.6 Free Energy diagrams

With the proviso that mechanism never can be proven, but be "well accepted" and "established", we begin this chapter. To gain a thorough understanding of enzymatic reactions, bound substrates, intermediates, products and transition states on the reaction pathway have to be characterized and described along with the interconversion energies between them. In order to suitable outline these reaction pathways, free energy diagrams have been employed. As we build up a free energy diagram in Article II, this chapter serves the purpose of introducing a classical example of a textbook energy diagram, where the concept behind energy diagrams in enzyme reactions are emphasized and to introduce the reader to a free energy diagram which describes the inner catalytic cycle.

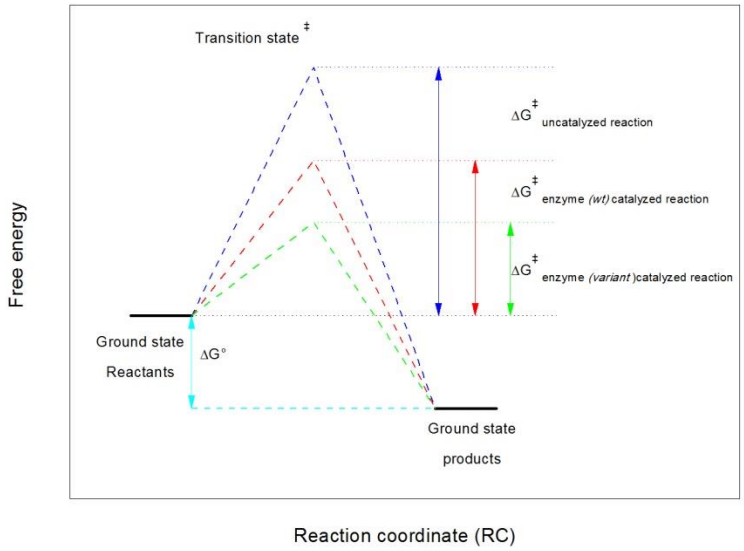


Figure 7: Gibbs free energy diagram for an uncatalysed reaction (blue curve) compared to enzyme catalysed reaction (red and green curves).

Figure 7 illustrates a textbook example of a free energy reaction coordinate profile for an uncatalyzed reaction and for a simple enzyme catalyzed reaction. The free energy difference between reactants and products is indicated by the cyan double arrow (ΔG°). The blue, red and green dashed lines illustrate the potential energy the reaction must overcome to proceed. According to Transition state theory (TST) the highest potential energy along the reaction pathway correspond to the transition state complex and the free energy difference between reactants and the transition state is referred to as Gibbs free activation energy $\Delta G^{\ddagger}$. The free energy diagrams are often used to emphasize that enzymes function by offering an alternative pathway for the reaction with lower Gibbs free activation energy. This is also partly our intention and the higher Gibbs free activation energy ($\Delta G^{\ddagger}$) of uncatalyzed reaction compared with the enzyme catalyzed reaction is illustrated

with the larger blue double arrow and the smaller red and green double arrows. The energy diagram in *Figure 7*, however, does not represent enzymatic reactions very well as energy diagrams for even the simplest reaction consisting of one enzyme, one substrate and one product, may display several pinnacles (transition states) and valleys due to the numerous intermediate complexes. An example is given in *Figure 8* which shows a detailed energy diagram for the hydrolytic steps of the inner catalytic cycle involving glycosylation and deglycosylation for *Hj*Cel7A acting on a cellulose chain.

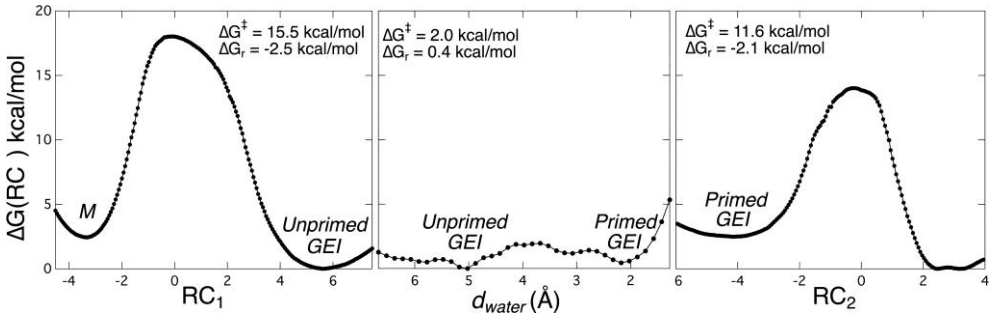


Figure 8: Hydrolytic free energy barriers for TrCel7A. Free energy barriers for the hydrolytic steps of glycosylation (left) and deglycosylation (right) for. M denotes the Michaelis complex and GEI the glycosyl-enzyme intermediate. Unprimed and primed GEI denotes two product-binding modes, where the primed GEI binding mode is suggested to allow a water molecule (the nuclephile water) to approach the anomeric carbon of GEI. The middle section shows the free energy as a function of the distance from the nucleophilic water to the active site (Knott et al., 2014).

The two-step double displacement mechanism (also referred to as ping-pong mechanism) results in two transition states (*Figure 8* and *Figure 9*). The first occurs along the reaction coordinate of the first reaction, where the nucleophile attack on the anomeric carbon results in the glycosyl-enzyme intermediate. Because of the glycosyl-enzyme intermediate, this step has also been referred to as the glycosylation step. Accordingly, the next reaction is referred to as the deglycosylation step as the substrate, the glycosyl-enzyme complex, is degraded through the reaction with a water molecule. In addition to the nucleophile attack of Glu212 in the glycosylation step, the carboxylic acid (Glu217) on the opposite site acts as a general acid and deprotonates the leaving group. The same carboxylic acid (Glu217) also plays a crucial role in the second reaction, where it deprotonates the water, which then attacks the glycosyl-enzyme intermediate, leading to the final degradation of this complex. The second transition state can be found between these reactions.

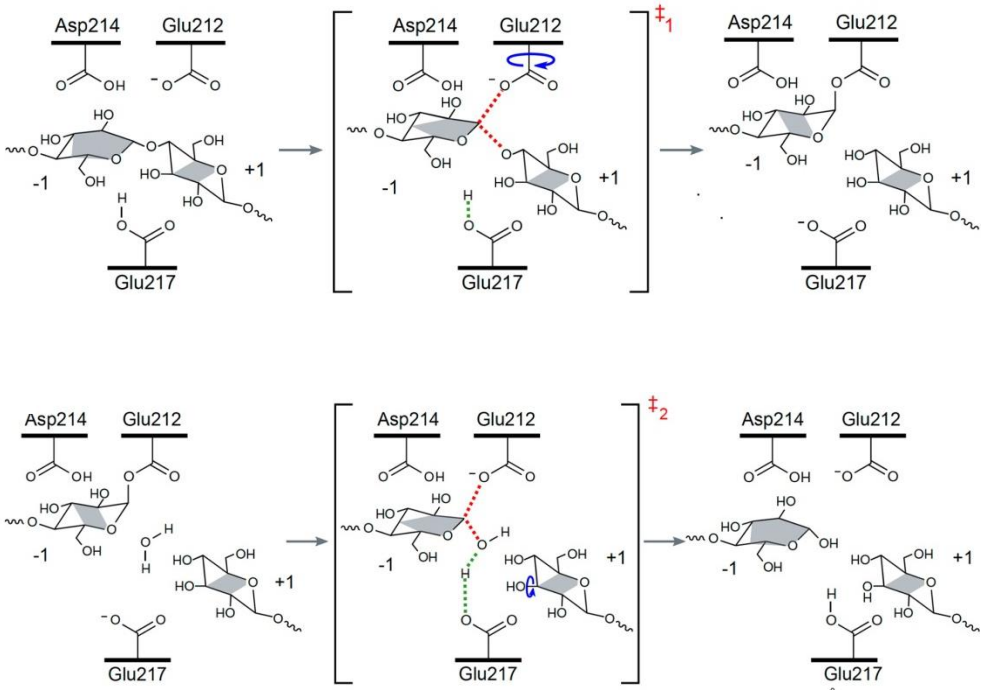


Figure 9. Schematic illustration of the retaining mechanism and the two transition states. In the first step (glycosylation), Glu212 makes a nucleophile attack on the anomeric carbon, which results in the glycosyl-enzyme intermediate. In the second step (deglycosylation) Glu217 deprotonates the water, which then attack the glycosyl-enzyme. Modified from (Knott et al., 2014).

As mentioned above this free energy diagram provides a very detailed description of the hydrolytic cycle. This free energy diagram for inner catalytic cycle is based on a high resolution crystal structure and transition path sampling (Knott et al., 2014) and we recognize that the free energy landscape depicted in this thesis in comparison is coarse, exemplified by the single peak for k_{cat} (section 2.1), which captures all the elementary steps presented in *Figure 8*. TST is particularly useful for analyzing structure-reactivity relationships and by using the Hammond postulate (Hammond, 1955), which states that if there is an unstable intermediate on the reaction pathway, the transition state for the reaction will resemble the structure of this intermediate, this permits us to discuss the structure of the intermediates in section 2.1 on a superficial level. Formulated differently, the Hammond postulate states that species with similar energies along the reaction path will also have similar structures. In the current work we present an energy profiles for the enzymatic hydrolysis of cellulose, with special emphasis on activation parameters for adsorption and desorption of enzyme on the insoluble substrate. Specifically we apply TST to calculate the changes in activation energies. For this thesis, the importance of TST is that it relates the experimentally obtained rate of reaction to the differences in Gibbs free energy between the ground state and the transitions state. According to the Eyring equation which confers to TST, the

activation free energy $\Delta G^{\neq}$ (the pinnacles in the free energy diagram) may be calculated from rate constant.

$$\Delta G^{\neq} = -RT \ln\left[\frac{h_p k_i}{k_B T}\right] \quad \text{(eq. 7)}$$

Where, ${}_pk_i$ is the rate constant, T the temperature, R the gas constant, k_B and h are Boltzmann's and Planck's constants respectively. To separate $\Delta G^{\neq}$ into entropic ($-T\Delta S^{\neq}$) and enthalpic ($\Delta H^{\neq}$) contributions we applied an Eyring plot, $\ln\left[h_p k_i / k_B T\right]$ against temperature, which has a slope of $-\Delta H^{\neq}/R$, where $\Delta H^{\neq}$ (i.e the standard enthalpy change of going from the ground state to the transition state) is the activation enthalpy as defined in transition state theory. Based on the above estimates of free energies- and enthalpies of activation, the entropic contributions to activation $T\Delta S^{\neq} = \Delta H^{\neq} - \Delta G^{\neq}$ can be calculated. For a complete elucidation of the free energy reaction coordinate profile, thermodynamic information on adsorption and hydrolysis is also required for the energy diagram (the valleys). Standard free energy changes of association are calculated as

$$\Delta G^{o}_{assoc} = -RT \ln\left[1/K_d\right] \quad \text{(eq. 8)}$$

where K_d is the dissociation constant. We however used the parameter ${}_pK_M$. The values of ${}_pK_M$ were compared with the partition coefficient K_p, which is model independent and often used as a measure of the overall affinity of a cellulase for the cellulose surface (Palonen et al., 1999, Linder and Teeri, 1996). It should be noted that K_p corresponds to a binding constant, and hence we compare the reciprocal $1/K_p$ with ${}_pK_M$. The elucidation of K_p will be addressed in chapter 2. As for activation free energy $\Delta G^{\neq}$ we also separated the standard free energy changes of association $\Delta G°_{assoc}$ into ΔH°_{assoc}, which can be estimated by a Van't Hoff analysis (*i.e.* plots of ln[K(T)] vs 1/T) using either ${}_pK_M$ or K_p. As we now have a say over ΔG°_{assoc} and ΔH°_{assoc}, ΔS°_{assoc} can be inferred.

1.7 Protein engineering

The main objective of the RESAB project is to design variants of cellulases with higher reaction rates. In doing so the rate limiting step has to be identified and the activation free energy $\Delta G^{\neq}$ of this step lowered (*Figure 7*). The difference between uncatalyzed and enzymes catalyzed reaction is related to the complementary of the enzymes structure to the structures of the substrates[4]. As the complementary part of the enzyme and substrate match up, various molecular interactions are formed and

[4] J.B.S. Haldane 1930

the sum of all these interactions makes up the binding energy, which then can be translated into rate enhancement and/or specificity. The contributions of each single contact point can essentially be probed by removing or exchanging these point for instance by site-directed mutagenesis (Hutchison et al., 1978). For reviews on protein engineering efforts on cellulase in general, see (Schulein, 2000, Bommarius et al., 2014).

As Divne et al revealed, the crystal structure of *H.jecorina* Cel7A they also identified a range of amino acids involved in the substrate binding (Divne et al., 1998) . Of the substrate interacting amino acids, the aromatic tryptophan residues have attracted the most attention due to the high appearance of these amino acids in sugar binding sites as well as the accepted functions of hydrophobic stacking interactions with sugar moity (Audette et al., 2003). Specifically, two of the four tryptophan residues lining the tunnel (W38 and W40) have been assessed by by mutagenesis studies (Kari et al., 2014, Nakamura et al., 2013) (*Table 2*). In addition, other amino acids which display direct contact with the oligosaccharides have been assessed (von Ossowski et al., 2003, Becker et al., 2001). Amino acids which do not have direct contact with the substrate can obviously also be targeted, and in actuality, most of the protein engineering conducted in Cel7A does not target the residues directly involved in the substrate interaction (*Table 2).*shows that there has been a profound interest in stabilizing Cel7A (Voutilainen et al., 2010, Dana et al., 2012, Komor et al., 2012, Heinzelman et al., 2010) as previously mentioned. Along with these studies, the role of N-linked glycosylation has also been probed (Adney et al., 2009). They found that changing the N-linked glycans in the tunnel forming loops had the greatest impact on both thermal stability and enzymatic activity on crystalline cellulose. The tunnel forming loops were also assigned a hotspot by Dana et al who generated a library of 469 chimeras and found that most of the 51 thermostabilized chimeras contained mutations in the loop network (Dana et al., 2012). Finally, Ossowski et al brought matters to a head and probed the role of the loop, which covers the active site. They concluded that this loop had evolved to facilitate the processive degradation of crystalline cellulose (von Ossowski et al., 2003). The loop areas in other GH enzymes have also attracted much attention, but are not selected for *Table 2*, as this table only depicts protein engineering in GH7 enzymes. Some of these loop variants for other GH families will however be introduced to the reader in section 2.2. *Table 2* does also not cover mutations where one of the three catalytic residues are exchanged (Hedeland et al., 1999, Divne et al., 1998).

Organism	Aim	Strategy	Mutation(s)	Hydrolysis condition	Temperature °C	Significance	ref
H.jecorina	Probing the role of stacking interactions	Rational mutagenesis	W40	0.1 % (w/v) High crystalline cellulose from *Cladophora* spp 1.4 µM enzyme 1 h and 2 h	27 °C	W40 is involved in recruiting individual substrate chains into the active site tunnel.	(Nakamura et al., 2013)
H.jecorina *Penicillium funiculosum*	Probing the role of N-linked glycosylation	Rational mutagenesis	*Hj*: N384A, N45A, N384 *Pf*: N45A, N388A, N430A, A196S	1.0 mg/ml BC 1 µM enzyme Continuous assay (120 h)	38 °C	Changing the N-linked glycans in the tunnel forming loops has the greatest impact on both thermal stability and enzymatic activity on crystalline cellulose. The mutations *Hj* N384A and *Pf* N388A and A196S improved hydrolysis of crystalline cellulosic substrate.	(Adney et al., 2009)
H. jecorina	Probing the role of the exo-loop	Rational mutagenesis	Y247F D241C/D249C D(G245-Y252)	0.7 mg/ml BMCC, 1.5 µM enzyme Continuous assay (50 h)	27 °C, 40 °C and 60 °C	The exo-loop contributes to the processive mechanism of *Hj* Cel7A.	(von Ossowski et al., 2003)
H. jecorina	Design and test of variant with lower substrate affinity.	Rational mutagenesis	W38A	0.5-60 g/l Avicel 100 nM enzyme Continuous assay (30 sec)	25 °C	The W38A variant showed reduced substrate affinity and increased in the maximum quasi-steady-state rate. This data suggests that that the dissociation is the rate-limiting step.	(Kari et al., 2014)
Rasamsonia emersonii	Improving thermostability	Rational mutagenesis	G4C/A72C N54C/P191C Q190C/I200C T243C/A375C G266C/D320C	50 mg/ml Avicel 1.4 µM enzyme Continuous assay (17 h)	70 °C, 75 °C and 80 °C	The variant N54C/P191C showed improved activity at 70 °C compared with *Re* Cel7A. All variants tested had improved activity at 75 °C compared with *Re* Cel7A	(Voutilainen et al., 2010)
R. emersonii	To test the effect of adding different types of CBM to *R.emersonii* Cel7A	Rational mutagenesis	Chimeric enzymes: different CBMs	10 mg/ml Avicel 1.4 µM enzyme Continuous assay (48 h)	45 °C - 75 °C	All the chimeric variants had improved activity compared with *Re* Cel7A at 45–65 °C The chimeric enzyme with a CBM3 derived from Clostridium thermocellum cellulosomal-scaffolding proteinCipA showed the highest activity.	(Voutilainen et al., 2014)
H.jecorina	Introduction of a more alkaline pH optimum	Rational mutagenesis	E223S/A224H/ L225V/T226A/D262G	0.7 mg/m BMCC 1.4 µM enzyme Continuous assay (3 h)	27 °C	The inserted histidine residue close to the active site residues changed the pH optimum. No improved activity compared with Hj wt at pH 3,5 and 7	(Becker et al., 2001)
H.jecorina	Probing the role of the linker	Rational mutagenesis	Deletions of linker length	Xylinum microcrystalline cellulose 2.1-10 µM enzyme Endpoint assay (3 h)	50 °C	Deletion of the first one-third of the linker reduces the binding capacity but does not affect enzymatic activity on crystalline cellulose, whereas the bigger deletion reduces the activity.	(Srisodsuk et al., 1993)
5 organisms	Improving thermostability	SCHEMA* Semi rational	32 variants Chimeric enzymes	60 mg/ml Lattice NT cellulose 14.6µg/ml enzyme Endpoint assay (16 h)	50 °C, 65 °C 68 °C and 70 °C	Complete prediction of thermostable chimeric enzymes. 5 tested variants had higher hydrolytic activity at 70 °C than *Re* wt.	(Heinzelman et al., 2010)
R.emersonii	Improving thermostability	FoldX[#] and Semi rational	43 variants Chimeric and point mutations	60 mg/ml Lattice NT cellulose 10 µg/ml enzyme Endpoint assay (20 h)	45 °C – 70 °C	5 tested variants had higher hydrolytic activity at 60 °C -75 °C than *Re* wt	(Komor et al., 2012)
11 organisms	Improving thermostability	Random mutagenesis	469 (library) Chimeric enzymes	4 mg/ml Avicel 0.57 µM enzyme Endpoint assay (22 h)	45 °C - 70 °C	3 tested variants had higher activity at 65 °C – 70 °C	(Dana et al., 2012)

Table 2. Literature survey on mutations introduced in Cel7As from different organisms.

2. PRESENT INVESTIGATIONS

2.1. Temperature effects on kinetic parameters and substrate affinity of Cel7A cellobiohydrolases (Article I, II, II)

2.1.1. Introduction

This chapter presents the combined results and discussion for papers I, II and III. On a very general level, temperature perturbations combine these three articles. More specifically to 1) directly monitor the effect of temperature on kinetic parameters (article I and III) and 2) obtain transition- and thermodynamic parameters, which can further enrich our mechanistic understanding of the Cel7A hydrolysis of cellulose (article II and III). As presented in section 1.5.2., temperature studies on cellulases fall into two experimental categories; either enzyme cocktails are tested on insoluble substrates or mono-component cellulases are tested on soluble substrates. Therefore, we are not aware of any previous studies which systematically investigate the temperature dependence of mono-component cellulase acting on insoluble cellulose and suggest that such investigations could contribute with important information.

2.1.2. Results and discussion

This result and discussion section is composed of 6 subsections. The *first* subsection describes kinetic and processivity data which form the basis for our steady state analysis. Following, the steady state inferred rate constants and their response to temperature will be presented in subsection *two*. The *third* subsection describes the temperature dependence of equilibrium parameters. Based on the kinetic and equilibrium parameters described in section *two* and *three* we form a free energy diagram (subsection *four*). The data described in the first four subsections are subsequently discussed in relation to the effect of CBMs (subsection 5) and the thermal adaption (subsection 6).

Basic experimental data.

We investigated the two native enzymes *H.jecorina* Cel7A and *R.emersonii* Cel7A which are different in the sense that *H. jecorina* Cel7A has two-module architecture while *R. emersoni* Cel7A only consists of a catalytic domain (Grassick et al., 2004, Divne et al., 1994). In addition to these two native enzymes, we also studied two enzyme variants; the catalytic domain of *H.jecorina* Cel7A without CBM and linker, and a chimeric protein composed of the linker and CBM from *H.jecorina* Cel7A and the *R. emersonii* enzyme. The choice of native or wild-type (wt) enzymes is based on the different temperature adaption of the fungi; *H.jecorina* is mesophilic while *R.emersonii* is a thermophile fungus and the choice of variants was to investigate the role of the CBM. We will refer to these 4 enzymes by their origin (*Hj* or *Re*) followed by subscript "CORE" or "CBM" for one- and two domain variants (*e.g.* Hj_{CBM} for the *H. jecorina* wild type).

We measured steady state rates, ${}_pv_{ss}$, for the four investigated enzymes by estimating the ratio of the final cellobiose concentration and the contact time (1 hr.). The rates (${}_pv_{ss}$) were normalized with respect to the total enzyme concentration (E_0=0.40 µM) to obtain the specific activity, ${}_pv_{SS,}/E_0$, in units of s^{-1}. For more details and reflections upon the experimentally obtained steady state rates, consult section 2.1.3. The specific steady state rates ${}_pv_{ss}/E_0$ were plotted as function of substrate load (*Figure 10*).

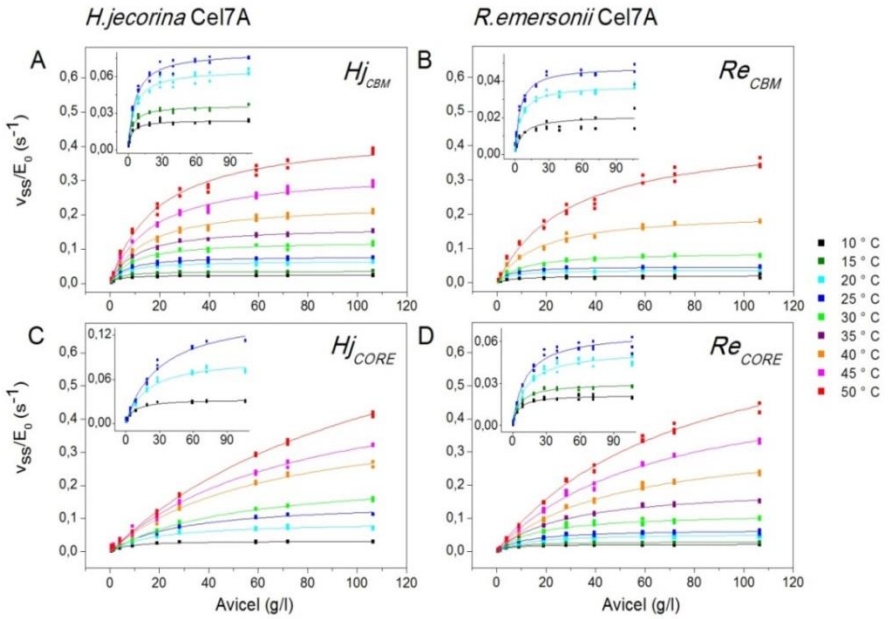


Figure 10: Specific enzyme activity (${}_pv_{ss}/E_0$) for Hj_{CBM} (panel A), Re_{CBM} (B), Hj_{CORE} (C) and Re_{CORE} (D) plotted as a function of Avicel load (0-106 g/l) at between 10 °C and 50 °C. Symbols represent all experimental data from triplicate measurements and lines are best fit of eq. 4. Insets show enlargements of results at the lower temperatures, as they are hard to assess on the main figures. The figure is adapted from article I.

