



DI Patricia Bubner

Probing the paradigms of carbohydrate active enzymes

Mechanistic investigations of
glycosyltransferases and
glycosyl hydrolases

DISSERTATION

zur Erlangung des akademischen Grades einer
Doktorin der technischen Wissenschaften

erreicht an der

Technischen Universität Graz

betreut durch:

Univ.-Prof. DI Dr. Bernd Nidetzky
Institut für Biotechnologie und Bioprozesstechnik
Technische Universität Graz

2011

EIDESSTATTLICHE ERKLÄRUNG

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

Graz, am

.....

(Unterschrift)

STATUTORY DECLARATION

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

.....

date

.....

(signature)

De Profundis Clamavi

Zu dir · du einzig teure · dringt mein schrei
Aus tiefster schlucht darin mein herz gefallen·
Dort ist die gegend tot · die luft wie blei
Und in dem finstern fluch und schrecken wallen.

Sechs monde steht die sonne ohne warm.
In sechsen lagert dunkel auf der erde.
Sogar nicht das polarland ist so arm·
Nicht einmal bach und baum noch feld noch herde.

Erreicht doch keine schreckgeburt des hirnes

Das kalte grausen dieses eis-gestirnes
Und dieser nacht · ein chaos riesengross!

Ich neide des gemeinsten tieres los
Das tauchen kann in stumpfen schlafes schwindel.
So langsam rollt sich ab der zeiten spindel!

(Charles Baudelaire – Les Fleurs du Mal)

Acknowledgements

Even though it is only my name which appears on the cover of this dissertation, a great number of people have contributed to the production of this work. It is them to whom I wish to express my sincere gratitude.

First and foremost, I would like to thank my supervisor Univ.-Prof. Dr. Bernd Nidetzky who provided the academic, scientific and financial backing of this dissertation. His guidance, constructive criticism, scientific knowledge and patience were invaluable and enabled me not only to complete this thesis but also to develop a stronger confidence in my skills as a scientist. I have been amazingly fortunate to work with an advisor who gave me the freedom to develop my thesis projects independently.

I also want to thank Christiane Luley and Harald Pichler, both members of my thesis committee, for their scientific as well as personal support. I appreciate your time and all of your helpful advice. To my colleagues Regina Kratzer, Sigrid Egger, Simone Pival and Caterina Boniello – thank you for your support throughout my time at the institute. I am grateful for having made your acquaintance and for the fun times we had – at work and after work. Karin Longus, Margret Schiller and Jacqueline Harg helped with hands-on work in the laboratory. In general, I always appreciated the supportive, engaging working environment at the Institute of Biotechnology and Biochemical Engineering (IBB) and the helpful attitude of the entire IBB staff. I also owe a big “thank you” to Elena Pietrich and Doris Ribitsch who taught me so much about the importance of teamwork and the productivity that comes from it.

I am truly grateful to my friends who have been my emotional anchors and my advisors throughout many years. I am a fortunate person to know all of you and I want to express my sincere appreciation for your invaluable contributions, your support and your encouragement: Waltraud Gröbacher-Eicher, Monika Hitter, Bernadette Müller Kmet', Gernot Percht, Martina Ruch, Dominika Stiger, and Renate Winkler.

It is my immediate family who, early in my childhood, sparked my interest in natural sciences. I especially wish to express my gratitude to my father, Claudius Bubner, who instilled in me the inspiration to set high goals and the confidence to reach them. My mother, Elisabeth Bubner, showed me, among many other things, how to be hard-working and persistent and thus taught me how to achieve what I aim for.

Most importantly, I want to thank Andreas Gredler who has been a constant source of love, care, support and strength during all these years. Andy, you have shared the many uncertainties, challenges and sacrifices on the long way up to the completion of this dissertation and I deeply appreciate your patience and your constant encouragement.

Abstract

Carbohydrates are fundamental for life and ubiquitous in nature, e.g. as nutrients, osmolytes, biopolymers (i.e. cellulose) and in cellular energy storage (i.e. starch). Therefore, enzymes involved in formation and breakdown of glycosidic bonds – “carbohydrate-active enzymes” (CAZymes) – are of high scientific interest. The first chapter of this thesis deals with the enzymatic hydrolysis of cellulose by glycosyl hydrolases. This is a key process for the production of second generation biofuels which represent a long-standing leading theme in the field of sustainable energy. Despite the wealth of knowledge about the structure and function of cellulases, the elusive mechanism by which these enzymes disintegrate the complex structure of their insoluble substrate is still elusive. Microscopy was recognized as a powerful tool in research on cellulose degradation as early as in the late 1970s, where transmission electron microscopy was employed to study the binding of cellulases to their substrate. In the last decade, the methodology to visualize biological processes on the micro-, meso-, and nanoscale has improved significantly.

We here present a study, where we used atomic force microscopy (AFM) to investigate the structural dynamics of enzymatic cellulose degradation. AFM has the necessary resolution to observe even single enzymes, does not destroy the substrate or the enzymes and measurements in liquids are possible. We studied a complete *Trichoderma reesei* cellulase system on a nano-flat substrate using AFM. As a first step in substrate disintegration, elongated fissures emerge which develop into coniform cracks as disintegration continues. The dynamics of crack morphology reflects the interplay between surface degradation inside and outside of the crack. Hindered diffusion presumably leads to product inhibition and loss of cooperative interaction between the enzymes, thus transiently limiting cellulase activity inside the growing crack. These novel insights can help us to gain a better understanding of enzymatic cellulose degradation, which, in turn, might aid in making biofuels from lignocellulosic feedstock economically feasible.

In the second chapter of this thesis, we elucidate the determinants of donor specificity in glycosyltransferases (GTs). The majority of glycosidic bonds in nature is formed by GTs. GTs catalyze the transfer of a glycosyl moiety from an activated donor sugar to an acceptor thereby tying a glycosidic bond between them. We here present our efforts in elucidating the donor specificity in *Schizophyllum commune* trehalose phosphorylase which span from rational design to directed evolution.

Due to the fact that chemical synthesis of complex carbohydrates is extremely difficult, GTs are of high interest for tasks requiring glycosidic bond formation. As GTs are highly specific enzymes and crystallographic data is rare, directed evolution is widely employed to improve their substrate promiscuity. However, the bottleneck in many directed evolution approaches of GTs is the lack of an efficient high-throughput screening (HTS). We here present a novel screening method for GTs and compare it to previously published ones in terms of sensitivity, HTS-suitability (including HTS from

crude cell extracts), and unambiguousness. Developing a general HTS assay for GTs suffers from the intrinsic restrictions of the reaction to screen: glycosidic bond formation and reaction products cannot be measured directly. Also, false positives due to increased error-hydrolysis of the NDP-sugar donor is an issue. This can only be circumvented by screening in degradation direction, as transfer of the glycosyl moiety on the NDP must occur for a positive signal. Here, we present a dehydrogenase-coupled assay which is an example for a screening in degradation direction. So far, it suffers from a strict substrate specificity restricting the screening to UDP-glucose. However, it is a highly sensitive, reliable and stable continuous assay.

Zusammenfassung

Kohlenhydrate sind äußerst wichtig für das Leben und allgegenwärtig, zum Beispiel als Nährstoffe, Osmolyten, Biopolymere (z.B. Zellulose) und im zellulären Energiespeicher (z.B. Stärke). Daher sind die Enzyme, die glykosidische Bindungen knüpfen und spalten – sogenannte „kohlehydrataktive Enzyme“ (CAZymes) – von großem wissenschaftlichem Interesse. Das erste Kapitel dieser Dissertation beschäftigt sich mit der enzymatischen Hydrolyse von Zellulose durch Glykosylhydrolasen, einem Schlüsselfaktor bei der Biotreibstoffproduktion. Trotz der Fülle an verfügbaren Informationen über Struktur und Funktion der Zellulasen ist der genaue Mechanismus, mit dem sie ihr komplexes Substrat zerlegen, unklar. Schon in den späten 1970ern wurde die Mikroskopie als vielseitiges Hilfsmittel in der Zelluloseforschung entdeckt und es gab erste Transmissionselektronenmikroskopiestudien. Im letzten Jahrzehnt haben sich die möglichen Methoden, mit denen man biologische Prozesse im Mikro-, Meso- und Nanobereich visualisieren kann, vervielfacht.

Wir präsentieren hier eine Studie, in der wir mit Rasterkraftmikroskopie (atomic force microscopy, AFM) die strukturelle Dynamik des enzymatischen Zelluloseabbaus untersucht haben. AFM kann die nötige Auflösung erzielen, um Einzelenzyme zu beobachten ohne Substrat oder Enzyme zu zerstören. Man kann auch unter Flüssigkeit messen. Wir haben ein *Trichoderma reesei* Cellulasekomplettsystem auf einem nano-flachen Substrat untersucht. Zunächst entstehen längliche Fissuren, die sich mit fortschreitendem Substratabbau zu tiefen, konischen Rissen entwickeln. Die Dynamik dieser Risse spiegelt das Zusammenspiel des Abbaus der Substratoberfläche in den Rissen sowie außerhalb dieser wider. Erschwerte Diffusion führt wahrscheinlich zu Produktinhibierung und Verlust an kooperativer Interaktion der Enzyme, was die Aktivität der Zellulasen in einem wachsenden Riss zeitweise vermindert. Diese neuen Erkenntnisse erweitern unser Verständnis des enzymatischen Zelluloseabbaus und könnten dadurch ein Schritt in eine Zukunft sein, in der Biotreibstoffe aus Lignozellulose ökonomisch sind.

Im zweiten Kapitel dieser Dissertation erkunden wir die Determinanten der Donorspezifität in Glykosyltransferasen (GTs). Ein Großteil der natürlich vorkommenden glykosidischen Bindungen werden von GTs geknüpft. GTs katalysieren den Transfer eines Glykosylrestes von einem aktivierten Donorzucker auf einen Akzeptor und knüpfen dabei eine glykosidische Bindung zwischen den beiden. Wir zeigen unsere Untersuchungen an der Donorspezifität der Trehalose Phosphorylase aus *Schizophyllum commune*, die rationales Design sowie auch gerichtete Evolution beinhalten.

Da die chemische Synthese von komplexen Kohlenhydraten sehr schwierig ist, sind GTs von hohem Interesse wenn eine glykosidische Bindung geknüpft werden soll. Um promiskuitivere GTs zu bekommen, verwendet man oft gerichtete Evolution. Dies wird erschwert, weil es keinen generell geeigneten Hochdurchsatz-Screening (HTS) Assay für GTs gibt. Wir präsentieren eine neue Screeningmethode für GTs und vergleichen sie mit anderen Methoden hinsichtlich ihrer Sensitivität,

HTS Eignung (auch aus Zellrohextrakten), und Eindeutigkeit. Die Entwicklung eines generellen HTS-Assays ist schwierig da die Bildung einer glykosidischen Bindung und die Reaktionsprodukte nicht direkt gemessen werden können. Auch falsch positive Hits durch „Error-Hydrolysis“ des NDP-Donorzuckers können ein Problem sein. Dies kann umgangen werden indem man in die Degradierungsrichtung der Reaktion screent, weil der Transfer eines Glykosylrestes auf das NDP passieren muss, damit man hier ein positives Signal bekommt. Wir präsentieren hier einen Dehydrogenase-gekoppelten Assay der ein Beispiel für solch ein Screening in Degradationsrichtung ist. Derweil ist allerdings die Substratspezifität des gekoppelten Enzyms auf das Screening mit UDP-Glukose als Donor beschränkt. Es ist dies allerdings ein sehr sensitiver, zuverlässiger und stabiler kontinuierlicher Assay.

Table of Contents

Chapter One: Visualizing cellulase activity in situ	1
1.1. Introduction	2
1.1.1. (Ligno)Cellulose	4
1.1.2. Cellulases	5
1.2. Review: Visualizing cellulase activity	11
1.3. Cellulases dig deep: in situ observation of the mesoscopic structural dynamics of enzymatic cellulose degradation	45
1.3.1. Appendix to “Cellulases dig deep: in situ observation of the mesoscopic structural dynamics of enzymatic cellulose degradation”	59
1.4. Conclusion and outlook	66
1.5. Abbreviations	67
1.6. References	68
Chapter Two: Finding the “synthase switch” in <i>Schizophyllum commune</i> trehalose phosphorylase	72
2.1. Introduction	73
2.1.1. Trehalose phosphorylase from <i>Schizophyllum commune</i>	75
2.1.2. In quest for the “synthase switch”	75
2.2. Rational approach	78
2.2.1. Background and strategy	78
2.2.2. Results and discussion	79
2.2.3. Conclusion and outlook	86
2.2.4. Materials and methods	87
2.3. Evolutionary approach: High-throughput assays for the directed evolution of glycosyltransferases	93
2.3.1. Appendix to “High-throughput assays for the directed evolution of glycosyltransferases”	101
2.4. Conclusion and outlook	104
2.5. Abbreviations	106
2.6. References	107
2.7. Additional acknowledgements	109
Appendix	110
List of publications	111

Chapter One

Visualizing cellulase activity in situ

1.1. Introduction

The impending world-wide energy shortage is the Sword of Damocles above our heads urging us to tap new energy sources. It was the global oil crisis in 1973 that triggered mankind's quest for alternative energies. In the early eighties, the emerging biotech raised avid hope for a solution to the impending energy problem by offering the production of fuel ethanol from biological feedstock such as lignocellulosic (waste) material. Due to its abundance, sustainability and "nearly zero" net CO₂ balance cellulosic ethanol offers a promising perspective as a green and economic energy source. Today, 30 years later, the energy issue is more urgent than ever, but a solution is still out of reach. Additionally, the ethanol need is expected to increase in the future, and grain biofuel alone will not satisfy this demand, especially as it has raised severe concerns in terms of economical efficiency and social benefit (Lynd *et al.* 2008).

As a promising means to resolve these problems using lignocellulosic (waste) material as a feedstock for biofuel production has attracted intense scientific interest in recent years and so has energy biotechnology as a whole (Arantes and Saddler 2010, Lynd *et al.* 2008, Solomon *et al.* 2007). According to Lynd *et al.* (2008), however, the intrinsic recalcitrance of the insoluble cellulose to its saccharification is still the key bottleneck. Therefore, this recalcitrance needs to be overcome so as to make energy from lignocellulosic feedstock economically justifiable. In order to efficiently exploit lignocellulosic biomass for fuel production, the degradation of plant cell wall microfibrils into fermentable sugars needs to be highly efficient. So far, this is impeded by the intrinsic recalcitrance of lignocellulose, which roots in its complex morphology. Enzymatic hydrolysis of cellulose is key for the production of second generation biofuels; however, cellulose recalcitrance manifests itself in sluggish hydrolysis rates during enzymatic breakdown of cellulose, even at high enzyme concentrations (Himmel *et al.* 2007, Lynd *et al.* 2008). Overcoming cellulose recalcitrance therefore constitutes a central aim in biofuels development.

The intensity of research on enzymatic cellulose degradation first peaked in the late 1970s and early 1980s and has become very popular again in the first decade of the new millennium. This early research provided a great deal of today's basis of knowledge about enzymatic cellulose hydrolysis, including the finding that the actual cellulase is, in fact, an enzyme complex comprising three major hydrolytic enzyme activities which act jointly to break their insoluble substrate down to glucose: cellobiohydrolase (CBH), endoglucanase (EG), and β -glucosidase (cellobiase, BGL). Also, first studies on cellulase synergism (i.e., the interaction between different enzymes of the cellulase complex with each other) were performed in that time. By the end of the 1980s, the gross of the structural characterization of cellulases was achieved using small angle X-ray scattering (SAXS), confirming the organization of the "tadpole-shaped" CBH and EG into a cellulose binding domain

(CBD¹) and a catalytic domain (Abuja *et al.* 1988a, Abuja *et al.* 1988b, Pilz *et al.* 1990). In the 1990s, crystallization of various cellulases and nuclear magnetic resonance (NMR) studies of CBDs provided detailed structural information about the enzymes involved in cellulose degradation (Divne *et al.* 1994, Dominguez *et al.* 1995, Kleywegt *et al.* 1997, Mattinen *et al.* 1997, Varrot *et al.* 1999). Together with various other studies on complexed and non-complexed systems as well as on cellulose itself, this resulted in a huge amount of knowledge available about the enzymes (structure, function, interaction, synergism etc.) and the substrate (structure, allomorphs, pretreatment etc.). However, despite extensive research spanning over more than four decades so far and the resultant wealth of knowledge, processes for cellulose hydrolysis are still advanced empirically due to insufficient understanding of the mechanisms underlying insoluble substrate deconstruction by enzyme action. Translating the large amount of knowledge available about structure and catalytic function of cellulose degrading enzymes into comprehension of cellulase activity on the insoluble substrate has proven to be remarkably difficult. This is essentially due to the two following reasons: firstly, the heterogeneous composition of the insoluble substrate; and secondly, the paucity of methodology suitable of visualizing the action of cellulases on the cellulose surface at the nanometer scale. Schematic views of enzymatic hydrolysis of cellulose published nowadays have hence changed little compared to the ones decades ago (Arantes and Saddler 2010, Bansal *et al.* 2009, Igarashi *et al.* 2009, Lynd *et al.* 2002), which is the best evidence of the slow progress made in the field. Multiple factors concerning the enzyme (i.e. mode of adsorption to the substrate, individual components of cellulase systems showing synergy, product inhibition, stability) and the cellulosic substrate (i.e. crystallinity, available surface area, pore size, degree of polymerization) certainly play a crucial role in hydrolysis (Lynd *et al.* 2002). The relative importance and interdependence of these factors will nevertheless remain elusive, pending clarification of fundamental questions.

The first chapter of this thesis firstly aims at reviewing the history and possibilities of cellulase visualization, thus providing a common ground and outlook for future visualization approaches and secondly at giving an insight into the structural dynamics occurring during the elusive process of saccharification of the insoluble cellulose by cellulases. To this end, a study employing atomic force microscopy (AFM) is presented here: “Cellulases dig deep: A mesoscopic view on the structural dynamics of cellulose degradation“. In this study we achieved a time-resolved in situ visualization of the effect of enzyme action on the surface of cellulose at a nanoscale resolution. By using a special nano-flat preparation of cellulose we were able to directly monitor the complete process of deconstruction of the substrate from early fissuring events on the cellulose surface to the progressing surface erosion at longer hydrolysis times. We extracted the dynamic alteration of the cellulose surface structure quantitatively from AFM data. In the study presented here, we want to show how it reflects the interplay between surface degradation inside and outside of cracks, which

¹ CBDs are now often referred to as cellulose- or carbohydrate binding module (CBM) due to the fact that substrate binding modules binding to other carbohydrates than cellulose were identified (McLean *et al.* 2000).

develop as a result of the laterally processive and penetrating actions of the adsorbed cellulases occurring simultaneously. This mesoscopic structural view on cellulose degradation provides an immediate explanation for the peculiar kinetics of enzyme adsorption and soluble sugar formation which have long been the source of controversial interpretations in literature (Arantes and Saddler 2010, Lynd *et al.* 2002, Mansfield *et al.* 1999, Zhang and Lynd 2004). The results presented here possibly serve as a basis to opening up new strategies for making biofuel production from lignocellulosic feedstock economic.

1.1.1. (Ligno)Cellulose

Cellulose is an abundant natural biopolymer of β -1,4 linked glucose which is exclusively produced biosynthetically: either by photosynthesis as in higher plants and algae or non-photosynthetically as in certain microorganisms. The repeating unit in the unbranched cellulose polymer chain is D-anhydrocellobiose – resulting from the 180° twist occurring through the β -1,4 linkage of two D-anhydroglycopyranose units. The planar sugar rings of different chains are aligned in one plane and engage in stacking interactions with other layers of cellulose chains, allowing tight assembly of the material. These characteristics – linearity, rotation, and stacking interactions – favor a high degree of organization between single cellulose chains and the formation of a stable intra- and interchain hydrogen bond network. Eventually, this leads to tightly packed and highly ordered (“crystalline”) regions in cellulose (Cuissinat and Navard 2006, Lynd *et al.* 2002, Zhang and Lynd 2004). Thus, cellulose is insoluble in water: the crystalline regions of cellulose are barely accessible even for small molecules such as water. In amorphous regions, the entry of water can cause substantial swelling. Therefore, to dissolve cellulose harsh conditions are required. Only in the last decade a promising novel class of cellulose solvents has been opened up for non-derivatizing dissolution of cellulose: ionic liquids (Cuissinat and Navard 2006, Froix and Nelson 1975, Heinze and Koschella 2005).

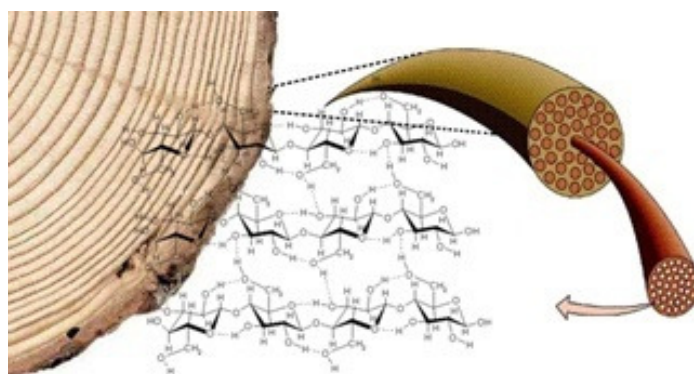


Figure I. Cellulose – an abundant natural biopolymer. Shown is the natural organization of cellulose into micro- and macrofibrils, exemplified on higher plants (wood).

Native cellulose crystals occur as a variable mixture of two different forms, cellulose I_α and cellulose I_β , the first one being the primary form in algae and the latter one in plants. Cellulose I_α and cellulose I_β differ in symmetry and directional formation of the hydrogen bond network; the chains themselves, nevertheless, are aligned in parallels in each microfibril, meaning that the reducing ends

of each chain point in the very same direction. In plants and algae the synthesized linear cellulose crystals (“microfibrils”) range from approximately 4 nm (plants) to approximately 20 nm (algae) in size. Several microfibrils are combined to a macrofibril, which, in plants, is often heavily intertwined with lignin at its outer microfibrils. Lignin is a heavily cross-linked, non-carbohydrate macromolecule that is covalently bound to hemicellulose and substantially contributes to cellulose breakdown being arduous. In plants, lignin constitutes up to 30% and hemicelluloses up to 35% of the dry weight (Jarvis 2003, Lynd *et al.* 2002, Zhang and Lynd 2004).

The influence of the cellulose structure on its recalcitrance to enzymatic hydrolysis has long been discussed. It is, nevertheless, still largely unclear at which level and to which extent substrate related factors act “rate retarding” in enzymatic cellulose disintegration (Jarvis 2003, Zhang and Lynd 2004). Ever since researchers have addressed this problem, the complexity of the process has posed an insurmountable obstacle. Regarding the substrate, for instance, multiple factors need to be taken into consideration: i) accessibility; ii) available surface area; iii) crystallinity; iv) degree of polymerization; v) lignin and hemicellulose content; vi) change of substrate features during disintegration, vii) pretreatment of the lignocellulosic materials (Zhang and Lynd 2004). In order to improve the accessibility of the substrate itself, several different approaches have been taken, e.g. steam explosion, acid pretreatment, and dissolution in ionic liquids (Lee *et al.* 2009).

1.1.2. Cellulases

Several fungi and bacteria are able to degrade cellulosic materials and can thus use it as an energy source. While fungi usually secrete cellulases to do so (“non-complexed cellulase systems”), bacteria harbor the majority of their cellulolytic enzymes in the cellulosome (“complexed cellulase systems”), an exocellular organelle located on the cell surface, which attaches to cellulose using a so-called dockerin. The term “cellulase” refers to a system of different proteins, primarily glycosyl hydrolases, rather than to a single enzyme. It is generally understood that enzymatic breakdown of cellulose requires the synergistic interplay of several different enzymes and probably also the support of non-hydrolytic proteins. The three major types of glycosyl hydrolases constituting the principal components of a complete cellulase system are: i) the exo-acting CBH, endo-acting EG, and BGL (see Figure II). They are working synergistically to degrade cellulose as outlined in Figure II. The exo-enzymes cleave off cellobiose units from chain ends and are further classified into CBH I – which starts at the reducing chain end – and CBH II – which starts at the non-reducing chain end. EGs specifically work on amorphous regions, cleaving within cellulose chains to generate new chain ends (Arantes and Saddler 2010, Kurasin and Valjamae 2011, Lynd *et al.* 2002, Zhang and Lynd 2004).

The majority of fungal CBHs and EGs feature an additional cellulose binding domain (CBD, often also referred to as cellulose binding module, CBM), which serves to adsorb the enzyme onto its

insoluble substrate. It is a noncatalytic, small domain² connected via a glycosylated short peptide linker to the catalytic domain. CBDs are wedge-shaped with a flat face harboring two tyrosines crucial for non-covalent binding onto the plane sugar rings of cellulose (Linder *et al.* 1995). According to their function CBDs have been divided into three types: i) surface-binding CBDs specific for insoluble crystalline areas; ii) CBDs binding to the chains of soluble cellulose; iii) CBDs binding to chain ends of soluble cellulose. The role(s) of the non-catalytic CBDs in enzymatic cellulose saccharification is intensively discussed: it has long been known that these modules are crucial for hydrolysis of insoluble crystalline cellulose by bringing the catalytic unit close to its insoluble substrate and by increasing the concentration of cellulases on the cellulose surface. In addition to these important tasks there have been speculations about possible other peculiar skills, especially non-hydrolytic disruption of the cellulose, often referred to as “amorphogenesis” (Arantes and Saddler 2010, Dagel *et al.* 2011, Lee *et al.* 2000, Linder *et al.* 1995, Nigmatullin *et al.* 2004). The CBD’s supposed role in the amorphogenesis of cellulose (Figure III) was – inter alia – supported by the findings in an AFM study showing that catalytically inactivated CBH I causes defects in the surface of cotton fibers (Lee *et al.* 2000). In absence of catalytic activity, the formation of these defects was accounted to the CBDs.

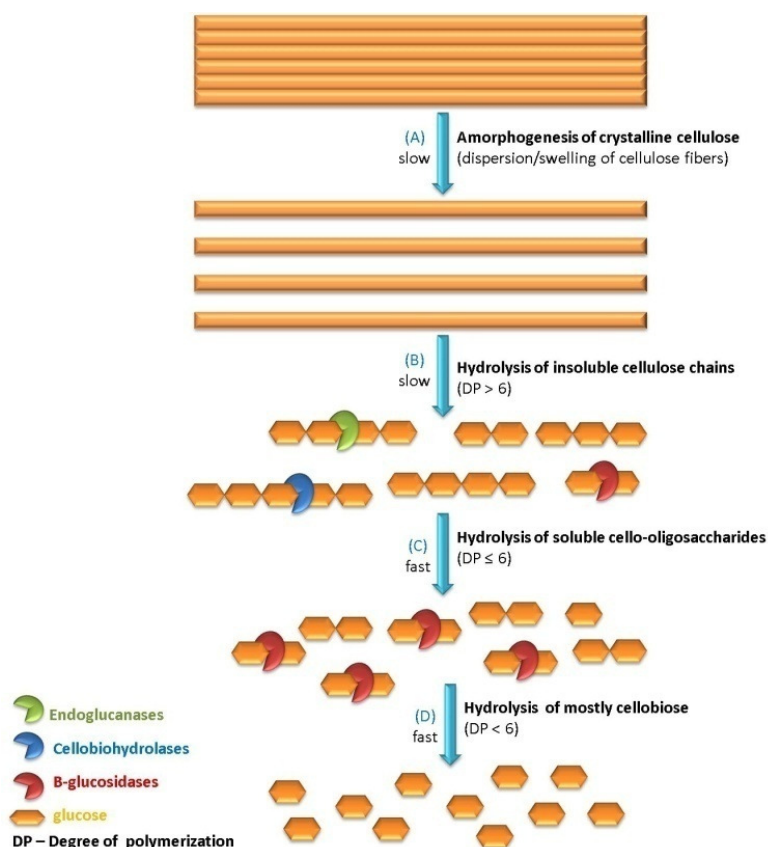


Figure II. Current model of cellulose degradation (Arantes and Saddler 2010).