The symbols in *Figure 10* represent the experimental data and the lines are best fits of the Michaelis-Menten (section 1.4.2. eq. 4). The lines seem to fit the experimental data very well under all investigated conditions and the kinetic parameters, $_pV_{max}/E_0$ and $_pK_M$ derived from the non-linear regression are listed in *Table 3*. These data along with processivity measurements (see below) will form the basis of our further analysis.

	T (°C)	$_pV_{max}/E_0$ (s^{-1})x10^3	$_pK_M$ (g L^{-1})
Hj$_{CBM}$	10	24 ± 1	2 .8 ± 0.3
	15	36 ± 1	3.5 ± 0.3
	20	65 ± 1	5.5 ± 0.5
	25	80 ±1	6.2 ± 0.5
	30	123 ± 2	8.6 ± 0.6
	35	165 ± 2	10 ± 0.4
	40	234 ± 4	14 ± 0.9
	45	328 ± 7	17 ± 1.1
	50	442 ± 10	19 ± 1.3
Hj$_{CORE}$	10	33 ± 1	7.1 ± 0 .5
	20	90± 3	18 ± 1 .7
	25	152 ± 5	30 ± 2 .6
	30	232 ± 6	50 ± 2 .6
	40	425 ± 16	63 ± 4 .7
	50	895 ± 35	122 ± 8
Re$_{CBM}$	10	21 ± 2	2.8 ± 2.4
	20	38 ± 1	4.5 ± 0.4
	25	48 ± 1	4.5 ± 0.3
	30	89 ± 1	10 ± 0.7
	40	206 ± 5	17 ± 1.4
	50	439 ± 15	29 ± 2.7
Re$_{CORE}$	10	22 ± 1	4.9 ± 0.6
	15	30 ± 1	6.5 ± 0.5
	20	54 ± 1	12 ± 1.2
	25	67 ± 2	12 ± 1.3
	30	115 ± 4	17 ±1.7
	35	196 ± 3	30 ± 1.3
	40	347 ± 9	48 ± 2.6
	45	534 ± 15	63 ± 3.5
	50	755 ± 28	76 ± 5.1

Table 3: Temperature dependence of kinetic Cel7A wild types, Hj_{CBM} and Re_{CORE}, and the corresponding constructs Hj_{CORE} and Re_{CBM}. Maximal specific rates, $_pV_{max}/E_0$ (± standard error), and processive Michaelis constants, $_pK_M$ (± standard error), were derived from the regression analysis shown in Figure 10.

The processivity number can be estimated from separate measurements as the relative concentrations of soluble products by chromatographic measurements (Horn et al., 2012). Specifically, we used the ratio [cellobiose]/[cellotriose], as a

measure of n, where the concentration of cellotriose indicate the initial cuts, while the concentration of cellobiose reflects consecutive cuts (Horn et al., 2012, Divne et al., 1998). The processivity estimated as the ratio of soluble sugar is an indirect measure and the meaning of the exact values remains debatable (Horn et al., 2012), however, for comparative studies this method offers an easy and robust way to estimate n. The n-values presented in *Figure 11* range from ~10 to ~30, where the two-domain enzymes possess a higher processivity than their one-domain analogs. The processivity numbers obtained here correspond well to previously reported values (9 and 23) for the Hj_{CORE} and Hj_{CBM} *(Cruys-Bagger et al., 2013b).*These results are in close alignment with conclusions reached in previous studies; Zheng and Ding showed that deletion of the family 1 CBM appended to the family 5 endoglucanase from *Volvariella volvaricea*, resulted in decreased processivity (Zheng and Ding, 2013). Similarly, deletion of a family 3 CBM attached to a EG derived from *Thermomonospora fusca* also resulted in lowered processivity (Irwin et al., 1998). In contrast to these studies, a native two-domain GH5 from *Saccharophagus degradans* showed a similar processivity to its truncated version (Watson et al., 2009) indicating that the role of CBM for the processivity might not be straight forward. Results in *Figure 11* suggest a limited temperature dependence of n, as it is essential constant over the range of temperature tested. A similar result was found for the exocellulase Cel48A from *Thermobifida fusca*, which showed moderate decreased processivity at elevated temperatures (Kostylev et al., 2014).

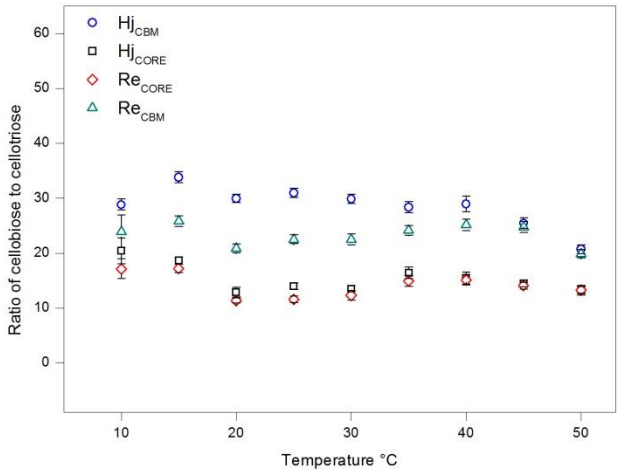


Figure 11: Processivity measured at different temperatures. The average processivity was estimated as the product ratio n~[cellobiose]/[cellotriose]. The concentrations were measured by ion chromatography after 1h hydrolysis trials. Circles (○) illustrate n for $H.jecorina_{CBM}$ Cel7A, squares (□) for $H.jecorina_{core}$, triangles (Δ) for $R.emersonii_{CBM}$ Cel7A and diamonds (◊) for $R.emersonii_{core}$ Cel7A. Data shown as mean ± standard errors of triplicates. The symbols describing the four different enzymes in this figure apply to all figures in this section.

Temperature dependence of kinetic parameters.

To highlight the effect of temperature on the kinetic parameters, we plotted the derived parameters ${}_pV_{max}/E_0$ and ${}_pK_M$ (table 3) and the specificity constant ${}_p\eta$ (eq. 6) as function of temperature (*Figure 12* A, B and C). *Figure 12* shows that that ${}_pV_{max}$, ${}_pK_M$ and ${}_p\eta$ increase with temperature in an exponential manner.

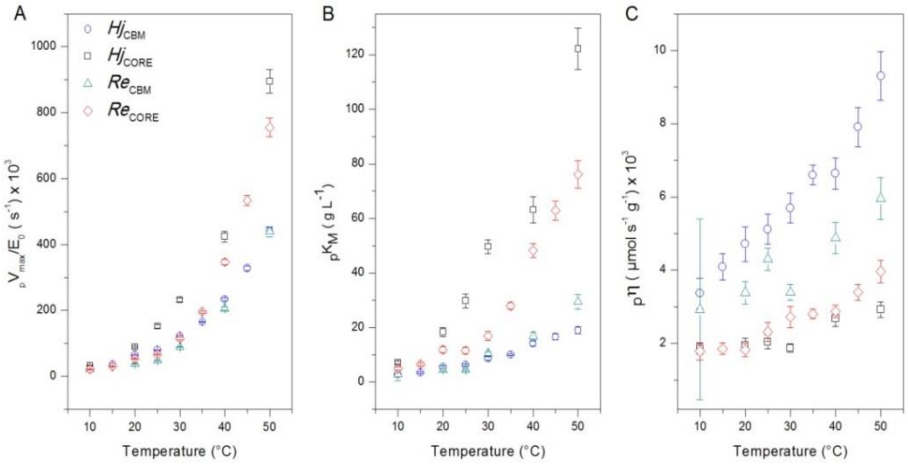


Figure 12. The estimated maximum specific rate (${}_pV_{max}/E_0$, panel A), Michaelis constant ${}_pK_M$, (panel B) and specificity constant (${}_p\eta$, panel C) plottet as a function of temperature.

The temperature dependence of these kinetic parameters, are expressed as Q_{10} values. The Q_{10} value is a factor by which the rate changes as the temperature increase by 10 °C. The intuitive Q_{10} value provides us with a single parameter which we use for further discussions. We estimated the Q_{10} by $\ln(Q_{10}) = {}^{10E_a}/_{RT^2}$ where the activation energies (E_a) is the slopes appearing from plots of the natural logarithm of the kinetic parameters against the reciprocal of the absolute temperature (*Figure 13*).

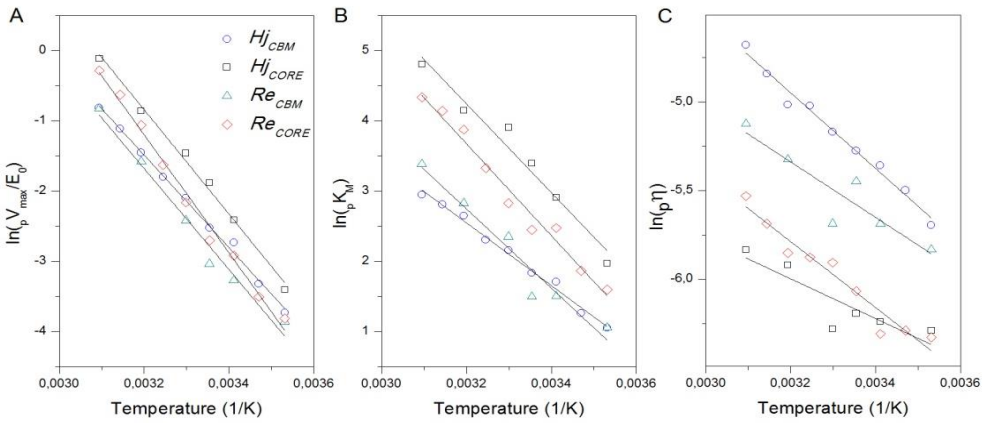


Figure 13. Arrhenius plots. The natural logarithm of the maximum specific rate ($_pV_{max}$, panel A), the Michaelis constant ($_pK_M$, panel B) and specificity constant ($_p\eta$, panel C) plotted against the reciprocal of the absolute temperature.

The results in *Figure 13* shows linear relations and the E_a values (obtained from the slopes) were 60 - 70 kJ/mol for $_pV_{max}$ (panel A), 37 - 52 kJ/mol for $_pK_M$ (panel B) and 9 -16 kJ/mol for $_p\eta$ (panel C). These values translate into Q_{10} values of 2.1-2.6 for $_pV_{max}$, 1.6-2.1 for $_pK_M$ and 1.1 – 1.2 for $_p\eta$. Comparing the Q_{10} values for $_pV_{max}$ and $_p\eta$, we find that the reaction rate is much more temperature activated[5] at high substrate loads than low substrate loads. This conclusion can be made as $_p\eta$ may be interpreted as an apparent second order rate constant (at low substrate loads), which governs the kinetics, $v_{ss} \approx {}_p\eta\,[E]_0[S]$ (Fersht, 1998). As an illustration $_p\eta$ is the linear slope of eq. 4 (shown in *Figure 10*) at very low substrate concentrations, $[S] << {}_pK_M$.

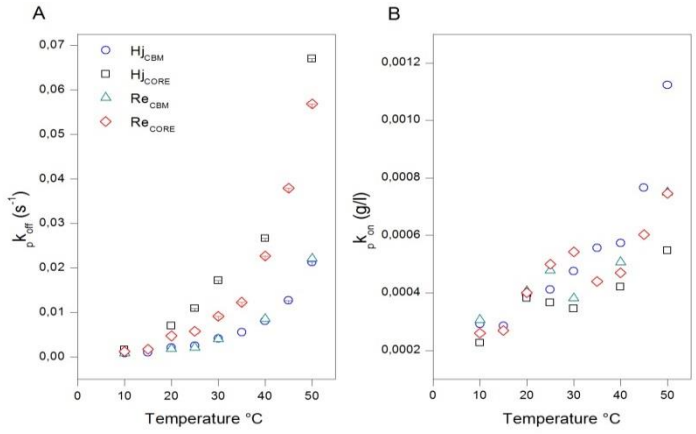


Figure 14. The estimated $_pk_{off}$ (panel A) and $_pk_{on}$ (panel B) plotted as function of temperature.

[5] In other words, the rate is more increased at elevated temperatures.

We have obtained measurements of n, $_pV_{max}$ and $_pK_M$ and may estimate $_pk_{off}$ (eq. 5) and $_pk_{on}$ (eq 6) at all the temperatures tested. As previously mentioned, the majority of the rate constants reported for GH7 are obtained in the range of 20 °C – 30 °C (Igarashi et al., 2011, Igarashi et al., 2012, Kurasin and Väljamäe, 2011, Jalak and Väljamäe, 2010, Cruys-Bagger et al., 2013b, Gruno et al., 2004) but as SHF and SSF have a higher process temperature, information at industrial relevant condition seems pertinent. *Figure 14* shows a plot of $_pk_{off}$ (panel A) and $_pk_{on}$ (panel B) as function of temperature. In continuation of the data analysis for $_pV_{max}$ for $_pK_M$ and to highlight the temperature dependence we find the Q_{10} values for $_pk_{off}$ and $_pk_{on}$ (*Figure 15*). The E_a values derived from the slopes in *Figure 15* were 60 - 70 kJ/mol for $_pk_{off}$ (panel A) and 14-23 kJ/mol for $_pk_{on}$ (panel B) which yield Q_{10} values of 2.1-2.6 for $_pk_{off}$ and 1.2–1.3 for $_pk_{on}$. In relation to these data we found the Q10 value of 1.5 for the catalytic event (see article II). Thus these data indicate a much higher temperature dependence of $_pk_{off}$ compared with $_pk_{on}$ and $_pk_{cat}$. It is worth noticing that the same conclusion can be reached without assuming that $_pk_{cat} >> {}_pk_{off}$ (see section 2.1.3).

This much higher temperature sensitivity of the dissociation, naturally lead to speculation whether $_pk_{off}$ is still the rate limiting constant at increasing temperatures. This however cannot be revealed by our steady state approach and we suggest that pre-steady state measurements should be performed in order to have a final say over the rate-limiting step at increased temperature. Literature values for the kinetic constants also do not hint a possible change in rate determining step, as they are obtained in the range of 20 °C – 30 °C (Igarashi et al., 2011, Igarashi et al., 2012, Kurasin and Väljamäe, 2011, Jalak and Väljamäe, 2010, Cruys-Bagger et al., 2013b, Gruno et al., 2004). Change in the rate limiting step with increasing temperatures has previously been suggested for GH enzymes due to a breakdown in an Arrhenius plot (Weber and Fink, 1980). Breakpoints in an Arrhenius plot may indicate a change the rate-limiting step for the reaction and therefore the linearity of the Arrhenius plot (*Figure 13*, panel A) do not support an interpretation of a change. However the absence of a break does not provide the final proof. As seen in *Figure 13* and *Figure 15* $_pk_{on}$ and $_pk_{off}$ conform to a good approximation to the Arrhenius equation. The linear relationship of chosen kinetic parameter in an Arrhenius plot is seen as a good indicator for a valid model (Brown et al., 2010). Thus, this linear relationship of $_pk_{on}$ and $_pk_{off}$ K_M and $_pV_{max}$ further suggest the validity of scheme 1.

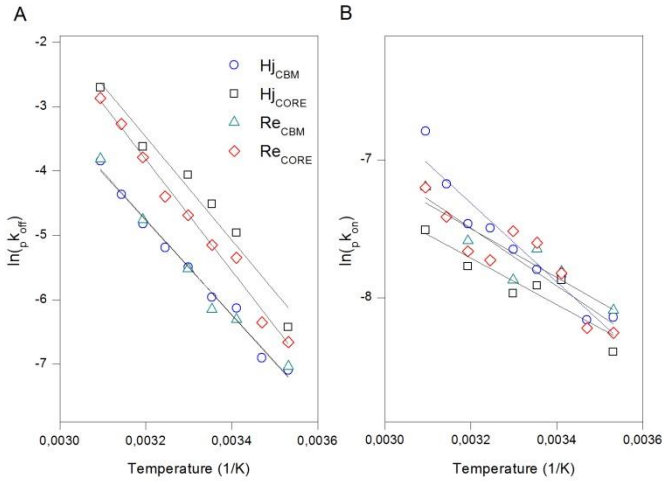


Figure 15. Arrhenius plots. The natural logarithm of $_pk_{off}$ (panel A) and $_pk_{on}$ (panel B) plotted against the reciprocal of the absolute temperature.

Temperature dependence of equilibrium parameters: The affinity of Cel7A to cellulose can be measured in a separate experiment, were the free enzyme concentration is the observable. These measurements are independent of the steady state model and can be seen as an additional method to underpin the conclusion made for the steady state model. Various studies have previously found that the adsorbed population of both two-domain Cel7A and isolated CBM decreases with increasing temperature (Palonen et al., 1999, Moran-Mirabal et al., 2011, Linder and Teeri, 1996, Ooshima et al., 1983) and actually, according to Le Chatelier's principle, lowered affinity at higher temperature is inevitable for an exothermic reaction.

To find these equilibrium parameters we found the distributions between free and adsorbed enzymes by measuring the fluorescence of the bulk enzymes after hydrolysis experiments. Results of these measurements are presented in *Figure 16*, which shows the fraction of bound enzyme as a function of the substrate load. To quantify the affinity, the data were translated into substrate coverage (Γ) in μmol/g, using the relationship, $\Gamma=(E_0 - [E])/S_0$, where E_0 and S_0 are the initial concentration of enzyme and substrate respectively, and [E] is the measured free enzyme concentration. We plotted Γ against the free enzyme concentration, [E] (not shown) and determined the slope for $[E]\rightarrow 0$. This is the so-called partitioning coefficient, K_P, which is often used as a measure of the overall affinity of a cellulase for the cellulose surface (Palonen et al., 1999, Linder and Teeri, 1996).

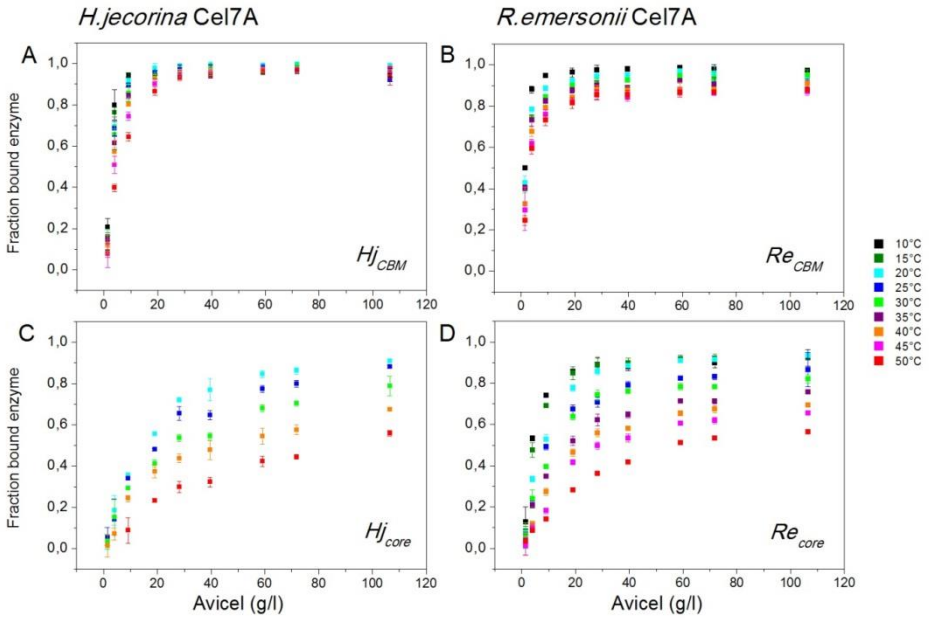


Figure 16. Fraction of bound Cel7A in the hydrolysis samples as function of substrate load at temperatures from 10 to 50 °C. The total enzyme concentration was 400 nM in all samples and the measurements we made after 1 hr. contact time. All points are average ± standard deviation for triplicate measurements. The figure is adapted from article I.

The data above gives a nice overall description of the affinity and its temperature dependence but does not offer detailed information of the affinity contribution of individual amino acid present on either the core domain or the CBM. As Divne et al revealed the crystal structure of *H.jecorina* Cel7A they also identified a range of amino acids involved in the substrate binding (Divne et al., 1998). To probe the affinity throughout the tunnel, we tested the binding of *Hypocrea jecorina* Cel7A to cello oligosaccharide (COS) of different degree of polymerization (DP). Isothermal titration calorimetry (ITC) offers all the parameters ΔG°, ΔS°, ΔH° and the dissociation constant Kd in one single experiment. For the binding of Cel7A to cello oligosaccharide by ITC, we generated a catalytic impeded variant were the nucleophile Glutamic acid residue in position 212 was exchanged with a Glutamine residue henceforth referred to by the one letter code E212Q. This variants has previously been generated with for the purpose of determining the crystal structure of *H.jecorina* complexed with oligo saccharides (Divne et al., 1998). For the purpose of determining an oligo saccharide complexed crystal structure, the activity of E212Q was sufficient low, but for ITC measurements the residual activity complicated the analysis due to a decreasing concentration of COS at room temperature. This is reflected as a shift towards the left of the inflexion point in (*Figure 17*). As previously stated, the simplest way to increase the reaction rate of

an enzymatic reaction is to increase the reactions temperature and obviously the inverse applies too, so by lowering the temperature to 10 °C no observable reduction in the COS concentration was measured within the timeframe of an ITC experiment. Therefore the parameters ΔG°, ΔS°, ΔH and the dissociation constant Kd could be extracted at 10 °C. As the data in *Figure 16* and the various studies (Palonen et al., 1999, Moran-Mirabal et al., 2011, Linder and Teeri, 1996, Ooshima et al., 1983) have established that the amount of bound enzyme decreases with increasing temperature and that higher temperature is preferred in the industrial saccharification process, it seems pertinent to assess ligand binding at higher temperature. However, as just accounted for the hydrolytic activity of E212Q was not sufficient low at higher temperature. It appeared however that the binding of the E212Q-COS was strong, which results in a plateau in the enthalpograms (A graph of enthalpy plotted against time during a reaction) before the curve rises (*Figure 17*). The plateau represents a scenario where all the injectants bind and therefore ΔH° is simply the measured heat normalized with respect to the amount of injectant.

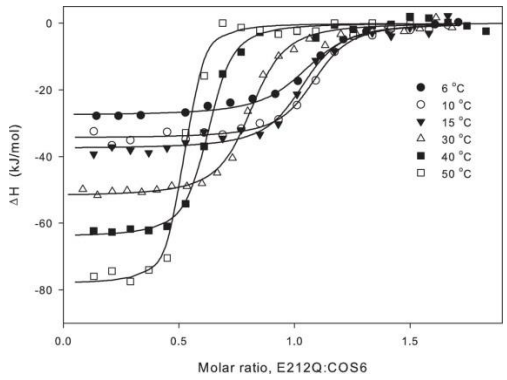


Figure 17. Example of a temperature series for the binding of COS to the E212Q variant. The enthalpy change upon injecting E212Q into a COS6 solution in the calorimetric cell is plotted against the E212Q:COS6 molar ratio.

Changes in temperature now serve two purposes for in the experimental setup. In addition to being one of two primary variable factors (the other being the COS length) it works as a tool to obtaining parameters which we otherwise could not determine. A concise survey of the latter follows; the binding enthalpies were read from the plateau levels in the enthalpograms before the curve rises. Next the heat capacity change ΔC_p was determined as the slope of a linear fit to ΔH (T). Subsequently the thermodynamic data (obtained at 10 °C) and the heat capacity can be inserted into Van't Hoff equation (Clarke and Glew, 1966, Schonbeck et al., 2012) which gives $K_B(T)$, and hence the standard free energy change ($\Delta G^0(T) = -RT \ln[K_B(T)]$) and standard entropy change ($T\Delta S^0(T) = \Delta H(T) - \Delta G^0(T)$). These three parameters were then plotted against temperature (*Figure 18*). As for the

stable thermodynamic parameters described in *Figure 16*, *Figure 18* also shows a reduced affinity with increasing temperature illustrated by the less negative ΔG°. The increase might be difficult to discern from *Figure 18* and we give the example that ΔG° changes from -37 kJ/mol at 10 °C to -32 kJ/mol (ΔΔG° 5 kJ/mol) at 50 °C for COS8 and this results corresponded with data obtained from *Figure 16*, where the difference in standard free energy changes at 10 °C and 50 °C was ΔΔG°= 5 kJ/mol. *Figure 18* shows that the binding enthalpy for COS8 became more exothermic (decreasing slope) with temperature. The two major contributions to this exothermic reaction are 1) polar interactions, which are governed by the substrate-enzyme interacting hydrogen bonds and (Divne et al., 1998) 2) hydrophobic interactions, which are associated by release of structured water from non-polar surfaces. These contributions are opposite, meaning that the polar interactions are exothermic (ΔH < 0) while the release from of structured water from non-polar surfaces are endothermic (ΔH>0). Thus at higher temperature the binding enthalpy is dominated by the hydrogen bonds. This favorable change in binding enthalpy with temperature is counterbalanced by an unfavorable contribution of $T\Delta S^{\circ}$, which in the end results in lower affinity at increasing temperatures.

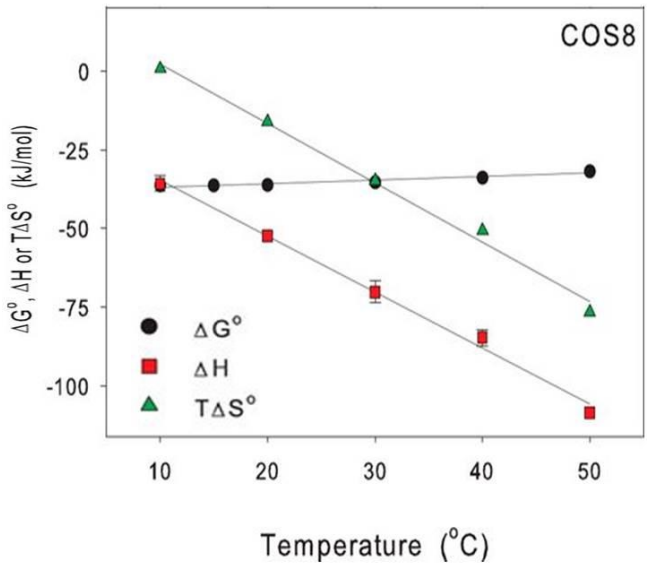


Figure 18. ΔG°,ΔTS°, and ΔH for Cel7A-COS8 binding plotted as function of temperature.

Free energy diagram.

Based on the simplified reaction scheme, presented in section 1.4.2., we aim for describing a free energy diagram (i.e. a plot that shows changes in free energy through the reaction path) (section 1.6). In order to achieve this we have measured the temperature sensitivity of the rate constants ${}_pk_{on}$, ${}_pk_{off}$ and ${}_pk_{cat}$. The three rate

constants were inserted into the Eyring equation (section 1.6, eq. 7) to calculate the activation free energy $\Delta G^{\neq}$ of the rate of association ($\Delta G^{\neq}_{assoc}$), catalysis ($\Delta G^{\neq}_{cat}$) and dissociation ($\Delta G^{\neq}_{dissoc}$). We note that while $\Delta G^{\neq}_{assoc}$ and $\Delta G^{\neq}_{dissoc}$ were obtained for the four enzymes Hj_{CBM}, Hj_{CORE}, Re_{CBM} *and* Re_{CORE}, we only have $\Delta G^{\neq}_{cat}$ values for Hj_{CBM}. *Figure 19* shows the transition state free energies for the rate of association ($\Delta G^{\neq}_{assoc}$) and dissociation ($\Delta G^{\neq}_{dissoc}$) plotted as function of temperature for each of the four investigated enzymes. The data point in *Figure 19* show a fairly constant $\Delta G^{\neq}_{dissoc}$ when plotted against temperature, whereas $\Delta G^{\neq}_{assoc}$ increase steadily though the temperature interval. Thus, these results suggest a stronger temperature dependence of $\Delta G^{\neq}_{assoc}$ compared to $\Delta G^{\neq}_{dissoc}$. Comparisons of the 4 enzymes (Hj_{CBM}, Hj_{CORE}, Re_{CBM} and Re_{CORE}) reveal that $\Delta G^{\neq}_{assoc}$ does not differ with the modularity of the enzymes, whereas a systematic increase in activation free energy for dissociation $\Delta G^{\neq}_{dissoc}$ for enzymes with CBM compared to one-domain variants was observed.

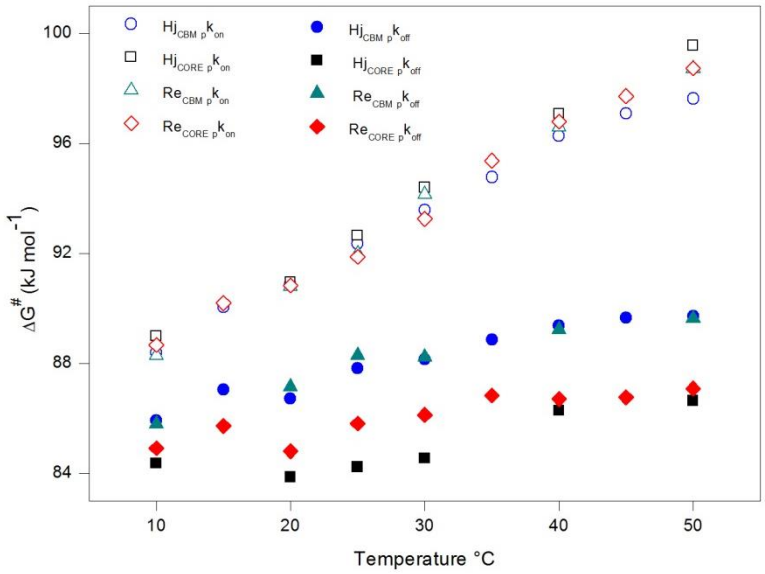


Figure 19: Transition state free energies determined from eq 7(section 1.6) and plotted as a function of temperature. Open symbols show values for the association step (calculated from $_pk_{on}$), while filled symbols are for dissociation (calculated from from $_pk_{off}$).

The activation free energies ($\Delta G^{\neq}$) were dissected into contribution from $\Delta H^{\neq}$ and $\Delta S^{\neq}$ (section 1.6). Thus, we plotted $\ln\left[h_p k_i / k_B T\right]$ against 1/T in *Figure 20* which shows a slope which equals $-\Delta H^{\neq}/R$. Given the estimates of respectively free energies- and enthalpies of activation, the entropic contributions to activations $T\Delta S^{\neq}=H^{\neq} - \Delta G^{\neq}$ were calculated. In summary, we have obtained transition state parameters for $\Delta^{\ddagger}G$, $\Delta^{\ddagger}H$ and $\Delta^{\ddagger}S$ and that for the association, catalysis (data not shown) and dissociation as defined in the simplified reaction scheme (*scheme 1*).

All 9 transitions state parameters are found for *Hj$_{CBM}$* Cel7A and for Hj$_{CORE}$, Re$_{CBM}$ and Re$_{CORE}$ we have 6 parameters; association- and dissociation- transition state parameters.

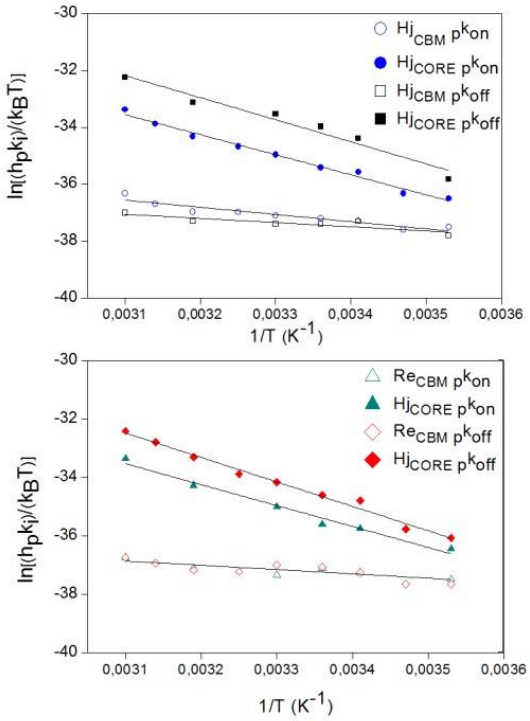


Figure 20: Eyring plots for association (open symbols; calculated from $_pk_{on}$) and dissociation (filled symbols; calculated from $_pk_{off}$) for the four investigated enzymes. Squares represent two-domain enzymes (i.e. with a CBM) and triangles are for one-domain enzymes. The slope, α, specifies the activation enthalpy, $\alpha=-\Delta H^{\neq}/R$.

Thermodynamic information on adsorption and hydrolysis is also required for the energy diagram (the depression in the energy diagram). Using (eq. 8) yields standard free energy changes of association, ΔG°_{assoc}, of – 40 to -45 kJ/mol. The enthalpy change for the association process, ΔH°_{assoc}, can be then estimated by a Van't Hoff analysis using either $_pK_M$. This approach suggested ΔH°_{assoc} values of -40 to – -50 kJ/mol, and a similar value was found by direct (calorimetric) measurements of ΔH° for the binding of cellooctaose to *Hj*Cel7A around room temperature (paper III). As the values of ΔG°_{assoc} and ΔH°_{assoc} are comparable in size, it follows that the entropic contribution to association, $T\Delta S^{\circ}_{assoc}= \Delta H^{\circ}_{assoc}-\Delta G^{\circ}_{assoc}$ at room temperature must be close to zero. We additionally need to consider the standard free energy change (ΔG°_{cat}) for the conversion of cellulose into cellobiose. As we are not aware of any data which can provide us this value, we use data for the β-1,4 glycosidic bond in cellobiose, which was found to be about -16 kJ/mol at room temperature (Tewari et al., 2008). Based on this, we use

ΔG°_{cat} =-16 kJ/mol in the energy diagram (*Figure 21*) The enthalpy change for the hydrolysis of glycosidic bonds in Avicel conversely, has previously been measured and is weakly exothermic, with a ΔH°_{cat} of about -4 kJ/mol (Murphy et al., 2010b). Note that this value appears twice in the free energy diagram (*Figure 22*, *Figure 23* and

Figure 24). We now have in total of 18 parameters of which 15 are experimentally obtained, and we can form the free energy diagram for *H.jecorina* Cel7A.

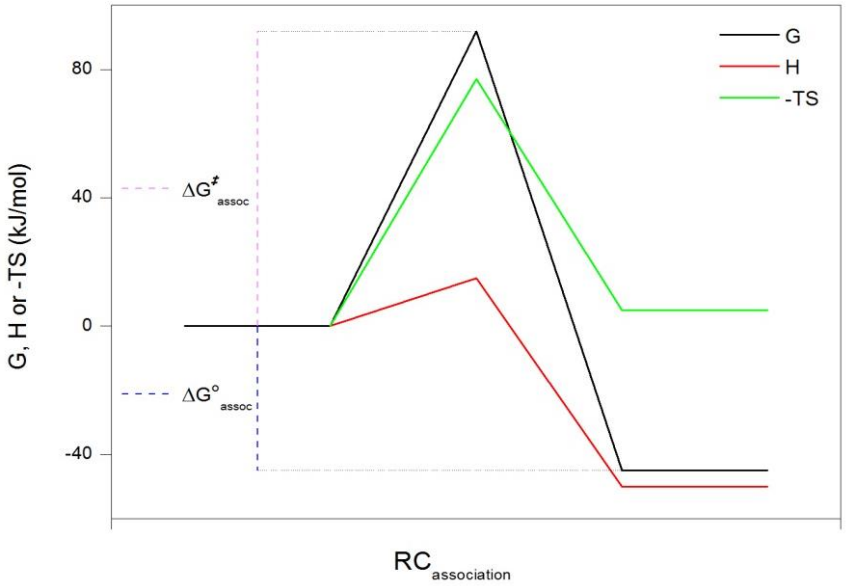


Figure 21: Energy diagram for the Cel7A-cellulose association. The abscissa is the reaction path and the ordinate shows changes in ΔG, (black curve), ΔH (red curve) and –TΔS (green curve) along the reaction coordinate (RC). The entropic contribution is given as its negative value, –TΔS. The colored dashed lines identifies the Transition state free energies for association, $\Delta G^{\ddagger}_{assoc}$ (pink) or the standard free energy changes of association, ΔG°_{assoc} (blue).

Association: The necessity of adsoption of cellulase to the solid cellulose through either the CBM and/or the catalytic domain is given by the simple fact that only surface adsorbed enzyme is potentially active. As can be seen in *Figure 21*, the rate of association is governed by a sizable activation free energy ($\Delta G^{\neq}_{assoc}$ ~ 92 kJ/mol) (black track) which is dominated by the entropic contribution ($-T\Delta S^{\neq}_{assoc}$ ~77 kJ/mol) (green track) and with a small contribution from the activation enthalpy, $\Delta H^{\neq}_{assoc}$ (15 kJ/mol) (red track). As previously stated, the Hammond postulate permits us to discuss the structure of the intermediate on a superficial level. As the entropy is the dominant part and as the exact values ($-T\Delta S^{\neq}_{assoc}$ ~77 kJ/mol) has actually been suggested to be comparable to the penalty of restraining the rotational and translational degrees of freedom of the enzyme (Page and Jencks,

1971), we surmise that the transition state is located early in the path of complexation, where these associated enzyme has lost its degrees of freedom, but not yet established extensive contacts with the cellulose strand in the binding tunnel. The interpretation is also supported by the low activation enthalpy, as few changes in the interactions (e.g. hydrogen bond) would occur at this point. To take this interpretation one step further, the transition state of association could be the initial threading of the cellulose strand into the tunnel. This interpretation would be in alignment with previous studies that find the work required to abstract a glucopyranose unit from crystalline cellulose is quite high (5 to 14 kJ/mol depending on position and crystal structure), although the affinity in the first part of the catalytic tunnel only contributes -3 to -7 kJ/mol per glucopyranose unit to the standard free energy of binding (Beckham et al., 2011, Colussi F. et al., 2015). Thus, these observations hint that early complexes in the association path could be rather unstable and may constitute the transition state of the rate of association.

Following the transition state, we find that the intermediate EC_m complex has almost the same entropy, but a much lower enthalpy (*Figure 19*) and thus we suggest that the association at room temperature is mainly dominated by enthalpy. In contrast to this suggestion adsorption of proteins from the bulky phase on solid surfaces is generally thought to be dominated by an entropy increase (Norde, 1986, Nakanishi et al., 2001) and the same conclusion was reached for binding of two different CBMs to microcrystalline cellulose (Creagh et al., 1996, Georgelis et al., 2012). They argued that the large positive changes in entropy for CBM binding to cellulose is consistent with freeing constrained water at the hydrophobic surfaces of the protein and cellulose (Creagh et al., 1996). The divergent conclusions may be explained by the composite nature of ${}_pk_{on}$ as it covers several steps, such as adsorption/binding of the enzyme to the cellulose surface, recognition of an accessible free end and threading of the cellulose strand. However, as we will see later (page 38), the distribution of $\Delta H°$ and $\Delta TS°$ are much temperature dependent and structural interpretation should be put carefully forward.

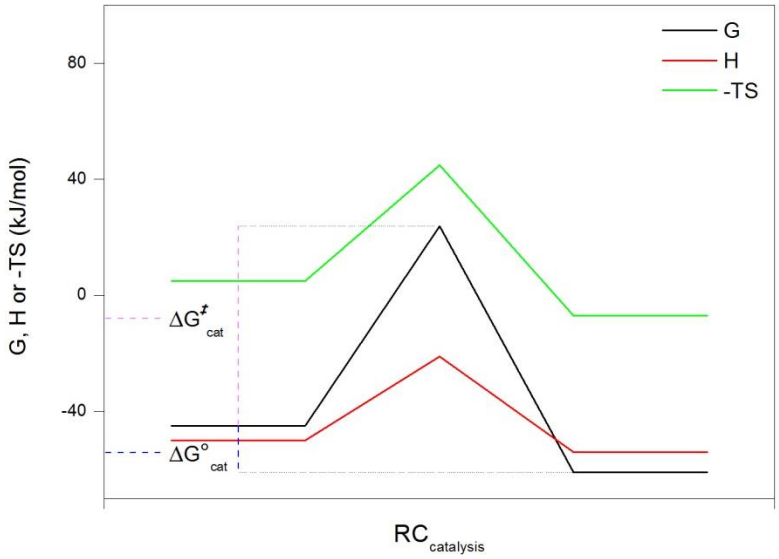


Figure 22: Energy diagram for the Cel7A hydrolysis.

Catalysis: Data for the rate constant governing the inner catalytic cycle (the hydrolysis of β-1,4 glycosidic bond, expulsion of the formed product and sliding of the cellulose chain forward in the tunnel to form a new activated complex) are not described in this thesis (see paper II). In contrast to the ${}_pk_{on}$ and ${}_pk_{off}$ rate constant, ${}_pk_{cat}$ cannot be derived from steady state data as we assume that ${}_pk_{cat} >> {}_pk_{off}$.
Figure 22 shows the energy diagram for the inner catalytic cycle. In contrast to the very detailed energy diagram presented in section 1.6 (*Figure 8*), which depicted the glycosylation- and the de-glycosylation step along with two product binding modes (Knott et al., 2014), the energy diagram in *Figure 22* gives a rather coarse description, where these different steps are lumped together. Despite these differences, Knott et al find an energy barrier of 65 kJ/mol (the glycosylation step), corresponding very well with our extracted data ($\Delta G^{\neq}_{cat}$ = 69 kJ/mol). The enthalpy of activation was estimated on the basis of Eyring plots using data at 5 °C, 25 °C and 50 °C (paper II). This analysis showed $\Delta H^{\neq}_{cat}$ = 29 kJ/mol and entropic contributions $-T\Delta S^{\neq}_{cat}$ was therefore 40 kJ/mol. As the entrophy contribution is fairly higher, this could indicate that restriction of the involved components responsible for formation of the glycosyl-enzyme intermediate (the first transition state *Figure 8*) would be a significant contribution to the free energy barrier of catalysis.

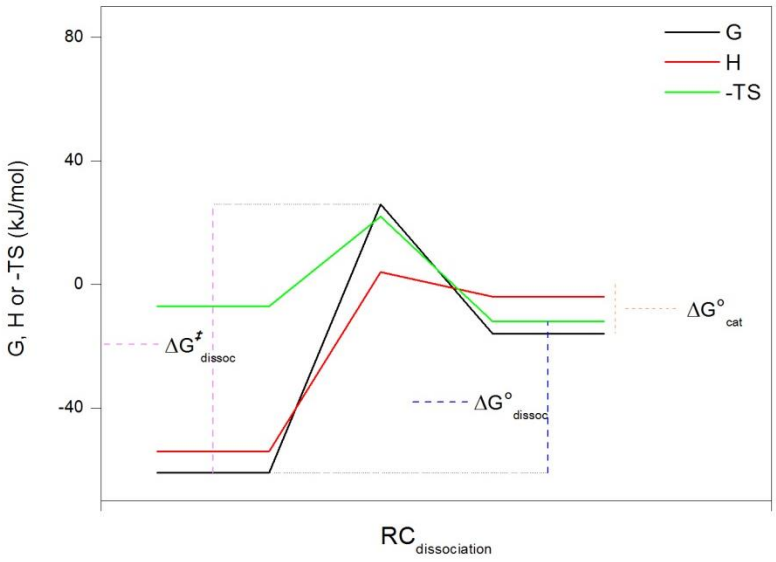


Figure 23: Energy diagram for the dissociation of Cel7A from the insoluble cellulose.

Dissociation. Opposite to the association, *Figure 23* reveal that $\Delta G^{\neq}_{dissoc}$ ~ 87 kJ/mol, is dominated by the enthalpy contribution, $\Delta H^{\neq}_{dissoc}$ ~ 58 kJ/mol. This, together with the observation that the enthalpy of the transition state for dissociation is on the same level as the dissociated components, suggests that the transition state is located late in the dissociation path with few remaining enzyme-substrate contacts. We note that $\Delta H^{\neq}$ is the standard enthalpy change of going from the ground state to the transition state and is analogues to the activation energy E_a. As discussed on page 33 reactants with high activation energies E_a have higher temperature activation and this obviously also apply for the activation enthalpy quantity $\Delta H^{\neq}$. Thus, the comparably high $\Delta H^{\neq}_{dissoc}$ has the consequence that dissociation is accelerated much more by increasing temperatures than the two other steps.

As highlighted in the General introduction (section 1.6) and in the discussion for the inner catalytic cycle on the previous page, the structure of the transition state for the catalytic event has been discussed in more details. In contrast, experimental and computational data is absence for elucidating the structures of the transitions state of the rate of association and dissociation. Additionally, we are not aware of any publications which consider possible structures of the transitions states of the rate of association and dissociation and we postulate that the structural interpretations given in this thesis (and in publication II) are the first suggestions regarding these structures.