² For example, the size of the CBD of the major *Trichoderma reesei* CBHI was determined with 30 x 10 x 18 Å (Linder *et al.* 1995).

Supposedly, there are also other non-hydrolytic proteins which assist the enzymes in making the recalcitrant substrate more accessible (“amorphogenesis”). One class of these non-hydrolytic proteins is similar to plant expansins³ and was found to be expressed in different bacteria and fungi. *Trichoderma reesei*, a soft rot fungus, and *Bjerkandera adusta*, for instance, express two expansin-like proteins which were given telling names: *T. reesei* swollenin (Saloheimo *et al.* 2002) and the recently discovered *Bjerkandera adusta* loosenin (Quiroz-Castaneda *et al.* 2011). Both swollenin and loosenin were shown to significantly disrupt cellulose without releasing reducing sugars. Additionally, subsequent hydrolysis of cellulose pretreated with either of them was significantly facilitated. Also, in light micrographs the disruptive action of loosenin caused visible balloon-like swelling of a cellulose fiber (acetate buffer, pH 5) (Quiroz-Castaneda *et al.* 2011). The mechanism of these expansin-like proteins, however, is completely unknown and thus can only be hypothesized (Arantes and Saddler 2010, Mansfield *et al.* 1999, Quiroz-Castaneda *et al.* 2011, Saloheimo *et al.* 2002).

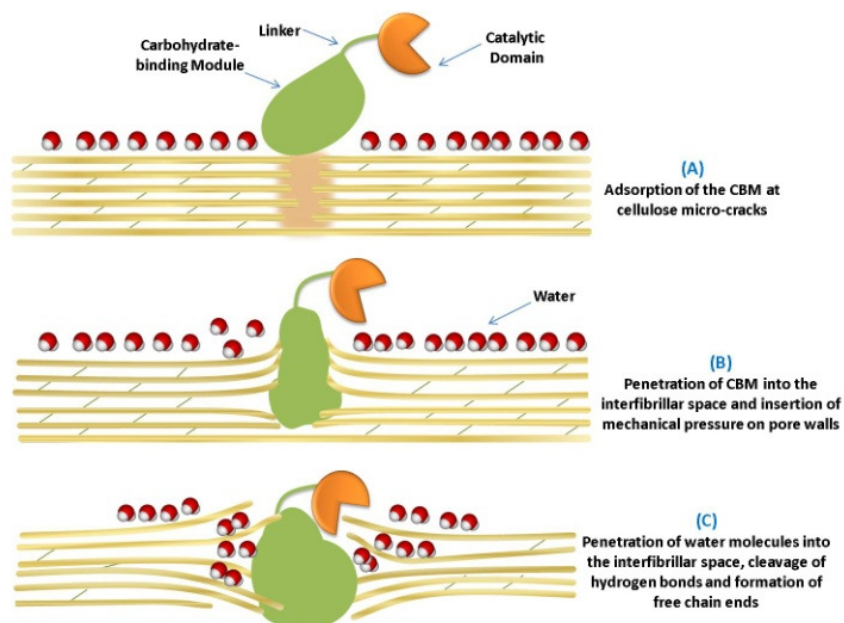


Figure III. Scheme of possible cellulose amorphogenesis by the carbohydrate binding domain (Arantes and Saddler 2010).

Complexed cellulase systems – or “cellulosomes” – are sophisticated multienzyme complexes bound to the bacterial cell. As outlined in Figure IV, cellulosomes generally serve to keep the cellulose degrading enzymes close to the bacterial cell. They are approximately 18 nm in size (Zverlov *et al.* 2008). Bacteria producing cellulosomes for enzymatic cellulose degradation are anaerobic, e.g. *Clostridium thermocellum*, *C. cellulovorans*, *C. cellulolyticum* and *Ruminococcus flavefaciens*. Compared to non-complexed systems produced by aerobic species such as fungi and several bacteria

³ Plant expansins play an important role in plant cell enlargement and several other processes and are thought to be capable of disrupting the hydrogen bond network of cellulose. Usually between 250 – 275 amino acids in length, expansins are made up of two domains with the C-terminal domain being a putative CBD (Quiroz-Castaneda *et al.* 2011, Saloheimo *et al.* 2002).

(e.g. *Cellulomonas sp.*), the process by which complexed cellulase systems degrade cellulose is even less understood. *C. thermocellum* is the most extensively studied among the organisms producing cellulosomes. It grows well on cellobiose and Avicel and its cellulosome can easily be isolated from the culture supernatant (Lamed and Bayer 1988, Lynd *et al.* 2002, Morag *et al.* 1991, Zhang and Lynd 2005, Zverlov and Schwarz 2008). The organization of its cellulosome has been studied using electron microscopy (Garcia-Alvarez *et al.* 2011, Madkour and Mayer 2003, Mayer *et al.* 1987) and crystallization studies (Adams *et al.* 2010). Within the cellulosome the cellulolytic enzymes are arranged on the “scaffoldin” – a non-catalytic protein – by interactions between its cohesin modules and the dockerin modules of the catalytic proteins of the cellulosome (see Figure IVB) (Garcia-Alvarez *et al.* 2011, Lynd *et al.* 2002).

1.1.2.1. State of the art models of cellulase action and their historical background

Generations of scientists have spent an immense effort to elucidate the mechanisms underlying the synergistic breakdown of cellulose by cellulases which led to a detailed understanding of structure and function of the enzymes involved. Nevertheless, despite of decades of intensive research and its findings, the overall process of enzymatic cellulose breakdown is still elusive. This is reflected in the sketchy models and graphs published, which have barely changed in the past three decades: while our knowledge of structure and function of the enzymes has significantly increased, our understanding of the exact mechanism by which they disintegrate cellulose has hardly improved.

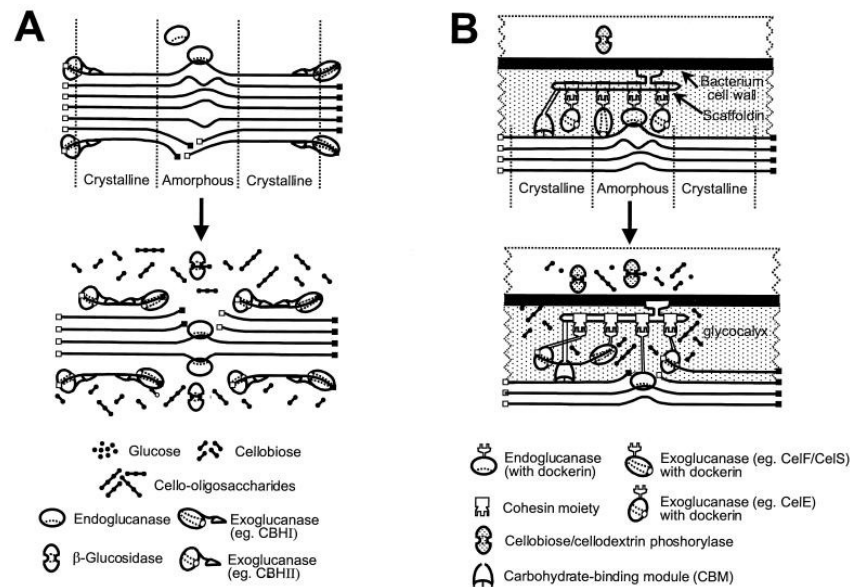


Figure IV. Model of enzymatic cellulose degradation. A: non-complexed systems. B: complexed systems (Lynd *et al.* 2002).

Generally, non-complexed – mostly fungal – cellulase systems have been studied much more extensively than the cellulosomes. In 1950, a first model (“C₁-C_x-model”) of synergistic behavior among cellulases postulated the cooperation between an unidentified non-hydrolytic factor C₁ and a hydrolytic C_x component of the cellulase system (Reese *et al.* 1950). In the late 1970s, the supposed non-hydrolytic factor C₁ was found to be a CBH II (Wood and McCrae 1978), which led to a change

in cellulase nomenclature. In the following years of intense research, the model of synergistic behavior among cellulases was gradually refined (Chanzy *et al.* 1983, Henrissat *et al.* 1985, White and Brown 1981, Wood and McCrae 1978). By 1985, it was well established that enzymatic cellulase degradation is accomplished by synergistic action of endo- and exoglucanases as well as β -glucosidase, as shown in Figure II. From the mid-1990s onwards – with the additional knowledge of the overall shape and structure of cellulases as well as their organization into domains (Abuja *et al.* 1988a, Gilkes *et al.* 1991, Rouvinen *et al.* 1990) – CBH and EG have often been displayed as an elliptic core with a wedge-shaped tail in various models (Figure IV) (Lynd *et al.* 2002, Zhang and Lynd 2004).

The intense efforts undertaken to understand cellulose hydrolysis at the nanoscale level – enzyme structure determination and modeling approaches, probing synergistic behavior, detailed examination of adsorption and hydrolysis, modeling of adsorption and hydrolysis kinetics (Bansal *et al.* 2009, Bommarius *et al.* 2008, Levine *et al.* 2010, Zheng *et al.* 2009, Zhong *et al.* 2009) – were not able to explain the exact mechanism by which cellulases successfully disintegrate their insoluble substrate. The current models (Figure II and Figure IV) are therefore incomplete and do not advance our understanding of the process in order to overcome the main bottleneck in enzymatic cellulose degradation: the recalcitrance of cellulose (Lynd *et al.* 2008).

1.1.2.2. A short history of cellulase visualization

In order to capitalize on the fact that a solid substrate is used, visualization is the method of choice to gain a deeper understanding of how cellulose breakdown is accomplished. In the early 1980s, when research on cellulose hydrolysis had its first zenith, some of the research groups most active in this field started to employ transmission electron microscopy (TEM) to study the enzymatic breakdown of cellulose (Chanzy *et al.* 1984, Chanzy *et al.* 1983, White and Brown 1981). However, a direct observation of the process using TEM is impossible: for visualization, auxiliary means such as gold particles or negative staining are required and, moreover, the substrate as well as the catalyst are delicate matters and therefore easily destroyed by the high-voltage electron beams employed. This may give rise to artifacts. The same is true for (environmental) scanning electron microscopy [(E)SEM]: even though significantly less energy is applied, electron damage can not be ruled out on delicate matters. A decade ago, atomic force microscopy (AFM) – then a novel method in biophysical applications – was applied to observe the impact of cellulases on cotton fibers. In this study, Lee *et al.* showed how *T. reesei* CBH I leaves “paths” or “tracks” on the surface of cotton fibers, whereas EG II just “smoothed” the fibers (Lee *et al.* 2000). The non-homogenous and rough surface of the cotton fibers, however, does not allow for quantification of cellulase action on a nanoscale. In the following years, a couple of other groups also applied different AFM methods, i.e. using a liquid cell, high speed imaging, or phase imaging on different substrates and with different enzyme preparations (Igarashi *et al.* 2009, Liu *et al.* 2009, Liu *et al.* 2011, Nigmatullin *et al.* 2004, Quirk *et al.* 2010, Santa-Maria and Jeoh 2010). Recently, studies on the action of CBH I on strictly crystalline

cellulose have been published by two different groups. They found out that CBH I hydrolyzes cellulose on the hydrophobic faces of the crystal and that it – as proposed – moves along its substrate (“sliding action”) (Igarashi *et al.* 2009, Liu *et al.* 2011). As EGs, which preferably act on amorphous cellulose, are barely active with highly crystalline cellulose, they could not test them. In all these previous AFM studies, the employed substrates limited the results: either they used complex substrates which have insufficient surface flatness in order to extract relevant data or to directly visualize cellulases, such as cotton fibers (Lee *et al.* 2000) or the cellulosic substrates used were chemically modified (Nigmatullin *et al.* 2004) or strictly crystalline – thus allowing only the observation of distinct cellulase activities (Igarashi *et al.* 2009, Liu *et al.* 2011).

Recently, total internal reflection fluorescence microscopy (TIRF-M), confocal Raman microscopy (CRM) together with secondary ion mass spectrometry (SIMS) and X-ray tomography have been used to attempt visualization of enzymatic cellulose degradation (Li *et al.* 2010, Liu *et al.* 2010, Penttilä *et al.* 2010). These methods, however, can only be used to supplement in situ AFM visualization of enzymatic action as they are limited either in resolution or in in situ possibilities. AFM renders visualization of the enzymatic breakdown of cellulose possible without destroying substrate or enzymes. Also, measurements in liquids are possible when using a so-called liquid cell. Thus, AFM is currently the method of choice for in situ imaging of (single) cellulases and their action on their substrate.

1.2. Review: Visualizing cellulase activity

in preparation for Biotechnology and Bioengineering

Visualizing cellulase activity

Patricia Bubner¹, Harald Plank², Bernd Nidetzky¹

¹Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12, A-8010 Graz, Austria; telephone: 43-316-873-8400; fax: 43-316-873-8434; email: bernd.nidetzky@tugraz.at; ²Institute for Electron Microscopy, Graz University of Technology, Steyrergasse 17, A-8010 Graz, Austria

Correspondence to: Bernd Nidetzky

Running Title: Visualizing cellulase activity

ABSTRACT

On the 30th anniversary of the first study that aimed at visualizing cellulase activity we here review the work done in this field from the early transmission electron microscopy (TEM) studies – which provided much of the evidence on cellulase mechanism we are taking for granted today – to the sophisticated combined biophysical techniques available nowadays. Together with biochemical methods, visualization of the activity of cellulases on their solid, recalcitrant substrate has provided the broad knowledge on enzymatic cellulose degradation that we hold today. Additionally, this review shall give an overview of the limitations and pitfalls of visualization and the conclusions that can be drawn from such studies. Finally, an outlook on work to be done in the field is given.

KEYWORDS: cellulose degradation; cellulase; AFM; TEM; renewable energy

Introduction

In the last decade, the methodology for visualization of biological processes on the micro-, meso- and nanoscale has improved significantly. This is firstly due to advance in the respective techniques leading to enhanced possibilities and secondly because of a tighter connection of biology and physics as a whole – which has also led to the emergence of the interdisciplinary field of biophysics. Today we are able to image single biomolecules as well as to generate videos of proteins in action: impressive results have been published recently, showing for example the shape of DNA nanostructures (Sander and Golas 2011), the way a restriction enzyme interacts with DNA (Crampton et al. 2007), or how the intensively studied myosin V “walks” along an actin filament (Kodera et al. 2010). The respective methods are crucial for disciplines such as (bio)nanotechnology, where visualization is a powerful tool to check the structures generated. Also, they open up novel perspectives in basic and applied research.

(Ligno)cellulose has been recognized as a promising alternative energy source to ease the impending world-wide energy shortage. As a feedstock for biofuel production it has also attracted intense scientific interest. Due to its abundance, sustainability and “nearly zero” net CO₂ balance, cellulosic ethanol offers a promising perspective as a green and economic energy source (Arantes and Saddler 2010; Lynd et al. 2008; Solomon et al. 2007). In order to efficiently exploit lignocellulosic biomass for fuel production, the degradation of plant cell wall microfibrils into fermentable sugars needs to be highly efficient. This, however, is hindered by the intrinsic properties of the substrate which is extremely recalcitrant to its degradation. Firstly, the strong intra- and interchain hydrogen bond network in its crystalline regions prevents even small molecules such as water from entering. Secondly, cellulose in plants is heavily intertwined with lignin at its outer microfibrils thus further limiting accessibility. Additionally, the intrinsic properties of the substrate – such as the degree of polymerization and crystallinity index – change during hydrolysis. And finally, concerning the enzymatic saccharification of cellulose, little is known about the exact mechanism of their hydrolytic cellulose breakdown, even though an incredible amount of information is available about structure and function of cellulases. It is expected that ethanol from enzymatic cellulose hydrolysis will be economically and ecologically favorable over ethanol derived by acid hydrolysis or from corn (Bansal et al. 2009). However, enzymatic cellulose hydrolysis suffers from several problems, such as the hydrolysis rate decrease even at low degrees of substrate conversion, high enzyme cost and the high enzyme loadings required (Himmel et al. 2007; Lynd et al. 2008). According to Lynd et al. (2008), the intrinsic recalcitrance of the insoluble cellulose to its saccharification is still the key bottleneck to be overcome in order to make energy from lignocellulosic feedstock economically justifiable.

One effect of the world energy crisis in the 1970s was that it raised an immense interest in biofuels research. In the following decades, a large knowledge basis on cellulose (structure, allomorphs, pretreatment methods etc.) and cellulose degrading enzymes

(structure, function, interaction, synergism, etc.) was built through research in varying intensity. However, kinetics and synergism of cellulases on their substrate are still poorly understood. A cellulase is actually an enzyme complex comprising three major hydrolytic enzymes which act jointly (“synergistically”) to break their insoluble substrate down to glucose: cellobiohydrolases (CBHs), endoglucanases (EGs), and β -glucosidases (BGLs). CBH and EG bind to their insoluble substrate by mediation of a cellulose binding domain (CBD), which may also assist in substrate disruption non-hydrolytically (“amorphogenesis”) (Arantes and Saddler 2010; Mansfield et al. 1999). The current model of enzymatic cellulose degradation is shown in Figure 1. Generally, today’s model barely differs from the one in the 1980s: a schematic depiction of the tadpole-like shape of the enzymes attacking the substrate and – in a “black box” mechanism – releasing soluble cellobiose which is eventually cleaved by BGL to glucose. For a better understanding of how the enzymes disrupt the solid substrate, visualization is a promising opportunity.

30 years of cellulase visualization: the history of cellulase visualization in a nutshell

Conventional kinetic characterization fails for cellulases as it requires a soluble substrate; therefore, Michaelis Menten kinetics cannot be determined for cellulases (Bansal et al. 2009). The adsorption of cellulases to their substrate further complicates characterization: firstly, because it is a process which is hard to describe mathematically as it does not follow Langmuir adsorption laws which are defined as a monolayer-type adsorption; secondly, because adsorption serves as a means to increase the concentration of enzymes at the substrate surface and is possibly also reversible; and finally, because adsorption is partly random and cellulases are also competing for binding sites (Bansal et al. 2009; Valjamae et al. 1998). However, in order to capitalize on the fact that cellulose is an insoluble substrate for the hydrolytic enzymes, visualization is the method that is expected to clarify fundamental questions about the factors leading to the decrease in hydrolysis rates. Moreover, it will help

to connect and translate the existing knowledge about the structure and the catalytic function of cellulases into comprehension of their activity on the insoluble substrate. In investigations on the hydrolysis of insoluble starch granules, for example, synchrotron UV fluorescence microscopy was applied successfully: taking advantage of protein autofluorescence on a non-fluorescent solid substrate, the disintegration of a granule by amylases was visualized at 283 nm resolution (Tawil et al. 2011).

Microscopy was recognized as a powerful tool in research on cellulose degradation as early as in the late 1970s, where transmission electron microscopy (TEM) was employed to study the binding of cellulases to their substrate. The nanoscale visualization methods available back then were basically limited to TEM. In the early 1980s, when research on cellulose hydrolysis had its first zenith, some of the research groups most active in this field started to employ TEM to study the enzymatic breakdown of cellulose (Chanzy et al. 1984; Chanzy et al. 1983; White and Brown 1981). However, a direct observation of the process using TEM is impossible: for visualization, auxiliary means such as gold particles or negative staining are required and, moreover, substrate as well as catalyst are delicate matters and therefore easily destroyed by the high-voltage electron beams employed. This may give rise to artifacts. The same is true for scanning electron microscopy (SEM) as well as for the newer environmental SEM (ESEM) which was developed in the late 1980s: even though in ESEM significantly less energy is applied, electron damage can not be ruled out on delicate matters. Additionally, the vacuum required in electron microscopy can harm biological samples, especially when a certain solvent content is present. Also, sample preparation can lead to artifacts and damage to sensitive material: TEM requires some sort of labeling or staining, and ESEM sputtering (surface coating with conductive material).

Fortunately, these limitations do not apply for atomic force microscopy (AFM) which was applied to observe the impact of cellulases on cotton fibers a decade ago for the first time. At this time, it was a novel method in biophysical applications. In the study mentioned,

Lee *et al.* showed how *T. reesei* CBH I leaves “paths” or “tracks” on the surface of cotton fibers, whereas EG II just “smoothed” the fibers (Lee *et al.* 2000). The non-homogenous and rough surface of the cotton fibers, however, does not allow for quantification of cellulase action on a nanoscale. In the following years, a few other groups also applied different AFM methods, i.e. using a liquid cell, high speed imaging, or phase imaging on different substrates and with different enzyme preparations (Igarashi *et al.* 2009; Liu *et al.* 2009; Liu *et al.* 2011; Nigmatullin *et al.* 2004; Quirk *et al.* 2010; Santa-Maria and Jeoh 2010). Recently, studies on the action of CBH I on strictly crystalline cellulose have been published by two different groups. They found out that CBH I hydrolyzes cellulose on the hydrophobic faces of the crystal and that it – as proposed – moves along its substrate (“sliding action”) (Igarashi *et al.* 2009; Liu *et al.* 2011). Recently, also total internal reflection fluorescence microscopy (TIRF-M), confocal Raman microscopy (CRM) together with secondary ion mass spectrometry (SIMS) and X-ray tomography have been used to attempt visualization of enzymatic cellulose degradation (Li *et al.* 2010; Liu *et al.* 2010; Penttilä *et al.* 2010). These methods, however, are limited either in resolution or in in situ possibilities.

This review aims firstly at summarizing and comparing the diverse visualization methods employed and visualization studies done so far – from the early TEM studies to the various methods used today. Secondly, we want to generate a state of the art basis of knowledge about cellulase visualization, identify limitations of the methods and provide an outlook of what one can expect from future studies. Visualization of cellulase action is anticipated to enhance our knowledge about the interactions between the enzymes (“synergism”) as well as between the enzymes and their substrate. It should also provide insight into – and possibly quantification of – the changes occurring in the substrate during hydrolysis and their influence on the drop in hydrolysis rate which occurs at 20-40 % conversion (Bansal *et al.* 2009).

Visualization methods

The visualization of biological processes on delicate matters requires a system which is not only capable of producing high resolution images but also does so in a “natural” environment. Roughly, this applies to a buffered aqueous system in ambient pressure and temperature. Additionally, we need to carefully select the substrate we use: it should be as close to a natural substrate as possible. Only in such an environment we can guarantee that what we observe is not biased. Now, laboratory environments have many advantages but mostly they are quite artificial, and, additionally, advanced imaging techniques often require quite harsh procedures, e.g. highly pressurized vacuum chambers, low temperatures or even freezing, dry surfaces, modifications to visualize biomolecules (on the substrate or the molecule site), or intense sample processing. To observe cellulases and their action on cellulose we thus need i) a natural (preferably aqueous) environment, ii) a natural substrate, iii) high resolution (Ding et al. 2008). In the following, the methods used so far to visualize enzymatic cellulose degradation are briefly described with a focus on their applicability and limitations (see also Table I).

Electron microscopy

Transmission electron microscopy (TEM)

The first attempts to visualize cellulases were accomplished using TEM, which has been the key technique in elucidation of the structure of biological materials, e.g. eukaryotic and prokaryotic cells, plant cell walls, and, of course, of diverse cellulosic materials (Peckys et al. 2011). In TEM, an accelerated electron beam is exerted onto a specimen that needs to be thin enough to allow the gross of the electrons to pass through. Within the specimen, the electrons interact with the sample thus yielding an image which is then focused and detected (Denk and Horstmann 2004). The resolution of TEM achieves atomic scale, e.g. in cryo-electron microscopy (cryo-EM). Its major disadvantage, however, is that biological material can

neither be imaged in its natural environment nor is real-time imaging possible. Additionally, the sample preparation procedure required for biological samples, such as labeling or heavy metal staining, might alter any features of the samples, and the high electron radiation doses required might further damage delicate specimens (Allison et al. 2010; Peckys et al. 2011). Actually, the highly crystalline cellulose produced by the algae *Cladophora* was reported to be sensitive to beam damage (Imai et al. 2003). Therefore, one can imagine that cellulosic materials with higher amorphous regions or other fragile structural features are even more susceptible to intense electron radiation.

In order to visualize proteins using TEM, labeling is required. The methods used for visualization of cellulases employed mainly two methods to this end: direct colloidal gold labeling of cellulases or detection of already bound cellulases with specific antibodies which are either labeled with gold or carry an antigen for another – gold-labeled – antibody (immuno-TEM) (Blanchette et al. 1989; Chanzy et al. 1984; Chanzy et al. 1983; Gilkes et al. 1988; Imai et al. 1998; Lehtiö et al. 2003; Nieves et al. 1991; Penttilä et al. 2010; Pinto et al. 2008; White and Brown 1981).

Recently, the applicability of TEM has been successfully advanced to allow for measurements in liquids: by employing scanning transmission electron microscopy (STEM), imaging fully hydrated native cells is possible with a resolution of 32 nm – without the need for heavy metal staining. Intracellular structures were identified according to their size, shape, location and difference in density. To characterize proteins within cells or organelles, however, fluorescent or other labels are required. Even though the electron dose applied in STEM is significantly lower than in TEM, radiation damage cannot be ruled out completely (de Jonge et al. 2009; Peckys et al. 2011).

Scanning electron microscopy (SEM)

In SEM, which is a raster scan method used to visualize surfaces, a focused high-energy electron beam scans the surface of interest. Depending on the detection mode, the signal used

to create the image can be produced by secondary or backscattered electrons or characteristic X-rays (Denk and Horstmann 2004). SEM is widely employed to image various surfaces, also biological ones. However, non-conductive surfaces need to be sputtered for image acquisition, for example with gold. As in TEM, biological samples need to be fixed, e.g. by freezing, dehydration or chemical treatment. Due to its lower energy electron beams – which minimize the possibility of surface damage – and the possibility to measure in liquids, environmental scanning microscopy (ESEM) has gained popularity (Collins et al. 1993).

High-resolution SEM (HRSEM) was successfully applied to investigate the external and internal structures of plant cells in several studies at a sufficiently high resolution. The improvement in resolution is achieved by fitting a field emission gun to a SEM (Brecknock et al. 2011).

Although numerous electron microscopy studies have provided us with high-resolution information on substrate structures [e.g. (ligno)cellulosic material, plant cell walls, etc.] and were also employed to investigate the macromolecular structure of cellulosomes, they are not suited to investigate the dynamics of protein–protein or protein–substrate interactions in real time as required to elucidate the mechanisms of enzymatic cellulose hydrolysis (Allison et al. 2010; Ding et al. 2008): albeit high resolution images are possible, high contrast images depend on staining (usually with heavy metals), proteins and/or substrate often need to be labeled for visualization, extensive specimen preparation is required and might lead to artifacts, three-dimensional information cannot be obtained, and measurements in aqueous systems are generally limited to ESEM and STEM, which so far lags behind TEM in terms of resolution. To avoid deflection of the electrons, the microscopes are operated in vacuum. Thus, the general environment within an electron microscope and, of course, the electron beam might damage the substrate, the proteins or change their interaction (Allison et al. 2010; Denk and Horstmann 2004).

Atomic force microscopy (AFM)

AFM was developed in 1986 from the scanning tunneling microscope (STM), a raster scanning microscope requiring a conductive sample. In the beginning, the primary modus operandi of AFM was contact mode where the deflections of a flexible cantilever are determined. These deflections result from the interactions of the tip with the surface. However, contact mode measurement exerts strong lateral forces onto the surface, which bears the risk of artifacts: features might be moved across the surface by the tip and delicate (biological) surfaces might be damaged. Also, it is hard to visualize very soft surfaces in contact mode due to artifacts generated by the applied force or damage to the surface. Tapping (also referred to as “intermittent contact”) mode is now state of the art in investigating biological samples. In tapping mode, the tip is oscillated at its resonance frequency and scans in close proximity to the sample. The interaction between the surface and the tip affects the oscillation frequency, which then generates the image (Allison et al. 2010).

All the limitations which apply to electron microscopy are overcome by AFM: measurements are possible in air as well as in liquid, pressure and temperature can be chosen as required, and no special sample pretreatment – such as sputtering, labeling or staining – is required. Thus, it is possible to observe biomolecules – in situ and continuously – in a physiological environment. Tapping mode AFM allows for direct observation of dynamic biological processes. Another advantage in AFM imaging is that three-dimensional images of the surface can be generated from the information measured. One of the restrictions in AFM, however, is the substrate surface – in contrast to TEM, which allows to look inside the observed material. Also, the surface investigated needs to be reasonably flat in order to achieve a high resolution and to be able to differentiate between substrate features and bound biomolecules. Another drawback is the small scan size: laterally, one scan is restricted to approximately 100 μm and vertically to just 10 μm (tip limitation) (Allison et al. 2010; Ding

et al. 2008). When aiming at visualizing fast dynamic biological processes in real time, ample measurement speed is required to capture events which cannot be met by a conventional AFM with its relatively slow scan speed. In recent years, high speed (also referred to as “fast scan”) AFM (HS-AFM) was developed and shown to be able to obtain 100 x 100 pixel images in 80 ms in reasonable resolution thus generating a video of dynamic processes (Casuso et al. 2009).

Further development of AFM involves coupling with other scanning probe microscopy (SPM) techniques: scanning ion-conductance microscopy (SICM) to monitor changes in ion currents on the surface; scanning electrochemical microscope (SECM), which senses electrochemical changes (reactions); Kelvin probe force microscopy (KFM), which senses variations in the surface electric potential; scanning near field ultrasonic holography (SNFUH), which is capable of resolving structural features underneath the surface (e.g., in cells) by using acoustic waves (Allison et al. 2010).

Visualization of cellulase activity

Cellulosic substrates employed

Cellulose has regions of high order (crystalline) and of low order (amorphous), respectively. Crystalline cellulose can be organized in different allomorphs. The crystalline areas of native cellulose are usually of allomorph I, either I_{α} (triclinic unit cell) or I_{β} (monoclinic unit cell). The metastable I_{α} phase can be transformed into the thermodynamically more stable I_{β} by thermal treatment (Klemm et al. 2005). The other allomorphs, cellulose II, III and IV, are thermodynamically favored over allomorph I. Crystalline cellulose II can be produced by precipitation and regeneration of dissolved cellulose (i.e. in ionic liquids), or mercerization (alkalization). Mercerization is a relevant technical process, for example in fiber spinning and

also as a hydrolysis pretreatment method, where native cellulose is processed with aqueous sodium hydroxide to increase hydrolysis. Hence, investigations on the cellulose allomorphs I and II are of relevance (Ahola et al. 2008; Klemm et al. 2005).

The choice of the substrate is crucial when studying cellulase activity. In a study on the adsorption behavior of a *T. viride* CBH I (Cel7A), for example, the authors demonstrated that the maximal adsorption capacity of the I_β allomorph is higher; however, the hydrolysis rate was lower compared to the I_α allomorph, probably due to enzyme overcrowding (Igarashi et al. 2006). This shows that the specific activity of the enzyme as well as the adsorption depend on the type of cellulose allomorph. Not only does the substrate need to meet specific criteria depending on the biophysical method used, but it also impacts the conclusions one can draw from the entire study. EGs, for example, preferably act on amorphous cellulose and are barely active on highly crystalline cellulose such as *Valonia* cellulose. On the other hand, CBH I has a strong affinity towards such highly crystalline substrates and is even capable of digesting entire cellulose crystals (Chanzy et al. 1984). Generally speaking, differences in crystallinity, morphology and structural organization of the substrate might influence cellulase activity. When using a well-characterized substrate, substrate-related effects can be dissected from cellulase activity (Ahola et al. 2008). Another substrate feature which is often neglected – especially in cellulose films – is the thickness of the substrate. In order to study a possible three-dimensional action of the enzymes, the substrate needs to be of sufficient thickness. Otherwise, only surface phenomena can be evaluated. A thickness of 20 nm, for instance, or less is only slightly more than the length of a CBH I (12 nm) (Abuja et al. 1988).

Bacterial and algal cellulose

Bacterial cellulose (BC) and algal cellulose share the advantage that they are not intertwined with lignin – in contrast to cellulose in higher plants. BC from *Gluconacetobacter xylinus* (previously: *Acetobacter xylinum*), was used in several visualization studies (see Table I). The cellulose particles secreted by *G. xylinus* are small, homogeneous and of high purity. In

the first TEM study on cellulases, White and Brown (1981) used *G. xylinus* cellulose particles to investigate the adsorption of cellulases using colloidal gold labeling. Prior to visualization of enzyme binding, they investigated the structural organization of the BC. They found out that the cellulose microfibrils (~3 nm) self-assembled into periodically twisted ribbons (width 40–60 nm). The microfibrils are dense and to a large extent – but not completely! – made up of crystalline cellulose, mostly of allomorph I (80%). They are encased by less densely packed cellulose sheets thus forming twisted ribbons (Astley et al. 2001; Klemm et al. 2005).

Cellulose from algae, such as *Cladophora* and *Valonia* cellulose, is highly crystalline cellulose I. The algal cellulose microfibrils are made of the two allomorphs I_α and I_β with the I_α-allomorph as predominant form (Imai et al. 2003). *Valonia* cellulose can easily be prepared from vesicles of *V. macrophysa* or *V. ventricosa*. Each microfibril is an individual cellulose crystal (Chanzy and Henrissat 1983); hence, it is possible to obtain single cellulose crystals of high purity. *Valonia* cellulose is widely used as standard crystalline cellulose substrate in various studies on cellulases (Chanzy et al. 1984; Liu et al. 2011). Single *Valonia* microfibrils have monocrystalline character (Chanzy and Henrissat 1983). In their study on CBH I action on cellulose, Liu *et al.* (2011) used cross-sections of *Valonia* cell walls embedded in resin and single *Valonia* crystals (15–40 nm in height and 0.2–200 nm long) and characterized them using AFM. They showed that single *Valonia* microfibrils are hexagonal in shape (Liu et al. 2011).

Both, bacterial and algal celluloses, are well characterized and easy to obtain in a pure form without the need of delignification. Thus, they lend themselves as substrates for cellulase visualization studies. Algal celluloses offer the advantage that single crystallites with low surface roughness can be obtained easily, which makes them well suited for AFM applications. Due to their high crystallinity and their well-defined, homogeneous structure, the dynamics of enzymes acting on crystalline regions, i.e. CBH I, can be examined well. However, EGs preferably act on amorphous substrates, thus highly crystalline algal cellulose

is not suitable for investigating the action of a complete cellulase system. BC, on the other hand, is less homogeneous; it is also irregular in shape and considered semi-crystalline, which makes it a very special substrate. Additionally, it has a rough and irregular surface in contrast to the smooth surface of algal cellulose (Santa-Maria and Jeoh 2010). Hence, BC is not an ideal model substrate for studying cellulase activity.

Cellulosic substrates from higher plants

In all plants, cellulose is the major cell wall constituent providing stiffness and elasticity at the same time. It is organized in fibrils and is intertwined with hemicelluloses and lignin at varying degrees, depending on the type of plant. In cotton fibers, for example, the degree of crystallinity is lower than in hardwood (Klemm et al. 2005). They typically contain 90–95 % cellulose which is nearly exclusively of the I_{α} -allomorph (Lee et al. 2000). Visualization trials with entire sections of wood aimed at either locating different cellulolytic and hemicellulolytic enzymes in plant cells (Evans et al. 1991), locating different (hemi)cellulose types (Berg et al. 1988; Blanchette and Abad 1988), and/or imaging the degradation of entire pieces of native wood (Blanchette et al. 1989). Thus, depending on the study design, the specificity of enzymes or the presence of a certain type of (hemi)cellulose in the plant cell ultrastructure could be shown. Besides, the different degradation mechanisms of cellulase systems from different organisms could be demonstrated with this approach.

Kraft pulp is produced by treating wood with sodium hydroxide and sodium sulfide and leads to very pure cellulose fibers. During pulping, low-order areas and the I_{α} -fraction of the wood cellulose is converted to the I_{β} allomorph thus increasing order (Hult et al. 2000). The harsh conditions applied break microfibrils thereby producing new chain ends and also expose amorphous regions which were before buried within the substrate. Treatment with sulfuric acid salts is also used for pulping and yields sulfite pulp which contains a larger amount of by-products (Klemm et al. 2011). Microfibrillar cellulose can be produced from

pulp by dispersion and homogenization. The substrate thus obtained contains enzyme accessible amorphous and crystalline regions (Nieves et al. 1991).

Artificial cellulosic substrates

Visualization of cellulase action requires a homogeneous, well-defined substrate that allows for measuring all of the activities of the cellulase system. Additionally, raster scanning methods need a sufficiently flat surface. In order to achieve this, cellulose films were used to generate substrates which meet these demands. These films can be achieved by different approaches, one being regeneration from solution. To this end, cellulose is dissolved, e.g. in *N*-methylmorpholine-*N*-oxide (NMMO) or an ionic liquid and subsequently precipitated with an antisolvent. This procedure yields the allomorph cellulose II (Ahola et al. 2008). Eriksson *et al.* published two papers in 2005 dealing with the enzymatic digestion of cellulose model films (Eriksson et al. 2005a; Eriksson et al. 2005b). The spin-coated regenerated cellulose (SC) films used were approximately 30 nm thick and had a root mean square (RMS) surface roughness of 4.8 nm in dry state. Upon swelling in water for 15 hours they were 70 nm thick; unfortunately, the RMS surface roughness was not determined after swelling as the authors did not aim at visualizing the degradation but determined it by weight loss of the cellulose II films.

Native cellulose I films featuring crystalline and amorphous regions – and some residual hemicellulose – can be obtained by spin-coating of cellulose nanofibril suspensions onto silica. This yields 10 nm thick films of nanofibrillar cellulose (NFC) with a RMS surface roughness of 3 nm as determined using AFM (Ahola et al. 2008). Cellulose nanocrystal cast (NC) films also consist entirely of crystalline cellulose I. They are deposited from nanocrystal dispersions in water and subsequently freeze-dried (RMS roughness 10 nm, thickness 20–30 nm). Despite a relatively low overall RMS roughness, random aggregates reach heights up to 50 nm. The nanocrystals are formed by acid hydrolysis of Whatman filter paper with HCl at elevated temperatures. SC films can be produced by spin-coating NMMO solubilized

Avicel and subsequent drying at 80°C. The films were 30–35 nm thick, had a RMS roughness of 5.1 nm (Ahola et al. 2008). Films of non-native cellulose are for example Langmuir-Schaefer (LS) films, and SC films. LS and SC films consist of mainly amorphous cellulose with some cellulose II crystalline regions. LS films are obtained by spin-coating of trimethylsilyl cellulose (TMSC) and subsequent desilylation with HCl vapor resulting in 10–15 nm thick films with layered morphology. They showed a RMS roughness of approximately 1 nm (Ahola et al. 2008).

Milestones in visualizing cellulase activity

Until the mid-1990s, electron microscopy was the prevailing technique to visualize biological matter. Even though dynamic processes cannot be captured using electron microscopy (EM) amazing results were obtained, which aided the understanding of biological processes substantially. The first electron microscopy studies on cellulases primarily aimed at visualizing the mere presence of the enzymes and their adsorption to the substrate. Because contrast in TEM images can only be obtained through staining, the specimens were heavily pretreated – fixed and stained – and, additionally, the enzymes often were coupled to colloidal gold in order to visualize them. In 1981, White and Brown published a study where they investigated adsorption and consequence of action of cellulases on bacterial cellulose from *G. xylinus*. They incubated the substrate which was immobilized on a micronet or thin carbon grid with purified cellulase components (CBH I and EG IV), as well as the complete cellulase system. Then, they used uranyl acetate for negative staining of the samples. In this way, they also visualized sole enzyme particles on a grid. However, determining adsorbed enzymes on the substrate is not that straightforward because it is hard to distinguish enzyme particles from other possible features without any label present on the enzymes. Anyway, they observed changes in the structure of the substrate upon enzymatic degradation: whereas CBH alone did not affect the ribbon structure even after long incubation times, EG produced

visible splaying which was accounted to cleavage of intramolecular glycosidic bonds thus weakening the supramolecular ribbon structure. Incubation with solely CBH did not show any change on the substrate, which is probably due to the lack of chain ends protruding from *G. xylinus* ribbons. The complete system produced visible splaying and splitting of fibers and occasionally, small fragments broke off (White and Brown 1981). Furthermore, the negative staining with uranyl acetate that they employed might have distorted the obtained image as it might alter the interaction of proteins and the substrate.

In the following years, Chanzy and Henrissat did comprehensive work on visualization of cellulases using TEM. They investigated the action of fungal cellulases on *Valonia* and *G. xylinus* cellulose. In their first study on *Valonia* microcrystals and intact cell walls they found that hydrolysis with complete cellulase systems from *T. reesei* and *S. commune* resulted in profound morphological changes visible on the microscale – such as fragmented vesicles – and on the nanoscale as a loss of integrity of microfibrils which broke apart in single, small fibrils (width: 2–10 nm). They also found small fragments who broke off the crystals as a result of hydrolysis. They concluded that enzymatic hydrolysis is a two step process: firstly, microcrystals are broken down in smaller fibrils which are hydrolyzed completely in a second step. Sample treatment after hydrolysis involved the following subsequent procedure: washing with buffer, storage in ethanol, deposition on carbon coated grids, drying and negative staining with uranyl acetate (Chanzy and Henrissat 1983).

A subsequent study investigated the role of *T. reesei* CBH I in hydrolysis of *Valonia* microcrystals. They proved that CBH I alone can degrade crystalline cellulose completely. Concomitant HPLC analysis revealed that the major product is cellobiose, although also a minor amount of glucose was present. Direct visualization of cellulases on a grid revealed spherical particles with a diameter of 7–8 nm (Chanzy et al. 1983). This size differs from what was later determined as the size of intact CBH I in aqueous environment (length: 18 nm, radius of gyration 4.27 nm) (Abuja et al. 1988). A possible explanation is the loss of a

hydrating shell, which can lead to severe structural changes in the enzyme. Of course, also the sample preparation necessary for TEM might influence size and shape of the biomolecules, and probably also the W/Ta alloy the employed in this study for shadowing (Chanzy et al. 1983). As a next step, in order to determine if there are any preferences in CBH I adsorption, they labeled CBH I with gold particles (4–6 nm), which are electron dense markers that can easily be distinguished in TEM images. The original activity of CBH I decreased by 40 % upon labeling (Chanzy et al. 1984). This loss of activity can be explained by the mechanism of conjugation of the colloidal gold to proteins which depends on charge attraction, hydrophobic adsorption and dative binding to thiol groups. As the face of the CBD responsible for adsorption to the substrate is highly hydrophobic (Linder et al. 1995), preferred binding of colloidal gold at this site might prevent adsorption of the enzyme to its substrate. In the study of Chanzy *et al.*, colloidal gold labeling of CBH I enabled them to visualize the enzyme directly (i.e. without further shadowing or staining) using the minimum dose technique. This technique uses a lower acceleration of the electrons and a reduced beam intensity. TEM images of *Valonia* cellulose incubated with labeled CBH I showed that individual particles were bound preferably at the edges of the microfibril. This was the first visualization study to prove the preferential binding of CBH I to the (110) face of the crystal. In some cases, however, aggregates which do not show the preferred binding had been formed (Chanzy et al. 1984). Now, formation of aggregates could be attributed to the protein as well as to the colloidal gold, as one particle might have conjugated with more than one enzyme. Additionally, unproductively bound or generally inactivated CBH I cannot be distinguished. It therefore remains elusive whether or not the aggregates consist of active or inactive enzymes and if aggregation is in consequence of labeling or an intrinsic feature of the enzymes themselves. However, as the amount of clustering was decreasing slightly with lower enzyme loadings (Chanzy et al. 1984), aggregation might as well have been a consequence of high enzyme/substrate ratios also occurring naturally.

In a subsequent study, Chanzy and Henrissat (1985) investigated the synergistic behavior of endo- and exo-enzymes using an EG II and a CBH II and *Valonia* cellulose. This time, they waived gold-labeling and focused on the effects of enzymatic action, using negative staining with uranyl acetate to this end. The results confirmed the synergism of these enzymes: incubation with CBH II alone showed that attack occurred only at one of the two tips of a microfibril, whereas incubation with both enzymes also showed degradation at rare non-perfect regions in the microfibril (“kinks”), where EG II was able to attack and generate new non-reducing chain ends. In comparison with their previous study on CBH I (Chanzy et al. 1984), they showed that CBH I and CBH II attack crystalline cellulose at distinctively different sites of the crystal. Additionally, this study showed that *Valonia* cellulose microcrystals consist of parallel packed chains. Also, specific action of CBH II (Cel6A) from one end of the microcrystal – the non-reducing end – was demonstrated nicely for the first time (Chanzy and Henrissat 1985). More than a decade later, J. Sugiyama’s group published a TEM study of the unidirectional activity of a CBH I (Cel7A). By selectively silver-staining the reducing ends of *Valonia* cellulose they demonstrated that CBH I in fact attacks at the reducing end thereby producing characteristic sharp and “pointed” edges (Imai et al. 1998) as was observed previously for CBH II from the opposite end of the crystal (Chanzy and Henrissat 1985). They also claimed evidence for processivity of the exo-enzyme (Imai et al. 1998), however, as in situ monitoring of enzyme action is not possible with TEM, this could not be verified until more than 10 years later by using AFM (Igarashi et al. 2009).

Today, colloidal gold labeling is a widely employed standard technique in protein biochemistry as biomacromolecules can easily be coupled to the colloidal gold particles (Geoghegan and Ackerman 1977). Nowadays, colloidal gold particles are available commercially in different sizes, from an average diameter of approximately 1 up to 250 nm. There are also several colloidal gold-labeled proteins available for purchase, for example antibodies. Silver enhancement is often used in TEM studies with colloidal gold to increase

contrast (Humbel et al. 1995). Being a convenient and reliable method, gold-labeling was used in several studies from the mid-eighties onwards. It was employed, for example, to locate different (hemi)cellulose types (Berg et al. 1988; Blanchette and Abad 1988). Also, the degradation of wood with white rot fungi was investigated this way (Blanchette et al. 1989). Specificity of binding of different cellulolytic and hemicellulolytic enzymes in plant cells was also determined with this method (Evans et al. 1991).

In their 1991 publication, Nieves *et al.* expanded the colloidal gold coupling to the so-called “immuno-EM” method. They coupled gold particles of 10 and 15 nm to two different monoclonal antibodies (mAbs) which were specific for *T. reesei* CBH I and EG I, respectively. With this elaborate design they could easily distinguish between the two enzymes according to the size of the gold particle. As a substrate, microfibrillar Aspen cellulose was produced by dilute sulfuric acid treatment, subsequent alkali extraction and a final homogenization step in deionized water. From the resulting suspension, they then prepared copper grids for TEM with adsorbed Aspen cellulose and treated them with the enzymes. Next, they incubated the samples with the immunogold solution. After washing with water and drying, they performed negative staining with uranyl acetate. In their experiments, they also observed preferred binding of CBH I to the ends of fibrillar substrate structures. As a result of the acid treatment, which breaks up the supramolecular structure of the microcrystals thus generating new chain ends, plenty binding sites for CBH I were available. Again, frequent aggregate formation of CBH I was observed, even though apparently enough unoccupied sites were available. EG I also aggregated to some extent; however, whereas CBH I exhibited binding to amorphous as well as to crystalline regions, EG I exclusively bound to amorphous cellulose, which the authors demonstrated nicely in different approaches with either the single or the combined activities. In the combined dual labeling experiments they also found that CBH I was apparently bound to split microfibers above the plane and sometimes exhibited apparent association with EG I (Nieves et al. 1991).

These might represent new chain ends generated by EG I. When the two enzymes appeared associated, the stoichiometry was interesting: CBH I always outnumbered EG I. Not only did this corroborate the theory that EG I initially attacks the substrate (Nieves et al. 1991), it was also a first hint on binding site competition between these two cellulase activities. However, one needs to consider that the actual cellulose morphology was not determined at the EG I binding sites, hence it is not confirmed that these are really amorphous regions.

In 1993, H. Chanzy and B. Henrissat tied in to their previous work: using immuno-EM, they investigated the adsorption of a *Cellulomonas fimi* EG (CenA) and its sole CBD on BC and *Valonia* cellulose in cooperation with J. Sugiyama and R. A. J. Warren's group (Gilkes et al. 1993). The catalytic domain of this two-domain EG was shown to be catalytically active even without the CBD, however, it did not show affinity for cellulose. This was the first time that binding of a CBD alone was visualized. For the immuno-EM they employed antibody gold conjugates with a particle size of 5 nm. In contrast to *T. reesei* EG, which showed exclusive preference for supposedly amorphous regions (Nieves et al. 1991), CenA exhibited uniform binding on *Valonia* cellulose and BC (Gilkes et al. 1993). Its isolated CBD adsorbs similarly, but with a slight preference for the edges of a particular crystal face, comparable to what was reported from *T. reesei* CBH I (Chanzy et al. 1984). It remains unclear if this preference is specific targeting or an artifact due to disruption of the microfibrils by TEM sample pretreatment or probably even by the CBDs themselves ("amorphogenesis").

Immuno-EM is advantageous in a way that the cellulases themselves do not need to be labeled with gold particles which might alter their biological behavior, i.e. adsorption, penetration and/or degradation. The complex experimental setup of immuno-EM, however, has also severe disadvantages: for instance, non-specific interaction of antibodies was observed to varying extents, but to a minimum of 10% (Gilkes et al. 1993). Additionally, antibodies are huge proteins, and sometimes even a second protein have been used for the

reaction in some studies. In the end, the cellulase would be coupled to one or two proteins and a colloidal gold particle of 5–15 nm which can significantly influence the accurateness of the investigation. Besides, in case cellulases penetrate into the substrate structure, the antibodies might be too big and interactions in confined areas might also be unfavorable for them to gain access to small pores, cracks or channels. In 2009, Donohoe et al. published a study on evaluating the effectiveness of corn stover pretreatment for enzymatic hydrolysis with *T. reesei* CBH I (Cel7A) and an EG I (Cel7B) employing immuno-EM. They showed that pretreatment enabled cellulases to penetrate the highly recalcitrant sclerenchyma cells more easily (Donohoe et al. 2009). Immuno-EM might be more useful in comparative approaches like this, however, no quantitative results can be obtained due to the limitations mentioned.

The 1996 publication of Lee *et al.* was a milestone in visualization of cellulase activity and at the same time the starting point for AFM investigations on this topic. In this publication, they showed the results of the effect of CBH I-treatment on cotton fibers and compared it to an inactive form of the enzyme (Lee et al. 1996). This was the first AFM study published on visualization of enzymatic cellulose degradation. It was also the first study visualizing cellulase action on cotton fibers, which would probably be a matter too delicate for TEM investigation. In a following study they also used tapping-mode AFM to examine how a *T. reesei* CBH I, EG II and a thermostable cellulase from *Thermotoga maritima* cellulase, which is unable to bind to cellulose, attack cotton fibers (Lee et al. 2000). To this end they did sequential and simultaneous incubation and showed how it affected the microfibril surface: whereas sole incubation with CBH I caused indentive paths and pits, sole incubation with EG II obviously just “smoothed” the surface. Upon incubation of the cotton fibers with both enzymes, now individual microfibrils could not be seen anymore, probably due to complete disruption of the fibers by the synergistic action of the cellulases. Using hexachloropalladate-inactivated CBH I they found formation of distinct holes in the surface

of the fibers which they accounted to a supposed penetration by the CBDs of the catalytically inactive enzyme. Their sample preparation involved buffer washing and vacuum drying, which they claimed did not alter the surface (Lee et al. 2000). After these first AFM imaging studies by Lee et al. several followed; for the next decade, they primarily focused on observable substrate changes rather than on direct visualization of single enzymes.

In 2004, Nigmatullin *et al.* studied the binding of a *T. reesei* CBD with regenerated cellulose from a cellulose acetate solution. The CBDs formed aggregates of varying size (up to 80 nm) which were found all over the surface thus increasing the RMS roughness (from 0.4 to 2.6 nm). Aggregation is possibly a result of the hydrophobic character of one face of the CBDs. However, they did not describe any surface disruption in their study. In a study published by Ahola *et al.* in 2008, the authors compared the degradation of different types of model films to each other (see also 3.1.3) by combining AFM imaging and quartz crystal microbalance with dissipation (QCM-D) measurements (Ahola et al. 2008). With this setting, they found out that higher enzyme/substrate ratios showed markedly faster adsorption until the surface area was saturated. The thin films were quickly degraded by the optimized complete *T. reesei* cellulase system used (CelluclastTM) under the chosen experimental settings. AFM imaging was used mainly to follow the morphological changes in the substrate. NFC films apparently have a high and well accessible surface area due to their fibrillar structure as opposed by SC or LS films. The NFC film is nearly completely digested with only some small, probably highly crystalline, remainders. LS, SC, and NC film degradation proceeded significantly slower, especially in the highly crystalline NC film where undigested crystals remained after hydrolysis. The LS film, which consists to a large extent of amorphous cellulose but also of crystalline cellulose II, was not degraded completely, whereas the SC film – which is of similar structure – was (Ahola et al. 2008). The enzymatic digestion of Kraft pulp by complete and single cellulases from *Humicola insolens* and *Aspergillus* spp. was followed using AFM imaging. This study is interesting

because they claimed to have visualized single EG V particles on the heterogeneous pulp substrate (Liu et al. 2009).

The next milestone in the visualization of cellulase action was a study published in 2009 by Igarashi *et al.* Not only was this the first study employing high-speed AFM (HS-AFM) for that purpose, but it was also the first one that actually observed cellulase action in a natural environment. In this elaborate work, they were able to determine the linear processivity of CBH I on *Valonia* cellulose with a rate of 3.5 nm/s (Igarashi et al. 2009). In a similar setting, Liu et al (2011) showed specific binding of CBH I to the hydrophobic faces of *Valonia* cellulose. They also observed what was probably a CBH I molecule sliding along a microfibril (Liu et al. 2011).

Recently, combination of AFM with other biophysical methods was applied to study cellulase action. In a sophisticated setting combining “conventional” AFM and confocal fluorescence microscopy (CFM), Santa Maria and Jeoh (2010) showed that CBH I untwists bacterial cellulose ribbons, probably by attacking the crystalline moieties of the ribbons thus decreasing order. However, although this is an interesting observation, these ribbons are an exclusive feature of BC and hence this is not necessarily transferable onto disintegration of cellulose from other sources (Santa-Maria and Jeoh 2010).