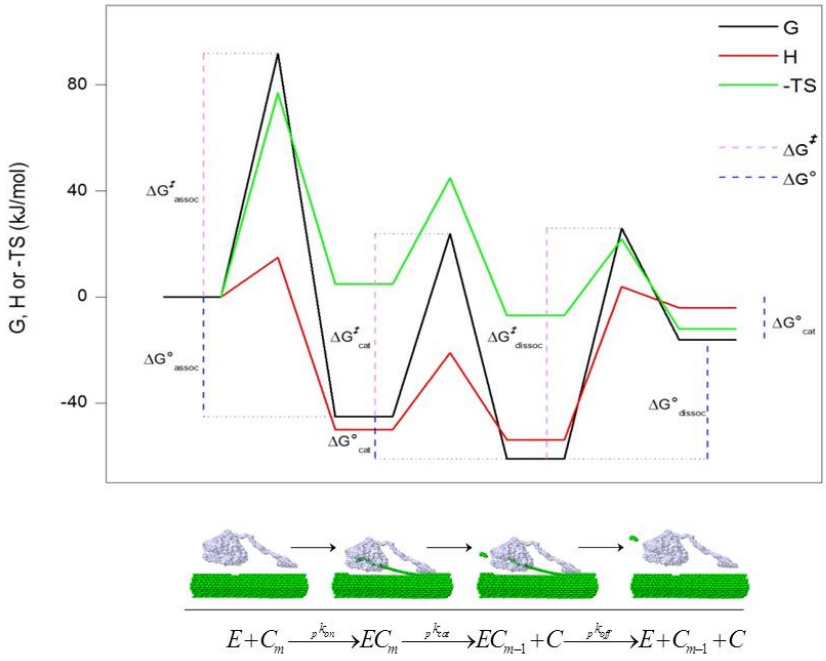


Figure 24: Complete free Energy diagram for Cel7A hydrolysis of insoluble cellulose. The diagram defines a total of five changes (three activated complexes and two equilibria.

In summary, Gibbs free activation energy for the rate of association $\Delta G^{\neq}_{assoc}$ is dominated by entropic contributions while enthalpy contributions mainly govern $\Delta G^{\neq}_{dissoc}$. For $\Delta G^{\neq}_{cat}$ the entrophy contribution is fairly higher (*Figure 24*). Based on the fundamental description of Cel7A hydrolysis of cellulose given by the elucidation of the free energy diagram and the known architecture of Cel7A (Divne et al., 1998) , we suggest some general rational alteration of the enzyme structure, which may lower the transition states free energies and eventually lead to a higher catalytic rate. The early transition state offset by entropy for association implies that $_pk_{on}$ is insusceptible to interaction in the active site tunnel as all the interactions between the active site of the enzyme and the cellulose strand are established subsequent to the transition state. In this case protein engineering which targets the interactions at the entrance of tunnel may have only a moderate if any effect on $_pk_{on}$ One of the most distinctive differences of the structure of *H.jecorina* Cel7A and *R.emersonii* Cel7A is that the latter possesses an extra tyrosine residue (Tyr47) located at the entrance of the tunnel (*Figure 25*), which has been suggested to constitute an additional binding site (Grassick et al., 2004). In line with the insensitivity of $_pk_{on}$, *H.jecorina* Cel7A and *R.emersonii* Cel7A do possess similar $\Delta G^{\neq}_{assoc}$ values (*Figure 19*) although *R.emersonii* Cel7A may possess extra cellulose binding sites. Another consequence of the uniform $\Delta G^{\neq}_{assoc}$ values is that the CBM does not contribute to $_pk_{on}$, which further highlights the insusceptibility of

the association. We suggest that protein engineering aiming at lowering the transition state barrier for $_pk_{on}$ by modifying the entrance or tunnel interactions possible result in no or only moderate changes.

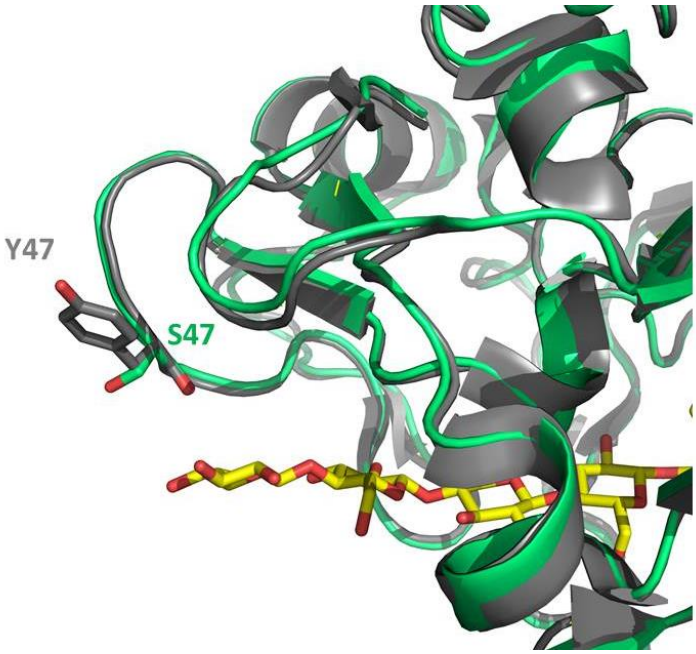


Figure 25. Superposition of H.jecorina Cel7A CEL7 (green), and R.emersonii 3PL3 (gray). The illustration highlights the tunnel entrance area around subsite -7. R,emersonii Cel7A contains a tyrosine at position 47 whereas H.jecorina contains a serine at the equivalent position. The side chain of Y47 and S47 are shown (PyMol).

In contrast to $_pk_{on}$, we surmised that the transition state for dissociation is located late in the reaction pathway where most enzyme-substrate interactions are broken. This means that all the enzyme substrate interactions are upstream the dissociation and that $_pk_{off}$ are much dependent on these contacts. The prediction that modification of the enzyme substrate interactions would have a moderate effect on $_pk_{on}$ but primarily influence $_pk_{off}$, was recently demonstrated for *H.jecorina* Cel7A, where the removal of an enzyme substrate interaction was modified by exchange of a tryptophan with an alanine (Kari et al., 2014). This substitution led to reduced affinity for the substrate which was mainly exerted through faster rate of dissociation. Additionally, we found a small but significant increase in activation free energy $\Delta G^{\neq}_{dissoc}$ for enzymes with CBM, implying that the CBM exerts its effect by lowering the dissociation (discussed more thoroughly below) (*Figure 19*). Altogether these data suggest that $_pk_{off}$ may be tuned in a controlled way by targeting enzyme substrate interactions or by engineering the CBM.

The effect of CBM. The generally accepted function of the CBM is to facilitate higher enzyme affinity for the solid substrate surface (Várnai et al., 2013, Boraston et al., 2004). In accordance with these findings, we find the difference in the standard free energy of association (ΔG°_{assoc}) for one- and two domain enzymes was about 4 kJ/mol (derived on the basis of ${}_pK_M$ and Kp values at 25 °C inserted into eq. 8) and we note that this value is in agreement with earlier reported values (Kari et al., 2014, Stahlberg et al., 1991).

Although accumulated knowledge links affinity and CBM, it is not known how the increased affinity can be dissected into contribution from association and dissociation. Returning to *Figure 19*, we recall that the transition state free energies for association were the same for the four variants (Hj_{CBM}, Hj_{CORE}, Re_{CBM} and Re_{CORE}). Additionally, the similar slopes were found for all the enzymes (*Figure 20*), suggested that the CBM did not also not affect the contribution of $\Delta H^{\neq}_{assoc}$ and $-T\Delta S^{\neq}_{assoc}$ to transition state free energies for association. Cumulatively, this data clearly indicates that the CBM does not affect the transition state of association and this means that the rate of association ${}_pk_{on}$ is not affected by a CBM. *Figure 19* showed that the CBM increased the free energy barrier for dissociation by 3-4 kJ/mol and this difference could be attributed to changes in the $\Delta H^{\neq}_{dissoc}$ (*Figure 20*); where for enzymes with a CBM was 5-10 kJ/mol higher than for the corresponding one-domain enzymes. Accordingly, our data suggests that the higher affinity of CBM appended enzymes is due to slower dissociation. As mentioned in the General Introduction, the linker connecting the CBM with the core domain has been assigned several roles and one of them is that the highly glycosylated linker assist the binding of the enzyme to its substrate during the enzymatic hydrolysis (Payne et al., 2013b). As our experiments obviously do not distinguish between CBM and linker, slower enzyme detachment (${}_pk_{off}$) could also stem from some combination of the two. As an addendum to this comment, Payne et al actually concluded that the CBM-linker significantly enhanced the binding over the CBM alone (Payne et al., 2013b).

The first study which successfully removed the CBM by proteolysis, found a 90 % decrease in activity upon removal of the linker and CBM (Tomme et al., 1988). Consequently, it was concluded that the CBM was required for efficient hydrolytic activity and this result and statement had been repeated several times. More recently, Vikarii and co-workers (Le Costaouëc et al., 2013, Pakarinen et al., 2014, Várnai et al., 2013) conducted more systematic investigations of the CBM effects on activity and found the same activity for one- and two domain enzymes at high solid concentration. Similar effects of the CBM have been observed in less comprehensive studies (Srisodsuk et al., 1993, Kari et al., 2014).

The significance of our data is best described by re-plotting the steady state rates ${}_pv_{ss}$ of Hj_{CBM} and Hj_{CORE} in the same plot (*Figure 26*). Accordingly, the

experimental data in *Figure 26* represent the specific steady state rates $_pv_{ss}/E_0$ for Hj_{CBM} and Hj_{CORE} at 30 °C and the lines are the best fits of eq. 4 (section 1.4.2.) At around 35 g/L Avicel, the two lines intersect, illustrating that at this substrate load the two enzymes are catalytically equal. At lower substrate load it is catalytically advantageous to have a CBM appended (the area to the left of the vertical dashed line), although the opposite is true at high substrate concentrations (the area to the right of the dashed line). The effect is quit distinctive. At 5 g/l, for example, where many earlier studies have been conducted, the rate of the two-domain enzyme is twice as high; while at saturated substrate concentration found the opposite ratio ($_pV_{max}$ of Hj_{CORE} is twice as high as $_pV_{max}$ for Hj_{CBM}).

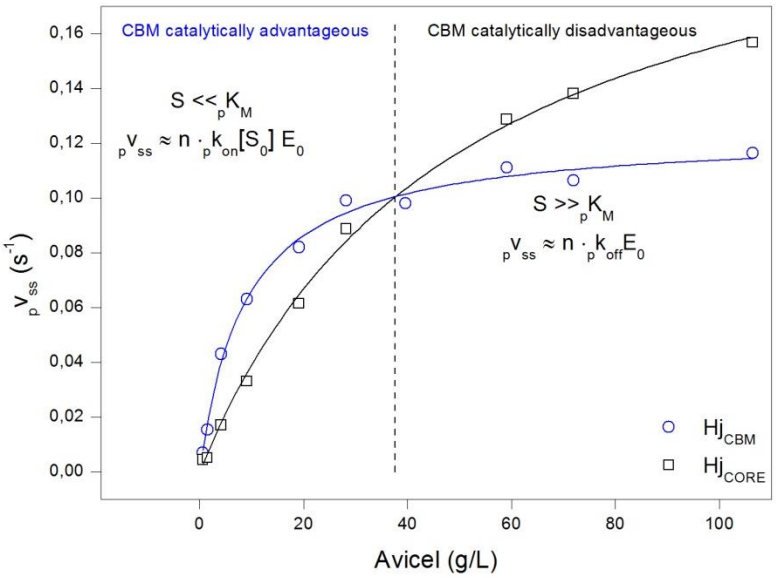


Figure 26. The specific rate for Hj_{CBM} (black) and Hj_{CORE} (blue) plotted against substrate load at 30 °C. It appears that the one-domain variant was slower at low substrate, but became faster than the two-domain enzyme above approximately 35 g/l. The vertical dashed line separates the plot into two sections; the area to the left indicates substrate concentrations where it is advantageous to have an appended CBM, whereas the area to the right indicates substrate concentrations where it is disadvantageous to have an appended CBM.

Focusing on the high substrate loads when the rate approached $_pV_{max}$, we recall from the steady state theory (Cruys-Bagger et al., 2013a) (section 1.4.2.) that the specific rate $_pv_{ss}$ equals the product of n, $_pk_{off}$ and E_0 (eq. 5) (*Figure 26*). As the processivity number (n) (*Figure 11*) is higher for the two-domain enzymes compared with single domain enzymes, n does obviously not prompt the higher rate of the single-domain enzyme. Nonetheless, we have earlier arrived at the conclusion that the CBM hampers dissociation ($_pk_{off}$) and this observation could indeed explain the catalytical superiority of the single-domain Cel7A at high substrate loads.

Now we turn into low substrate loads. Additionally, an approximate expression for the steady state rate have been proposed, ${}_pv_{ss} \approx n \cdot {}_pk_{on} \cdot S_0 \cdot E_0$ (Cruys-Bagger et al., 2013a), and reveals that either n or ${}_pk_{on}$ may offer the answer for the higher hydrolytic activity of the two-domain enzyme. We expected *a priori* that ${}_pk_{on}$ would be the causal factor and rationalized that the CBM promotes the adsorption at low substrate loads where "finding" the substrate is critical. However, as we previously concluded that the CBM does not affect the rate of association, it follows that n must be the parameter of interest. Besides this rationale, the processivity number (n) is 2 times higher for the two-domain enzyme compared with the core version (*Figure 11*) which may result in the two-fold increase in the specific steady state rates at low substrate load (0.5-10 g/l). As a result, we suggest that the CBM facilitates the processive movement, thereby increasing the rate at low substrate concentrations. Further aspects of this suggestion are discussed in section 2.2.

As with the pairwise comparison between Hj_{CBM} and Hj_{CORE} at 30 °C (*Figure 26*), analogous plots were conducted at temperature in the range of 10 °C - 50 °C. All these comparisons revealed an intersecting point, which, when plotted against temperature, showed a straight line (*Figure 27*). As in *Figure 26* the dashed line stipulates two regions; at conditions (temperature and substrate load) above the dashed line it is beneficial to possess a CBM, whereas the CBM is disfavored below the line. In other words, at low temperatures the single –domain enzymes are favored in a broader interval (substrate load), than at high temperatures.

Due to the important finding that CBM impedes the hydrolysis at high substrate load, it is interesting to note that most characterized cellulases as present do possess a CBM, although more cellulases are identified to be of single domain origin (Várnai et al., 2013). As early studies reported found a 90 % decrease in activity upon removal of the linker and CBM (Tomme et al., 1988) it may not seem surprising that more two-domain cellulases have been characterized. Our data does not contradict these findings, but stresses that the superiority of the two-domain enzymes is limited to a certain (low) substrate range.
Varnai et al. noted that the current knowledge of the CBM is obtained by experiments carried out at low substrate concentration, and that no kinetic experiments have been performed at higher substrate concentration. In relation to this statement they suggested that the substrate concentration in the cellulase screening process should be performed at high solids as screening at dilute conditions will favor the two-domain enzymes and underestimate, for instance, the single domain enzymes (Várnai et al., 2013). Our data supplements this suggestion as it was observed that the single domain enzymes had twice as high maximum velocity as the two-domain enzymes. Regarding the selection of new wt enzymes for screening, Varnai et al. speculated that the amount of water in the natural

habitat of the cellulose producing fungi could be the key factor in determining the evolutionary development of microbial cellulolytic systems, as the these fungi inhabit a broad range of habitats, ranging from aqueous (e.g. springs) to dry (e.g. rotting wood) environments. On the basis of this, they suggested selecting new wild type in accordance with the natural habitat of cellulose degrading fungi and bacteria. We, too, find these selection criteria interesting.

In addition to the higher maximum velocity, single domain enzymes are characteristic in that they are less lignin inhibited (Rahikainen et al., 2013a), an issue which becomes more important with increasing temperature (Rahikainen et al., 2013b) and the absence of the CBM makes the cellulases more prone to dissociate, which is a quality in the recovery of the enzymes (Várnai et al., 2013).

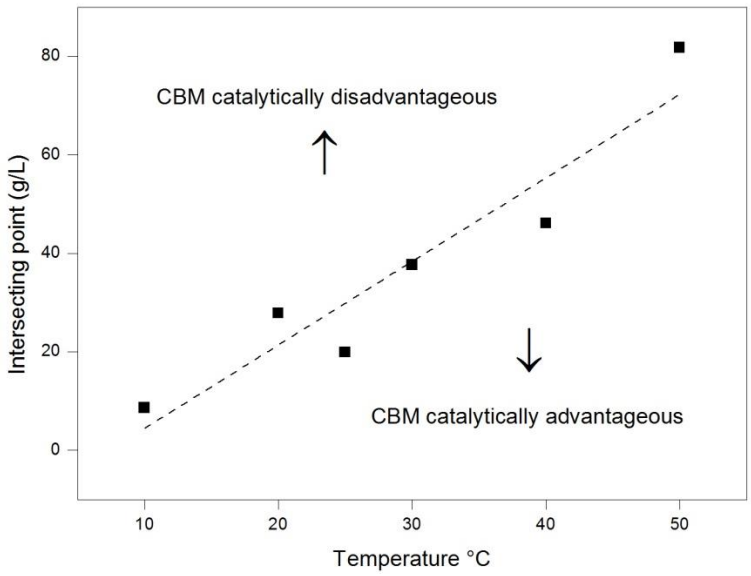


Figure 27. Effects of temperature and binding module on the activity of Cel7A from H. jecorina. The main panel shows the location (i.e. substrate load) of this crossover as a function of temperature. The line separates the plane into regions where the CBM respectively promoted (upper-left) and reduced (lower right) enzyme activity

Effects on thermophile and mesophile origin. In addition to the different modularity of *R.emersonii* and *Hjecorina* Cel7A, comparisons of these enzymes may also highlight the differences of their adaption to temperature, as *R.emersonii* is a thermophilic and thermotolerant fungus (Maheshwari et al., 2000, Houbraken et al., 2012), whereas *H.jecorina* is a mesophile fungus. One of the general trends which applies for enzymes with different adaptions to temperature is that enzymes adapted to higher temperatures display stronger thermoactivation (*i.e.* higher activation energies) (Garsoux et al., 2004). This is also the case for *R.emersonii*

Cel7A which displayed Q_{10} values for ${}_pV_{max}$ that were higher by 0.2-0.3 units compared with *H.jecorina* Cel7A.

2.1.3. Additional note on ${}_pk_{on}$ and ${}_pk_{off}$

In the description given above of the temperature sensitivity of the association rate constant ${}_pk_{on}$ and the dissociation constant ${}_pk_{off}$, we make the assumptions that $k_{cat} >> k_{off}$ at high substrate concentrations. Given the emerging knowledge (Kurasin and Väljamäe, 2011, Cruys-Bagger et al., 2013b), this seems to be a valid assumption. However, it is actually possible to obtain information on the temperature sensitivity of these two rate constants without this assumption. As mentioned in Chapter 1 and section 2.1.2, separate rate constants will, to a good approximation, obey the Arrhenius equation. According to steady state theory, ${}_pK_M$ is given as;

$$ {}_pK_M \equiv \frac{{}_pk_{off}}{{}_pk_{on}} $$

By constraining the constants (${}_pk_{on}$ and ${}_pk_{off}$) by the Arrhenius equation, this yields;

$$ \begin{aligned} {}_pK_M(T) &\equiv \frac{A_{k_{off}} e^{-\frac{E_{k_{off}}}{RT}}}{A_{k_{on}} e^{-\frac{E_{k_{on}}}{RT}}} \\ &= \frac{A_{k_{off}}}{A_{k_{on}}} e^{\frac{E_{k_{on}} - E_{k_{off}}}{RT}} \\ &= \alpha e^{\frac{\beta}{T}} \end{aligned} $$

where $\alpha = \frac{A_{k_{off}}}{A_{k_{on}}}$ and $\beta = \frac{E_{k_{on}} - E_{k_{off}}}{R}$. It can be seen that temperature dependence on ${}_pK_M$ is given by a single-exponential function and that the difference in the activation energies, ΔE, can be calculated when β is known. The difference in the activation energies is then given by:

$$ \begin{aligned} \beta &= \frac{E_{k_{on}} - E_{k_{off}}}{R} \Leftrightarrow \\ E_{k_{on}} &= E_{k_{off}} + \beta R \end{aligned} $$

It can be seen that if: β is positive then $Ek_{on} > Ek_{off}$ and if β is negative then $Ek_{on} < Ek_{off}$. One approach is to fit the temperature-dependence of ${}_pK_M$ to the single-exponential function or use the more classical approach:

$$ {}_pK_M\left(T\right) = \alpha e^{\frac{\beta}{T}} \Leftrightarrow \ln\left({}_pK_M\left(T\right)\right) = \ln\left(\alpha\right) + \beta\frac{1}{T} $$

Thus, by plotting the natural logarithm of the Michaelis constant ($_pK_M$) for the four enzymes (*Hj_{CBM}*, *Hj_{CORE}*, *Re_{CBM}* and *Re_{CORE}*) against the reciprocal of the absolute temperature (*Figure 13*, *panel B*), we obtained slopes (β·R) with values ranging between -36 to -53 kJ mol^{-1} (see also page 33). As the β·R term is negative for all the enzymes this means that the activation energy for $_pk_{on}$ is lower than for $_pk_{off}$ ($Ek_{on} < Ek_{off}$).The most important aspects of this approach is that it is independent of the assumption that $_pk_{cat}$ should be larger than $_pk_{off}$.

2.1.3 Critique, reservations and choices for the experimental setup and data analysis

Steady state rate. The kinetic experiments presented in this thesis aims to describe the initial rates at quasi steady state. This implies for example that the rate has to be measured at short time intervals before significant changes (i.e. substrate degradation and product formation) occur and that the EC complex (enzyme-substrate) (se scheme 1) is constant over time. With regards to these demands, our choice of estimating the reaction time around an hour by endpoint measurements may appear overly simplified, at least compared with classic kinetics. However, the reaction time of 1 hour was carefully selected. Firstly, as we aimed for measuring the activity at low temperature (10 °C) and at low substrate concentration (0.5 g/l), the reaction time had to be one hour in order to be above experimental detection limit. With regards to the initial rates, the substrate conversion never exceeded 1%, due to the inherent slow cellulose degradation by Cel7A and hence assumption of constant substrate load appears reasonable. Moreover, the reaction time was also suitable in the sense that product concentrations in samples with the highest substrate loads were much lower than published inhibition constants for Cel7A (Murphy et al., 2013, Teugjas and Valjamae, 2013, Gruno et al., 2004). As Cel7A hydrolysis of insoluble cellulose is characterized by nonlinear kinetics and has a characteristic burst phase, the choice of making endpoint measurements and estimating the reaction rate by the slope between t = 0 and the chosen reaction's time seems a bit problematic as the occurrence of the burst is substrate dependent (Cruys-Bagger et al., 2013b). To avoid too much influence of the burst phase, the longer time interval is desirable. Additionally, we note that experimental measurements of Michaelis Menten like curves for numerous temperatures and enzymes requires a rather high throughput, where compromises are unavoidable.

2.2. Loop engineered variants (publication IV and V)

2.2.1. Introduction

This chapter combines a presentation of the results from publications IV and V along with data which are not covered by any of the publications. This chapter describes point mutations in one of the major tunnel forming loop (B2) in *Hypocrea jecorina* or *Rasamsonia emersonii* Cel7A (*Figure 28*).

Based on comprehensive site-directed mutagenesis studies in *Thermomonospora fusca* E2, *Thermobifida fusca* Cel6B and *Thermobifida fusca* Cel6A, Wilson and co-workers concluded that non-catalytic amino acids play a crucial role in the function of cellulases (Zhang and Wilson, 1997, Zhang et al., 2000b, Zhang et al., 2000a). Many of the point mutations in these studies were placed in the loop areas and, although no uniform conclusion was reached, it appeared that changing amino acids in the loops have a pronounced effect upon the activity and function of the cellulases (Zhang and Wilson, 1997, Zhang et al., 2000b, Zhang et al., 2000a).