Conclusion and outlook

Today we have a broad basis of knowledge about the substrate as well as the enzymes available, much of which was found or validated using visualization of the process. Electron microscopic methods used in the beginning lack the ability to in situ visualize biological processes and the comprehensive and harsh sample treatment required is not suited to this end as it might damage the substrate, the enzymes or features generated through enzymatic action. Also, colloidal gold labeling or immuno-EM are not suited due to the size of the labels

and hence possible restriction of access to substrate cavities or otherwise altering enzyme behavior. AFM, especially in a liquid cell and/or as HS-AFM, on the other hand, is the current state of the art method to address the current questions in cellulose hydrolysis in situ and under “natural” conditions, i.e. aqueous environment, controlled (ambient) temperature and pressure, no sample modification. AFM renders visualization of the enzymatic breakdown of cellulose possible without destroying substrate or enzymes. Also, measurements in liquids are possible when using a so-called liquid cell. Thus, AFM is currently the method of choice for in situ imaging of (single) celluloses and their action on their substrate. Combined with other biophysical methods, such as QCM-D or confocal spectroscopy, great progress and interesting insights have been achieved recently. Another important issue which needs to be addressed is the question of a standard model substrate for such studies. As the morphology and deep structure of cellulose greatly affects enzyme dynamics – and thus what we see – this is a question of particular interest.

Much of our knowledge about the mechanism of enzymatic cellulose degradation was detected in or validated through elaborate visualization studies. From the first TEM studies 30 years ago, the biophysical toolbox was expanded greatly providing us with a wide variety of visualization and supplementary methods to elucidate the structural dynamics of enzymatic cellulose degradation. This has prepared the ground for many exciting studies on cellulase action to be done in the next years.

References

- Abuja PM, Schmuck M, Pilz I, Tomme P, Claeysens M, Esterbauer H. 1988. Structural and functional domains of cellobiohydrolase I from *Trichoderma reesei*. Eur. Biophys. J. 15:339-342.
- Ahola S, Turon X, Osterberg M, Laine J, Rojas OJ. 2008. Enzymatic hydrolysis of native cellulose nanofibrils and other cellulose model films: effect of surface structure. Langmuir 24(20):11592-9.
- Allison DP, Mortensen NP, Sullivan CJ, Doktycz MJ. 2010. Atomic force microscopy of biological samples. Wiley Interdiscip Rev Nanomed Nanobiotechnol 2(6):618-34.

- Arantes V, Saddler JN. 2010. Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. *Biotechnol. Biofuels* 3:4.
- Astley OM, Chanliaud E, Donald AM, Gidley MJ. 2001. Structure of *Acetobacter* cellulose composites in the hydrated state. *Int J Biol Macromol* 29(3):193-202.
- Bansal P, Hall M, Realf MJ, Lee JH, Bommarius AS. 2009. Modeling cellulase kinetics on lignocellulosic substrates. *Biotechnol. Adv.* 27(6):833-48.
- Berg RH, Erdos GW, Gritzali M, Brown RDJ. 1988. Enzyme-gold affinity labelling of cellulose. *J. Electron Microsc. Tech.* 8:371-379.
- Blanchette RA, Abad AR. 1988. Ultrastructural localization of hemicellulose in birch wood (*Betula papyrifera*) decayed by brown and white rot fungi. *Holzforschung* 42:393-398.
- Blanchette RA, Abad AR, Cease KR, Lovrien RE, Leathers TD. 1989. Colloidal Gold Cytochemistry of Endo-1,4-beta-Glucanase, 1,4-beta-D-Glucan Cellobiohydrolase, and Endo-1,4-beta-Xylanase: Ultrastructure of Sound and Decayed Birch Wood. *Appl. Environ. Microbiol.* 55(9):2293-2301.
- Boisset C, Chanzy H, Henrissat B, Lamed R, Shoham Y, Bayer EA. 1999. Digestion of crystalline cellulose substrates by the *Clostridium thermocellum* cellulosome: structural and morphological aspects. *Biochem J* 340 (Pt 3):829-35.
- Boisset C, Fraschini C, Schulein M, Henrissat B, Chanzy H. 2000. Imaging the enzymatic digestion of bacterial cellulose ribbons reveals the endo character of the cellobiohydrolase Cel6A from *Humicola insolens* and its mode of synergy with cellobiohydrolase Cel7A. *Appl Environ Microbiol* 66(4):1444-52.
- Brecknock S, Dibbayawan TP, Vesik M, Vesik PA, Faulkner C, Barton DA, Overall RL. 2011. High resolution scanning electron microscopy of plasmodesmata. *Planta*.
- Casuso I, Koder N, Le Grimellec C, Ando T, Scheuring S. 2009. Contact-mode high-resolution high-speed atomic force microscopy movies of the purple membrane. *Biophys J* 97(5):1354-61.
- Chanzy H, Henrissat B. 1983. Electron microscopy study of the enzymic hydrolysis of *Valonia* cellulose. *Carbohydr Polym* 3:161-173.
- Chanzy H, Henrissat B. 1985. Unidirectional degradation of *Valonia* cellulose microcrystals subjected to cellulase action. *FEBS Lett* 184(2):285-288.
- Chanzy H, Henrissat B, Vuong H. 1984. Colloidal gold labelling of 1,4-β-D-glucan cellobiohydrolase adsorbed on cellulose substrates. *FEBS Lett.* 172(2):193-197.
- Chanzy H, Henrissat B, Vuong H, Schulein M. 1983. The action of 1,4-β-D-glucan cellobiohydrolase on *Valonia* cellulose microcrystals. An electron microscopy study. *FEBS Lett.* 153(1):113-118.
- Collins SP, Pope RK, Scheetz RW, Ray RI, Wagner PA, Little BJ. 1993. Advantages of environmental scanning electron microscopy in studies of microorganisms. *Microsc Res Tech* 25(5-6):398-405.
- Crampton N, Yokokawa M, Dryden DT, Edwardson JM, Rao DN, Takeyasu K, Yoshimura SH, Henderson RM. 2007. Fast-scan atomic force microscopy reveals that the type III restriction enzyme EcoP15I is capable of DNA translocation and looping. *Proc Natl Acad Sci U S A* 104(31):12755-60.
- de Jonge N, Peckys DB, Kremers GJ, Piston DW. 2009. Electron microscopy of whole cells in liquid with nanometer resolution. *Proc Natl Acad Sci U S A* 106(7):2159-64.
- Denk W, Horstmann H. 2004. Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. *PLoS Biol* 2(11):e329.
- Ding SY, Xu Q, Crowley M, Zeng Y, Nimlos M, Lamed R, Bayer EA, Himmel ME. 2008. A biophysical perspective on the cellulosome: new opportunities for biomass conversion. *Curr Opin Biotechnol* 19(3):218-27.
- Donohoe BS, Selig MJ, Viamajala S, Vinzant TB, Adney WS, Himmel ME. 2009. Detecting cellulase penetration into corn stover cell walls by immuno-electron microscopy. *Biotechnol Bioeng* 103(3):480-9.
- Eriksson J, Malmsten M, Tiberg F, Callisen TH, Damhus T, Johansen KS. 2005a. Enzymatic degradation of model cellulose films. *J. Colloid Interface Sci.* 284(1):99-106.

- Eriksson J, Malmsten M, Tiberg F, Callisen TH, Damhus T, Johansen KS. 2005b. Model cellulose films exposed to H. insolens glucoside hydrolase family 45 endo-cellulase--the effect of the carbohydrate-binding module. *J. Colloid. Interface. Sci.* 285(1):94-9.
- Evans CS, Gallagher IM, Atkey PT, Wood DA. 1991. Localisation of degradative enzymes in white-rot decay of lignocellulose. *Biodegradation* 2:93-106.
- Geoghegan WD, Ackerman GA. 1977. Adsorption of horseradish peroxidase, ovomucoid and anti-immunoglobulin to colloidal gold for the indirect detection of concanavalin A, wheat germ agglutinin and goat anti-human immunoglobulin G on cell surfaces at the electron microscopic level: a new method, theory and application. *J. Histochem. Cytochem.* 25(11):1187-200.
- Gilkes NR, Kilburn DG, Miller RC, Jr., Warren RA, Sugiyama J, Chanzy H, Henrissat B. 1993. Visualization of the adsorption of a bacterial endo-beta-1,4-glucanase and its isolated cellulose-binding domain to crystalline cellulose. *Int J Biol Macromol* 15(6):347-51.
- Gilkes NR, Warren RA, Miller RC, Jr., Kilburn DG. 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(21):10401-7.
- Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD. 2007. Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* 315(5813):804-7.
- Hult EL, Larsson PT, Iversen T. 2000. A comparative CP/MAS ¹³C-NMR study of cellulose structure in spruce wood and kraft pulp. *Cellulose* 7:35-55.
- Humbel BM, Sibon OC, Stierhof YD, Schwarz H. 1995. Ultra-small gold particles and silver enhancement as a detection system in immunolabeling and in situ hybridization experiments. *J Histochem Cytochem* 43(7):735-7.
- Igarashi K, Koivula A, Wada M, Kimura S, Penttila M, Samejima M. 2009. High speed atomic force microscopy visualizes processive movement of *Trichoderma reesei* cellobiohydrolase I on crystalline cellulose. *J. Biol. Chem.* 284(52):36186-90.
- Igarashi K, Wada M, Hori R, Samejima M. 2006. Surface density of cellobiohydrolase on crystalline celluloses. A critical parameter to evaluate enzymatic kinetics at a solid-liquid interface. *FEBS J* 273(13):2869-78.
- Imai T, Boisset C, Samejima M, Igarashi K, Sugiyama J. 1998. Unidirectional processive action of cellobiohydrolase Cel7A on *Valonia* cellulose microcrystals. *FEBS Lett* 432(3):113-6.
- Imai T, Putaux J-L, Sugiyama J. 2003. Geometric phase analysis of lattice images from algal cellulose microfibrils. *Polymer* 44:1871-1879.
- Klemm D, Heublein B, Fink HP, Bohn A. 2005. Cellulose: fascinating biopolymer and sustainable raw material. *Angew Chem Int Ed Engl* 44(22):3358-93.
- Klemm D, Kramer F, Moritz S, Lindstrom T, Ankerfors M, Gray D, Dorris A. 2011. Nanocelluloses: a new family of nature-based materials. *Angew Chem Int Ed Engl* 50(24):5438-66.
- Kodera N, Yamamoto D, Ishikawa R, Ando T. 2010. Video imaging of walking myosin V by high-speed atomic force microscopy. *Nature* 468(7320):72-6.
- Lee I, Evans BR, Lane LM, Woodward J. 1996. Substrate-enzyme interactions in cellulase systems. *Biores. Technol.* 58:163-169.
- Lee I, Evans BR, Woodward J. 2000. The mechanism of cellulase action on cotton fibers: evidence from atomic force microscopy. *Ultramicroscopy* 82(1-4):213-21.
- Lehtiö J, Sugiyama J, Gustavsson M, Fransson L, Linder M, Teeri TT. 2003. The binding specificity and affinity determinants of family 1 and family 3 cellulose binding modules. *Proc. Natl. Acad. Sci. U.S.A.* 100(2):484-9.
- Li Z, Chu LQ, Sweedler JV, Bohn PW. 2010. Spatial correlation of confocal Raman scattering and secondary ion mass spectrometric molecular images of lignocellulosic materials. *Anal. Chem.* 82(7):2608-11.
- Linder M, Mattinen ML, Kontteli M, Lindeberg G, Stahlberg J, Drakenberg T, Reinikainen T, Pettersson G, Annala A. 1995. Identification of functionally important amino acids in the cellulose-binding domain of *Trichoderma reesei* cellobiohydrolase I. *Protein. Sci.* 4(6):1056-64.

- Liu H, Fu S, Zhu JY, Li H, Zhan H. 2009. Visualization of enzymatic hydrolysis of cellulose using AFM phase imaging. *Enzyme and Microbial Technology* 45:274-281.
- Liu Y-S, Luo Y, Baker JO, Zeng Y, Himmel ME, Ding S-Y. 2010. Single molecule study of cellulase hydrolysis of crystalline cellulose. Society of Photo-Optical Instrumentation Engineers (SPIE) Photonics West 2010. San Francisco, CA, USA.
- Liu YS, Baker JO, Zeng Y, Himmel ME, Haas T, Ding SY. 2011. Cellobiohydrolase hydrolyzes crystalline cellulose on hydrophobic faces. *J. Biol. Chem.* 286(13):11195-11201.
- Lynd LR, Laser MS, Bransby D, Dale BE, Davison B, Hamilton R, Himmel M, Keller M, McMillan JD, Sheehan J and others. 2008. How biotech can transform biofuels. *Nat. Biotechnol.* 26(2):169-72.
- Mansfield SD, Mooney C, Saddler JN. 1999. Substrate and Enzyme Characteristics that Limit Cellulose Hydrolysis. *Biotechnol. Prog.* 15(5):804-816.
- Nieves RA, Ellis RP, Todd RJ, Johnson TJ, Grohmann K, Himmel ME. 1991. Visualization of *Trichoderma reesei* Cellobiohydrolase I and Endoglucanase I on Aspen Cellulose by Using Monoclonal Antibody-Colloidal Gold Conjugates. *Appl Environ Microbiol* 57(11):3163-70.
- Nigmatullin R, Lovitt R, Wright C, Linder M, Nakari-Setälä T, Gama M. 2004. Atomic force microscopy study of cellulose surface interaction controlled by cellulose binding domains. *Colloids Surf. B Biointerfaces* 35(2):125-35.
- Peckys DB, Mazur P, Gould KL, de Jonge N. 2011. Fully hydrated yeast cells imaged with electron microscopy. *Biophys J* 100(10):2522-9.
- Penttilä PA, Várnai A, Leppänen K, Peura M, Kallonen A, Jääskeläinen P, Lucenius J, Ruokolainen J, Siika-aho M, Viikari L and others. 2010. Changes in submicrometer structure of enzymatically hydrolyzed microcrystalline cellulose. *Biomacromolecules* 11(4):1111-7.
- Pinto R, Amaral AL, Ferreira EC, Mota M, Vilanova M, Ruel K, Gama M. 2008. Quantification of the CBD-FITC conjugates surface coating on cellulose fibres. *BMC Biotechnol.* 8:1.
- Quirk A, Lipkowski J, Vandenende C, Cockburn D, Clarke AJ, Dutcher JR, Roscoe SG. 2010. Direct visualization of the enzymatic digestion of a single fiber of native cellulose in an aqueous environment by atomic force microscopy. *Langmuir* 26(7):5007-13.
- Sander B, Golas MM. 2011. Visualization of bionanostructures using transmission electron microscopical techniques. *Microsc Res Tech* 74(7):642-63.
- Santa-Maria M, Jeoh T. 2010. Molecular-scale investigations of cellulose microstructure during enzymatic hydrolysis. *Biomacromolecules* 11(8):2000-7.
- Singh S, Simmons BA, Vogel KP. 2009. Visualization of biomass solubilization and cellulose regeneration during ionic liquid pretreatment of switchgrass. *Biotechnol Bioeng* 104(1):68-75.
- Solomon BD, Barnes JR, Halvorsen KE. 2007. Grain and cellulosic ethanol: History, economics, and energy policy. *Biomass and Bioenergy* 31:416 - 425.
- Tawil G, Jamme F, Refregiers M, Vikso-Nielsen A, Colonna P, Buleon A. 2011. In situ tracking of enzymatic breakdown of starch granules by synchrotron UV fluorescence microscopy. *Anal Chem* 83(3):989-93.
- Tébeka IR, Silva AG, Petri DF. 2009. Hydrolytic activity of free and immobilized cellulase. *Langmuir* 25(3):1582-7.
- Valjamae P, Sild V, Pettersson G, Johansson G. 1998. The initial kinetics of hydrolysis by cellobiohydrolases I and II is consistent with a cellulose surface-erosion model. *Eur. J. Biochem.* 253(2):469-75.
- White AR, Brown RM. 1981. Enzymatic hydrolysis of cellulose: Visual characterization of the process. *Proc. Natl. Acad. Sci. U.S.A.* 78(2):1047-51.
- Zhao L, Bulhassan A, Yang G, Ji HF, Xi J. 2010. Real-time detection of the morphological change in cellulose by a nanomechanical sensor. *Biotechnol. Bioeng.* 107(1):190-4.

Tables

Table I. Overview: visualization studies. *A. niger* – *Aspergillus niger*; *A. pullulans* – *Aureobasidium pullulans*; BMCC – bacterial microcrystalline cellulose; CD – catalytic domain; *C. fimi* – *Cellulomonas fimi*; CRM – confocal Raman microscopy; EX – endoxylanase; FITC – fluorescein isothiocyanate; *G. xylinus* – *Gluconacetobacter xylinus*; *H. insolens* - *Humicola insolens*; LS – Langmuir-Schaefer; NC - cellulose nanocrystal; QCM-D Quartz crystal microbalance with dissipation; SC – spin-coated regenerated cellulose; *S. commune* – *Schizophyllum commune*; SIMS – secondary ion mass spectrometry; *T. maritima* – *Thermotoga maritima*; *T. reesei* – *Trichoderma reesei*; TIRF-M – total internal reflection fluorescence microscopy; *V. macrophysa* – *Valonia macrophysa*; *V. ventricosa* – *Valonia ventricosa*;

Publication	Visualization methodology	Substrate	Enzyme(s) investigated	Additional techniques
White and Brown 1981	TEM	bacterial cellulose (<i>G. xylinus</i>)	<i>T. reesei</i> (QM9414) CBH I, EG IV, complete system	
Chanzy and Henrissat 1983	TEM	<i>V. macrophysa</i> cellulose	<i>T. reesei</i> complete system, <i>S. commune</i> complete system	
Chanzy et al. 1983	TEM	<i>V. macrophysa</i> cellulose	<i>T. reesei</i> CBH I (from Celluclast™)	HPLC analysis of released sugars
Chanzy et al. 1984	TEM / Au labeling	bacterial & algal cellulose (<i>G. xylinus</i> , <i>V. macrophysa</i>)	<i>T. reesei</i> CBH I (from Celluclast™)	
Chanzy and Henrissat 1985	TEM	<i>V. macrophysa</i> cellulose	<i>T. reesei</i> CBH II and EG II (from Celluclast™)	
Blanchette et al. 1989	TEM / Au labeling	birch wood (<i>B. papyrifera</i>)	<i>T. reesei</i> (QM9414) CBH I & EG II; <i>A. pullulans</i> EX	
Nieves et al. 1991	TEM / immuno-Au labeling (immuno-EM)	microfibrillar Aspen wood cellulose	<i>T. reesei</i> CBH I, EG I	

Publication	Visualization methodology	Substrate	Enzyme(s) investigated	Additional techniques
Gilkes et al. 1993	TEM / immuno-Au labeling (immuno-EM)	bacterial & algal cellulose (<i>G. xylinus</i> , <i>V. macrophysa</i>)	<i>C. fimi</i> CenA (EG), Cex (CBH) and their sole CBDs	protein adsorption measurement, BC flocculation test
Imai et al. 1998	TEM	<i>V. ventricosa</i> cellulose	<i>T. viride</i> CBH I (Cel7A)	silver staining of reducing ends
Boisset et al. 1999	TEM	<i>V. ventricosa</i> and bacterial cellulose	<i>C. thermocellum</i> cellulosome	FTIR, XRD
Boisset et al. 2000	TEM	bacterial cellulose	CBH I (Cel7A), CBH II (Cel6a) from <i>H. insolens</i>	
Lee et al. 2000	AFM	cotton fibers	<i>T. reesei</i> CBH I, EG II, <i>T. maritima</i> cellulase	
Lehtiö et al. 2003	TEM	BMCC, <i>V. ventriculosa</i> crystals	<i>T. reesei</i> and <i>C. thermocellum</i> CBD fusion proteins	CBDs: specific binding
Nigmatullin et al. 2004	AFM	regenerated cellulose from cellulose acetate	<i>T. reesei</i> CBD, double-CBD fusion protein	scanning probe microscopy-colloid probe
Ahola et al. 2008	AFM	cellulose model films (LS, SC, NC)	complete system (Celluclast TM NS50013, Novozyme)	QCM-D
Pinto et al. 2008	Confocal microscopy + immuno-TEM	Whatman CFI I, amorphous cellulose fibers, Sigmacell 20 fibers	CBD-FITC conjugates (prepared from Celluclast TM)	
Tébéka et al. 2009	AFM	Silica wafers	<i>T. reesei</i> complete system (ATCC 26921)	
Donohoe et al. 2009	immuno-EM	pretreated and native corn stover	<i>T. reesei</i> CBH I (Cel7A), EG I (Cel7B)	Compositional analysis of native and pretreated corn stover
Singh et al. 2009	Laser confocal microscopy, SEM, XRD	switchgrass (<i>Panicum virgatum</i>)	<i>T. reesei</i> complete system	
Liu et al. 2009	AFM	reed Kraft pulp	complete system, EG V w/ CBD, EG I w/o CBD (all: Novozyme)	

Publication	Visualization methodology	Substrate	Enzyme(s) investigated	Additional techniques
Igarashi et al. 2009	HS-AFM	cellulose from <i>Cladophora sp.</i>	<i>T. reesei</i> CBH I (Cel7A) and its sole CBD; mutants: E212Q, W40A	
Zhao et al. 2010	AFM	regenerated cellulose II	--	
Penttilä et al. 2010	X-ray microtomography; cryo-TEM;	Avicel	<i>T. reesei</i> complete system (Celluclast™ 10 FPU/g _{substrate}), <i>A. niger</i> BGL (Novozyme188)	WAXS, SAXS
Quirk et al. 2010	AFM	treated <i>G. xylinus</i> cellulose (MeOH dispersion)	<i>C. fimi</i> CenA	
Liu et al. 2010	HS-AFM + TIRF-M	<i>V. ventricosa</i> crystals, dispersed in H ₂ O	<i>T. reesei</i> CBH I, its CBD-GFP fusion	
Santa-Maria and Jeoh 2010	AFM	minimally processed cellulose microfibrils from <i>G. xylinus</i> and <i>Cladophora sp.</i>	<i>T. reesei</i> CBH I (Cel7A) labeled with Alexa Fluor 488	combined with confocal fluorescence microscopy
Li et al. 2010	SIMS, CRM	<i>Miscanthus x giganteus</i>	-	combined with LDI-MS
Liu et al. 2011	HS-AFM	<i>V. ventricosa</i> crystals	<i>T. reesei</i> CBH I	

Figures

Figure 1. Model of enzymatic cellulose disruption. A: Current model of enzymatic cellulose degradation and the unresolved mechanisms of amorphogenesis and hydrolysis of insoluble cellulose (Arantes and Saddler 2010). B: Proposed stepwise mechanism of a CBH: 1 – adsorption, 2 and 3 – productive binding to a chain end, 4 – bond hydrolysis. The mechanism of EGs is possibly similar with the exception that they act randomly within the chain (Bansal et al. 2009).

Figure 2. Milestones of early TEM studies. A: A. R. White and R. M. Brown, Jr., 1981 – the first visualization study on cellulase action. a – BC ribbons after 30 minutes of incubation with *T. reesei* CBH I (scale bar: 100 nm); b – BC ribbons after 30 minutes of incubation with *T. reesei* EG IV and some ribbon splaying (scale bar: 100 nm); c – fibril bundle with encircled possible cellulases (scale bar: 50 nm). B: H. Chanzy and B. Henrissat, 1985. a – typical *V. macrophysa* microcrystal with “kinks” highlighted (scale bar: 250 nm); b – the same as in a but after the incubation with CBH II for 16 h; R indicates the reducing, NR the non-reducing end (scale bar: 250 nm). C: H. Chanzy and B. Henrissat, 1985 – mode-of-action-scheme of exo-cellulases leading to thinned fibril ends. a – attack of the exo-enzyme at a non-reducing (NR) chain end located at the surface; b – exposure of new non-reducing chain ends by degradation of the surface layer; c – the non-reducing end is thus sharpened. D: R. A. Nieves *et al.*, 1991 – preference of CBH I [upper arrow, small particles (10 nm)] to crystalline regions and of EG I [lower arrow, bigger particles (15 nm)] to amorphous cellulose (scale bar: 200 nm).

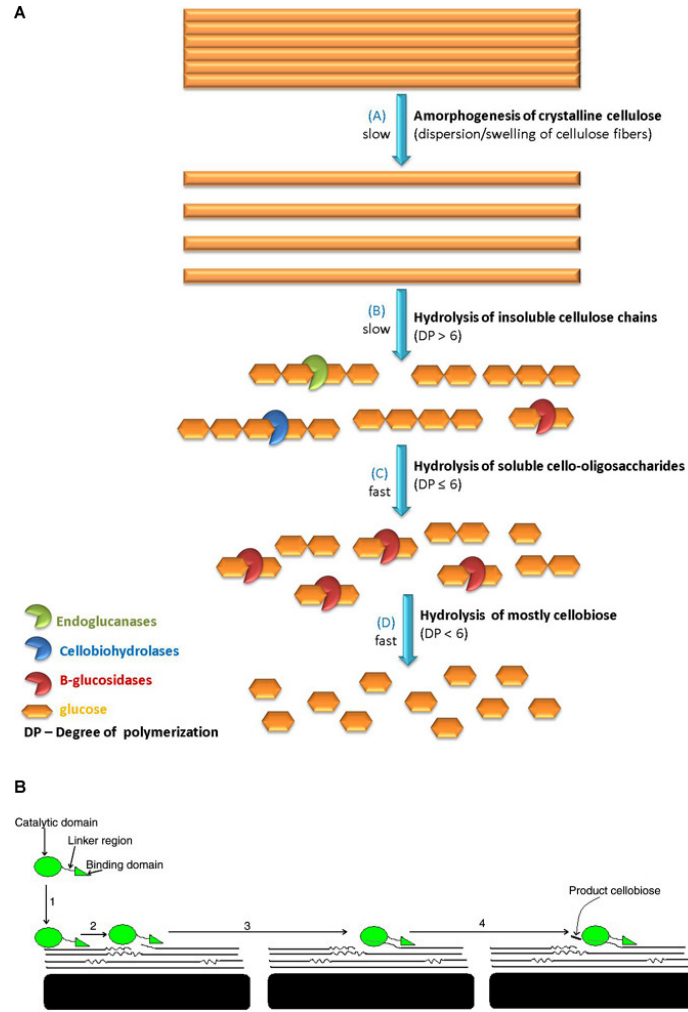


Figure 1.

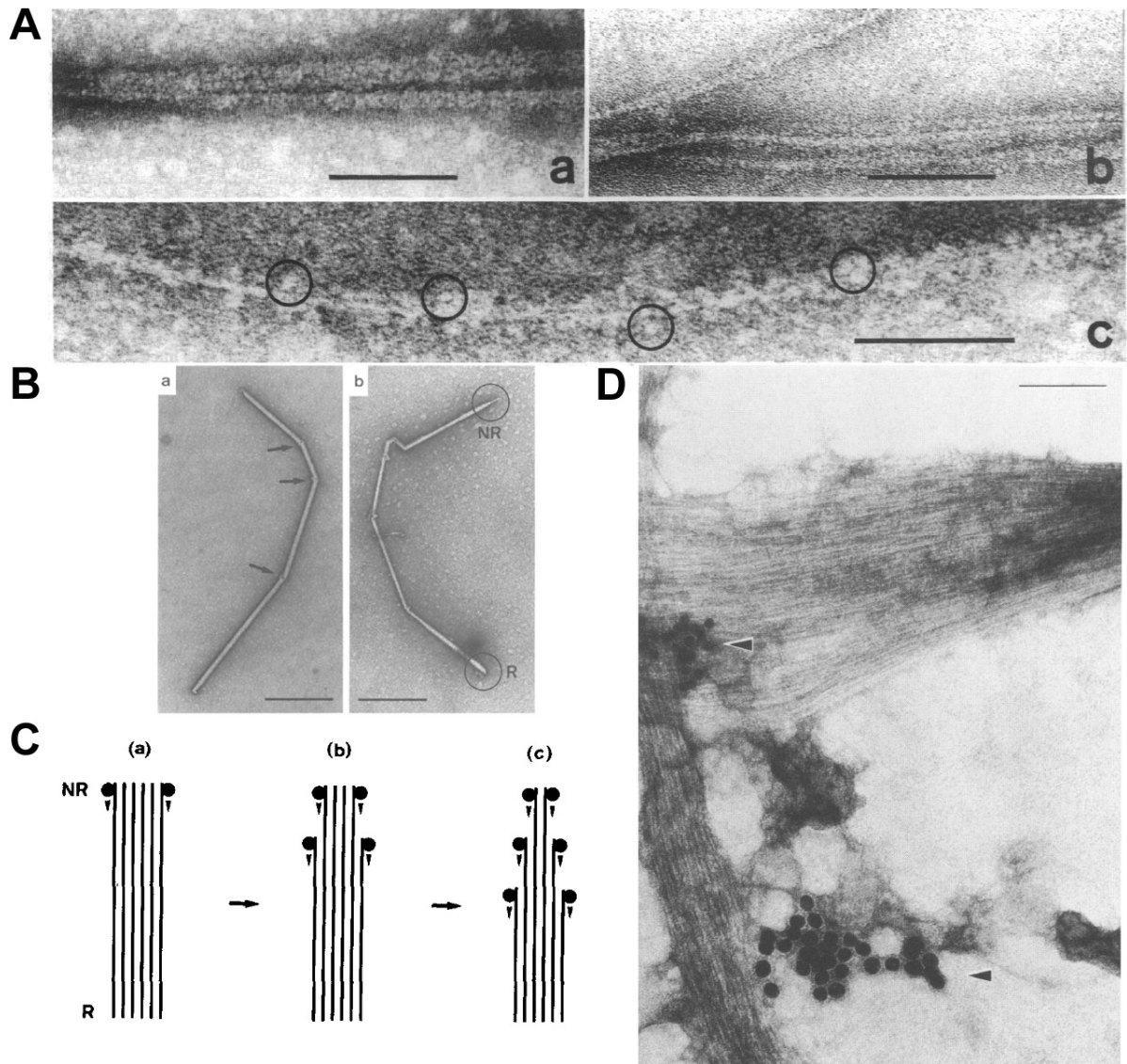


Figure 2.

1.3. Cellulases dig deep: in situ observation of the mesoscopic structural dynamics of enzymatic cellulose degradation

J. Biol. Chem., under review.

CELLULASES DIG DEEP: IN SITU OBSERVATION OF THE MESOSCOPIC STRUCTURAL DYNAMICS OF ENZYMATIC CELLULOSE DEGRADATION*

Patricia Bubner¹, Judith Dohr², Harald Plank², Claudia Mayrhofer², Bernd Nidetzky¹

¹Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12, A-8010 Graz, Austria; ²Institute for Electron Microscopy, Graz University of Technology, Steyrergasse 17, A-8010 Graz, Austria

Running title: A mesoscopic view on enzymatic cellulose degradation.

Address correspondence to: Bernd Nidetzky, Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12, A-8010 Graz, Austria. Phone: +43 316 873 - 8400; Fax: +43 316 873 - 8434; Email: bernd.nidetzky@tugraz.at.

ABSTRACT

Enzymatic hydrolysis of cellulose is key for the production of second generation biofuels which represent a long-standing leading theme in the field of sustainable energy. Despite the wealth of knowledge about cellulase structure and function, the elusive mechanism by which these enzymes disintegrate the complex structure of their insoluble substrate - which is the gist of cellulose saccharification - is still unclear. We herein present a time-resolved structural characterization of the action of cellulases on a nano-flat cellulose preparation, which enabled us to overcome previous limitations, using atomic force microscopy (AFM). As a first step in substrate disintegration, elongated fissures emerge which develop into coniform cracks as disintegration continues. Detailed data analysis allowed tracing the surface evolution back to the dynamics of crack morphology. This, in turn, reflects the interplay between surface degradation inside and outside of the crack: it is conceivable how hindered diffusion leads to product inhibition and loss of cooperative interaction between the enzymes, thus transiently limiting cellulase activity inside the growing crack.

The mesoscopic view presented here unifies the various mechanistic proposals from the past and immediately explains the dynamics of global process parameters such as protein adsorption and sugar release. Thus it serves to identify optimization targets for enhancing enzymatic cellulose saccharification.

Keywords: enzymatic cellulose degradation/ cellulase mechanism / cellulosic ethanol / AFM

INTRODUCTION

Long-term global efforts in the field of renewable energy aim at efficient utilization of lignocellulosic biomass for fuel production with the saccharification of cellulose being considered as the key bottleneck (1-5). In today's markets, a competitive commercial process requires the formation of soluble sugars from plant cell wall microfibrils being highly

efficient. This efficiency, however, is largely restrained by what is referred to as substrate recalcitrance: the structural resistance of the insoluble cellulose core to enzymatic conversion manifesting itself in sluggish rates of hydrolytic breakdown of cellulose, even at high enzyme concentrations (1,3). Overcoming cellulose recalcitrance therefore constitutes a central aim in biofuels development. Despite extensive research spanning over more than four decades, processes for cellulose hydrolysis are still advanced empirically due to insufficient understanding of the mechanisms underlying insoluble substrate deconstruction by enzyme action. Translating the wealth of knowledge about the structure and the catalytic function of cellulose degrading enzymes into comprehension of cellulase activity on the insoluble substrate has proven to be remarkably difficult, essentially because of the two following main complexities: firstly, the heterogeneous morphology of the cellulosic substrate; and secondly, the methodological difficulties in visualizing the action of cellulases on the cellulose surface at the nanometer scale. Schematic views of enzymatic hydrolysis of cellulose published nowadays have hence changed little compared to decades ago (4,6-8), which is the best evidence of the tedious progress made in the field. Multiple factors concerning the enzyme [i.e. mode of adsorption to substrate, individual components of cellulase systems showing synergy (cooperative interaction), product inhibition, stability] and the cellulosic substrate (i.e. crystallinity, available surface area, pore size, degree of polymerization) certainly play a crucial role in hydrolysis (7). The relative importance and interdependence of these factors will nevertheless remain elusive, pending clarification of the fundamental questions.

In this article, we present a study employing AFM in which we achieved a time-resolved in situ visualization of the effect of enzyme action on the surface of cellulose at a nanoscale resolution. By using a special nano-flat preparation of cellulose we were able to directly monitor the complete process of

deconstruction of the substrate from early fissuring events on the cellulose surface to the progressing surface erosion at longer hydrolysis times. We extracted the dynamic alteration of the cellulose surface structure quantitatively from AFM data, and we show how it reflects the interplay between surface degradation inside and outside of cracks, which develop as a result of the combined laterally processive and penetrating actions of the adsorbed cellulases. This mesoscopic structural view on cellulose degradation provides an immediate explanation for the peculiar kinetics of enzyme adsorption and soluble sugar formation which have long been the source of controversial interpretations in literature (4,7,9-10).

EXPERIMENTAL PROCEDURES

Materials

All materials were purchased of the highest purity available from Carl Roth (Karlsruhe, Germany) unless stated otherwise. Avicel PH-101 was from Sigma-Aldrich (St. Louis, MO, USA).

Preparation of the complete cellulase system.

We produced the complete cellulase system of *Trichoderma reesei* SVG 17 as described by Esterbauer et al. (11), supplemented it with 0.05 % sodium azide and stored it at 4 °C for the time course of the experiments. Routinely, we checked activity and protein content of the preparation to verify stability and consistency. As recommended by the IUPAC, we determined the cellulolytic activity using the well-established filter paper unit (FPU) assay (11-12). According to this assay, the cellulase system had a specific activity of 2.14 FPU/mg at a protein concentration of 0.5 mg/mL. We measured the protein concentration as described by Bradford (13) employing Roti®-Quant and Roti®-Nanoquant assays standardized with bovine serum albumin. We also used this method for determining the protein concentration in the reaction supernatant and the respective washing buffer solutions in our protein adsorption studies. For the preparation of appropriate dilutions and for the washing steps

during AFM experiments described below, we used a 50 mM sodium citrate buffer (pH 5.0) which is denoted as buffer.

Cellulose substrate and specimen preparation.

Firstly, we prepared a primary cellulose gel according to a previously published protocol (15). Secondly, in order to obtain a substrate preparation resembling a natural cellulosic substrate, we removed the solvent and loosely bound water using ethanol extraction and subsequently air-dried the substrate. This procedure resulted in a dehydrated gel. In contrast to the primary gel material, the dehydrated gel was stiff, non-transparent and substantially smaller. To confirm the partly crystalline character of the dehydrated gel, we performed X-ray diffraction (XRD) analysis (Fig. S1) of dehydrated gel samples using a Siemens D501 diffractometer (CuK α 1 radiation; Siemens, Munich, Germany). Additionally, we conducted combined differential scanning calorimetry (DSC) / thermogravimetric analysis (TGA) measurements with a Netzsch STA 449C (Netzsch Gerätebau GmbH, Selb, Germany) using a heating rate of 10 °C/min under a constant flow of helium (50.0 mL min⁻¹). We analyzed the gas products which evolved in the course of the measurement online with a Netzsch QMS 403C quadrupole mass spectrometer. Results of mass spectrometric (MS) and simultaneous thermal analyses (STA) confirmed the absence of ionic liquid in the dehydrated gel used for AFM data collection (14-15) (Fig. 1 and Table 1). Therefore, the substrate used herein was representative of the diversity of cellulose types applied in literature. Finally, to generate a stationary substrate for AFM investigation, we embedded the dehydrated gel in epoxy resin (Fig. S3). We confirmed the absence of epoxy diffusion into the substrate experimentally by energy dispersive X-ray spectroscopy (EDX) measurements in a variable pressure scanning electron microscope (VP-SEM Quanta 200 and Quanta 600).

Experimental strategy and AFM data collection.

Prior to cellulase exposure, we recorded AFM reference images of the dry substrate at defined

areas. Thereafter, we allowed the substrate to equilibrate in buffer at room temperature for two hours. Again, we recorded AFM reference images at defined areas. Following this, we incubated the substrate in the complete cellulase system at an appropriate dilution in order to achieve the desired enzyme / substrate ratio (36; 3.6; 0.36 mg/g_{substrate}) for the time indicated below at 30 °C without agitation. Then, we performed appropriate AFM control measurements of the substrate samples incubated at 30 °C in buffer continuously in order to affirm the absence of surface alteration by buffer alone after extended incubation times.

After each incubation step during the discontinuous strategy, we withdrew the substrate from the enzyme solution, washed three times in buffer, blow-dried by CO₂ spraying, measured it using the AFM and afterwards re-incubated it in a fresh enzyme dilution at the breakpoints indicated as follows (referring to total incubation time): 30 min, 60 min, 90 min (36 mg_{protein}/g_{substrate}); 15 min, 45 min, 105 min, 165 min, 225 min, 285 min (3.6 mg_{protein}/g_{substrate}); 30 min; 90 min; 210 min; 450 min (0.36 mg_{protein}/g_{substrate}). For continuous measurements, the incubation time was 90 minutes for 36 mg_{protein}/g_{substrate}; 285 minutes for 3.6 mg_{protein}/g_{substrate}; and 450 minutes for 0.36 mg_{protein}/g_{substrate}, respectively. Eventually, we dried these samples as described above for the discontinuous samples.

We used a commercial Dimension 3100 AFM equipped with a Hybrid scanner and Nanoscope IVa controller (Bruker AXS, Santa Barbara, USA) for all AFM measurements. We performed the imaging in tapping mode on dry substrate surfaces in air with an OMCL-AC 240 / 160 TS silicon probe (Olympus, Tokyo, Japan). In order to guarantee relevance and comparability of individual measurements, images were always obtained at identical areas using the same scan sizes. We chose the scan rates, setpoints and drive amplitudes appropriately in order to obtain stable scanning with the lowest energy dissipation possible. During image recording, we gave careful consideration to avoiding tip related artifacts by

permanently evaluating side wall angles and morphological features of the surface.

AFM image analyses.

We performed detailed image analysis using the software packages Research Nanoscope (V7.13, Bruker, AXS, Santa Barbara, USA) and Gwyddion (V2.2, <http://gwyddion.net>) to quantify observed features and confirm extracted data by cross-checking the results. In order to prove non-monotonic development of cracks, it was essential to gain information about the geometry of selected individual cracks, i.e. length, width and depth. To assure reliability, we thoroughly pre-investigated side walls of the cracks with respect to AFM tip related limitations.

We extracted data describing the changing topology of the overall substrate surface, i.e. root-mean-squared (RMS) surface roughness, crack footprint and number of cracks, with carefully adapted masking tools including height and curvature filters. To estimate the error of analyses, we did over- and underestimation of the applied masks on purpose, and we chose the maximum deviation as error. In order to follow the trend of the footprint over time, we compared the extracted footprint areas of identical sample areas with same scan size at specified points in time. Additionally, merging of cracks was proven both through observing the time-dependent evolution of cracks during enzymatic attack and by statistical methods. Finally, high resolution phase imaging, which yields the image of the dissipative interaction energy density, allowed for localization of enzymes on the substrate (Fig. 2).

RESULTS AND DISCUSSION

Specimen preparation and characterization.

Not unexpectedly, the reproducible preparation of a cellulosic substrate specimen suitable for AFM analysis presented a critical challenge: in order to make a general statement about cellulose degradation, it is, first of all, crucial that the substrate resembles a natural substrate and is hence chemically unmodified. Secondly, studying cellulase attack by means of AFM requires a homogeneously smooth, nano-flat

surface. In a recent AFM study which aimed at understanding the mechanism of cellobiohydrolase (CBH) I an important finding was published: CBH I preferably attacks crystalline cellulose at the hydrophobic faces (2). The types of native and processed celluloses applied in earlier AFM studies, however, were limiting in various ways (2,6,16-19). The use of highly crystalline cellulose alone, for instance, does not allow the investigation of an entire cellulase system featuring CBHs and endoglucanases (EGs) - with the latter ones being known to be unable to attack highly ordered areas (2,7,9,17). In contrast to previous research, we took a distinct approach involving the dissolution of the microcrystalline cellulose Avicel in the ionic liquid 1-butyl-3-methylimidazolium chloride (BMIMCl) at 100 °C which, upon cooling, resulted in a gel-like material (15). From this primary gel material we then removed the ionic liquid and excess water using ethanol which yielded a dehydrated gel of pure cellulose. We want to emphasize that this type of preparation does not involve precipitation of amorphous cellulose from the ionic liquid (20). It is rather a dissolution process involving formation of a gel consisting of BMIMCl, water – both loosely and tightly bound – and cellulose. Dehydration with ethanol removes a significant amount of water as well as all of the ionic liquid from this primary gel as confirmed by STA and MS analyses (Fig. 1, Table 1). Thus, the dehydrated gel is a pure cellulosic substrate. Evidence from XRD analysis revealed a mixed amorphous / crystalline character of the substrate preparation, similar to a natural cellulosic substrate (Fig. S1). The RMS surface roughness of the dehydrated gel was 8 - 10 nm after a two hour equilibration period in buffer. Compared to the size of the catalytic unit of a typical cellulase [~ 6 nm (21)], this clearly validates the above mentioned substrate for the purpose of this study in terms of surface smoothness. This special preparation enabled us not only to localize single enzymes in situ (Fig. 2), but also to observe the dynamic change of the surface during hydrolysis (7,9).

AFM investigation of enzymatic cellulose disintegration.

For the examination of substrate degradation we employed the cellulolytic enzyme system of the wood-rotting fungus *Trichoderma reesei* (7,11). This well-known cellulase contains all the individual activities required for an efficient breakdown of crystalline cellulose into soluble sugars, including CBHs, which cleave cellulose chains from the chain end (“exo”), and EGs, which cleave in the middle of a chain (“endo”). As the first step in our study, we incubated the cellulose specimen at 30 °C in sodium citrate buffer containing a defined amount of cellulase which we varied at three different levels reflecting limiting, intermediate and saturating adsorption of enzyme on the cellulose surface (36; 3.6; 0.36 mg_{protein}/g_{substrate}). At certain breakpoints we interrupted the enzyme action by removing the substrate, which we subsequently washed by dipping it into buffer. Upon blow-drying, we performed AFM imaging. Thereafter, we placed the sample into a fresh enzyme solution with the same cellulase concentration as before and incubated it until the next breakpoint. In order to rule out that the effects observed are biased by washing and re-incubation, we analyzed controls subjected to continuous incubation. Additionally, we kept track of protein adsorption onto the substrate concomitantly with the AFM measurements. The comprehensive data set thus obtained allows for quantitative monitoring of the disintegration of the cellulose surface upon enzymatic attack (Fig. 3). In this article, we discuss the results measured at intermediate enzyme loading as they gave the best resolution with respect to incubation time. The overall pattern of surface disruption, however, was independent of the amount of enzyme applied (Fig. S2).

Visualization results and implications.

Intriguingly, the first snapshot of the enzymatic reaction (Fig. 3A) shows elongated fissures which have instantly emerged all over the surface. They are aligned in their elongation direction, implying a preferred direction of enzymatic processing on the surface. This

finding correlates well with the observation of processive action of CBHI which was confirmed by AFM in recent publications (2,6). Cellulose surface fissuring was mirrored in an increase of overall RMS roughness. Interestingly, however, RMS roughness did not evolve continuously but rather stepwise, passing through a distinct plateau of constant RMS roughness (Fig. 3). We observed the appearance of an RMS roughness plateau at all three cellulase concentrations studied (36; 3.6; 0.36 mg_{protein}/g_{substrate}) (Fig. 4). Additionally, the development of RMS roughness and protein adsorption over time showed a highly similar trend for all three cellulase concentrations applied, suggesting that the disruption of the homogeneous surface structure results in additional surface area available for cellulase binding (Fig. 4). This provides substantial evidence that the hydrolysis rate limitation is not caused by decreasing enzyme-available surface area, contrary to what existing theories say (10).

Detailed analysis of the time dependence of the crack footprint gave further insight into the peculiar dynamics of RMS surface roughness: this footprint is a measure of the percentage of substrate surface that has become fissured by enzymatic attack. Strikingly, the RMS roughness reaching its plateau coincided with the decrease of the crack footprint, passing through its minimum value at the end of the plateau phase. Then the course of the crack footprint rises to a temporary maximum (Fig. 5A). Already fissured surface can obviously not be rebuilt in the enzymatic process; the observed decrease in crack footprint therefore inevitably implies that diminution in the cross-section of the formed cracks must occur. In fact, this is solely explained through advancing degradation of the surface outside the crack taking the lead over intra-crack hydrolysis, thereby causing an overall surface reduction (Fig. 6). This, in turn, can only be rationalized as the consequence of a severe limitation of the enzyme activity inside the cracks, resulting in selective slowing down of the *internal* reaction rate as compared to the corresponding reaction rate on the *external* surface. More precisely,

repetitive acceleration and deceleration processes are taking place within several cracks at the same time. Data analysis of various different cracks strongly support the notion that these localized events show a gradient over the length of the crack: while hydrolysis continues at a constant pace at the edges, it is subjected to severe restrictions in the middle of a crack. Generally, appearance of internal rate limitation is nonetheless local and temporary until limiting factors are neutralized by external hydrolysis promoting cross-section diminution and crack merging. The discontinuation of the internal rate limitation is clearly indicated by the increase in the crack footprint at longer reaction times. This signifies a further increase in cracked surface which - remarkably enough - was not reflected by an increasing total number of cracks, indicating the occurrence of merging events (Fig. 5B).

In figure 5C, we selected a typical small crack and extracted data about the evolution of its width and depth over time. Both parameters increased steeply at first, reaching their maximum at 105 minutes. Subsequently, width and depth decreased thus causing a cross-section diminution as indicated by the crack footprint. Unexpectedly, this occurred in spite of a simultaneously increasing overall RMS roughness (Fig. 3). These opposed trends of development of crack size parameters and RMS roughness can only be accounted for by an overall surface reduction, well consistent however, with the observation of a concurrent increase in the crack footprint (Fig. 5A). If smallish cracks lose rather than gain in size, the observed increase in RMS roughness must be due to three major reasons: firstly, constantly evolving new fissures; secondly, concomitant expansion of existing cracks at their edges; and finally, constant growth of the large cracks. Concerning the latter one, the decrease in crack number over time (Fig. 5B) indicates that crack coalescence is partially involved in the growth of these cracks. Figure 3 shows the time course of RMS roughness supported by AFM images. These images reveal the dynamics of the cross-sectional area of a rather small crack alongside

the unrestrained growth of a large crack. RMS roughness would seem to pass through the observed plateau at the point where hydrolysis in small cracks becomes limiting overall (“deceleration”), only to increase when surface hydrolysis has overcome the internal limitations (“acceleration”), which are caused by diffusional restrictions occurring in the confined space of the cracks.

These effects of hindered diffusion in small cracks will influence cellulolytic activity at two different levels: firstly, by product inhibition and secondly, through loss of synergy among individual cellulase components. As a known fact, cellobiose is the main product of cellulose hydrolysis and strongly inhibits CBH I, the major protein component and “pacemaker” exo-enzyme of the *T. reesei* cellulase system (22). According to the dynamics we observed, the formation of a concentration gradient for cellobiose will be promoted by restricted diffusivity within narrow cracks. This will a fortiori be favored by partial size exclusion of β -glucosidase, a natural companion enzyme of the actual cellulases. This enzyme, which lacks the ability to adsorb to cellulose, cleaves cellobiose to yield glucose. As a matter of fact, the cellobiose binding constant of the *T. reesei* cellulase system is greater by a factor of 6 than its glucose binding constant, which makes glucose considerably less inhibitory than cellobiose (7,23). The global importance of product inhibition is fully captured in an AFM experiment where continuous hydrolysis results in essentially the same surface disruptions but proceeds much more slowly than discontinuous hydrolysis. The discontinuous setup involved repeated removal of the supernatant, thus probably alleviating high local product concentrations. In addition to product inhibition, dynamic features of cellulase adsorption to cellulose - in particular the exchange of complementary endo- and exo-enzyme activities on the substrate surface - will become strongly attenuated within narrow cracks. In addition to this, a possible sliding movement of exo-acting enzymes will be hampered. The basis for synergistic interaction between cellulases is thus

gradually removed and deactivation of enzymes and pore clogging by protein aggregates might receive significance through phenomena such as irreversible adsorption, jamming, and aggregation (7-9,24-25).

A novel model of enzymatic cellulose disintegration.

From what was hitherto known about the mechanism of cellulose disruption, the mesoscopic view of enzymatic disintegration of cellulose presented in this article implies a paradigm shift in that the overall surface disruption caused by cellulase action is not merely a surface phenomenon, but is, in contrast, composed of an external (*surface erosion*) and internal (*surface penetration*) component. As visualized in figure 6, the evolution of surface structural features along the reaction timeline reflects the dynamics of surface erosion and penetration rates: in the initial phase of hydrolysis, the reaction proceeds “acceleratedly” (Fig. 6A). Then, the transient occurrence of internal limitations leads to a “deceleration”, meaning that surface hydrolysis predominates, thus causing cross-section diminution of those cracks where limiting factors prevail (Fig. 6B). As soon as these limitations are overcome by the ever progressing surface erosion, acceleration occurs again (Figs. 6B to C). Picturing the entire process of cellulose saccharification, it shows that this process is probably subjected to repetitive cycles of deceleration and acceleration caused by periodically emerging internal limitations and their overcoming. Eventually, this leaves the surface completely rugged, with essentially the entire external surface area consumed (Fig. 6D). At this point, where internal limitations prevail, acceleration is stalled. As a matter of fact, stagnation of the overall surface disruption rate would provide an immediate explanation for the puzzling observation which has been made with essentially all types of cellulosic substrates: the rate of soluble sugar formation displays a massive decline at typically 20-40 % cellulolytic conversion of the solid material (24-25).

We want to emphasize that the model we present here provides a common foundation for all of the different “rate-retarding factors“ which have been proposed in literature before: inhibition by cellobiose, (apparent) enzyme deactivation, loss of synergism, and decrease of substrate reactivity. By demonstrating causative relatedness of these factors, our model resolves complexities originating from the greatly varying emphases that researchers have previously put on the role of individual factors. Thus, our model serves to unify the diverse

mechanistic proposals made in the past. In terms of practical application, this model reveals how important it is that surface erosion and penetration activities of the cellulases are balanced. The advice “don’t dig too deep“ could present a new avenue for optimizing the hydrolysis performance of these enzymes. Future endeavors to improve industrial saccharification of cellulose should therefore focus on overcoming the herein described deceleration effects to achieve high reaction rates throughout the entire process.

ACKNOWLEDGEMENTS

We thank Karin Longus for producing the complete cellulase system; Nadejda Matsko for her assistance in substrate preparation; Gregor Trimmel and Thomas Haber for characterization of the substrate; Joachim Krenn and Ferdinand Hofer for helpful discussions; and Dominika Stiger for carefully reading the manuscript. Special thanks to Monika Hitter who contributed the professional artwork for the model of cellulose degradation (Fig. 6). The authors declare no competing financial interests.

AUTHOR CONTRIBUTIONS

P.B., J.D., H.P. and C.M. prepared the nano-flat cellulosic specimen. J.D. and H.P. designed, performed, optimized and analyzed AFM investigations. J.D., H.P., P.B. and B.N. analyzed and discussed the data and developed the proposed degradation model. P.B. and B.N. designed the study, conceptualized and performed experiments, provided the substrate and the enzymes, and wrote the paper. All authors commented on the manuscript.