Loop structures in Cel7A constitute around 1/3 of the catalytic domain and are mainly responsible for the formation of the active site tunnel (Divne et al., 1994). However, while whereas the catalytic mechanisms and the amino acids residues involved are well known, the role of the different loops are not yet fully understood. Loops which are part of the active site of enzymes are often highly conserved indicating a central role of these loops (Kokkinidis et al., 2012). The GH7 family partly deviates from this general observation, as the length of the loops determines the coarse division of this family into CBHs or EGs (Divne et al., 1994, Rouvinen et al., 1990, Spezio et al., 1993). Therefore, it might not come as a surprise that protein engineering in the loop network of Cel7As prompt significant changes (Adney et al., 2009, Dana et al., 2012, von Ossowski et al., 2003) and for other GH members a similar conclusion was reached (Chokhawala et al., 2015, Cheng et al., 2012).

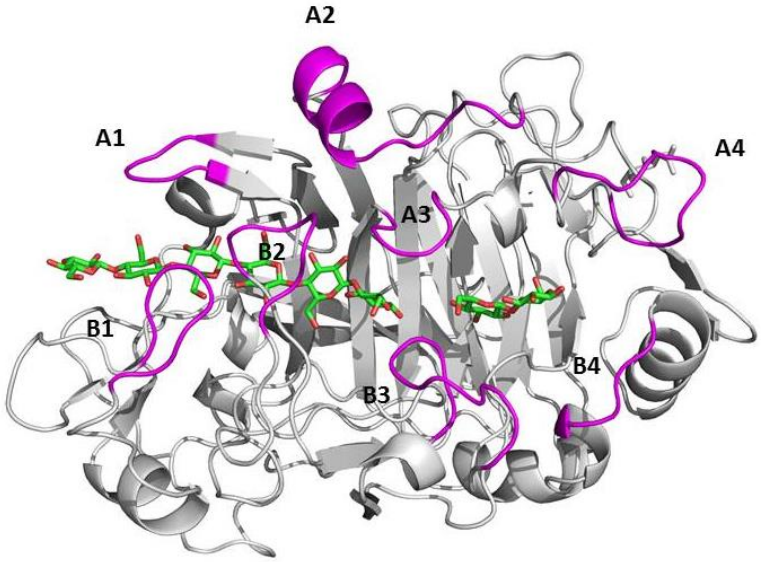
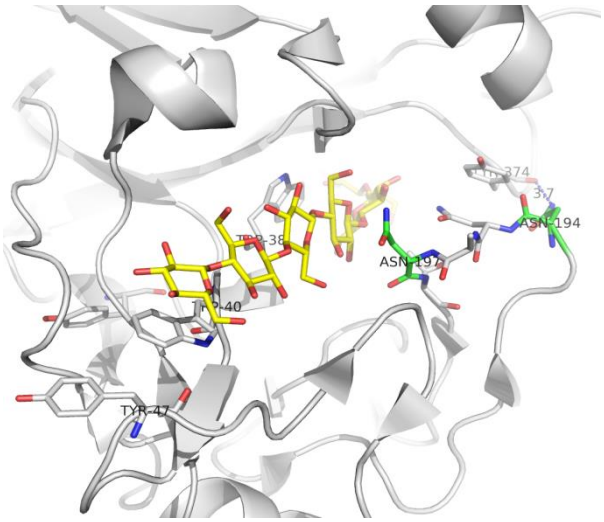


Figure 28: Structure of the major loops in TrCel7A enzymes (PDB code 4CEL). The loop nomenclature is adopted from (Momeni et al., 2013).

For describing of the major loops covering the active site tunnel, we adopted the loop nomenclature from Momeni et al. (Momeni et al., 2013) (*Figure 29Figure 28*). Loops A1 and B1 comprise the entrance loops (subsites -6/-7) whereas loop B4 and A4 are positioned near the product expulsion site. The A2 loop does not form contact with the cellulose strand and will, along with the entrance and expulsion site loop, not be further discussed. The remaining 3 loops (B2, B3 and A3) are positioned in the center of the core domain. The B3 loop, which has also been referred to as the exo-loop (von Ossowski et al., 2003) covers the catalytic residues along with the A3 loop. Protein engineering in the B3 loop indicated its importance for the processive degradation of crystalline cellulose (von Ossowski et al., 2003), an observation which was also seen for loops in *Thermobifida fusca* Cel6B (Zhang et al., 2000b).

Loops in proteins are notoriously flexible and the loops residing the Cel7A catalytic domains are no exception. However, compared with Cel6A from *Humicola insolens* for instance (Varrot et al., 1999), the flexibility of the loops of *Hj*Cel7A is low, with the exception being the loop B3, which shows backbone displacement of 2 Å (von Ossowski et al., 2003). Due to the higher flexibility of this loop, von Ossowski et al hypothesized that the B3 loop would be of special interest for the processive action (von Ossowski et al., 2003). In contrast, the B2 loop in *H. irregulare* and *H.grisea* has been identified as the most flexible region of the molecule (Momeni et al., 2013, Momeni et al., 2014). These workers argued that the high flexibility of these loops may not rely on the loops themselves but

stems from their interaction with the opposing loop (A3). More specifically, they suggested that the higher flexibility of both *HiCel7A* and *HgCel7A* compared with *Hj*Cel7A would be due to a tyrosine to histidine substitution in loop A3 (*Figure 29*). For *ReCel7A* high flexibility was also associated with the B2 loop, as Grassick et al. found the lowest electron density in this area, which may be interpreted as an indicator for the high loop dynamic (Grassick et al., 2004).



B2

ReCel7A	GWQPSSNNANTGIGDHGS
HjCel7A	GWEPSSNNANTGIGGHGS
HiCel7A	GWTPSSSDSNAGTGQYGS
HgCel7A	GWTGSTNDPNAGAGRYGT

A3

ReCel7A	VLVMSLWDDYAAQMLWLD
HjCel7A	VLVMSLWDDYYANMLWLD
HiCel7A	VLVMSVWDDHEANMLWLD
HgCel7A	VLVMSIWDDHASNMLWLD

Figure 29. A) Structure of the loops B2 and A3 of R.emersonii Cel7A. Side chains are drawn for the residues N194 and N197 (green) The side chain for the highly conserved tryptophan residues (W38, W40) lining the tunnel on the opposing site are also depicted. Alignment of sequences of the B2 and A3 loops. Rasamsonia emersonii Cel7A, Hypocrea jecorina Cel7A, Heterobasidion irregulare Cel7A and Humicola grisea Cel7A. Red letters indicate the exchanged amino acids.

Within the scope of RESAB, we aim for engineering more efficient enzymes. First we adopted the view that the rate of dissociation of the Cel7A from the cellulose ($_pk_{off}$) is the rate-limiting step (1.4.3.) and next we identified a possible structure of the transitions states of $_pk_{off}$, by virtue of the TST and the Hammond postulate (2.1). Based on this data we suggested that $_pk_{off}$ may be tuned in a controlled way by targeting enzyme substrate interactions. To further highlight this claim and with the aim of generating enzymes with higher $_pk_{off}$ (i.e.decreasing the activation free energy of $_pk_{off}$), we introduced mutations in the loop area which possible change or remove some substrate-enzyme interactions. Specifically, we introduced a point mutation in the B2 loop which encloses the cellulose chain around subsite -4 through interactions with the A3 loop (*Figure 29*). The point mutations were introduced in both *R.emersonii* and *H.jecorina* Cel7A (*Table 4*). More specifically, asparagine residues at position 194 and 197 were exchanged with alanine residues in *R.emersonii* Cel7A, whereas only the asparagine residue equivalent to N197 was exchanged in *H.jecorina* Cel7A. On the other hand this amino acid (N200) was

exchanged with nine different amino acids in (*Table 4*). The asparagine residue often involved in the active site is a suitable target as the polar side chain is particularly good for interaction with other polar or charged atoms. The active side tunnel of Cel7A is no exception, and the predominant interaction between protein and oligosaccharides at the first four binding sites are hydrogen bonds between the amide group of asparagine and glutamine residues and the hydroxyl groups of the oligosaccharides (Divne et al., 1998). To remove the possible asparagine side chain interactions we exchanged the asparagine residues with alanine residues, which are widely used since this residue possesses a small side chain and is not likely to perturb the protein structure.

***H.jecorina* Cel7A**	***R.emersonii* Cel7A**
N200A, N200D, N200E, N200F, N200G, N200H, N200Q, N200S, N200T, N200V, N200W, N200Y, ΔN200	N194A, N197A and N194A/N197A

Table 4. Substitutions introduced in H.jecorina Cel7A and R.emersonii Cel7A

N197 in *Hj*Cel7A is the equivalent asparagine to N194 in *R.emersonii* and this particular aspargine residue has been suggested to be important in the formation of the enclosed tunnel, as it interacts with a tyrosine residue (Y370) positioned in the A3 loop (Grassick et al., 2004, Divne et al., 1998). The asparagine at this position is however not fully conserved and *Heterobasidion irregulare* Cel7A, for instance, possesses a serine residue at the corresponding position (Momeni et al., 2013). The asparagine at positions equivalent to N197 (*ReCel7A*) (see supplementary material) are, on the other hand, fully conserved. As opposed to N194, no obvious interactions occur between the asparagine in position 197 and other internal parts of the enzyme or the complexed oligosaccharides. It may be, however that hydrogen bonds indirectly to the oligosaccharides via one of the water molecules which line the inner surface of the tunnel (Divne et al., 1998).

2.1.2. Results and discussion

Data for *ReCel7A* and *ReCel7A* variants were analyzed along the lines of the steady-state model (section 1.4.2.) and the experimental setup resembles those applied in publications I and II (chapter 2.1).

Rasamsonia emersonii Cel7A loop variants

ReCel7A, N194A, N197A and N194A/N197A

The specific steady-state rates ($_pv_{ss}/E_0$) were plotted against substrate load and the kinetic parameters $_pV_{max}$ and $_pK_M$ were obtained by non-linear regression analyses with respect to eq. 4. Kinetic parameters derived from this analysis are listed in *Table 5*. *Figure 30* shows data for the *ReCel7A* and the N194A/N197A double mutant at temperatures ranging from 25 °C – 50 °C. The results in *Figure 30* and

Table 5 reveal a systematic kinetic effect of the mutation; at low substrate load *ReCel7A* and the N194A/N197A variant are kinetically identical, whereas the variant is faster at higher substrate loads.

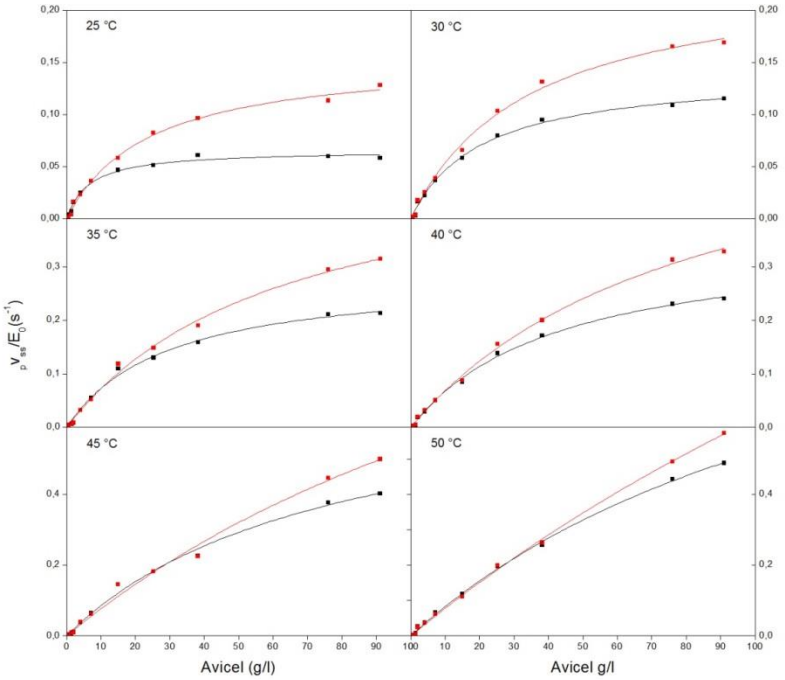


Figure 30. Specific enzyme activity ($_pv_{ss}/E_0$) for the ReCel7A wild type (black) double mutant N194A/N197A (red) at temperatures between 25 °C and 50 °C. Symbols represent experimental data with error bars for triplicate measurements (in many cases the error is comparable to the size of the symbol and thus hard to see). The curves show that best fit of eq. 4 The Specific enzyme activity ($_pv_{ss}/E_0$) are calculated as released soluble sugars after 0.5 hr

To elucidate the effects of each individual substitution, the two mutations were introduced singly and the kinetics was measured at 25 °C and 50 °C (*Figure 31* and *Table 5*). At 25 °C both single substitutions N194A and N197A led to increased maximal velocity $_pV_{max}/E_0$ by a factor of 1.3 and 1.5 compared with *wt ReCel7A*. However, they did not reach the same speed as the double mutant, which increased the maximal velocity by a factor 2. $_pK_M$ for the two single mutations were slightly increased, while the double mutation resulted in increased $_pK_M$ value by a factor 2.

	$_pV_{max}/E_0$ (s^{-1})x10^3				$_pk_M$ g/l			
	ReCel7A	***N194A***	***N197A***	***N194A/N197A***	***ReCel7A***	***N194A***	***N197A***	***N194A/N197A***
25 °C	66 ± 2 ***65 ± 2**	***84 ± 5**	***102 ± 6**	156 ± 6 ***135 ± 3**	13 ± 1 ***6.5 ± 1**	***15 ± 1**	***19 ± 3**	33 ± 1 ***24 ± 3**
30 °C	141 ± 5	NM	NM	238 ± 10	20 ± 2	NM	NM	34 ± 4
35 °C	286 ± 13	NM	NM	535 ± 33	29 ± 3	NM	NM	64 ± 7
40 °C	369 ± 14	NM	NM	658 ± 38	43 ± 4	NM	NM	83 ± 8
45 °C	747 ± 66	NM	NM	1560 ± 368	78 ± 12	NM	NM	193 ± 61
50 °C	1247 ± 83 ***651 ± 30**	***656 ± 29**	***940 ± 61**	2621± 380 ***911 ± 55**	141 ± 14 ***79 ± 6**	***81 ± 7**	***121 ± 13**	327 ± 58 ***117 ± 11**

Table 5 Maximal specific rates, $_pV_{max}/E_0$ (± standard error), and the Michaelis constants, $_pK_M$ (± standard error) for R.emersonii Cel7A wt, the single mutants N194A and N197A and the double mutant N194A/N197A at different temperatures.*the specific enzyme activities ($_pv_{ss}/E_0$) are calculated as released soluble sugars after1 hr.

At 50 °C the single mutant N194A and the *ReCel7A* show similar kinetics (*Figure 31*, B) and the double mutant N194A/N197A cannot be distinguished from N197A. This data indicates that the mutation N194A does modify the kinetics at 25 °C but not at 50 °C. In general the kinetic data showed that the N to A substitutions in both position 194 increased $_pK_M$ particularly at the lower temperatures. This suggests weaker substrate interactions in the mutants (*c.f.* eq 3), and to assess this more directly, we measured the adsorption behavior on Avicel of wild type *Re*Cel7A and all three variants at 25°C.

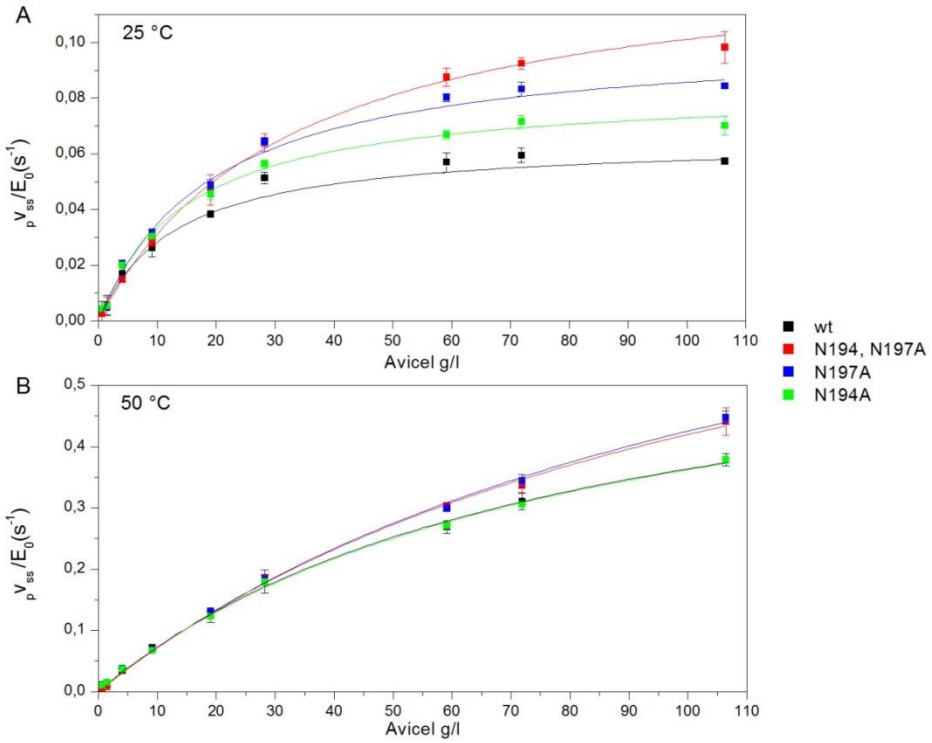


Figure 31. Specific enzyme activity ($_pv_{ss}/E_0$) for R.emersonii wt and the variants N194, N197 and N194A/N197A at 25 °C (panel A) and 50 °C (panel B). Symbols represent all experimental data from triplicate measurements and lines are best fit of eq.4. The Specific enzyme activity ($_pv_{ss}/E_0$) are calculated as released soluble sugars after 1 hr. hydrolysis and normalized with respect to the total enzyme concentration (E_0=0.40 µM) to obtain the specific activity, $_pv_{SS,}/E_0$, in units of s^{-1}.

Specifically, the substrate coverage, Γ, was calculated (see section 2.1.2.) Results in *Figure 32* show that replacement of N with A systematically reduced adsorption. The location of the adsorption isotherms in *Figure 32* wt>N194A>N197A>N194A/N197A matched the relative affinities suggested by the $_pK_M$ data in *Table 5*. The initial points in this plot (see figure 5) make up a linear slope, which has been referred to as the Partitioning coefficient, K_P and describes the partitioning of the enzyme between the solid and liquid phases at low surface coverage and has previously been used as a measure of the overall affinity of cellulases for the cellulose surface (Palonen et al., 1999, Linder and Teeri, 1996).

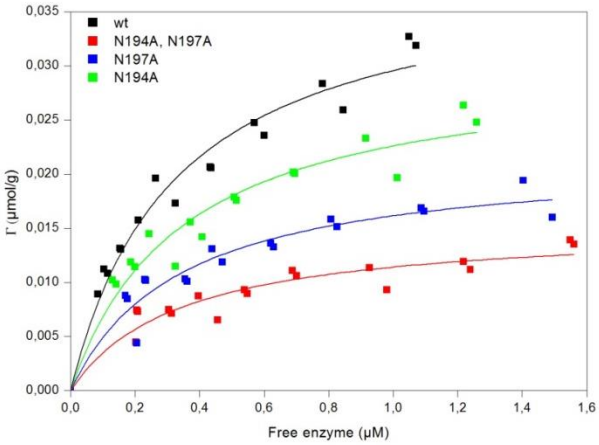


Figure 32. Binding isotherms on Avicel at 25°C for ReCel7A wild type (wt) and the three variants. Symbols are all experimental data and lines represent the best fit of a Langmuir isotherm.

The data confirmed that the investigated Asn-residues played a role in substrate interactions. Thus, around room temperature both N194A, N197A and the double mutant N194A/N197A showed lowered affinity for Avicel compared to the wild type. This lowering was reflected in both the kinetic parameter, ${}_pK_M$, and the independently measured partitioning coefficient, K_p. To quantify this weakening of substrate interactions, we calculated the difference (variant - wild type) in standard free energy of adsorption at 25°C, $\Delta\Delta G°=-RT\ln[K_p\,(var)/\,K_p(WT)]$. For the double mutant, we found $\Delta\Delta G°=2.4$ kJ/mol and the same value (2.3 kJ/mol) was obtained when inserting the analogous results for ${}_pK_M$ instead of K_p. The largest contribution to affinity was from N197; ($\Delta\Delta G°= 1.6$ kJ/mol) while the corresponding change for N194A was 0.8 kJ/mol. These latter numbers further show that the contributions from the two residues were essentially additive. *Figure 31* shows that the double mutant and the wt have identical initial slopes, a point which also can be illustrated by comparing the specificity constant, ${}_p\eta$ (eq. 6) The specificity constant ${}_p\eta$ can be inferred from the ${}_pV_{max}$ and ${}_pK_M$ values given in *Table 5*. In accordance with the identical slopes (*Figure 31*) the ratios of ${}_p\eta$ for wt and N194A/N197A, N194 or N197 were approximately 1 at all temperature tested. The significance of this information lies in the average processivity number, as n_p and ${}_p\eta$ will specify the values of k_{on}. (e.q. 6). We estimated the average processivity number as the relationship between cellobiose to cellotriose formed during hydrolysis and found an average processivity number around 14 for all the variants (data not shown). As the average processivity is unaltered, this implies that the introduced mutations did not affect the association rate constant (${}_pk_{on}$) and implies that the contribution to the lower affinity of the variants solely stems from an increased ${}_pk_{off}$. A result which is in agreement with our previous suggestion that ${}_pk_{off}$ may be modified a by cancelling the interaction which would be involved in substrate interactions.

Hypocrea jecorina Cel7A loop variants

In the study presented above, we generated three variants and compared their properties with the wt enzyme. Based on these comparisons, we claimed a possible relationship between low affinity and higher maximal velocity and speculated if the low affinity would be due to higher $_pk_{off}$. This approach of designing few variants, comparing their properties and claiming possible relationship appears often for rational mutagenesis studies and has been used applied in studies which probe the role of stacking interaction (Nakamura et al., 2013), N-linked glycosylation (Adney et al., 2009), the exo-loop (B3) (von Ossowski et al., 2003), the role of the linker (Srisodsuk et al., 1993) or for studies which designs variant with low substrate affinity (Kari et al., 2014), a more alkaline pH optimum (Becker et al., 2001). Collectively, these studies generate 1-5 different. In contrast, studies using a semi-rational approach often investigate a much higher number of variants (Dana et al., 2012, Heinzelman et al., 2010, Komor et al., 2012) and suggest some relationships based on a big data set. We adopted a semi-rational approach and exchanged an asparagine residues at position 200 in the B2 loop of *H.jecorina* Cel7A (equivalent to position 197 in *R.emersonii Cel7A*) with 13 different amino acid residues to further investigate the loop.

HjCel7A variants. The kinetic parameters $_pV_{max}$ and $_pK_M$ were measured by plotting the steady state rates as function of substrate loads. As shown in *Figure 33* which depicts $_pv_{ss}/E_0$ for all the N200 variants at 25 °C, non-linear regression analyses with respect to eq. 4 seems to describe the data very well (*Figure 33*). A conspicuous result which appears from *Figure 33* is the impressive variation of the kinetics caused by a single point mutation. This distinctive result is also quantitatively described in *Table 6* which summarizes the parameters $_pV_{max}$ and $_pK_M$ for data obtained at 25 °C, 30 °C, 40 °C and 50 °C.