REFERENCES

1. Lynd, L. R., Laser, M. S., Bransby, D., Dale, B. E., Davison, B., Hamilton, R., Himmel, M., Keller, M., McMillan, J. D., Sheehan, J., and Wyman, C. E. (2008) *Nat. Biotechnol.* **26**, 169-172
2. Liu, Y. S., Baker, J. O., Zeng, Y., Himmel, M. E., Haas, T., and Ding, S. Y. (2011) *J. Biol. Chem.* **286**, 11195-11201
3. Himmel, M. E., Ding, S. Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W., and Foust, T. D. (2007) *Science* **315**, 804-807
4. Arantes, V., and Saddler, J. N. (2010) *Biotechnol. Biofuels* **3**, 4
5. Solomon, B. D., Barnes, J. R., and Halvorsen, K. E. (2007) *Biomass and Bioenergy* **31**, 416 - 425
6. Igarashi, K., Koivula, A., Wada, M., Kimura, S., Penttila, M., and Samejima, M. (2009) *J. Biol. Chem.* **284**, 36186-36190
7. Lynd, L. R., Weimer, P. J., van Zyl, W. H., and Pretorius, I. S. (2002) *Microbiol. Mol. Biol. Rev.* **66**, 506-577
8. Bansal, P., Hall, M., Realff, M. J., Lee, J. H., and Bommarius, A. S. (2009) *Biotechnol. Adv.* **27**, 833-848
9. Mansfield, S. D., Mooney, C., and Saddler, J. N. (1999) *Biotechnol. Prog.* **15**, 804-816
10. Zhang, Y. H., and Lynd, L. R. (2004) *Biotechnol. Bioeng.* **88**, 797-824
11. Esterbauer, H., Steiner, W., Labudova, I., Hermann, A., and Hayn, M. (1991) *Biores. Technol.* **36**, 51-65

12. Ghose, T. K. (1987) *Pure & Applied Chemistry* **59**, 257-268
13. Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248-254
14. Ohtani, H., Ishimura, S., and Kumai, M. (2008) *Anal. Sci.* **24**, 1335-1340
15. Prasad, K., Kaneko, Y., and Kadokawa, J. (2009) *Macromol. Biosci.* **9**, 376-382
16. Nigmatullin, R., Lovitt, R., Wright, C., Linder, M., Nakari-Setala, T., and Gama, M. (2004) *Colloids Surf. B Biointerfaces* **35**, 125-135
17. Lee, I., Evans, B. R., and Woodward, J. (2000) *Ultramicroscopy* **82**, 213-221
18. Quirk, A., Lipkowski, J., Vandenende, C., Cockburn, D., Clarke, A. J., Dutcher, J. R., and Roscoe, S. G. (2010) *Langmuir* **26**, 5007-5013
19. Liu, H., Fu, S., Zhu, J. Y., Li, H., and Zhan, H. (2009) *Enzyme and Microbial Technology* **45**, 274-281
20. Zhao, H., Jones, C. L., Baker, G. A., Xia, S., Olubajo, O., and Person, V. N. (2009) *J Biotechnol* **139**, 47-54
21. Abuja, P. M., Schmuck, M., Pilz, I., Tomme, P., Claeysens, M., and Esterbauer, H. (1988) *Eur Biophys J* **15**, 339-342
22. Zhang, Y. H., and Lynd, L. R. (2004) *Biotechnol Bioeng* **88**, 797-824
23. Holtzapple, M., Cognata, M., Shu, Y., and Hendrickson, C. (1990) *Biotechnol. Bioeng.* **36**, 275-287
24. Levine, S. E., Fox, J. M., Blanch, H. W., and Clark, D. S. (2010) *Biotechnol. Bioeng.* **107**, 37-51
25. Bommarius, A. S., Katona, A., Cheben, S. E., Patel, A. S., Ragauskas, A. J., Knudson, K., and Pu, Y. (2008) *Metab. Eng.* **10**, 370-381

FOOTNOTES

The abbreviations used are: AFM, atomic force microscopy; BMIMCl, 1-butyl-3-methylimidazolium chloride; CBH, cellobiohydrolase; DSC, differential scanning calorimetry; EDX, energy dispersive X-ray spectroscopy; EG, endoglucanase; FPU, filter paper unit; MS, mass spectrometry; RMS, root-mean-squared; TGA, thermogravimetric analysis; STA, simultaneous thermal analysis; XRD, X-ray diffraction;

FIGURES

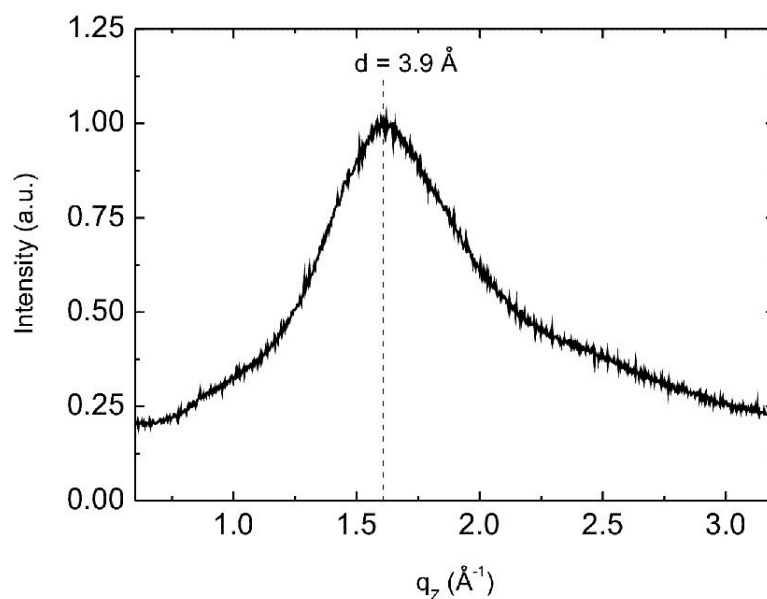


Figure 1. XRD profile of a dehydrated cellulose gel sample.

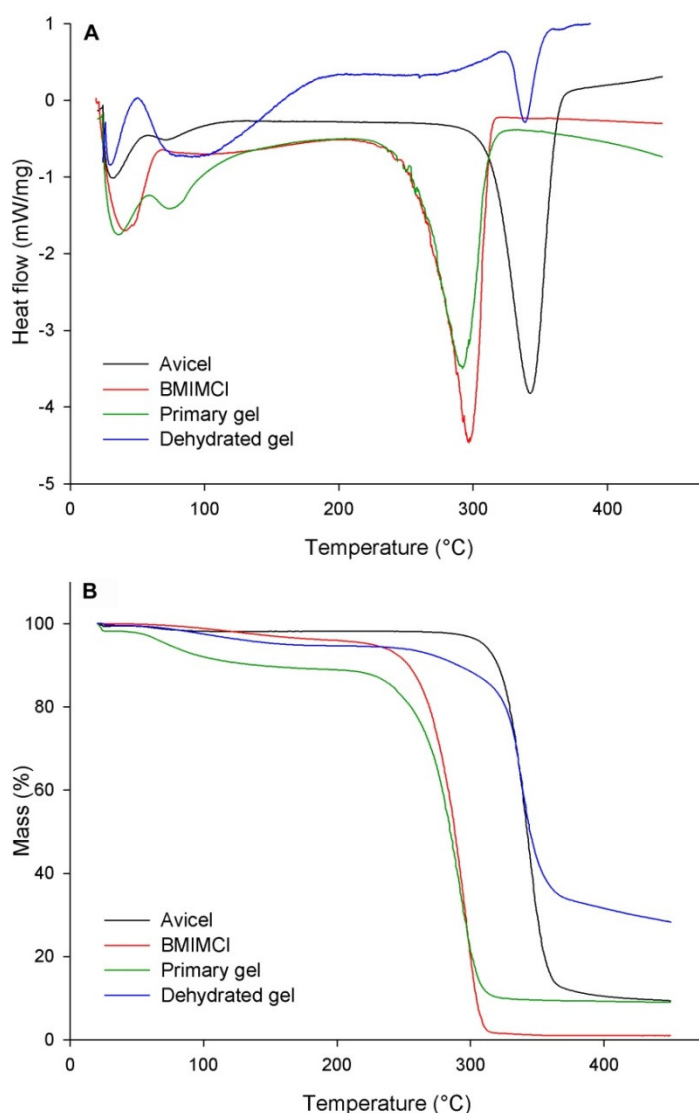


Figure 2. STA curves of microcrystalline cellulose (Avicel), the ionic liquid BMIMCl, the primary and the dehydrated gel. DSC (A) and TGA (B) measurements show remarkable differences in behavior of the dehydrated gel as compared to the primary gel. In both graphs the behavior of the dehydrated gel sample more closely resembles the one of the microcrystalline cellulose than the one of the primary gel. *B*: The difference in weight loss at 100°C indicates a pronounced loss of water in the primary gel but not in the dehydrated gel.

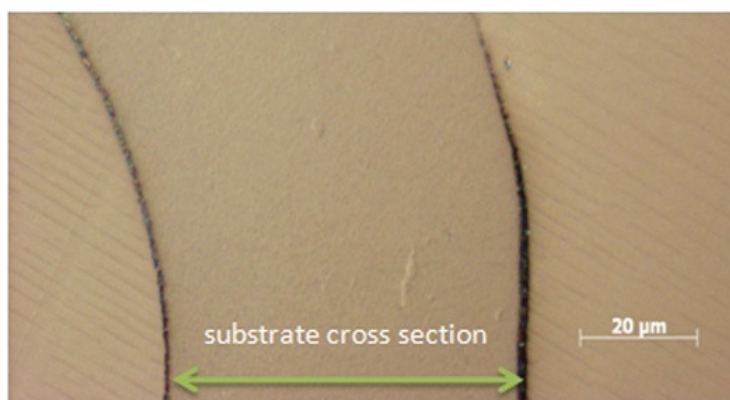


Figure 3. Light microscope image of the specimen cross section. In this image, the embedded dehydrated gel is clearly distinguishable from the epoxy resin it is embedded in order to enable AFM measurements.

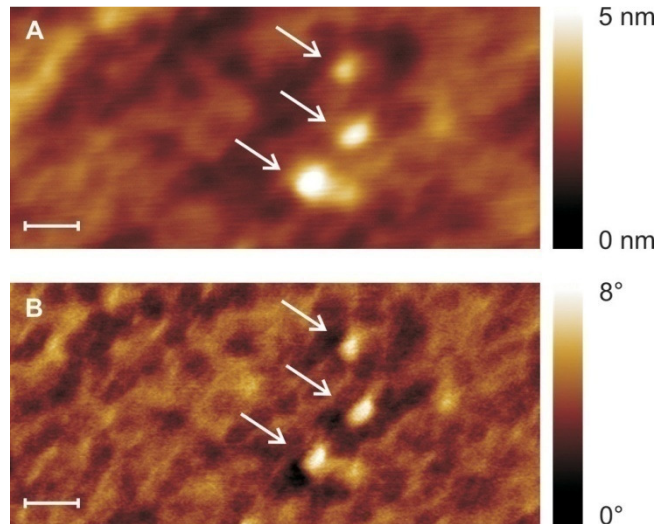


Figure 4. Visualization of single celluloses in an emerging crack. Height (A) and phase (B) image of a 200 nm scan, taken approximately 15 minutes after incubation with a cellulase concentration of $0.36 \text{ mg}_{\text{protein}}/\text{g}_{\text{substrate}}$, exhibit features not observed on the buffer-swollen image blank. These new features are approximately 5 nm in diameter and show a distinct phase signal (no slope effects). It is therefore very likely that they are single celluloses. The scale bar represents 20 nm.

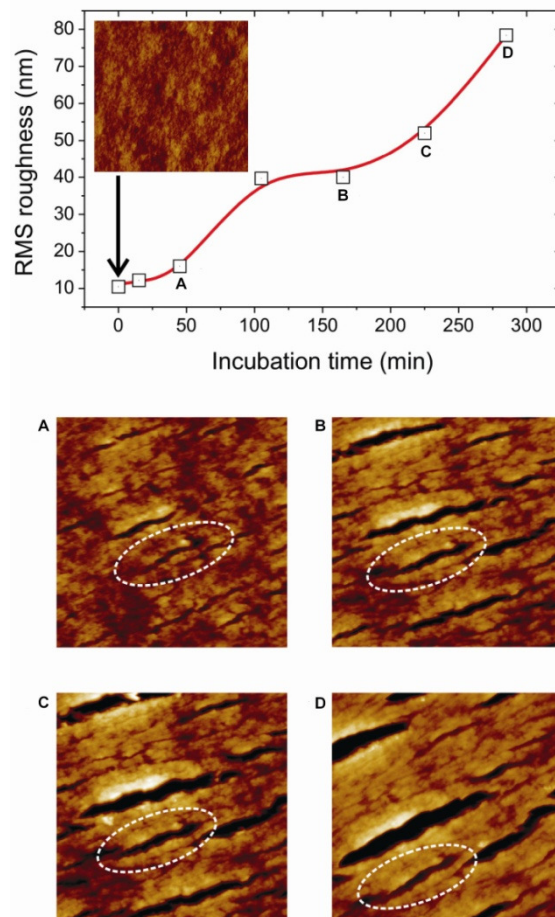


Figure 5. Enzymatic cellulose degradation reflected by RMS roughness variation. The image at 0 minutes shows the nano-flat surface after buffer swelling. Images A to D reveal the formation of elongated cracks with time at a cellulase loading of $3.6 \text{ mg}_{\text{protein}}/\text{g}_{\text{substrate}}$. As an example, a representative crack is circled and it is observable how it grows at first (A and B), then stagnates (C compared to B), and eventually becomes narrower (D compared to C).

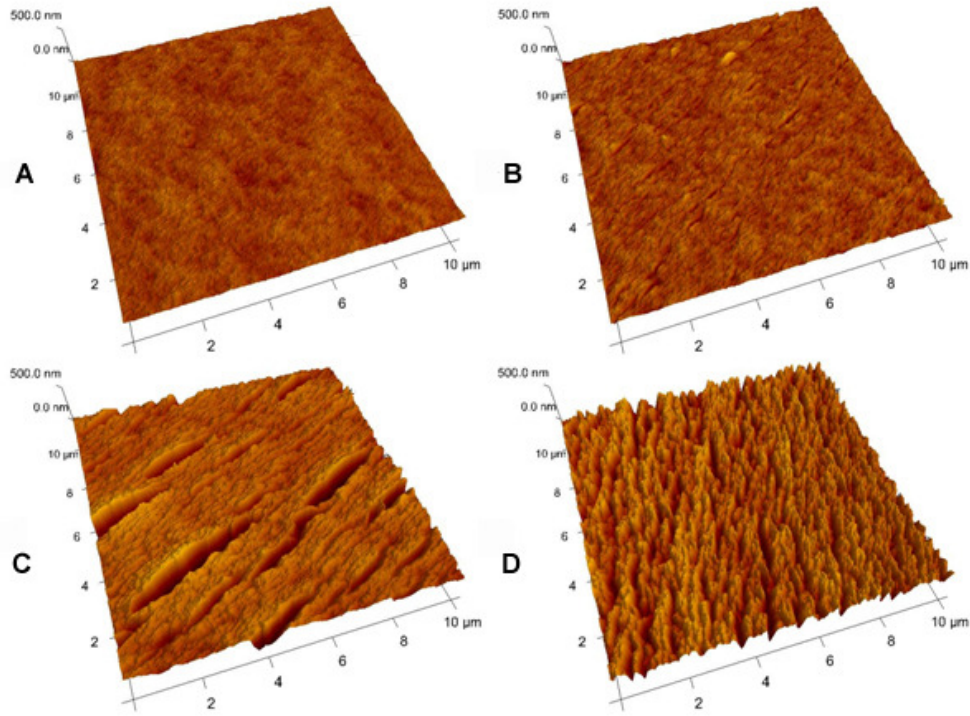


Figure 6. 3-D image of AFM investigated specimens at different enzyme loadings. A: Buffer blank, 120 min; B: 0.36 mg/g_{substrate}, 450 min incubation; C: 3.6 mg/g_{substrate}, 285 min incubation; D: 36 mg/g_{substrate}, 60 min incubation.

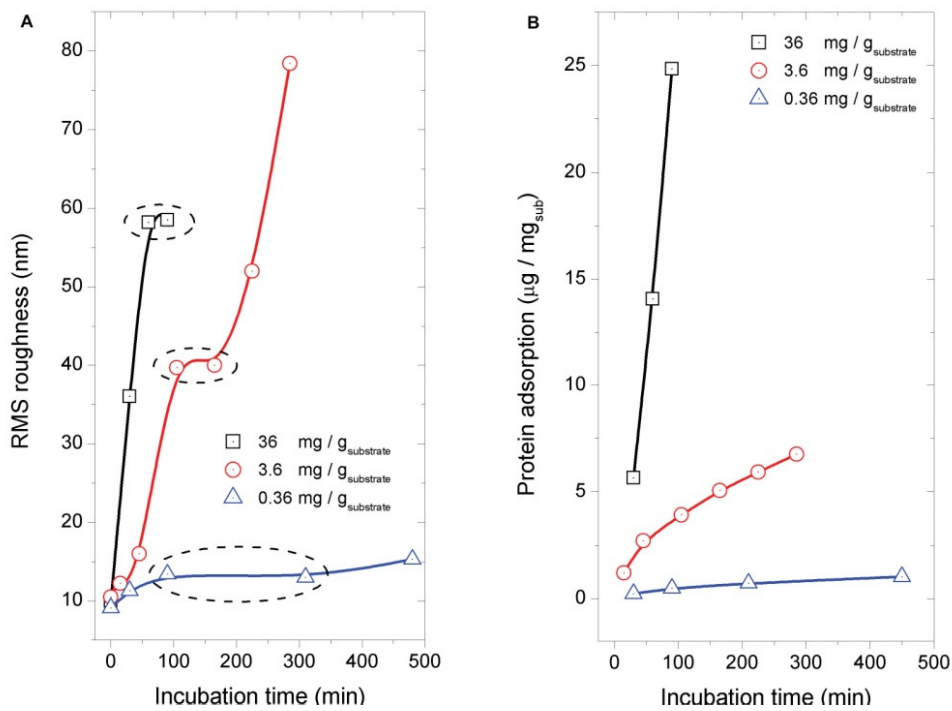


Figure 7. Course of RMS roughness and corresponding protein adsorption studies of saturated, intermediate, and limited cellulase concentration. The two graphs show a similar trend, indicating that celluloses readily adsorb to newly evolving surface during hydrolysis. Furthermore, the occurrence of a plateau is clearly visible in RMS roughness evolution at different enzyme loadings (36; 3.6; 0.36 mg_{protein}/g_{substrate}).

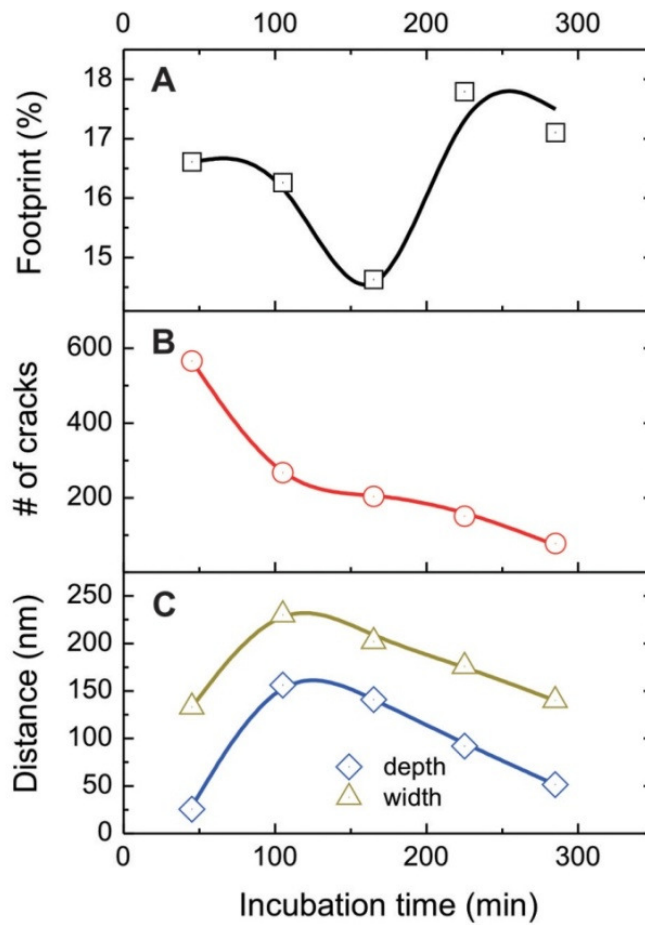


Figure 8. Quantitative analysis of AFM data. A: Course of the crack footprint over time. B: The total number of cracks is decreasing over time, in spite of constantly emerging new cracks. This is a result of numerous merging events which eventually lead to a network of valleys on the substrate surface. C: Development of depth and width of a representative small crack.

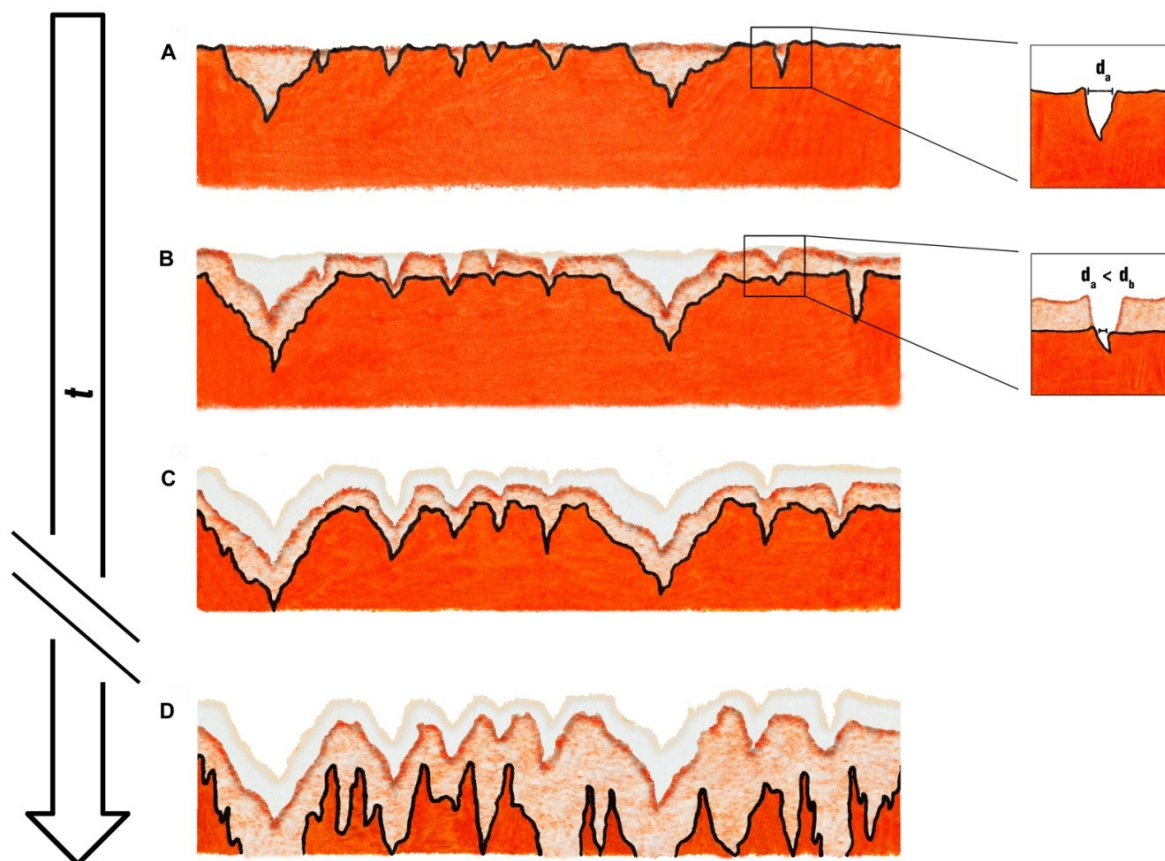
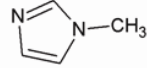
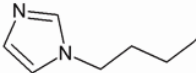


Figure 9. Model of enzymatic cellulose digestion based on quantitative AFM data. This schematic model aims at visualizing the mesoscopic perspective on enzymatic disruption of cellulose. Along the timeline we follow the cellulolytic degradation of the surface from random early fissuring events (A) to a completely fissured surface, where the entire external surface has been consumed and rate retarding factors become operative (D). The proportions in A through C reflect typical dimensions observed in relevant AFM measurements.

TABLE

Table 1. MS data of primary and dehydrated gel compared to microcrystalline cellulose (Avicel) and the ionic liquid BMIMCl. In this table, the intensity of the signal at a specific m/z ratio (“mass-to-charge ratio”) is represented by +++ (high), ++ (medium), + (low) and – (below the limit of detection). As compared to BMIMCl and the primary gel, signature peaks for the presence of BMIMCl were absent in Avicel and the dehydrated gel, indicating an ionic liquid free substrate. Thermal decomposition behavior of BMIMCl is described in *Ohtani et al.* (14).

m/z	fragment	chemical name	M_r [g/mol]	Avicel	BMIMCl	Primary gel	De-hydrated gel
82		1-methylimidazole	82.1	-	+++	+++	-
124		1-butylimidazole	124.2	-	++	+	-

1.3.1. Appendix to “Cellulases dig deep: in situ observation of the mesoscopic structural dynamics of enzymatic cellulose degradation”

1.3.1.1. Dissolution of cellulose in ionic liquids

Dissolving the water insoluble cellulose usually requires harsh conditions. Its crystalline regions are barely accessible even for small molecules such as water. Its amorphous regions, on the other hand, can swell substantially upon the entry of water. However, to completely dissolve cellulose, often uncommon and/or toxic solvents are necessary which cannot be recovered afterwards, e.g. *N,N*-dimethylacetamide/lithium chloride (DMA/LiCl). It was not until 2002 that a promising novel class of cellulose solvents was opened up for the non-derivatizing dissolution of cellulose: ionic liquids. Ionic liquids are salts with a very low melting point which is due to their composition of bulky, asymmetrical ions with a delocalized charge. The 1-butyl-3-methylimidazolium cation ($[BMIM]^+$), for example, is more asymmetric than the 1-butylpyridinium cation thus having a lower melting point. There are some ionic liquids which are liquid even at room temperature. Some of the general advantages of ionic liquids are that they have a low vapor pressure, are extremely polar, thermostable, non-derivatizing and easy to recycle after cellulose dissolution. Due to their recyclability they were often termed “green solvents”, which has been questioned recently as they are toxic and hazardous to the environment even in small amounts. Further disadvantages are the hygroscopic nature of ionic liquids, which leads to a gradual uptake of water from the air, and their viscosity. These water molecules then are loosely coordinated to the hydrophilic anion as the cation is rather hydrophobic. Increased water content significantly reduces the solubility of cellulose (Fidale *et al.* 2009, Pinkert *et al.* 2009).

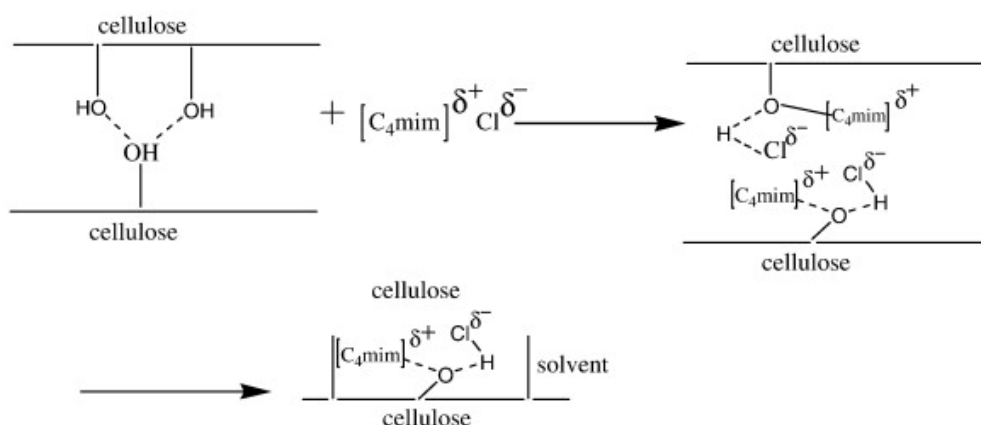


Figure V. The proposed dissolution mechanism of cellulose in BMIMCl (Feng and Zhong-Ian 2008).

The capability of ionic liquids to dissolve cellulose is primarily due to their ability to build strong hydrogen bonds with the substrate. Generally, imidazolium-based halide ionic liquids have a very strong hydrogen bond network that was even ascribed “covalent-like character” (Pinkert *et al.* 2009). The best cations to dissolve cellulose were found to be based on methylimidazolium and methylpyridinium structures with butyl side chains. The best anions are chloride, acetate, and formate. Dissolution of cellulose in BMIMCl presumably occurs via formation of electron donor – electron

acceptor complexes between cellulose (mainly C-6 and C-3 hydroxyl groups of neighbored cellulose chains) and both, the cation and the anion of the ionic liquid. These interactions weaken interchain hydrogen bonds thus leading to chain separation and, eventually, dissolution. The mechanism of dissolution of cellulose by BMIMCl is outlined in Figure V. Elevated temperatures assist in cellulose dissolution through better interaction of the chloride ion and the cellulose (Fidale *et al.* 2009, Pinkert *et al.* 2009).

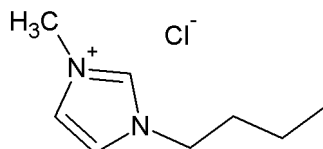


Figure VI. 1-butyl-3-methylimidazolium chloride (BMIMCl)

Even though ionic liquids do not derivatize cellulose⁴, they sometimes lower the degree of polymerization (DP) and they modify its morphology – if not upon dissolution in the ionic liquid, at least upon regeneration from it. In the study of Fidale *et al.*, SEM images of microcrystalline cellulose dissolved in 1-allyl-3-(1-butyl)imidazolium chloride showed a pronounced morphological change of the cellulose from an originally rough surface to a “smooth fibrous-like mass”, going hand in hand with a decrease in crystallinity (Fidale *et al.* 2009). From BMIMCl it is known that upon dissolution cellulose becomes amorphous without affecting its degree of polymerization and polydispersity. Besides, if the cellulose content dissolved in BMIMCl exceeds 10% (w/w), liquid crystalline solutions are formed. When precipitated (e.g. by addition of water) and subsequently regenerated from BMIMCl cellulose has a quite homogeneous macrostructure as described in several studies. However, it has also been indicated that during regeneration it could regain a certain degree of crystallinity depending on various factors, e.g. stress / strain (Swatloski *et al.* 2002, Zhao *et al.* 2009, Zhu *et al.* 2006).

1.3.1.2. Preparation of the nano-flat substrate specimen

A stringent requirement for a successful AFM study of enzyme action on cellulose is a nano-flat, chemically unmodified substrate that provides a homogeneously smooth surface for cellulase attack. In order to come to a general conclusion about cellulose saccharification, a system that resembles a natural cellulosic substrate – in having both, crystalline and amorphous regions – is desirable. The substrate specimen developed in our study “Cellulases dig deep: in situ observation of the mesoscopic structural dynamics of enzymatic cellulose degradation” is pure cellulose derived from a gel of cellulose and BMIMCl. This preparation features a nano-flat surface with an RMS roughness of 8–10 nm after equilibration in buffer for 2 hours. It also features amorphous as well as crystalline regions as confirmed by XRD-analysis. STA data (Figure 2) proved that the dehydrated gel we employed showed a behavior more similar to Avicel than to the primary gel material [prepared as published by Prasad *et al.* (Prasad *et al.* 2009)]. More precisely, the dehydrated gel has a lower water content and less of

⁴ Cellulose solvents are classified into derivatizing and non-derivatizing ones: derivatizing solvents form intermediates through chemical interaction with the hydroxyl groups of cellulose (Pinkert *et al.* 2009).

loosely coordinated water molecules compared to the primary gel. This is corroborated by the less pronounced mass loss at around 100°C in the TGA curve (Figure 2B) and by a lower MS peak for the loosely coordinated water. MS analysis was coupled to the STA. According to two previous publications (Kamavaram and Reddy 2008, Ohtani *et al.* 2008), BMIMCl will thermally decompose into the fragments as shown in Table i. As a matter of fact, we were able to identify two signature peaks⁵ of the ionic liquid: the peaks at *m/z* ratios typical for the BMIMCl – 1-methylimidazole and 1-butylimidazole – were present in the BMIMCl and the primary gel material (Table 1). Corresponding peaks at a relative molecular mass of 82 g/mol and 124 g/mol are entirely absent in Avicel as well as in the dehydrated gel. Additionally, the endothermic DSC peaks at about 300°C attributed to the decomposition of BMIMCl (Prasad *et al.* 2009) are not present in our cellulosic substrate and in Avicel (Figure 2A). This confirms the absence of BMIMCl in the dehydrated gel and thus the purity of the cellulosic substrate. Furthermore, the results of this extensive characterization lead to the conclusion that the novel substrate employed here is indeed a preparation that resembles a natural cellulosic substrate in featuring amorphous as well as crystalline regions (Levine *et al.* 2010, Zhang and Lynd 2004). Therefore, we lay claim that the results gained with this substrate allow general conclusions on the enzymatic degradation of natural cellulosic substrates.

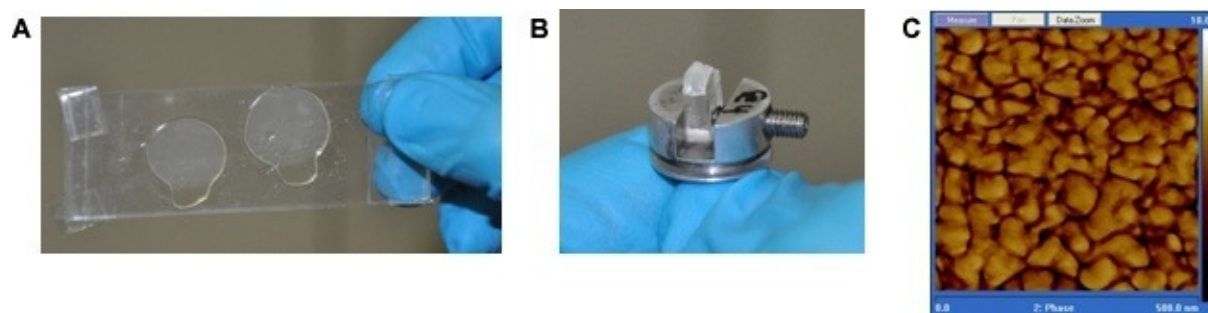
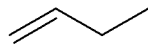

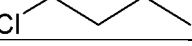
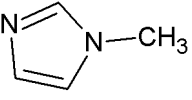
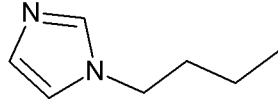
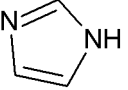


Figure VII. The cellulosic substrate. A: Primary gel. B: Substrate specimen embedded in epoxy and ready for AFM. C: Features of the substrate (size 25 – 50 nm) as measured in AFM (scan size: 500 nm) under air.

We want to emphasize that the type of preparation we used to generate our nano-flat substrate does not involve precipitation of amorphous cellulose from the ionic liquid. It is rather a dissolution process involving formation of a gel consisting of BMIMCl, water – both loosely and tightly bound – and cellulose (Pinkert *et al.* 2009, Prasad *et al.* 2009). Dehydration with ethanol removes a significant amount of water as well as all of the ionic liquid from this primary gel as confirmed by STA and MS analyses (Figure 2 and Table 1). In our study we also characterized the substrate using XRD which revealed our dehydrated gel as a mainly amorphous substrate with a certain degree of order (crystallinity) remaining (Figure 1).

⁵ Peaks of decomposition fragments with lower molecular mass cannot be distinguished from fragments of the cellulosic material.

Table i. Thermal decomposition fragments of BMIMCl [according to Ohtani *et al.* (2008) and Kavaram and Reddy (2008)] and their relative molar masses.

fragment	chemical name	M _r [g/mol]
	1-Butene	56.106
	1-Chloromethane	50.488
	1-Chlorobutane	92.567
	1-Methylimidazole	82.104
	1-Butylimidazole	124.184
	Imidazole	68.077

1.3.1.3. Additional discussion: a new model of cellulose degradation

In our study, we took a quantitative approach in analysis of AFM data and supplemented them with results from protein adsorption studies. We thus introduced a novel model of enzymatic cellulose degradation (Figure 9). For the first time, we showed that celluloses are not restricted to the surface of their solid substrate (Arantes and Saddler 2010, Mansfield *et al.* 1999) but, in fact, dig into their substrate. We want to emphasize that we observed cellulose hydrolysis as predominantly proceeding via crack formation and that a surface-restricted delamination, as postulated earlier (Arantes and Saddler 2010, Lee *et al.* 2000, Mansfield *et al.* 1999, Valjamae *et al.* 1998), is not occurring to an extent that influences the overall hydrolysis.

In the beginning of cellulose degradation, the hydrolysis rate is high due to the diffusion-controlled transport of macromolecules and products, i.e. short cellooligomers. When the first fissures develop, it is conceivable that the net flux between the intra-crack space and the bulk becomes a determinant of growing influence. In turn, this is dependent on the enzyme loading, the conversion and transport rates, the pore size and geometry as well as the intra-pore diffusivity. Besides, the synergistic action of celluloses involves multiple reactions whose reaction rates are likely to be influenced by varying diffusivities of the reactants thereby causing inhibition or facilitation of the reactivity of a single cellulase component. Therefore, hydrolysis will proceed at a high initial rate as long as the reaction is diffusion-controlled (van Roon *et al.* 2006). Consequentially, with progressing degradation of the substrate complex mass transport phenomena arise. It is evident that the diffusion coefficient in the micro-cracked substrate is different from bulk fluid and, moreover, is strongly dependent on the microstructure of the substrate, especially on size, space and distribution of cracks. If we approximate a crack as an elongated coniform pore whose apex can be expressed as a set of nano-cylinders of decreasing size, it is safe to conclude that molecular motion is subjected to severe restrictions in such a

confined geometry. Both, CBH I (65 kDa) and CBH II (58 kDa) of *T. reesei*, are of similar structure: they consist of an elliptic catalytic core domain and feature a C-terminal CBD that is connected to the core via a short linker. This yields the typical tadpole-like shape with an overall maximum diameter (D_{\max}) of 18 nm (core: 6.5 nm) and 21.5 nm (core: 6.0 nm) for CBH I and II, respectively (Abuja *et al.* 1988a, Abuja *et al.* 1988b). Accordingly, the estimated volume occupied by a CBH core domain is approximately 57–66 nm³, not including the protruding tail with a length of 12–16 nm and a diameter of about 3 nm. According to Cui (Cui 2005), the maximum accessible radius of a cylindrical pore with a diameter of 3 nm is 1.31 nm, which illustrates how the space in narrowing pores becomes limiting for macromolecules. On top of that, wall confinement plays an increasingly important role when the diffusing molecule is ten times or less smaller than the diameter of the pore. For example, for a cylindrical pore with a diameter of 1.5 nm, about five times bigger than a water molecule, the axial diffusion coefficient for water was calculated to be 3.43×10^{-5} cm²/s as compared to in bulk (5.19×10^{-5} cm²/s) (Cui 2005). In case of a typical cellulase with its ellipsoid core shape having a diameter of 4–6.5 nm and a total enzyme length of 18 – 21.5 nm (Abuja *et al.* 1988a, Abuja *et al.* 1988b, Kleywegt *et al.* 1997), this implies that every pore having a so called critical width of about 40–60 nm will already limit macromolecule diffusion to a high degree by wall confinement. Probably, due to the tadpole shape of celluloses, even bigger pores up to 200 nm limit diffusion to a certain extent. In narrowing coniform pores, the critical crack width can occur at any point within or at the surface of the crack. If we picture one of the small cracks that limit hydrolysis, and arbitrarily assume a surface width of 360 nm, which is about the maximum width of the chosen small pores, and in the same way assume a height of 200 nm, the critical width of 60 nm will be reached at a height of 33 nm from the bottom in a coniform geometry. At this point, celluloses will encounter diffusion restrictions. At a height of approximately 8 nm from the bottom, the pore diameter is approximately 14 nm, and a single macromolecule will not be able to move deeper into the pore than that, thereby probably clogging it. It is conceivable that at this point enzyme crowding occurs and, on top of that, high local product concentrations inhibit CBH I action, thereby disrupting synergism of celluloses (Levine *et al.* 2010). A cellobiose molecule, for example, has a length of approximately 1 nm (Levine *et al.* 2010, Zhang and Lynd 2004), hence soluble cellooligomers [degree of polymerization <6 (Mansfield *et al.* 1999)] have a length of about 1 – 2.5 nm. It is easily conceived how the diffusivity of reaction products is restricted in this confined geometry as well. This, along with diffusion-limited access of β -Glucosidase (BGL), will presumably lead to an accumulation of cellobiose and other cellooligomers in the apex of a coniform pore which in turn inhibit CBH. If diffusion is limited for CBH and EG, it is also feasible for BGL, which is a rather large enzyme having 71 (*T. reesei* SVG 17 BGL I) or even 114 kDa (*T. reesei* SVG 17 BGL II) (Chen *et al.* 1992). At this point, desorption and/or movement of celluloses, which might be crucial for further degradation, cannot happen due to the constraint space. Combined with a feasible “clogging” of the apex by macromolecules and inhibition by high local concentrations of cellooligomers, cellulose saccharification eventually halts in the diffusion limited parts of the pores.

As a consequence, the overall surface around the clogged crack is getting degraded until mass transfer is no longer limiting.

However, we want to emphasize that, according to the complexity of the overall process, the above scenario is always true for distinct cracks as this behavior is strongly dependent on the geometry of the fissure. As shown in Figure 5 the surface is covered with several interstices of different size and shape (i.e., length, depth, width), each having its very own distinct mass transfer problem. We presume a crucial role of the aspect ratio as well as the length of the fissures Figure 5: while the short cracks do not grow due to severe diffusion limitation, the longer ones do as the mass transfer in the upper regions of long cracks is not as limiting as in narrow pores that approximate the shape of a conus. Also, we want to emphasize that the geometry restrictions mentioned here are strictly confined to the minimum footprint of a cellulase and thus it is likely that mass transfer restrictions might occur much earlier. *Levine et al.* (*Levine et al.* 2010) approximated the footprint of a “typical cellulase” on the surface of cellulose with 21 nm², equaling 84 glucose units. However, they calculated the effective footprint by random sequential adsorption simulations to be as big as 39 nm² or 156 glucose units. In their modeling approaches, they also took account of enzyme crowding at low surface areas which leads to high cellulase surface concentrations. This in turn severely affects endoglucanase (EG) activity as CBH and EG compete for binding sites. Eventually, this phenomenon disrupts their synergistic interaction.

Comparing semi-continuous to continuous results further corroborates the conception of diffusion limitation. In principle, we observe analogous crack formation in both experimental setups, however, crack formation and overall degradation seem to advance more slowly in the continuously incubated samples. This is reasonable considering that in semi-continuous experiments the samples were washed and dried thus alleviating product inhibition by convective transport thereby promoting hydrolysis. As a matter of fact, it has been reported that cellulose conversion is higher in mixed and highly agitated reactors than in slowly agitated ones (*Mais et al.* 2002). Besides, in the semi-continuous setting the samples were incubated in fresh enzyme solutions afterwards, thus allowing higher loadings at newly generated binding sites.

In conclusion, we here present a novel model based on quantitative assessment of the substrate surface degradation during enzymatic cellulose hydrolysis. We saw that surface disintegration followed a sequence of steps: (1) *fissuring*: first of all, small fissures develop causing an increase of overall surface roughness. (2) *Limited diffusion*: next, hydrolysis slows down in the cracks having an aspect ratio that promotes diffusion limitation, i.e. small cracks in terms of width and length, due to two major facts: firstly, the increase in the local concentration of cellobiose which inhibits CBH action; and secondly, the blockade of enzymes clogging the narrow edges of the coniform pores. (3) *Constraint hydrolysis/ surface hydrolysis*: At this point, the hydrolysis rate on the surface outpaces the rate in the pores, conceivable by the observation that overall crack depth and width is leveling off or even decreasing in the small pores (Figure 9). Consequentially, hydrolysis proceeds on the surface,

where new fissures are constantly formed. Of course, hydrolysis also proceeds in the crevices having an aspect ratio that prevents mass transfer limitation, i.e. wide and long cracks. As a consequence of the various complex and individual mass transfer phenomena in cracks with different aspect ratios, we witness a morphology-induced acceleration and deceleration of the hydrolysis rate. (4) *Mass transfer regime / Hydrolysis slowdown*: ultimately, the cellulose surface is ragged completely, so that new fissures can hardly emerge and the existing clefts limit hydrolysis to an extent that slows down the overall hydrolysis rate. Due to the surface merely consisting of fissures, hydrolysis predominantly occurs in the coniform pores at this point. Thus, mass transfer limitation governs the hydrolysis rate. In contrast to previous theories stating that diffusion limitation cannot be the cause for the dramatic slowdown of hydrolysis at high degrees of substrate conversion (Bansal *et al.* 2009), we presume according to our observations that this is indeed the main cause. Overall, during the time we observed degradation, the formation of new interstices over time was constant, despite a decreasing overall number of cracks by merging events. It is conceivable that new fissures are also formed within the cracks, not only on the surface. Accordingly, we observed a permanent increase of surface area: the enzymes incessantly developed new surface. Moreover, corroborated by enzyme adsorption experiments, the vast majority of this newly created surface is accessible to celluloses (Figure 7B). Expanding on this idea, the fractal degrading of the substrate via elongated pores could lead to a final spongiform substrate where hydrolysis is strictly limited by multiple complex mass transfer phenomena. Figuring the degradation of cellulose proceeding via the formation of elongated coniform pores that strictly limit mass transfer by virtue of their geometry, it is plausible to conclude that hydrolysis is a function of the (ever-changing) aspect ratio of the interstices.

In their comprehensive modeling approach, Levine *et al.* (Levine *et al.* 2010) expressed that the hydrolysis rate decrease cannot solely be caused by limited surface area and crowding. To exactly model the drop in hydrolysis rate, they had to imply strong product inhibition and enzyme deactivation. They also stated that activity and synergistic interplay of celluloses are completely different at low or high surface area conditions, respectively. Strikingly, this is exactly what we observe in our study: while hydrolysis stalls in the crowded narrow pores and underlies severe limitations in terms of mass transfer and consequential product inhibition, hydrolysis on the surface of the substrate proceeds unhamperedly. Within these overcrowded pores, different scenarios are conceivable, such as enzyme deactivation, e.g. by loss of the ability to move or desorb; pore clogging, possibly by aggregate formation (Liu *et al.* 2009, Nigmatullin *et al.* 2004); diffusion limitation restricting enzyme access thereby causing severe product inhibition by accumulation of soluble cellooligomers and glucose as well as disruption of the synergistic interplay of celluloses; and severely hindered substrate access in so-called “dead-end pores”. However, as formation of pores that ultimately limit hydrolysis seems to be an intrinsic part of the mechanism of cellulose saccharification, this will be hard to circumvent.

1.4. Conclusion and outlook

Resolving the mechanism of enzymatic cellulose hydrolysis is a complex task whose accomplishment requires interdisciplinary endeavors at the frontier of biology, biotechnology, chemistry and physics. Combined efforts are required to resolve the long-remaining questions at the substrate as well at the enzyme side. Visualization of cellulase action will only be successful and lead to a better understanding of cellulose recalcitrance if combined with all the necessary methodology. Reviewing the variety of studies done on this matter is crucial to establish a common knowledge base.

Our AFM study on the structural dynamics of cellulase action presented in the second part unravels the complex mechanisms underlying the synergistic action of celluloses and may provide an explanation for the reason for the dramatic slowdown of hydrolysis during cellulose conversion that has been observed ever since (Bansal *et al.* 2009). By showing how the enzymes form deep cracks, we disrupt the paradigm of enzymatic cellulose degradation as a mere surface phenomenon (Arantes and Saddler 2010, Lee *et al.* 2000, Mansfield *et al.* 1999, Valjamae *et al.* 1998). On the other hand, we proved theories about various factors that were assumed to play a key role in the decrease of the hydrolysis rate, e.g. a gradual decrease of synergistic interplay, enzyme deactivation, decrease in substrate reactivity, and decrease in substrate accessibility (Bansal *et al.* 2009, Mansfield *et al.* 1999, Valjamae *et al.* 1998). With our study, we reconcile these theories. Because our substrate is extremely flat and resembles a natural substrates in its composition of amorphous and crystalline regions, it permits detailed analyses of very small structural changes during hydrolysis. We believe that it thus facilitates future research on the mechanisms of cellulose hydrolysis.

1.5. Abbreviations

AFM	atomic force microscopy
BC	bacterial cellulose
BGL	β -glucosidase
BMCC	bacterial microcrystalline cellulose
[BMIM]⁺	1-butyl-3-methylimidazolium cation
BMIMCl	1-butyl-3-methylimidazolium chloride
CBD	cellulose binding domains
CBH	cellobiohydrolase
CBM	carbohydrate binding module
CD	catalytic domain
CRM	confocal Raman microscopy
DMA/LiCl	<i>N,N</i> -dimethylacetamide/lithium chloride
DNA	deoxyribonucleic acid
DNS	3,5-Dinitrosalicylic acid
DP	degree of polymerization
DSC	differential scanning calorimetry
DTAF	dichlorotriazinylaminofluorescein
EDX	energy dispersive X-ray spectroscopy
EG	endoglucanase
EM	electron microscopy
ESEM	environmental scanning electron microscopy
EX	endoxylanase
FITC	fluorescein isothiocyanate
FPU	filter paper unit
GT	glycosyltransferase
HPLC	high pressure/performance liquid chromatography
HRSEM	high resolution scanning electron microscopy
HS-AFM	high speed AFM
immuno-EM	immunolabeling electron microscopy
KFM	Kelvin probe force microscopy
mAb	monoclonal antibody
LS	Langmuir-Schaefer
mAb	monoclonal antibody
MS	mass spectrometry
NC	nanocrystal cast
NDP	nucleoside diphosphate
NFC	nanofibrillar cellulose
NMMO	<i>N</i> -methylmorpholine- <i>N</i> -oxide
NMR	nuclear magnetic resonance
NR	non-reducing
pNPC	p-nitrophenyl cellobioside
QCM-D	quartz crystal microbalance with dissipation
RMS	root mean square
SAXS	small angle X-ray scattering
SC	spin-coated regenerated cellulose
SECM	scanning electrochemical microscope
SEM	scanning electron microscopy
SICM	scanning ion-conductance microscopy
SIMS	secondary ion mass spectrometry
SNFUH	scanning near field ultrasonic holography
SPM	scanning probe microscopy
STA	simultaneous thermal analysis
STEM	scanning transmission electron microscopy

STM	scanning tunneling microscope
TEM	transmission electron microscopy
TGA	thermogravimetric analysis
TIRF-M	total internal reflection fluorescence microscopy
TMSC	trimethylsilyl cellulose
UV	ultraviolet
XRD	X-ray diffraction

1.6. References

- Abuja, P. M., Pilz, I., Claeysens, M. & Tomme, P. Domain structure of cellobiohydrolase II as studied by small angle X-ray scattering: close resemblance to cellobiohydrolase I. *Biochem. Biophys. Res. Commun.* **156**, 180-185, (1988a).
- Abuja, P. M., Schmuck, M., Pilz, I., Tomme, P., Claeysens, M. & Esterbauer, H. Structural and functional domains of cellobiohydrolase I from *Trichoderma reesei*. *Eur. Biophys. J.* **15**, 339-342, (1988b).
- Adams, J. J., Currie, M. A., Ali, S., Bayer, E. A., Jia, Z. & Smith, S. P. Insights into higher-order organization of the cellulosome revealed by a dissect-and-build approach: crystal structure of interacting *Clostridium thermocellum* multimodular components. *J Mol Biol* **396**, 833-839, (2010).
- Arantes, V. & Saddler, J. N. Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. *Biotechnol. Biofuels* **3**, 4, (2010).
- Bansal, P., Hall, M., Realff, M. J., Lee, J. H. & Bommarius, A. S. Modeling cellulase kinetics on lignocellulosic substrates. *Biotechnol. Adv.* **27**, 833-848, (2009).
- Bommarius, A. S., Katona, A., Cheben, S. E., Patel, A. S., Ragauskas, A. J., Knudson, K. & Pu, Y. Cellulase kinetics as a function of cellulose pretreatment. *Metab. Eng.* **10**, 370-381, (2008).
- Chanzy, H., Henrissat, B. & Vuong, H. Colloidal gold labelling of 1,4- β -D-glucan cellobiohydrolase adsorbed on cellulose substrates. *FEBS Lett.* **172**, 193-197, (1984).
- Chanzy, H., Henrissat, B., Vuong, H. & Schülein, M. The action of 1,4- β -D-glucan cellobiohydrolase on *Valonia* cellulose microcrystals. An electron microscopy study. *FEBS Lett.* **153**, 113-118, (1983).
- Chen, H., Hayn, M. & Esterbauer, H. Purification and characterization of two extracellular beta-glucosidases from *Trichoderma reesei*. *Biochim. Biophys. Acta* **1121**, 54-60, (1992).
- Cui, S. T. Molecular self-diffusion in nanoscale cylindrical pores and classical Fick's law predictions. *J. Chem. Phys.* **123**, 054706, (2005).
- Cuissinat, C. & Navard, P. Swelling and dissolution of cellulose part 1: free floating cotton and wood fibres in *N*-methylmorpholine-*N*-oxide-water mixtures. *Macromol. Symp.* **244**, 1 - 18, (2006).
- Dagel, D. J., Liu, Y. S., Zhong, L., Luo, Y., Himmel, M. E., Xu, Q., Zeng, Y., Ding, S. Y. & Smith, S. In situ imaging of single carbohydrate-binding modules on cellulose microfibrils. *J Phys Chem B* **115**, 635-641, (2011).
- Divne, C., Stahlberg, J., Reinikainen, T., Ruohonen, L., Pettersson, G., Knowles, J. K., Teeri, T. T. & Jones, T. A. The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*. *Science* **265**, 524-528, (1994).
- Dominguez, R., Souchon, H., Spinelli, S., Dauter, Z., Wilson, K. S., Chauvaux, S., Beguin, P. & Alzari, P. M. A common protein fold and similar active site in two distinct families of beta-glycanases. *Nat Struct Biol* **2**, 569-576, (1995).
- Feng, L. & Zhong-Ian, C. Research progress on dissolution and functional modification of cellulose in ionic liquids. *J. Mol. Liq.* **142**, 1-5, (2008).
- Fidale, L. C., Possidonio, S. & El Seoud, O. A. Application of 1-allyl-3-(1-butyl)imidazolium chloride in the synthesis of cellulose esters: properties of the ionic liquid, and comparison with other solvents. *Macromolecular bioscience* **9**, 813-821, (2009).
- Froix, M. F. & Nelson, R. The interaction of water with cellulose from nuclear magnetic resonance relaxation times. *Macromolecules* **8**, 726-730, (1975).