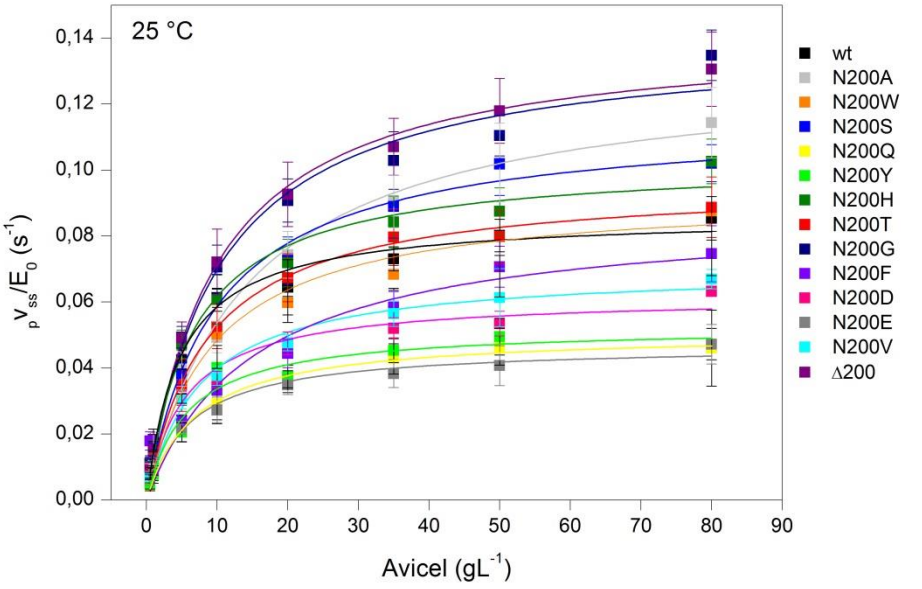


Figure 33. Specific enzyme activity ($_pv_{ss}/E_0$) for H.jecorina Cel7A and the substitutions at the N200 position at 25 °C. Symbols represent all experimental data from triplicate measurements and lines are best fit of eq. 4. The Specific enzyme activity ($_pv_{ss}/E_0$) is calculated as released soluble sugars after 1 hr. hydrolysis and normalized with respect to the total enzyme concentration (E_0=0.40 µM) to obtain the specific activity, $_pv_{SS,}/E_0$, in units of s^{-1}.

	25 °C		30 °C		40 °C		50 °C	
	$_pV_{max}/E_0$ $(s^{-1})x10^3$	$_pK_M$ $(g\ L^{-1})$	$_pV_{max}/E_0$ $(s^{-1})x10^3$	$_pK_M$ $(g\ L^{-1})$	$_pV_{max}/E_0$ $(s^{-1})x10^3$	$_pK_M$ $(g\ L^{-1})$	$_pV_{max}/E_0$ $(s^{-1})x10^3$	$_pK_M$ $(g\ L^{-1})$
ΔN200	142 ± 4	9.8 ± 0.9	302 ± 18	30 ± 4.2	400 ± 23	33 ± 4.3	740 ± 16	34 ± 1.6
N200G	140 ± 7	10 ± 1.8	256 ± 15	16 ± 3	400 ± 32	23 ± 3.8	618 ± 20	27 ± 2.1
N200A	113 ± 5	9.1 ± 1.6	178 ± 11	11 ± 2.4	308 ± 20	17 ± 3.2	462 ± 24	21 ± 3.0
N200H	103 ± 4	6.7 ± 1.1	186 ± 13	13 ± 2.9	307 ± 20	14 ± 2.9	460 ± 26	18 ± 3.0
N200S	101 ± 5	9.6 ± 1.4	198 ± 5	17 ± 1.3	411 ± 27	30 ± 4.6	513 ± 18	33 ± 2.6
N200T	97 ± 2	8.4 ± 0.6	190 ± 9	14 ± 2.2	410 ± 22	23 ± 3.2	586 ± 22	36 ± 3.0
N200W	93 ± 4	8.0 ± 1.6	150 ± 15	12 ± 4.0	293 ± 17	19 ± 3	438 ± 14	24 ± 1.9
N200F	87± 11	15 ± 5.8	150 ± 8	15 ± 2.4	429 ± 61	56 ± 15	505 ± 39	44 ± 6.9
wt	86 ± 3	4.7 ± 0.7	132 ± 8	5.6 ± 1.5	251± 19	12 ± 3.0	394 ± 19	14 ± 2.1
N200V	70 ± 3	7.6 ± 1.2	141 ± 8	16 ± 2.7	197 ± 18	22 ± 5.2	334 ± 12	19 ± 1.9
N200D	62 ± 3	5.4 ± 1.2	136 ± 9	9.3 ± 2.1	248 ± 19	15 ± 3.5	368 ± 15	18 ± 2.2
N200Y	53 ± 3	5.6 ± 1.5	112 ± 3	14 ± 1.3	177 ± 13	24 ± 4.6	232 ± 16	24 ± 4.3
N200Q	51 ± 1	6.1 ± 1.5	117 ± 4	12 ± 1.5	159 ± 12	20 ± 4.1	170 ± 17	13 ± 4.1
N200E	47 ± 2	6.2 ± 1.2	99 ± 7	7.9 ± 2.1	168 ± 13	13 ± 3.2	224 ± 9	14 ± 1.8

Table 6. $_pV_{max}$ og $_pK_M$ for Hjecorina Cel7A and 13 different asparagine substitutions at position 200.

To illustrate the effect of the mutations, *Figure 34* divides the group of residue substitutes into 4 different subgroups based on their physical, chemical and structural properties (Taylor, 1986).

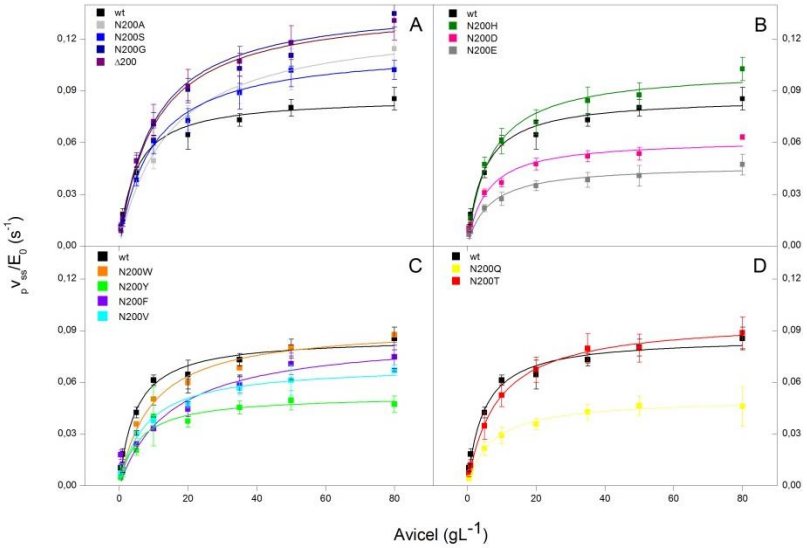


Figure 34. Specific enzyme activity ($_pv_{ss}/E_0$) for H.jecorina Cel7A and the substitutions at the N200 position at 25 °C. The residue substitutes are divided into 4 different subgroups based on their physical, chemical and structural properties (panel A: small amino acids, panel B: charged amino acids, panel C: hydrophobic amino acids and panel D: polar amino acid.

The classification of the mutations in *Figure 34* clearly shows that exchanging the asparagine to one of the amino acids belonging to the subcategory "small" (A, S, G) resulted in a higher maximum velocity (*Figure 34* A). Therefore it appears that exchanging the asparagine with small amino acids results in higher maximum velocity. Although threonine (T) in this context is categorized in the group of polar amino acid, it is small in size (*Figure 34*, D) and as with the other small amino acids substitution, N200T also display a higher $_pV_{max}$, although only moderately higher. A substitution which breaks this pattern is the introduction of a histidine or tryptophan residue, which does not fall within the group of small amino acids. Histidine is rather unique in its chemical properties as it has a pka value near to that of physiological pH, meaning that it is likely to be protonated at our experimental conditions. The hydrophobic aromatic residues W and F (panel C) show slightly higher values for $_pV_{max}$ compared with *Hj*Cel7A at 25 °C (*Table 6*). However, the kinetic footprint in form of $_pK_M$ differs. In particular, the N to F substitution lead to significantly altered kinetics compared with all the other substitutions. The phenylalanine residue differs only from the last substituted aromatic amino acid (tyrosine) in that it contains a hydroxyl group in place of the ortho hydrogen on the benzene ring. However, their kinetic footprints do not resemble each other and the substituted tyrosine result in decreased $_pV_{max}$. A decrease in $_pV_{max}$ was also noticed

for the mutants where the asparagine was exchanged by the negatively charged E and D, the polar Q or the hydrophobic V. A distinctive result which appears from (*Table 6*) is that HjCel7A has the lowest $_pK_M$, indicating a clear affinity preference of the asparagine in the B2 loop.

The variants of *HjCel7A* were tested at 4 different temperatures (*Table 6*). To compare the variants at these 4 temperatures, we found the relative activity ($_pV_{max}$ values) of the variants and compared them with the wt at 25 °C and 50 °C. Scatter plots of the relative $_pV_{max}$ values show a positive correlation, indicating that variants with high $_pV_{max}$ at 25 °C also have high $_pV_{max}$ at 50 °C.

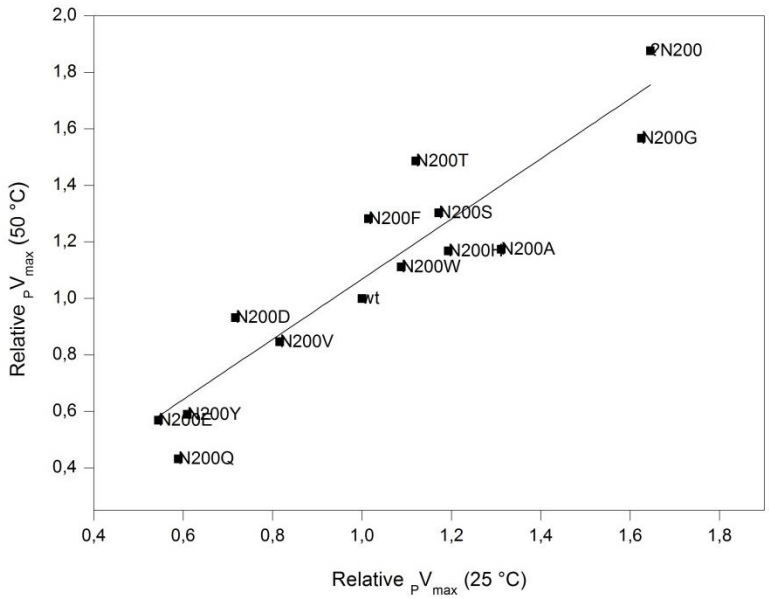


Figure 35. Scatter plots and regression line of $_pV_{max}$ at 25 °C and 50 ° C. The $_pV_{max}$ are described by their relative values. R^2 0.84, slope 1,06

Systematic exchanges of amino acids in position 200 in the loop area allow us to examine some possible relationships. For this purpose we made scatter plots of two measured variables e.g. $_pK_M$ and the processivity. The few dots in the scatter plots marked with red are excluded from the data underlying the linear regression. The choice of outliers (red dots) is not based on statistical evidence. They can, however, be explained by additional data or theoretical considerations.

Recently it has been suggested that the binding free energy of the cellulose chains to cellulases would be directly linked to the degree of processivity (Payne et al., 2013a). By comparing 5 different GH7 enzymes, each possessing a different degree of tunnel coverage and loop fluctuations, Payne et al reached the conclusion that more processive enzymes possess greater ability to associate tightly with the

substrate (Payne et al., 2013a). This link between affinity and processivity has also been suggested for another class of processive enzymes, namely the chitinases (Horn et al., 2006). We explored this relationship, by plotting the processivity on the ordinate and the values of $_pK_M$ on the abscissa (*Figure 36*).

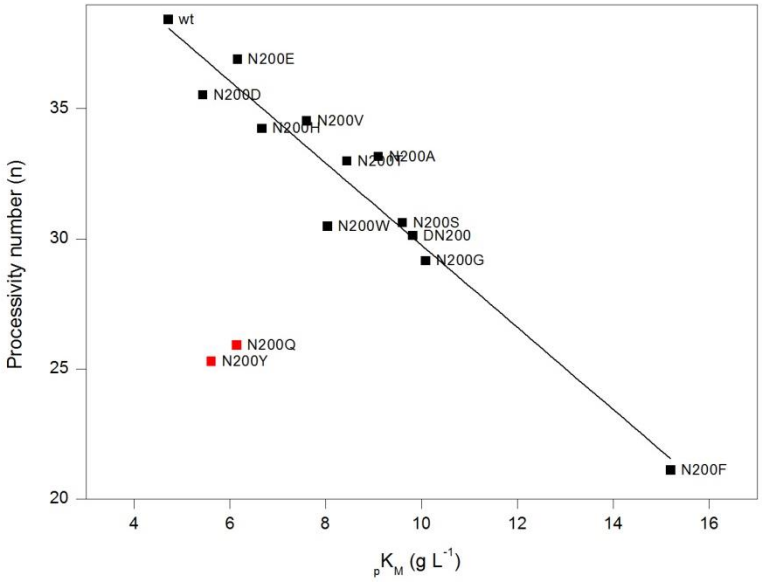


Figure 36. Relationship between $_pK_M$ and processivity of HjCel7A wt and 12 variants. R^2 0.82. The data points for N200Q and N200Y (red points) are not included for linear regression.

Figure 36 clearly suggests a linear negative correlation between $_pk_M$ and degree of processivity at 25 °C; variants with low cellulose affinity, illustrated by high $_pk_M$ values, also seem to be less processive. As accounted for above this relationship is in close alignment with the suggestion that higher affinity would promote higher degree of processivity. The chosen outliers N200Q and N200Y display a very low maximum velocity and the lowest kp values (data not shown). For both of these variants, amino acid residues with more bulky sidegroups substituted the asparagine and we suggest that these substitutions may cause more severe structural changes.

Various studies have examined the link between enzyme efficiency and processivity (von Ossowski et al., 2003, Horn et al., 2006, Zhang et al., 2000b). Zhang et al plotted the hydrolytic activity on filter paper and BMCC of various loop variants of *Thermobifida fusca* exocellulase Cel6A against their processivity and found good positive correlation (Zhang et al., 2000b). As previously mentioned, Ossowski et al. also found this connection for loop variants of

H.jecorina Cel7A and the same conclusions were reached for another group of processive enzymes; the chitinases (Horn et al., 2006, Zakariassen et al., 2009).
We too suggested a possible link between higher processivity at the hydrolytic rate (see page 49). However, this positive connection was only assumed at low substrate loads (*Figure 26*). To investigate this, we plotted the maximum velocity (*Figure 37*, panel A) and the specificity constant $_p\eta$ as function of the processivity (*Figure 37*, panel B). Both scatter plots do not reveal any obvious link between the degree of processivity and the maximum velocity or $_p\eta$.

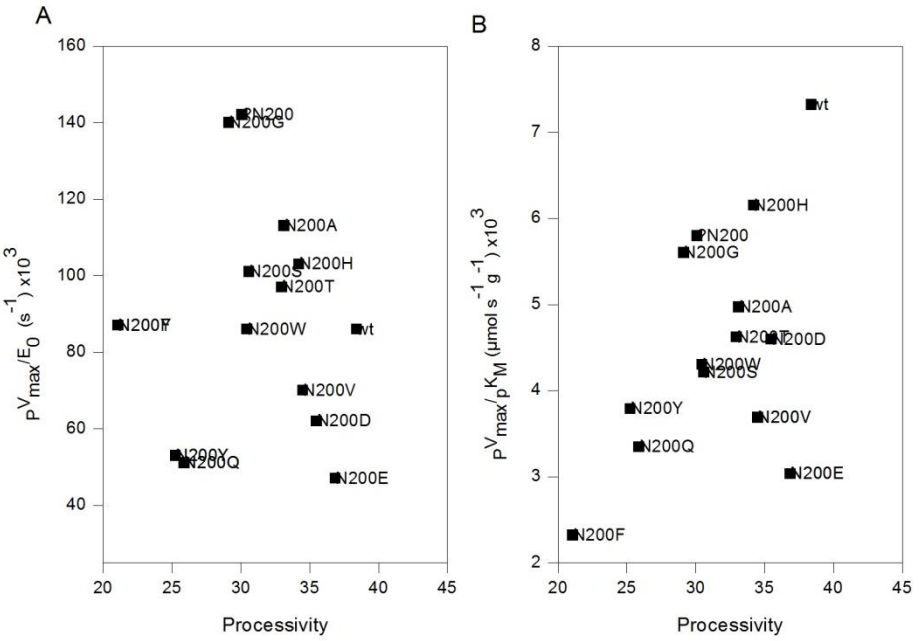


Figure 37. shows the scatterplots of pVmax (panel A) and $_p\eta$ (panel B) for the wt and variants plotted against the measured processivity.

As previously mentioned, work presented in the thesis adopts the premise that the dissociation of cellulase from cellulose is the rate-determining step. This cannot be tested directly as $_pk_{off}$ is inferred from $_pV_{max}$, and instead we plotted the maximum velocity as function of $_pK_M$, as it reflects the enzyme affinity to its substrate (*Figure 38*). The choice of the two outliers is based on the expression $_pV_{max} \approx n \cdot {}_pk_{off}$. The presence of n in the approximation signifies that the average processivity number additionally governs the maximum velocity. As the wt are a significantly more processive enzyme than the other variants and that N200F resides at the other end of the processivity scala, we marked these data points red. *Figure 38* shows a positive relationship between $_pV_{max}$ and $_pK_M$, indicating that variant with low affinity may be more catalytically efficient. This observation has indeed been previously observed for loop variants. By changing a single tyrosine residue in the tip of a loop to glycine, Cheng et al enhanced the activity of the hyperthermophile

bacteria *Thermotoga maritime* GH12 by 70 % β-glucan substrate and decreased the affinity (Cheng et al., 2012).

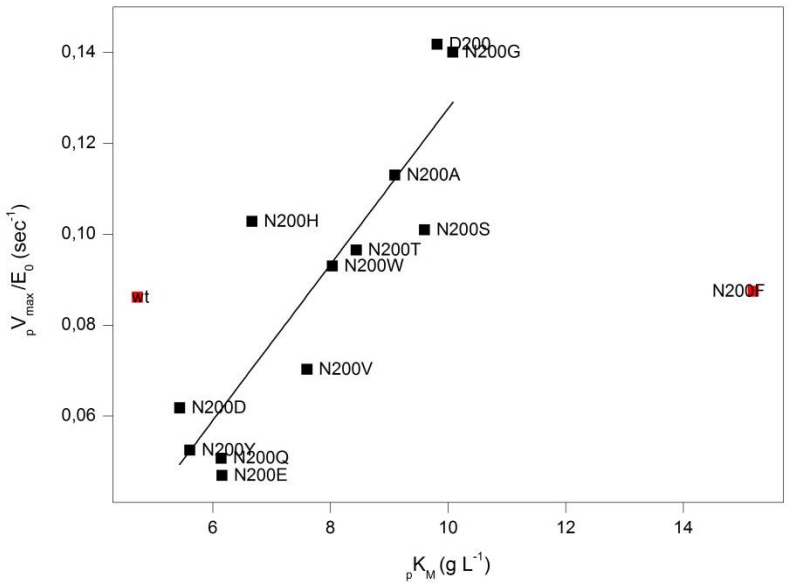


Figure 38. shows the scatterplot of $_pK_M$ and $_pV_{max}$ of HjCel7A .The scatterplots also show the regression line that best fits the data, R^2 0.75.. The data points for wt and N200F (red points) are not included for linear regression.

We additionally tested the enzymatic activity of the variants on soluble substrates e.g. p-nitrophenyl-beta-D-lactoside (pNPL) and found the kinetic parameters V_{max} and k_M. *Figure 39* plots V_{max} for pNPL against $_pV_{max}$ for crystalline cellulose. In perfect accord with previously studies (Zhang et al., 2006), no positive correlation is observed. An interesting observation is that the variants with low maximum velocity on Avicel are amongst the most efficient on pNPL.

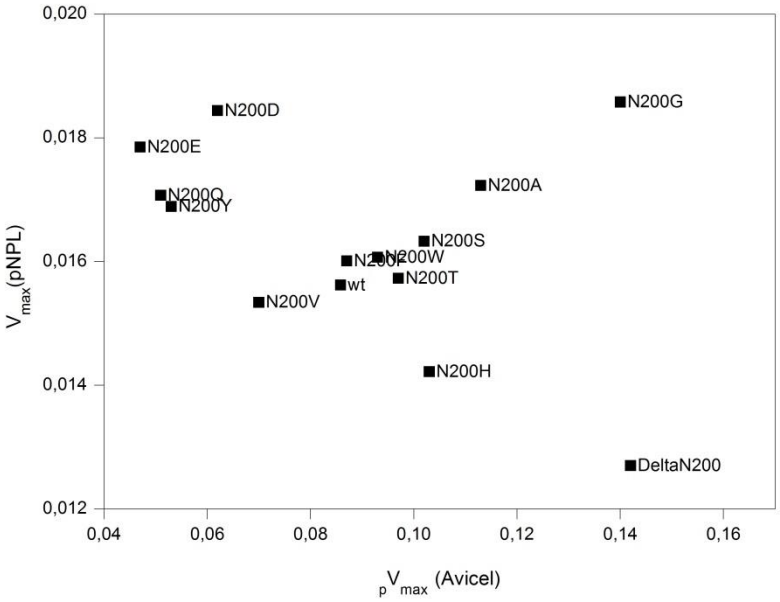


Figure 39. No positive correlation is seen between the maximum velocity of Cel7A wt and variants on insoluble crystalline cellulose (Avicel) and soluble substrate (pNPL).

The cornerstone of RESAB is to engineer more efficient enzymes under industrial conditions. We have just accounted for the efficiency of the engineered variants at industrial relevant conditions, but industrial conditions also imply prolonged hydrolysis time (~3 days), more heterologous substrate e.g. pretreated corn stower (PCS) and hydrolytic activity of Cel7A in a cocktail of divers GH enzymes. Patent WO2014138672 (VI) claims cellobiohydrolase variants comprising an alteration at one or more positions corresponding to positions 200 in *H. jecorina* or 194 and 197 in *R.emersonii*. The patent encompasses examples where the efficiency of both *R.emersonii* and *H.jecorina* Cel7A loop engineered variants are examined on pre-treated cornstower, at hydrolysis times of 24-72 hours and in the presence of a cocktail of GH enzymes. This data signifies that the faster rate obtained in some of the loop variants can be translated into more efficient enzymes at conditions mimicking the industrial settings.

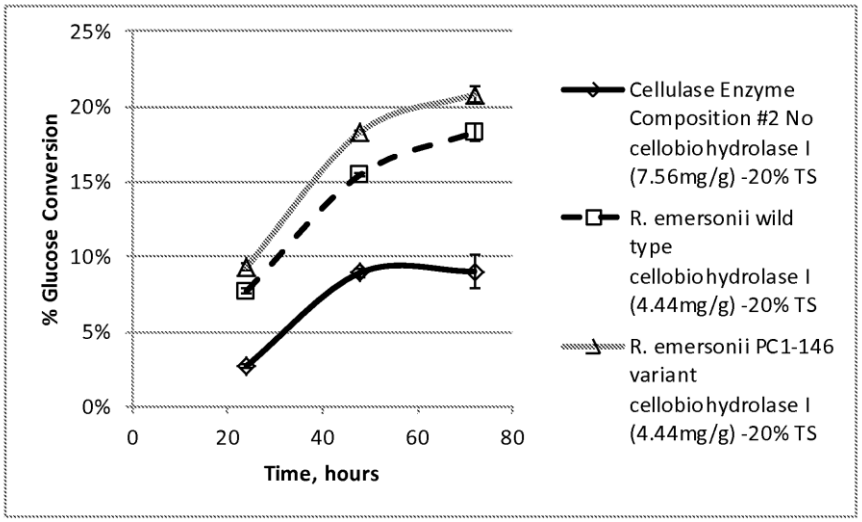


Figure 40. Comparison of the effect of the R. emersonii Cel7A wt and the double variant N194, N197 (PC1-146) on the hydrolysis of unwashed PCS (20% total solids) by a cellulase enzyme composition.

3. CONCLUSION

In summary, we have studied the kinetics for two native Cel7As; the mesophile two-domain *Hypocrea jecorina* Cel7A, consisting of a catalytic domain and a carbohydrate binding module, which are connected by a linker region, and the thermophile single domain *R. emersonii* Cel7A. In addition we also studied a truncated version of *H.jecorina* Cel7A and a chimeric protein composed of the linker and CBM from *H.jecorina* Cel7A and the *R. emersonii* enzyme. These four enzymes were chosen to highlight the effect of CBM and the natural adaption to temperature.

By comparing these enzymes at room temperature we found that while the CBM is beneficial at low substrate concentrations it is disadvantageous at higher substrate loads. Based on data presented in this dissertation we hypothesized that the increased processivity of two domain enzymes would result in higher hydrolytic activity at low substrate concentration, while the higher dissociation ($_pk_{off}$) of the single domain enzymes results in higher activity at high substrate concentration. Additionally, we find that the reaction temperature also affected the kinetics of the two-domain and single-domain enzymes; at elevated temperatures, the substrate load interval which defines the more catalytically efficient two-domain enzymes grew bigger. We monitored the temperature effect on the reaction rates at both high and low substrate concentrations and found that while the reaction rate was strongly temperature activated at high substrate concentration, the activity at low substrate concentration was barely increased at elevated temperatures.

We monitored the temperature effect of three rate constants $_pk_{on}$, $_pk_{cat}$ and $_pk_{off}$ which govern the association, catalytic activity and dissociation and find that $_pk_{off}$ is activated more than both $_pk_{cat}$ and $_pk_{on}$. The temperature sensitivity of the rate constants were additionally used for forming a free energy diagram for the Cel7A catalyzed hydrolysis of cellulose. The free energy reaction coordinate profile revealed that free activation energy for the enzyme substrate association is dominated by entropy contributions, while the free energy barrier for dissociation is dominated by enthalpy contributions.
By virtue of the Hammond postulate, we discussed the structure of the transition state intermediate along the reaction coordinates. We surmised that the transition state for $_pk_{on}$ appears early in the reaction pathway (before most enzyme-substrate contact are made) while the transition state for the $_pk_{off}$ appears late (subsequent to the removal of all enzyme-substrate interactions). We predict that $_pk_{off}$ depends strongly on the enzyme-substrate interactions as these contacts must be broken to

reach the transition state for dissociation. Therefore, we suggest that $_pk_{off}$ may be tuned in a controlled way by deleting/changing enzyme substrate interactions.

On the basis of this knowledge we aimed for engineering Cel7A variants with higher $_pk_{off}$. For these experiments, point mutations were introduced in a loop covering the cellulose strand at the -4 subsite in *H.jecorina* and *R.emersonii*. More specifically, two asparagine residues were exchanged with alanine residues in the B2 loop in position 194 and 197 of *R.emersonii* Cel7A. Replacement of one or both of the asparagine residues led to improved maximum velocity and lowered affinity for crystalline cellulose. Substituting the N197 corresponding asparagine residue in *H.jecorina*, N200 with 12 different amino acids additionally led to decreased affinity for cellulose. This site saturation approach in *H.jecorina* suggests that the role of the asparagine in this position is to facilitate 1) a high degree of processivity and 2) high affinity. In continuation thereof, plots of processivity and $_pK_M$ for all the N200 variants suggested a possible link between processivity and affinity. Furthermore the kinetic data for the N200 variants suggest a relationship between the maximum velocity and $_pK_M$.
I summary, we believe that modifying the kinetic properties of Cel7A by removing enzyme-substrate interactions is an useful approach to generate enzyme variants with higher activity at industrially relevant conditions.

References

ADNEY, W. S., JEOH, T., BECKHAM, G. T., CHOU, Y. C., BAKER, J. O., MICHENER, W., BRUNECKY, R. & HIMMEL, M. E. 2009. Probing the role of N-linked glycans in the stability and activity of fungal cellobiohydrolases by mutational analysis. *Cellulose,* 16**,** 699-709.

ARRHENIUS, S. 1889. Uber die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Sauren. Z. Phys. Chem.

ATALLA, R. H., BRADY, J. W., MATTHEWS, J. F., DING, S., HIMMEL, M. E. 2008. Chapter 6: Structures of plant cell wall cellulose.

AUDETTE, G. F., DELBAERE, L. T. J. & XIANG, J. 2003. Mapping protein : carbohydrate interactions. *Current Protein & Peptide Science,* 4**,** 11-20.

BANSAL, P., HALL, M., REALFF, M. J., LEE, J. H. & BOMMARIUS, A. S. 2009a. Modeling cellulase kinetics on lignocellulosic substrates. *Biotechnol Adv,* 27**,** 833-48.

BANSAL, P., HALL, M., REALFF, M. J., LEE, J. H. & BOMMARIUS, A. S. 2009b. Modelling cellulase kinetics on lignocellulosic substrates. *Biotechnology Advances,* 27**,** 833-848.

BECKER, D., BRAET, C., BRUMER, H., CLAEYSSENS, M., DIVNE, C., FAGERSTRÖM, B. R., HARRIS, M., JONES, T. A., KLEYWEGT, G. J., KOIVULA, A., MAHDI, S., PIENS, K., SINNOTT, M. L., STÅHLBERG, J., TEERI, T. T., UNDERWOOD, M. & WOHLFAHRT, G. 2001. Engineering of a glycosidase Family 7 cellobiohydrolase to more alkaline pH optimum: the pH behaviour of Trichoderma reesei Cel7A and its E223S/ A224H/L225V/T226A/D262G mutant. *Biochem J,* 356**,** 19-30.

BECKHAM, G. T., MATTHEWS, J. F., BOMBLE, Y. J., BU, L., ADNEY, W. S., HIMMEL, M. E., NIMLOS, M. R. & CROWLEY, M. F. 2010. Identification of amino acids responsible for processivity in a Family 1 carbohydrate-binding module from a fungal cellulase. *J Phys Chem B,* 114**,** 1447-53.

BECKHAM, G. T., MATTHEWS, J. F., PETERS, B., BOMBLE, Y. J., HIMMEL, M. E. & CROWLEY, M. F. 2011. Molecular-Level Origins of Biomass Recalcitrance: Decrystallization Free Energies for Four Common Cellulose Polymorphs. *Journal of Physical Chemistry B,* 115**,** 4118-4127.

BOMMARIUS, A. S., SOHN, M., KANG, Y. Z., LEE, J. H. & REALFF, M. J. 2014. Protein engineering of cellulases. *Current Opinion in Biotechnology,* 29**,** 139-145.

BORASTON, A. B., BOLAM, D. N., GILBERT, H. J. & DAVIES, G. J. 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J,* 382**,** 769-81.

BRETT, C. T. 2000. Cellulose microfibrils in plants: Biosynthesis, deposition, and integration into the cell wall. *International Review of Cytology - a Survey of Cell Biology, Vol 199,* 199**,** 161-199.

BROWN, R. F., AGBOGBO, F. K. & HOLTZAPPLE, M. T. 2010. Comparison of mechanistic models in the initial rate enzymatic hydrolysis of AFEX-treated wheat straw. *Biotechnol Biofuels,* 3**,** 6.

BUBNER, P., DOHR, J., PLANK, H., MAYRHOFER, C. & NIDETZKY, B. 2012. Cellulases dig deep: in situ observation of the mesoscopic structural dynamics of enzymatic cellulose degradation. *J Biol Chem,* 287**,** 2759-65.

CANTAREL, B. L., COUTINHO, P. M., RANCUREL, C., BERNARD, T., LOMBARD, V. & HENRISSAT, B. 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res,* 37**,** D233-8.

CARRARD, G., KOIVULA, A., SÖDERLUND, H. & BÉGUIN, P. 2000. Cellulose-binding domains promote hydrolysis of different sites on crystalline cellulose. *Proc Natl Acad Sci U S A,* 97**,** 10342-7.

CASPETA, L., CHEN, Y., GHIACI, P., FEIZI, A., BUSKOV, S., HALLSTROM, B. M., PETRANOVIC, D. & NIELSEN, J. 2014. Altered sterol composition renders yeast thermotolerant. *Science,* 346**,** 75-78.

CHEN, M. L. & WANG, F. S. 2010. Optimization of a Fed-Batch Simultaneous Saccharification and Cofermentation Process from Lignocellulose to Ethanol. *Industrial & Engineering Chemistry Research,* 49**,** 5775-5785.

CHENG, Y. S., KO, T. P., HUANG, J. W., WU, T. H., LIN, C. Y., LUO, W. H., LI, Q., MA, Y. H., HUANG, C. H., WANG, A. H. J., LIU, J. R. & GUO, R. T. 2012. Enhanced activity of Thermotoga maritima cellulase 12A by mutating a unique surface loop. *Applied Microbiology and Biotechnology,* 95**,** 661-669.

CHOKHAWALA, H. A., ROCHE, C. M., KIM, T. W., ATREYA, M. E., VEGESNA, N., DANA, C. M., BLANCH, H. W. & CLARK, D. S. 2015. Mutagenesis of Trichoderma reesei endoglucanase I: impact of expression host on activity and stability at elevated temperatures. *Bmc Biotechnology,* 15.

CLARKE, E. C. W. & GLEW, D. N. 1966. Evaluation of Thermodynamic Functions from Equilibrium Constants. *Transactions of the Faraday Society,* 62**,** 539-&.

COLUSSI F., SØRENSEN T.H., ALASEPP K., KARI J., CRUYS-BAGGER N., WINDAHL M.S., OLSEN J.P., BORCH K. & P., W. 2015. Probing substrate interactions in the active tunnel of a catalytically deficient cellobiohydrolase (Cel7). *Journal of Biological Chemistry,* 290**,** 2444-2454.

CREAGH, A. L., ONG, E., JERVIS, E., KILBURN, D. G. & HAYNES, C. A. 1996. Binding of the cellulose-binding domain of exoglucanase Cex from Cellulomonas fimi to insoluble microcrystalline cellulose is entropically driven. *Proc Natl Acad Sci U S A,* 93**,** 12229-34.

CRUYS-BAGGER, N., BADINO, S. F., TOKIN, R., GONTSARIK, M., FATHALINEJAD, S., JENSEN, K., TOSCANO, M. D., SORENSEN, T. H., BORCH, K., TATSUMI, H., VALJAMAE, P. & WESTH, P. 2014. A pyranose dehydrogenase-based biosensor for kinetic analysis of enzymatic hydrolysis of cellulose by cellulases. *Enzyme and Microbial Technology,* 58-59**,** 68-74.

CRUYS-BAGGER, N., ELMERDAHL, J., PRAESTGAARD, E., BORCH, K. & WESTH, P. 2013a. A steady-state theory for processive cellulases. *FEBS J,* 280**,** 3952-61.

CRUYS-BAGGER, N., ELMERDAHL, J., PRAESTGAARD, E., TATSUMI, H., SPODSBERG, N., BORCH, K. & WESTH, P. 2012a. Pre-steady-state kinetics for hydrolysis of insoluble cellulose by cellobiohydrolase Cel7A. *J Biol Chem,* 287**,** 18451-8.

CRUYS-BAGGER, N., REN, G. L., TATSUMI, H., BAUMANN, M. J., SPODSBERG, N., ANDERSEN, H. D., GORTON, L., BORCH, K. & WESTH, P. 2012b. An amperometric enzyme biosensor for real-time measurements of cellobiohydrolase activity on insoluble cellulose. *Biotechnology and Bioengineering,* 109**,** 3199-3204.

CRUYS-BAGGER, N., TATSUMI, H., REN, G. R., BORCH, K. & WESTH, P. 2013b. Transient Kinetics and Rate-Limiting Steps for the Processive Cellobiohydrolase Cel7A: Effects of Substrate Structure and Carbohydrate Binding Domain. *Biochemistry,* 52**,** 8938-8948.

DANA, C. M., SAIJA, P., KAL, S. M., BRYAN, M. B., BLANCH, H. W. & CLARK, D. S. 2012. Biased clique shuffling reveals stabilizing mutations in cellulase Cel7A. *Biotechnology and Bioengineering,* 109**,** 2710-2719.

DAY, A. G., GOEDEGEBUUR, F., GUALFETTI, P., MITCHINSON, C., NEEFE, P., SANDGREN, M., SHAW, A. & STAHLBERG, J. 2004. Novel variant hyprocrea jecorina cbh1 cellulases. Google Patents.

DE SOUZA, C. J. A., COSTA, D. A., RODRIGUES, M. Q. R. B., DOS SANTOS, A. F., LOPES, M. R., ABRANTES, A. B. P., COSTA, P. D., SILVEIRA, W. B., PASSOS, F. M. L. & FIETTO, L. G. 2012. The influence of presaccharification, fermentation temperature and yeast strain on ethanol production from sugarcane bagasse. *Bioresource Technology,* 109**,** 63-69.

DIVNE, C., STÅHLBERG, J., REINIKAINEN, T., RUOHONEN, L., PETTERSSON, G., KNOWLES, J. K., TEERI, T. T. & JONES, T. A. 1994. The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from Trichoderma reesei. *Science,* 265**,** 524-8.

DIVNE, C., STÅHLBERG, J., TEERI, T. T. & JONES, T. A. 1998. High-resolution crystal structures reveal how a cellulose chain is bound in the 50 A long tunnel of cellobiohydrolase I from Trichoderma reesei. *J Mol Biol,* 275**,** 309-25.

DRISSEN, R. E. T., MAAS, R. H. W., VAN DER MAAREL, M. J. E. C., KABEL, M. A., SCHOLS, H. A., TRAMPER, J. & BEEFTINK, H. H. 2007. A generic model for glucose production from various cellulose sources by a commercial cellulase complex. *Biocatalysis and Biotransformation,* 25**,** 419-429.

ERIKSSON, T., KARLSSON, J. & TJERNELD, F. 2002. A model explaining declining rate in hydrolysis of lignocellulose substrates with cellobiohydrolase I (cel7A) and endoglucanase I (cel7B) of Trichoderma reesei. *Appl Biochem Biotechnol,* 101**,** 41-60.

FERSHT, A. 1998. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*.

FOX, J. M., LEVINE, S. E., CLARK, D. S. & BLANCH, H. W. 2012. Initial- and processive-cut products reveal cellobiohydrolase rate limitations and the role of companion enzymes. *Biochemistry,* 51**,** 442-52.

GARSOUX, G., LAMOTTE, J., GERDAY, C. & FELLER, G. 2004. Kinetic and structural optimization to catalysis at low temperatures in a psychrophilic

cellulase from the Antarctic bacterium Pseudoalteromonas haloplanktis. *Biochemical Journal,* 384**,** 247-253.

GEORGELIS, N., YENNAWAR, N. H. & COSGROVE, D. J. 2012. Structural basis for entropy-driven cellulose binding by a type-A cellulose-binding module (CBM) and bacterial expansin. *Proceedings of the National Academy of Sciences of the United States of America,* 109**,** 14830-14835.

GERMAN, D. P., MARCELO, K. R. B., STONE, M. M. & ALLISON, S. D. 2012. The Michaelis-Menten kinetics of soil extracellular enzymes in response to temperature: a cross-latitudinal study. *Global Change Biology,* 18**,** 1468-1479.

GESSLER, K., KRAUSS, N., STEINER, T., BETZEL, C., SANDMANN, C. & SAENGER, W. 1994. Crystal-Structure of Beta-D-Cellotetraose Hemihydrate with Implications for the Structure of Cellulose-Ii. *Science,* 266**,** 1027-1029.

GILKES, N. R., WARREN, R. A., MILLER, R. C. & KILBURN, D. G. 1988. Precise excision of the cellulose binding domains from two Cellulomonas fimi cellulases by a homologous protease and the effect on catalysis. *J Biol Chem,* 263**,** 10401-7.

GRASSICK, A., MURRAY, P. G., THOMPSON, R., COLLINS, C. M., BYRNES, L., BIRRANE, G., HIGGINS, T. M. & TUOHY, M. G. 2004. Three-dimensional structure of a thermostable native cellobiohydrolase, CBH IB, and molecular characterization of the cel7 gene from the filamentous fungus, Talaromyces emersonii. *Eur J Biochem,* 271**,** 4495-506.

GRUNO, M., VÄLJAMÄE, P., PETTERSSON, G. & JOHANSSON, G. 2004. Inhibition of the Trichoderma reesei cellulases by cellobiose is strongly dependent on the nature of the substrate. *Biotechnol Bioeng,* 86**,** 503-11.

HALL, M., BANSAL, P., LEE, J. H., REALFF, M. J. & BOMMARIUS, A. S. 2011. Biological pretreatment of cellulose: Enhancing enzymatic hydrolysis rate using cellulose-binding domains from cellulases. *Bioresource Technology,* 102**,** 2910-2915.

HAMMOND, G. S. 1955. A Correlation of Reaction Rates. *Journal of the American Chemical Society,* 77**,** 334-338.

HE, D., BAO, L., LONG, Y., WEI, W. & YAO, S. 2000. A new study of the enzymatic hydrolysis of carboxymethyl cellulose with a bulk acoustic wave sensor. *Talanta,* 50**,** 1267-73.

HEDELAND, M., HOLMIN, S., NYGARD, M. & PETTERSSON, C. 1999. Chromatographic evaluation of structure selective and enantioselective retention of amines and acids on cellobiohydrolase I wild type and its mutant D214N. *Journal of Chromatography A,* 864**,** 1-16.

HEINZELMAN, P., KOMOR, R., KANAAN, A., ROMERO, P., YU, X. L., MOHLER, S., SNOW, C. & ARNOLD, F. 2010. Efficient screening of fungal cellobiohydrolase class I enzymes for thermostabilizing sequence blocks by SCHEMA structure-guided recombination. *Protein Engineering Design & Selection,* 23**,** 871-880.

HEINZELMAN, P., SNOW, C. D., WU, I., NGUYEN, C., VILLALOBOS, A., GOVINDARAJAN, S., MINSHULL, J. & ARNOLD, F. H. 2009. A family of thermostable fungal cellulases created by structure-guided

recombination. *Proceedings of the National Academy of Sciences of the United States of America,* 106**,** 5610-5615.

HON, D. N. S. 1994. Cellulose - a Random-Walk Along Its Historical Path. *Cellulose,* 1**,** 1-25.

HORN, S. J., SIKORSKI, P., CEDERKVIST, J. B., VAAJE-KOLSTAD, G., SORLIE, M., SYNSTAD, B., VRIEND, G., VARUM, K. M. & EIJSINK, V. G. H. 2006. Costs and benefits of processivity in enzymatic degradation of recalcitrant polysaccharides. *Proceedings of the National Academy of Sciences of the United States of America,* 103**,** 18089-18094.

HORN, S. J., SORLIE, M., VARUM, K. M., VALJAMAE, P. & EIJSINK, V. G. H. 2012. Measuring processivity. *In:* GILBERT, H. J. (ed.) *Cellulases.*

HOUBRAKEN, J., SPIERENBURG, H. & FRISVAD, J. C. 2012. Rasamsonia, a new genus comprising thermotolerant and thermophilic Talaromyces and Geosmithia species. *Antonie Van Leeuwenhoek,* 101**,** 403-21.

HU, G., HEITMANN, J. A. & ROJAS, O. J. 2009. In situ monitoring of cellulase activity by microgravimetry with a quartz crystal microbalance. *J Phys Chem B,* 113**,** 14761-8.

HUANG, S. Y. & CHEN, J. C. 1988. Ethanol-Production in Simultaneous Saccharification and Fermentation of Cellulose with Temperature Profiling. *Journal of Fermentation Technology,* 66**,** 509-516.

HUANG, Y., WEI, X. Y., ZHOU, S. G., LIU, M. Y., TU, Y. Y., LI, A., CHEN, P., WANG, Y. T., ZHANG, X. W., TAI, H. Z., PENG, L. C. & XIA, T. 2015. Steam explosion distinctively enhances biomass enzymatic saccharification of cotton stalks by largely reducing cellulose polymerization degree in G-barbadense and G-hirsutum. *Bioresource Technology,* 181**,** 224-230.

HUTCHISON, C. A., PHILLIPS, S., EDGELL, M. H., GILLAM, S., JAHNKE, P. & SMITH, M. 1978. Mutagenesis at a Specific Position in a DNA-Sequence. *Journal of Biological Chemistry,* 253**,** 6551-6560.

IGARASHI, K., UCHIHASHI, T., KOIVULA, A., WADA, M., KIMURA, S., OKAMOTO, T., PENTTILÄ, M., ANDO, T. & SAMEJIMA, M. 2011. Traffic jams reduce hydrolytic efficiency of cellulase on cellulose surface. *Science,* 333**,** 1279-82.

IGARASHI, K., UCHIHASHI, T., KOIVULA, A., WADA, M., KIMURA, S., PENTTILÄ, M., ANDO, T. & SAMEJIMA, M. 2012. Visualization of cellobiohydrolase I from Trichoderma reesei moving on crystalline cellulose using high-speed atomic force microscopy. *Methods Enzymol,* 510**,** 169-82.

IRWIN, D., SHIN, D. H., ZHANG, S., BARR, B. K., SAKON, J., KARPLUS, P. A. & WILSON, D. B. 1998. Roles of the catalytic domain and two cellulose binding domains of Thermomonospora fusca E4 in cellulose hydrolysis. *Journal of Bacteriology,* 180**,** 1709-1714.

JALAK, J. & VALJAMAE, P. 2010. Mechanism of Initial Rapid Rate Retardation in Cellobiohydrolase Catalyzed Cellulose Hydrolysis. *Biotechnology and Bioengineering,* 106**,** 871-883.

JOHNSON, K. A. 1992. *Transient-state kinetic analysis of enzymes reaction pathways in The Enzymes: Mechanisms of Catalysis* Academic Press Inc.: Salt Lake City, UT.

JORGENSEN, H., KRISTENSEN, J. B. & FELBY, C. 2007. Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities. *Biofuels Bioproducts & Biorefining-Biofpr,* 1**,** 119-134.

KANG, H. W., KIM, Y., KIM, S. W. & CHOI, G. W. 2012. Cellulosic ethanol production on temperature-shift simultaneous saccharification and fermentation using the thermostable yeast Kluyveromyces marxianus CHY1612. *Bioprocess and Biosystems Engineering,* 35**,** 115-122.

KARI, J., OLSEN, J., BORCH, K., CRUYS-BAGGER, N., JENSEN, K. & WESTH, P. 2014. Kinetics of Cellobiohydrolase (Cel7A) Variants with Lowered Substrate Affinity. *Journal of Biological Chemistry,* 289.

KERNOHAN, J. C. 1964. Activity of Bovine Carbonic Anhydrase in Imidazole Buffers. *Biochimica Et Biophysica Acta,* 81**,** 346-&.

KIPPER, K., VALJAMAE, P. & JOHANSSON, G. 2005. Processive action of cellobiohydrolase Cel7A from Trichoderma reesei is revealed as 'burst' kinetics on fluorescent polymeric model substrates. *Biochemical Journal,* 385**,** 527-535.

KLEMM, D., HEUBLEIN, B., FINK, H. P. & BOHN, A. 2005. Cellulose: Fascinating biopolymer and sustainable raw material. *Angewandte Chemie-International Edition,* 44**,** 3358-3393.

KNOTT, B. C., MOMENI, M. H., CROWLEY, M. F., MACKENZIE, L. F., GOTZ, A. W., SANDGREN, M., WITHERS, S. G., STAHLBERG, J. & BECKHAM, G. T. 2014. The Mechanism of Cellulose Hydrolysis by a Two-Step, Retaining Cellobiohydrolase Elucidated by Structural and Transition Path Sampling Studies. *Journal of the American Chemical Society,* 136**,** 321-329.

KOKKINIDIS, M., GLYKOS, N. M. & FADOULOGLOU, V. E. 2012. Protein Flexibility and Enzymatic Catalysis. *Structural and Mechanistic Enzymology: Bringing Together Experiments and Computing,* 87**,** 181-218.

KOMOR, R. S., ROMERO, P. A., XIE, C. B. & ARNOLD, F. H. 2012. Highly thermostable fungal cellobiohydrolase I (Cel7A) engineered using predictive methods. *Protein Engineering Design & Selection,* 25**,** 827-833.

KONSTANTINIDIS, A. K., MARSDEN, I. & SINNOTT, M. L. 1993. Hydrolyzes of Alpha-Cellobiosyl and Beta-Cellobiosyl Fluorides by Cellobiohydrolases of Trichoderma-Reesei. *Biochemical Journal,* 291**,** 883-888.

KOSTYLEV, M., ALAHUHTA, M., CHEN, M., BRUNECKY, R., HIMMEL, M. E., LUNIN, V. V., BRADY, J. & WILSON, D. B. 2014. Cel48A From Thermobifida fusca: Structure and Site Directed Mutagenesis of Key Residues. *Biotechnology and Bioengineering,* 111**,** 664-673.

KRAULIS, J., CLORE, G. M., NILGES, M., JONES, T. A., PETTERSSON, G., KNOWLES, J. & GRONENBORN, A. M. 1989. Determination of the three-dimensional solution structure of the C-terminal domain of cellobiohydrolase I from Trichoderma reesei. A study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. *Biochemistry,* 28**,** 7241-57.

KURASIN, M. & VALJAMAE, P. 2011. Processivity of Cellobiohydrolases Is Limited by the Substrate. *Journal of Biological Chemistry,* 286**,** 169-177.

LAIDLER, K. J. & PETERMAN, B. F. 1979. Temperature effects in enzyme kinetics. *Methods Enzymol,* 63**,** 234-57.

LARSEN, J., PETERSEN, M. O., THIRUP, L., LI, H. W. & IVERSEN, F. K. 2008. The IBUS process - Lignocellulosic bioethanol close to a commercial reality. *Chemical Engineering & Technology,* 31**,** 765-772.

LE COSTAOUËC, T., PAKARINEN, A., VÁRNAI, A., PURANEN, T. & VIIKARI, L. 2013. The role of carbohydrate binding module (CBM) at high substrate consistency: comparison of Trichoderma reesei and Thermoascus aurantiacus Cel7A (CBHI) and Cel5A (EGII). *Bioresour Technol,* 143**,** 196-203.

LINDER, M. & TEERI, T. T. 1996. The cellulose-binding domain of the major cellobiohydrolase of Trichoderma reesei exhibits true reversibility and a high exchange rate on crystalline cellulose. *Proc Natl Acad Sci U S A,* 93**,** 12251-5.

LONHIENNE, T., GERDAY, C. & FELLER, G. 2000. Psychrophilic enzymes: revisiting the thermodynamic parameters of activation may explain local flexibility. *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology,* 1543**,** 1-10.

MAHESHWARI, R., BHARADWAJ, G. & BHAT, M. K. 2000. Thermophilic fungi: Their physiology and enzymes. *Microbiology and Molecular Biology Reviews,* 64**,** 461-+.

MOMENI, M. H., PAYNE, C. M., HANSSON, H., MIKKELSEN, N. E., SVEDBERG, J., ENGSTROM, A., SANDGREN, M., BECKHAM, G. T. & STAHLBERG, J. 2013. Structural, Biochemical, and Computational Characterization of the Glycoside Hydrolase Family 7 Cellobiohydrolase of the Tree-killing Fungus Heterobasidion irregulare. *Journal of Biological Chemistry,* 288**,** 5861-5872.

MOMENI, M. H., STAHLBERG, J. & AL., E. 2014. Expression, crystal structure and cellulase activity of the thermostable cellobiohydrolase Cel7A from the fungus Humicola grisea var. thermoidea

Acta Crystallogica, 70**,** 2356-66.

MORAN-MIRABAL, J. M., BOLEWSKI, J. C. & WALKER, L. P. 2011. Reversibility and binding kinetics of Thermobifida fusca cellulases studied through fluorescence recovery after photobleaching microscopy. *Biophys Chem,* 155**,** 20-8.

MURPHY, L., BAUMANN, M. J., BORCH, K., SWEENEY, M. & WESTH, P. 2010a. An enzymatic signal amplification system for calorimetric studies of cellobiohydrolases. *Analytical Biochemistry,* 404**,** 140-148.

MURPHY, L., BOHLIN, C., BAUMANN, M. J., OLSEN, S. N., SORENSEN, T. H., ANDERSON, L., BORCH, K. & WESTH, P. 2013. Product inhibition of five Hypocrea jecorina cellulases. *Enzyme and Microbial Technology,* 52**,** 163-169.

MURPHY, L., BORCH, K., MCFARLAND, K. C., BOHLIN, C. & WESTH, P. 2010b. A calorimetric assay for enzymatic saccharification of biomass. *Enzyme and Microbial Technology,* 46**,** 141-146.

MUTTURI, S. & LIDEN, G. 2013. Effect of Temperature on Simultaneous Saccharification and Fermentation of Pretreated Spruce and Arundo. *Industrial & Engineering Chemistry Research,* 52**,** 1244-1251.

NAKAMURA, A., TSUKADA, T., AUER, S., FURUTA, T., WADA, M., KOIVULA, A., IGARASHI, K. & SAMEJIMA, M. 2013. The Tryptophan Residue at the Active Site Tunnel Entrance of Trichoderma reesei Cellobiohydrolase Cel7A Is Important for Initiation of Degradation of Crystalline Cellulose. *Journal of Biological Chemistry,* 288**,** 13503-13510.

NAKANISHI, K., SAKIYAMA, T. & IMAMURA, K. 2001. On the adsorption of proteins on solid surfaces, a common but very complicated phenomenon. *J Biosci Bioeng,* 91**,** 233-44.

NIDETZKY, B. & STEINER, W. 1993. A New Approach for Modeling Cellulase Cellulose Adsorption and the Kinetics of the Enzymatic-Hydrolysis of Microcrystalline Cellulose. *Biotechnology and Bioengineering,* 42**,** 469-479.

NISHIYAMA, Y., LANGAN, P. & CHANZY, H. 2002. Crystal structure and hydrogen-bonding system in cellulose 1 beta from synchrotron X-ray and neutron fiber diffraction. *Journal of the American Chemical Society,* 124**,** 9074-9082.

NORDE, W. 1986. Adsorption of Proteins from Solution at the Solid-Liquid Interface. *Advances in Colloid and Interface Science,* 25**,** 267-340.

OOSHIMA, H., SAKATA, M. & HARANO, Y. 1983. Adsorption of cellulase from Trichoderma viride on cellulose. *Biotechnol Bioeng,* 25**,** 3103-14.

OSULLIVAN, A. C. 1997. Cellulose: the structure slowly unravels. *Cellulose,* 4**,** 173-207.

PAGE, M. I. & JENCKS, W. P. 1971. ENTROPIC CONTRIBUTIONS TO RATE ACCELERATIONS IN ENZYMIC AND INTRAMOLECULAR REACTIONS AND CHELATE EFFECT. *Proceedings of the National Academy of Sciences of the United States of America,* 68**,** 1678-&.

PAKARINEN, A., OSTERGAARD HAVEN, M., DJAJADI, D. T., VÁRNAI, A., PURANEN, T. & VIIKARI, L. 2014. Cellulases without carbohydrate-binding modules in high consistency ethanol production process. *Biotechnol Biofuels,* 7**,** 27.

PALONEN, H., TENKANEN, M. & LINDER, M. 1999. Dynamic interaction of Trichoderma reesei cellobiohydrolases Cel6A and Cel7A and cellulose at equilibrium and during hydrolysis. *Applied and Environmental Microbiology,* 65**,** 5229-5233.

PARK, S., BAKER, J., HIMMEL, M., PARILLA, P. & JOHNSON, D. 2010. Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. *Biotechnology For Biofuels,* 3.

PAYNE, C. M., JIANG, W., SHIRTS, M. R., HIMMEL, M. E., CROWLEY, M. F. & BECKHAM, G. T. 2013a. Glycoside Hydrolase Processivity Is Directly Related to Oligosaccharide Binding Free Energy. *Journal of the American Chemical Society,* 135**,** 18831-18839.

PAYNE, C. M., KNOTT, B. C., MAYES, H. B., HANSSON, H., HIMMEL, M. E., SANDGREN, M., STAHLBERG, J. & BECKHAM, G. T. 2015. Fungal Cellulases. *Chemical Reviews,* 115**,** 1308-1448.

PAYNE, C. M., RESCH, M. G., CHEN, L., CROWLEY, M. F., HIMMEL, M. E., TAYLOR, L. E., SANDGREN, M., STÅHLBERG, J., STALS, I., TAN, Z. & BECKHAM, G. T. 2013b. Glycosylated linkers in multimodular lignocellulose-degrading enzymes dynamically bind to cellulose. *Proc Natl Acad Sci U S A,* 110**,** 14646-51.

PRAESTGAARD, E., ELMERDAHL, J., MURPHY, L., NYMAND, S., MCFARLAND, K. C., BORCH, K. & WESTH, P. 2011. A kinetic model for the burst phase of processive cellulases. *FEBS Journal,* 278**,** 1547-1560.

RAHIKAINEN, J. L., EVANS, J. D., MIKANDER, S., KALLIOLA, A., PURANEN, T., TAMMINEN, T., MARJAMAA, K. & KRUUS, K. 2013a. Cellulase-lignin interactions-the role of carbohydrate-binding module and pH in non-productive binding. *Enzyme Microb Technol,* 53**,** 315-21.

RAHIKAINEN, J. L., MOILANEN, U., NURMI-RANTALA, S., LAPPAS, A., KOIVULA, A., VIIKARI, L. & KRUUS, K. 2013b. Effect of temperature on lignin-derived inhibition studied with three structurally different cellobiohydrolases. *Bioresour Technol,* 146**,** 118-25.

REESE, E. T. 1976. History of the cellulase program at the U.S. army Natick Development Center. *Biotechnol Bioeng Symp***,** 9-20.

ROUVINEN, J., BERGFORS, T., TEERI, T., KNOWLES, J. K. C. & JONES, T. A. 1990. 3-Dimensional Structure of Cellobiohydrolase-Ii from Trichoderma-Reesei. *Science,* 249**,** 380-386.

SCHONBECK, C., HOLM, R. & WESTH, P. 2012. Higher Order Inclusion Complexes and Secondary Interactions Studied by Global Analysis of Calorimetric Titrations. *Analytical Chemistry,* 84**,** 2305-2312.

SCHULEIN, M. 2000. Protein engineering of cellulases. *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology,* 1543**,** 239-252.

SHANG, B. Z., CHANG, R. & CHU, J. W. 2013. Systems-level Modeling with Molecular Resolution Elucidates the Rate-limiting Mechanisms of Cellulose Decomposition by Cellobiohydrolases. *Journal of Biological Chemistry,* 288**,** 29081-29089.

SHIBAFUJI, Y., NAKAMURA, A., UCHIHASHI, T., SUGIMOTO, N., FUKUDA, S., WATANABE, H., SAMEJIMA, M., ANDO, T., NOJI, H., KOIVULA, A., IGARASHI, K. & IINO, R. 2014. Single-molecule Imaging Analysis of Elementary Reaction Steps of Trichoderma reesei Cellobiohydrolase I (Cel7A) Hydrolyzing Crystalline Cellulose I (alpha) and III (I). *Journal of Biological Chemistry,* 289**,** 14056-14065.

SOUSA, R., JR., CARVALHO, M. L., GIORDANO, R. L. C. & GIORDANO, R. C. 2011. Recent trends in the modeling of cellulose hydrolysis. *Brazilian Journal of Chemical Engineering,* 28**,** 545-564.

SPEZIO, M., WILSON, D. B. & KARPLUS, P. A. 1993. Crystal-Structure of the Catalytic Domain of a Thermophilic Endocellulase. *Biochemistry,* 32**,** 9906-9916.

SRISODSUK, M., REINIKAINEN, T., PENTTILÄ, M. & TEERI, T. T. 1993. Role of the interdomain linker peptide of Trichoderma reesei cellobiohydrolase I in its interaction with crystalline cellulose. *J Biol Chem,* 268**,** 20756-61.

STAHLBERG, J., JOHANSSON, G. & PETTERSSON, G. 1991. A New Model for Enzymatic-Hydrolysis of Cellulose Based on the 2-Domain Structure of Cellobiohydrolase-I. *Bio-Technology,* 9**,** 286-290.

STONE, M. M., WEISS, M. S., GOODALE, C. L., ADAMS, M. B., FERNANDEZ, I. J., GERMAN, D. P. & ALLISON, S. D. 2012. Temperature sensitivity of soil enzyme kinetics under N-fertilization in two temperate forests. *Global Change Biology,* 18**,** 1173-1184.

STÅHLBERG, J., DIVNE, C., KOIVULA, A., PIENS, K., CLAEYSSENS, M., TEERI, T. T. & JONES, T. A. 1996. Activity studies and crystal structures of catalytically deficient mutants of cellobiohydrolase I from Trichoderma reesei. *J Mol Biol,* 264**,** 337-49.

STÅHLBERG, J., HENRIKSSON, H., DIVNE, C., ISAKSSON, R., PETTERSSON, G., JOHANSSON, G. & JONES, T. A. 2001. Structural basis for enantiomer binding and separation of a common beta-blocker: crystal structure of cellobiohydrolase Cel7A with bound (S)-propranolol at 1.9 A resolution. *J Mol Biol,* 305**,** 79-93.

TAHERZADEH, M. J. & KARIMI, K. 2008. Pretreatment of lignocellulosic wastes to improve ethanol and biogas production: A review. *International Journal of Molecular Sciences,* 9**,** 1621-1651.

TAYLOR, W. R. 1986. The Classification of Amino-Acid Conservation. *Journal of Theoretical Biology,* 119**,** 205-&.

TEERI, T. T. 1997. Crystalline cellulose degradation: New insight into the function of cellobiohydrolases. *Trends in Biotechnology,* 15**,** 160-167.

TETER, S., CHERRY, J., WARD, C., JONES, A., HARRIS, P. & YI, J. 2005. *Variants of glycoside hydrolases*.

TEUGJAS, H. & VALJAMAE, P. 2013. Product inhibition of cellulases studied with C-14-labeled cellulose substrates. *Biotechnology for Biofuels,* 6.

TEWARI, Y. B., LANG, B. E., DECKER, S. R. & GOLDBERG, R. N. 2008. Thermodynamics of the hydrolysis reactions of 1,4-beta-D-xylobiose, 1,4-beta-D-xylotriose, D-cellobiose, and D-maltose. *Journal of Chemical Thermodynamics,* 40**,** 1517-1526.

TOMME, P., VAN TILBEURGH, H., PETTERSSON, G., VAN DAMME, J., VANDEKERCKHOVE, J., KNOWLES, J., TEERI, T. & CLAEYSSENS, M. 1988. Studies of the cellulolytic system of Trichoderma reesei QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis. *Eur J Biochem,* 170**,** 575-81.

TOMME, P., WARREN, R. A., MILLER, R. C., JR., K., D.G. & GILKES, N. R. 1995. Cellulose-binding domains: classification and properties. *Enzymatic Degradation of Insoluble Polysaccharides* American Chemical Society, Washington.

VÁRNAI, A., SIIKA-AHO, M. & VIIKARI, L. 2013. Carbohydrate-binding modules (CBMs) revisited: reduced amount of water counterbalances the need for CBMs. *Biotechnol Biofuels,* 6**,** 30.

VARROT, A., SCHULEIN, M. & DAVIES, G. J. 1999. Structural changes of the active site tunnel of Humicola insolens cellobiohydrolase, Cel6A, upon oligosaccharide binding. *Biochemistry,* 38**,** 8884-8891.

VERVERIS, C., GEORGHIOU, K., CHRISTODOULAKIS, N., SANTAS, P. & SANTAS, R. 2004. Fiber dimensions, lignin and cellulose content of

various plant materials and their suitability for paper production. *Industrial Crops and Products,* 19**,** 245-254.

VON OSSOWSKI, I., STÅHLBERG, J., KOIVULA, A., PIENS, K., BECKER, D., BOER, H., HARLE, R., HARRIS, M., DIVNE, C., MAHDI, S., ZHAO, Y., DRIGUEZ, H., CLAEYSSENS, M., SINNOTT, M. L. & TEERI, T. T. 2003. Engineering the exo-loop of Trichoderma reesei cellobiohydrolase, Cel7A. A comparison with Phanerochaete chrysosporium Cel7D. *J Mol Biol,* 333**,** 817-29.

VOUTILAINEN, S. P., MURRAY, P. G., TUOHY, M. G. & KOIVULA, A. 2010. Expression of Talaromyces emersonii cellobiohydrolase Cel7A in Saccharomyces cerevisiae and rational mutagenesis to improve its thermostability and activity. *Protein Eng Des Sel,* 23**,** 69-79.

VOUTILAINEN, S. P., NURMI-RANTALA, S., PENTTILÄ, M. & KOIVULA, A. 2014. Engineering chimeric thermostable GH7 cellobiohydrolases in Saccharomyces cerevisiae. *Appl Microbiol Biotechnol,* 98**,** 2991-3001.

VOUTILAINEN, S. P., PURANEN, T., SIIKA-AHO, M., LAPPALAINEN, A., ALAPURANEN, M., KALLIO, J., HOOMAN, S., VIIKARI, L., VEHMAANPERÄ, J. & KOIVULA, A. 2008. Cloning, expression, and characterization of novel thermostable family 7 cellobiohydrolases. *Biotechnol Bioeng,* 101**,** 515-28.

WANG, L. S., ZHANG, Y. Z. & GAO, P. J. 2008. A novel function for the cellulose binding module of cellobiohydrolase I. *Science in China Series C-Life Sciences,* 51**,** 620-629.

WATSON, B. J., ZHANG, H. T., LONGMIRE, A. G., MOON, Y. H. & HUTCHESON, S. W. 2009. Processive Endoglucanases Mediate Degradation of Cellulose by Saccharophagus degradans. *Journal of Bacteriology,* 191**,** 5697-5705.

WEBER, J. P. & FINK, A. L. 1980. Temperature-dependent change in the rate-limiting step of beta-glucosidase catalysis. *J Biol Chem,* 255**,** 9030-2.

WOLFENDEN, R., LU, X. D. & YOUNG, G. 1998. Spontaneous hydrolysis of glycosides. *Journal of the American Chemical Society,* 120**,** 6814-6815.

WOLFENDEN, R. & SNIDER, M. J. 2001. The depth of chemical time and the power of enzymes as catalysts. *Accounts of Chemical Research,* 34**,** 938-945.

WU, I. & ARNOLD, F. H. 2013. Engineered thermostable fungal Cel6A and Cel7A cellobiohydrolases hydrolyze cellulose efficiently at elevated temperatures. *Biotechnology and Bioengineering,* 110**,** 1874-1883.

YANG, B., WILLIES, D. M. & WYMAN, C. E. 2006. Changes in the enzymatic hydrolysis rate of avicel cellulose with conversion. *Biotechnology and Bioengineering,* 94**,** 1122-1128.

YE, Z. L. & BERSON, R. E. 2014. Factors affecting cellulose hydrolysis based on inactivation of adsorbed enzymes. *Bioresource Technology,* 167**,** 582-586.

ZAKARIASSEN, H., EIJSINK, V. G. H. & SORLIE, M. 2010. Signatures of activation parameters reveal substrate-dependent rate determining steps in polysaccharide turnover by a family 18 chitinase. *Carbohydrate Polymers,* 81**,** 14-20.

ZAKARIASSEN, H., AAM, B. B., HORN, S. J., VARUM, K. M., SORLIE, M. & EIJSINK, V. G. H. 2009. Aromatic Residues in the Catalytic Center of

Chitinase A from Serratia marcescens Affect Processivity, Enzyme Activity, and Biomass Converting Efficiency. *Journal of Biological Chemistry,* 284**,** 10610-10617.

ZHANG, S., BARR, B. K. & WILSON, D. B. 2000a. Effects of noncatalytic residue mutations on substrate specificity and ligand binding of Thermobifida fusca endocellulase Cel6A. *European Journal of Biochemistry,* 267**,** 244-252.

ZHANG, S., IRWIN, D. C. & WILSON, D. B. 2000b. Site-directed mutation of noncatalytic residues of Thermobifida fusca exocellulase Cel6B. *Eur J Biochem,* 267**,** 3101-15.

ZHANG, S. & WILSON, D. B. 1997. Surface residue mutations which change the substrate specificity of Thermomonospora fusca endoglucanase E2. *Journal of Biotechnology,* 57**,** 101-113.

ZHANG, S., WOLFGANG, D. E. & WILSON, D. B. 1999. Substrate heterogeneity causes the nonlinear kinetics of insoluble cellulose hydrolysis. *Biotechnology and Bioengineering,* 66**,** 35-41.

ZHANG, Y., XU, J. L., QI, W., YUAN, Z. H., ZHUANG, X. S., LIU, Y. & HE, M. C. 2012. A fractal-like kinetic equation to investigate temperature effect on cellulose hydrolysis by free and immobilized cellulase. *Appl Biochem Biotechnol,* 168**,** 144-53.

ZHANG, Y. H. P., HIMMEL, M. E. & MIELENZ, J. R. 2006. Outlook for cellulase improvement: Screening and selection strategies. *Biotechnology Advances,* 24**,** 452-481.

ZHENG, F. & DING, S. J. 2013. Processivity and Enzymatic Mode of a Glycoside Hydrolase Family 5 Endoglucanase from Volvariella volvacea. *Applied and Environmental Microbiology,* 79**,** 989-996.