- Garcia-Alvarez, B., Melero, R., Dias, F. M., Prates, J. A., Fontes, C. M., Smith, S. P., Romao, M. J., Carvalho, A. L. & Llorca, O. Molecular architecture and structural transitions of a *Clostridium thermocellum* mini-cellulosome. *J Mol Biol* **407**, 571-580, (2011).
- Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C., Jr. & Warren, R. A. Domains in microbial beta-1, 4-glycanases: sequence conservation, function, and enzyme families. *Microbiol Rev* **55**, 303-315, (1991).
- Heinze, T. & Koschella, A. Solvents applied in the field of cellulose chemistry - a mini review. *Polímeros* **15**, 84-90, (2005).
- Henrissat, B., Driguez, H., Viet, C. & Schülein, M. Synergism of cellulases from *Trichoderma reesei* in the degradation of cellulose. *Nature Biotechnology* **3**, 722-726, (1985).
- Himmel, M. E., Ding, S. Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W. & Foust, T. D. Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* **315**, 804-807, (2007).
- Igarashi, K., Koivula, A., Wada, M., Kimura, S., Penttila, M. & Samejima, M. High speed atomic force microscopy visualizes processive movement of *Trichoderma reesei* cellobiohydrolase I on crystalline cellulose. *J. Biol. Chem.* **284**, 36186-36190, (2009).
- Jarvis, M. Chemistry: cellulose stacks up. *Nature* **426**, 611-612, (2003).
- Kamavaram, V. & Reddy, R. G. Thermal stabilities of di-alkylimidazolium chloride ionic liquids. *International Journal of Thermal Sciences* **47**, 773-777, (2008).
- Kleywegt, G. J., Zou, J. Y., Divne, C., Davies, G. J., Sinning, I., Stahlberg, J., Reinikainen, T., Srisodsuk, M., Teeri, T. T. & Jones, T. A. The crystal structure of the catalytic core domain of endoglucanase I from *Trichoderma reesei* at 3.6 Å resolution, and a comparison with related enzymes. *J. Mol. Biol.* **272**, 383-397, (1997).
- Kurasin, M. & Valjamae, P. Processivity of cellobiohydrolases is limited by the substrate. *The Journal of biological chemistry* **286**, 169-177, (2011).
- Lamed, R. & Bayer, E. A. Cellulosomes from *Clostridium thermocellum*. *Meth. Enzymol.* **160**, 472-482, (1988).
- Lee, I., Evans, B. R. & Woodward, J. The mechanism of cellulase action on cotton fibers: evidence from atomic force microscopy. *Ultramicroscopy* **82**, 213-221, (2000).
- Lee, S. H., Doherty, T. V., Linhardt, R. J. & Dordick, J. S. Ionic liquid-mediated selective extraction of lignin from wood leading to enhanced enzymatic cellulose hydrolysis. *Biotechnol Bioeng* **102**, 1368-1376, (2009).
- Levine, S. E., Fox, J. M., Blanch, H. W. & Clark, D. S. A mechanistic model of the enzymatic hydrolysis of cellulose. *Biotechnol. Bioeng.* **107**, 37-51, (2010).
- Li, Z., Chu, L. Q., Sweedler, J. V. & Bohn, P. W. Spatial correlation of confocal Raman scattering and secondary ion mass spectrometric molecular images of lignocellulosic materials. *Anal. Chem.* **82**, 2608-2611, (2010).
- Linder, M., Mattinen, M. L., Kontteli, M., Lindeberg, G., Stahlberg, J., Drakenberg, T., Reinikainen, T., Pettersson, G. & Annala, A. Identification of functionally important amino acids in the cellulose-binding domain of *Trichoderma reesei* cellobiohydrolase I. *Protein. Sci.* **4**, 1056-1064, (1995).
- Liu, H., Fu, S., Zhu, J. Y., Li, H. & Zhan, H. Visualization of enzymatic hydrolysis of cellulose using AFM phase imaging. *Enzyme and Microbial Technology* **45**, 274-281, (2009).
- Liu, Y.-S., Luo, Y., Baker, J. O., Zeng, Y., Himmel, M. E. & Ding, S.-Y. in *Society of Photo-Optical Instrumentation Engineers (SPIE) Photonics West 2010* (San Francisco, CA, USA, 2010).
- Liu, Y. S., Baker, J. O., Zeng, Y., Himmel, M. E., Haas, T. & Ding, S. Y. Cellobiohydrolase hydrolyzes crystalline cellulose on hydrophobic faces. *J. Biol. Chem.* **286**, 11195-11201, (2011).
- Lynd, L. R., Laser, M. S., Bransby, D., Dale, B. E., Davison, B., Hamilton, R., Himmel, M., Keller, M., McMillan, J. D., Sheehan, J. & Wyman, C. E. How biotech can transform biofuels. *Nat. Biotechnol.* **26**, 169-172, (2008).
- Lynd, L. R., Weimer, P. J., van Zyl, W. H. & Pretorius, I. S. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* **66**, 506-577, (2002).
- Madkour, M. & Mayer, F. Structural organization of the intact bacterial cellulosome as revealed by electron microscopy. *Cell Biol Int* **27**, 831-836, (2003).

- Mais, U., Esteghlalian, A. R. & Saddler, J. N. Influence of mixing regime on enzymatic saccharification of steam-exploded softwood chips. *Appl. Biochem. Biotechnol.* **98-100**, 463-472, (2002).
- Mansfield, S. D., Mooney, C. & Saddler, J. N. Substrate and Enzyme Characteristics that Limit Cellulose Hydrolysis. *Biotechnol. Prog.* **15**, 804-816, (1999).
- Mattinen, M. L., Kontteli, M., Kerovuo, J., Linder, M., Annala, A., Lindeberg, G., Reinikainen, T. & Drakenberg, T. Three-dimensional structures of three engineered cellulose-binding domains of cellobiohydrolase I from *Trichoderma reesei*. *Protein Sci* **6**, 294-303, (1997).
- Mayer, F., Coughlan, M. P., Mori, Y. & Ljungdahl, L. G. Macromolecular Organization of the Cellulolytic Enzyme Complex of *Clostridium thermocellum* as Revealed by Electron Microscopy. *Applied and environmental microbiology* **53**, 2785-2792, (1987).
- McLean, B. W., Bray, M. R., Boraston, A. B., Gilkes, N. R., Haynes, C. A. & Kilburn, D. G. Analysis of binding of the family 2a carbohydrate-binding module from *Cellulomonas fimi* xylanase 10A to cellulose: specificity and identification of functionally important amino acid residues. *Protein Eng* **13**, 801-809, (2000).
- Morag, E., Halevy, I., Bayer, E. A. & Lamed, R. Isolation and properties of a major cellobiohydrolase from the cellulosome of *Clostridium thermocellum*. *J Bacteriol* **173**, 4155-4162, (1991).
- Nigmatullin, R., Lovitt, R., Wright, C., Linder, M., Nakari-Setälä, T. & Gama, M. Atomic force microscopy study of cellulose surface interaction controlled by cellulose binding domains. *Colloids Surf. B Biointerfaces* **35**, 125-135, (2004).
- Ohtani, H., Ishimura, S. & Kumai, M. Thermal decomposition behaviors of imidazolium-type ionic liquids studied by pyrolysis-gas chromatography. *Anal. Sci.* **24**, 1335-1340, (2008).
- Penttilä, P. A., Várnai, A., Leppänen, K., Peura, M., Kallonen, A., Jääskeläinen, P., Lucenius, J., Ruokolainen, J., Siika-aho, M., Viikari, L. & Serimaa, R. Changes in submicrometer structure of enzymatically hydrolyzed microcrystalline cellulose. *Biomacromolecules* **11**, 1111-1117, (2010).
- Pilz, I., Schwarz, E., Kilburn, D. G., Miller, R. C., Jr., Warren, R. A. & Gilkes, N. R. The tertiary structure of a bacterial cellulase determined by small-angle X-ray-scattering analysis. *The Biochemical journal* **271**, 277-280, (1990).
- Pinkert, A., Marsh, K. N., Pang, S. & Staiger, M. P. Ionic liquids and their interaction with cellulose. *Chem Rev* **109**, 6712-6728, (2009).
- Prasad, K., Kaneko, Y. & Kadokawa, J. Novel gelling systems of kappa-, iota- and lambda-carrageenans and their composite gels with cellulose using ionic liquid. *Macromol. Biosci.* **9**, 376-382, (2009).
- Quirk, A., Lipkowski, J., Vandenende, C., Cockburn, D., Clarke, A. J., Dutcher, J. R. & Roscoe, S. G. Direct visualization of the enzymatic digestion of a single fiber of native cellulose in an aqueous environment by atomic force microscopy. *Langmuir* **26**, 5007-5013, (2010).
- Quiroz-Castaneda, R. E., Martinez-Anaya, C., Cuervo-Soto, L. I., Segovia, L. & Folch-Mallol, J. L. Loosenin, a novel protein with cellulose-disrupting activity from *Bjerkandera adusta*. *Microb Cell Fact* **10**, 8, (2011).
- Reese, E. T., Sui, R. G. H. & Levinson, H. S. The biological degradation of soluble cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis. *J. Bacteriol.* **59**, 485-497, (1950).
- Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J. K. & Jones, T. A. Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. *Science* **249**, 380-386, (1990).
- Saloheimo, M., Paloheimo, M., Hakola, S., Pere, J., Swanson, B., Nyssonen, E., Bhatia, A., Ward, M. & Penttilä, M. Swollenin, a *Trichoderma reesei* protein with sequence similarity to the plant expansins, exhibits disruption activity on cellulosic materials. *European journal of biochemistry / FEBS* **269**, 4202-4211, (2002).
- Santa-Maria, M. & Jeoh, T. Molecular-scale investigations of cellulose microstructure during enzymatic hydrolysis. *Biomacromolecules* **11**, 2000-2007, (2010).
- Solomon, B. D., Barnes, J. R. & Halvorsen, K. E. Grain and cellulosic ethanol: History, economics, and energy policy. *Biomass and Bioenergy* **31**, 416 - 425, (2007).
- Swatloski, R. P., Spear, S. K., Holbrey, J. D. & Rogers, R. D. Dissolution of cellulose [correction of cellose] with ionic liquids. *J Am Chem Soc* **124**, 4974-4975, (2002).

- Valjamae, P., Sild, V., Pettersson, G. & Johansson, G. The initial kinetics of hydrolysis by cellobiohydrolases I and II is consistent with a cellulose surface-erosion model. *Eur. J. Biochem.* **253**, 469-475, (1998).
- van Roon, J. L., Arntz, M. M., Kallenberg, A. I., Paasman, M. A., Tramper, J., Schroen, C. G. & Beftink, H. H. A multicomponent reaction-diffusion model of a heterogeneously distributed immobilized enzyme. *Appl. Microbiol. Biotechnol.* **72**, 263-278, (2006).
- Varrot, A., Hastrup, S., Schulein, M. & Davies, G. J. Crystal structure of the catalytic core domain of the family 6 cellobiohydrolase II, Cel6A, from *Humicola insolens*, at 1.92 Å resolution. *The Biochemical journal* **337** (Pt 2), 297-304, (1999).
- White, A. R. & Brown, R. M. Enzymatic hydrolysis of cellulose: Visual characterization of the process. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 1047-1051, (1981).
- Wood, T. M. & McCrae, S. I. The cellulase of *Trichoderma koningii*. Purification and properties of some endoglucanase components with special reference to their action on cellulose when acting alone and in synergism with the cellobiohydrolase. *The Biochemical journal* **171**, 61-72, (1978).
- Zhang, Y. H. & Lynd, L. R. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. *Biotechnol. Bioeng.* **88**, 797-824, (2004).
- Zhang, Y. H. & Lynd, L. R. Regulation of cellulase synthesis in batch and continuous cultures of *Clostridium thermocellum*. *J. Bacteriol.* **187**, 99-106, (2005).
- Zhao, H., Jones, C. L., Baker, G. A., Xia, S., Olubajo, O. & Person, V. N. Regenerating cellulose from ionic liquids for an accelerated enzymatic hydrolysis. *J. Biotechnol.* **139**, 47-54, (2009).
- Zheng, Y., Pan, Z., Zhang, R. & Jenkins, B. M. Kinetic modeling for enzymatic hydrolysis of pretreated creeping wild ryegrass. *Biotechnol. Bioeng.* **102**, 1558-1569, (2009).
- Zhong, L., Matthews, J. F., Hansen, P. I., Crowley, M. F., Cleary, J. M., Walker, R. C., Nimlos, M. R., Brooks, C. L., 3rd, Adney, W. S., Himmel, M. E. & Brady, J. W. Computational simulations of the *Trichoderma reesei* cellobiohydrolase I acting on microcrystalline cellulose Ibeta: the enzyme-substrate complex. *Carbohydr. Res.* **344**, 1984-1992, (2009).
- Zhu, S., Wu, Y., Chen, Q., Yu, Z., Wang, C., Jin, S., Ding, Y. & Wu, G. Dissolution of cellulose with ionic liquids and its application: a mini-review. *Green Chemistry* **8**, 325-327, (2006).
- Zverlov, V. V., Klupp, M., Krauss, J. & Schwarz, W. H. Mutations in the scaffoldin gene, cipA, of *Clostridium thermocellum* with impaired cellosome formation and cellulose hydrolysis: insertions of a new transposable element, IS1447, and implications for cellulase synergism on crystalline cellulose. *J. Bacteriol.* **190**, 4321-4327, (2008).
- Zverlov, V. V. & Schwarz, W. H. Bacterial cellulose hydrolysis in anaerobic environmental subsystems--*Clostridium thermocellum* and *Clostridium stercorarium*, thermophilic plant-fiber degraders. *Ann. N. Y. Acad. Sci.* **1125**, 298-307, (2008).

Chapter Two
**Finding the “synthase switch”
in *Schizophyllum commune*
trehalose phosphorylase**

2.1. Introduction

“Glycosyltransferases are central to all synthetic processes involving carbohydrates”

Davies, GJ. *Nature Structural Biology* 8, 98-100 (2001)

Carbohydrates are fundamental for life and ubiquitous in nature: as nutrients, osmolytes, cell wall scaffolding biopolymer (i.e. cellulose), in cellular energy storage (i.e. starch) and signaling pathways, in glycoproteins and glycolipids. Therefore, enzymes involved in formation and breakdown of glycosidic bonds – so-called “carbohydrate-active enzymes” (CAZymes) – are of great importance and of high scientific as well as industrial and medical interest (Breton *et al.* 2006, Davies *et al.* 2005). The majority of glycosidic bonds in nature is formed by glycosyltransferases (GTs), an enzyme class that catalyzes the transfer of a glycosyl moiety from an activated donor sugar to a specific acceptor and ties a glycosidic bond between them (Figure VIII). Activated GT donor sugars feature a good leaving group, and are for example sugar phosphates, nucleoside diphosphate (NDP) sugars, and lipid-phospho sugars. With respect to their reaction mechanism, GTs can either retain or invert the anomeric configuration. According to their mechanism, they are classified retaining or inverting GTs (Breton *et al.* 2006, Davies 2001). Named after Nobel laureate Luis F. Leloir, who discovered the first sugar nucleotide GTs using nucleotide sugars as donors, they are also termed Leloir enzymes (Lairson *et al.* 2008).

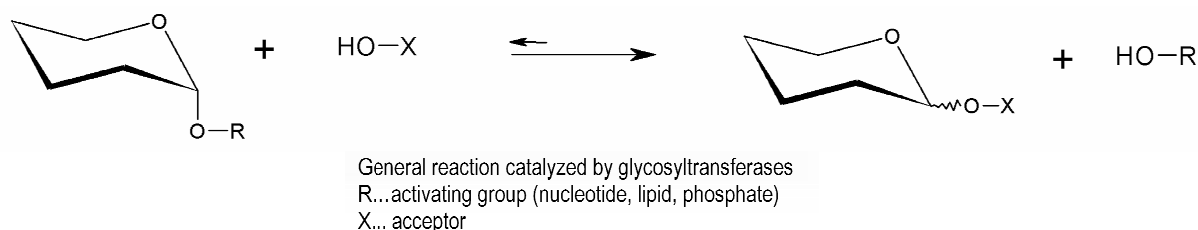


Figure VIII. General reaction scheme of glycosyltransferases (GTs). Note that the reaction equilibrium usually is on the right side (synthesis direction) with the exception of phosphorylases which employ sugar phosphates as donors. In phosphorylases, the degradation half reaction (“phosphorolysis” direction) is preferred.

Regarding to their sequence, GTs have been assigned to 94 different families (as with July, 2011)⁶ (Davies 2001, Davies *et al.* 2005, Lairson *et al.* 2008). It has been demonstrated that the sequence based family classification is very useful in grouping GTs as the overall fold and the reaction mechanism – inverting or retaining – can be safely deduced from it for all members of the family (Davies and Henrissat 1995). GTs which have been classified in a family all take on the same fold, however, substrate specificity, function and origin can vary greatly (Breton *et al.* 2006).

Crystallography-based structural information on GTs is scarce due to general difficulties in sufficient expression levels, purity, stability and crystallization of these enzymes (Breton *et al.* 2006).

⁶ A comprehensive database of CAZymes can be found at www.cazy.org.

Additionally, modeling the structure of a certain GT with the crystal structure of a family member with low sequence identity is often unsatisfactory. This, in turn, makes rational design very difficult. GTs can take on three different folds called GT-A-, GT-B-, and GT-C-fold. All of the three folds feature $\alpha/\beta/\alpha$ sandwiches and strongly resemble each other. The two main folds, which are also found in nucleotidyltransferases and sugar epimerases, are GT-A and GT-B. All families adopting the GT-A fold have a common DxD motif which interacts with the donor phosphate group via a coordinated divalent cation (usually Mn^{2+}) which is required for catalytic activity. The rare GT-C fold (GT42 family) is similar to the GT-A fold but lacks the DxD motif (Breton *et al.* 2006). The general fold similarity among a family makes predictions about the secondary structure a bit easier, even if sequence identity is low.

While the reaction mechanism of inverting GTs is well known, the mechanism of retaining GTs is still elusive. Inverting GTs employ a S_N2 -type mechanism in which in the first place, a general base abstracts a proton from the incoming nucleophile of the acceptor. Via direct nucleophilic displacement of the leaving group, the anomeric configuration is then inverted in a single step. Retaining glycosyltransferases, by contrast, presumably employ a double displacement mechanism where an active site nucleophile stabilizes the first of two oxocarbenium ion-like transition states. Then, a covalent glycosyl-enzyme-intermediate is formed. In the next step, the second oxocarbenium ion-like transition state is formed through activation of the acceptor hydroxyl by a general base that eventually attacks the covalent intermediate as a nucleophile. However, evidence to prove the S_N2 -like mechanism for retaining GTs, for example isolation of a covalent intermediate, has not been put forth yet. Another mechanism proposed for retaining GTs is an S_Ni -like or “internal return”-like mechanism. This type of reaction mechanism would proceed via a single transition state. It would require positioning of donor and acceptor in close proximity to each other to allow for leaving group departure and nucleophilic attack in a synchronous manner. It is, however, very difficult to prove this reaction type as no covalent intermediate is formed here (Goedl *et al.* 2006, Lairson and Withers 2004).

The general importance of complex carbohydrates in a wide variety of relevant biological processes as well as potential drugs makes GTs a target of high interest for diverse tasks requiring glycosidic bond formation. In order to craft designer carbohydrates, a wide repertoire of GTs with tailored specificities is required. Chemical synthesis of complex carbohydrates is extremely difficult and involves multiple steps with disappointing yields (Aharoni *et al.* 2006, Kittl and Withers 2010). The term “glycorandomization” describes the requirement for flexible enzymes which are able to synthesize complex carbohydrates from a variety of building blocks. More precisely, the enzymes need to be promiscuous in their substrate specificity. This task is of special importance for the production of therapeutically or industrially relevant products (Park *et al.* 2009). As GTs are highly specific enzymes and crystallographic data is rare, directed evolution is widely employed to improve substrate promiscuity. However, due to the lack of structural information and, more importantly, of

effective high-throughput screenings, increasing the promiscuity of GTs is still a tedious process (Kittl and Withers 2010). Another semi-rational approach which led to some success is the building of chimeras (“domain swapping”): as the acceptor- (N-terminus) and the donor binding site (C-terminus) in GT-B fold GTs are located in separate domains, the combination of two different domains can alter the substrate specificity (Chang 2009, Park *et al.* 2009).

2.1.1. Trehalose phosphorylase from *Schizophyllum commune*

Phosphorylases are a class of its own within the CAZYmes, with some of them counted to the glycosidases and some to the GTs. They use sugar phosphates as donors and release phosphate as a leaving group (Lairson and Withers 2004). The degradation half-reaction, termed phosphorolysis, is usually favored in contrast to synthases using NDP-sugar donors. Phosphorylases are found throughout several families and are proposed to be the evolutionary link between glycosidases and GTs. The greater part of them are classified as glycosidases (e.g. sucrose phosphorylase, family GH13) but some are counted among GTs (Eis *et al.* 2001, Goedl *et al.* 2008, Lairson and Withers 2004). Trehalose phosphorylase from *Schizophyllum commune* (*ScTPase*⁷, EC 2.4.1.231) belongs to the GT4 family. The members of this family are retaining GTs which take on the GT-B fold and the majority of them are NDP-sugar dependent synthases. However, some trehalose phosphorylases are among them, e.g. *ScTPase* which catalyzes synthesis and phosphorolysis of α,α -trehalose. The preferred physiological reaction direction of *ScTPase* is phosphorolysis which yields α -D-glucose 1-phosphate (G1P) and α -D-glucose (Gluc). An internal return like S_Ni mechanism has been proposed for *ScTPase* (Goedl *et al.* 2006).

The non-reducing disaccharide trehalose (α -D-glucopyranosyl- α' -D-glucopyranoside) is used by prokaryotes as well as eukaryotes as a carbon source, storage compound, or protection against environmental stress (Argüelles 2000). Trehalose has many favorable properties, such as a highly hydrophilic character, peculiar chemical stability and no internal hydrogen bonds (Argüelles 2000). It is a so called “compatible solute” which denotes a heterogeneous group of small organic molecules serving as natural protectants of cells against stress. These compatible solutes are, for example, able to stabilize protein structures or to protect from hypertonic stress by acting as osmolytes. Chemically, many of them are based on glycosidic structures (Sawangwan *et al.* 2010).

2.1.2. In quest for the “synthase switch”

Bacterial trehalose phosphorylases employ an inverting mechanism and are classified into family GH65. By contrast, *ScTPase* along with some other trehalose phosphorylases of fungal origin are retaining enzymes which have been grouped in the GT4 family. This family mainly consists of synthases of various origins and different substrate preferences. They all share the GT-B fold. Among

⁷ *ScTPase*, GenBank accession number ABC84380.1

them, also trehalose phosphorylases are found, which suggests a close evolutionary relation between these phosphorylases and the synthases. It is indeed a very interesting question what exactly distinguishes a synthase from a phosphorylase: which residues, for example, are responsible for the difference in donor specificity between a GT4 trehalose synthase and a GT4 trehalose phosphorylase? In contrast to trehalose synthases, *ScTPase* synthesizes trehalose from G1P and glucose but not from NDP-donor sugars, such as uridine 5'-diphospho- (UDP) glucose (see Figure IX). In this study we investigate the determinants of substrate specificity in trehalose synthases and trehalose phosphorylases.

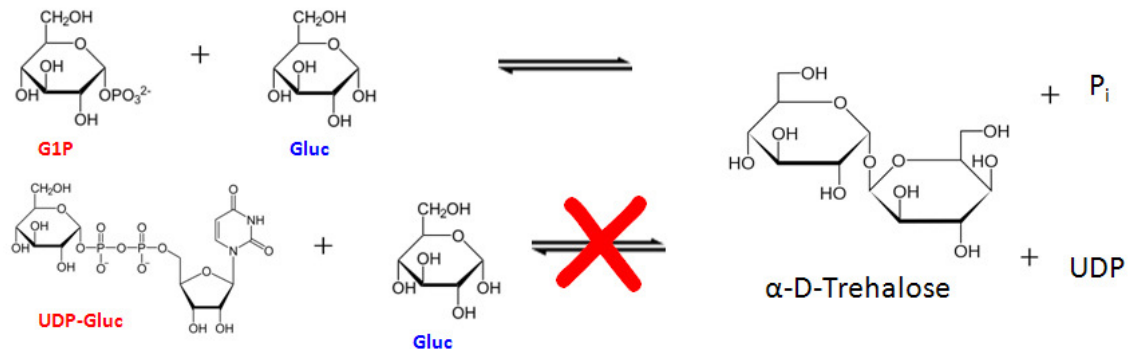


Figure IX. Reaction catalyzed by *ScTPase* and reaction of a trehalose synthase which is not catalyzed by *ScTPase*.

Active-site residues crucial for activity are highly conserved among GTs in the fold GT-B family (Goedl *et al.* 2008). This is also the case for GT4 family members which take on the GT-B fold and show a high degree of conservation in the residues responsible for substrate binding. Another interesting fact is that the GT4 trehalose synthase from the hyperthermophilic archeon *Pyrococcus horikoshii*⁸ (*PhTreT*) does not only share high sequence similarity with the GT4 trehalose synthases, but also with GT4 trehalose phosphorylases (Ryu *et al.* 2005). For example, it shares 77 % overall sequence identity with the trehalose synthase from *Pyrococcus furiosus*⁹ (*PfTs*), but still 28 % with *ScTPase*, and 26 % with a trehalose phosphorylase from *Grifola frondosa*¹⁰ (*GfTPase*) (Ryu *et al.* 2005).

To elucidate which residues are responsible for a synthase-like or a phosphorylase-like activity, mutagenesis of relevant residues is the method of choice. To rationally design such mutants, structural information is required to some extent. However, until now (July 2011), only 13 crystal structures have been published within the GT4 family, which has more than 17000 members based on the sequence classification. The resolved structures reveal the typical GT-B fold domain organization with the active site located in a cleft between two Rossmann-fold domains (Goedl *et al.* 2008). For *ScTPase*, no crystal structure is available so far. This restricts rational design to sequence alignments

⁸ *PhTreT*, GenBank accession number BAA30133.1

⁹ *PfTs*, GenBank accession number AAL81866.1

¹⁰ *GfTPase*, GenBank accession number BAA31350.1

and homology modeling, which is expected to generate a basis for structural predictions. Due to the high degree of structural conservation with GTs in general and the subfamilies in particular, homology modeling should render reliable rational design possible.

However, usually, if no crystal structure is available and rational design thus fails, directed evolution leads to success. Screening assays for GTs are hard to design and an effective assay suitable for high-throughput screening (HTS) is still the bottleneck in the directed evolution of most GTs. Direct measurement of enzymatic activity would be highly desirable in a HTS assay. However, it is not possible to determine GT activity directly using fluorescence or absorbance as no change in these parameters occurs during cleavage or formation of the glycosidic bond. Hence, assaying glycosyl transfer requires coupling in some way. In the following chapters, our approaches towards rational and random engineering of *ScTPase* towards synthase activity are described and our results in assay development are discussed.

2.2. Rational approach

2.2.1. Background and strategy

The recombinant *ScTPase* gene has 2247 base pairs (bp) and the full-length protein has 748 amino acids (aa) including its N-terminal His-tag and a calculated molecular mass of 82.8 kDa. According to size-exclusion experiments, its main active form is a dimer (Goedl *et al.* 2006). Interestingly, the *PhTreT* also occurs as a dimer in solution as well as in its crystallized form (Woo *et al.* 2010). In contrast to the original fungal *ScTPase* where 1 mol Mg^{2+} was bound per mol of *ScTPase* monomer, the recombinant enzyme did not contain significant amounts of Mg^{2+} . During storage at 4 °C, the recombinant *ScTPase* breaks into two parts giving rise to a shorter, catalytically inactive N-terminal part (approximately 16 kDa) and a longer C-terminal part (approximately 65 kDa) which remains fully active (Eis *et al.* 2001, Goedl *et al.* 2006). Various attempts to produce the sole truncated C-terminus recombinantly in *E. coli* failed: expression resulted in plenty of inclusion body formation. The molecular mass of the original fungal *ScTPase* was determined with 60 kDa, which probably comprises the entire C-terminal part of the recombinant enzyme (Eis *et al.* 2001, Goedl *et al.* 2006).

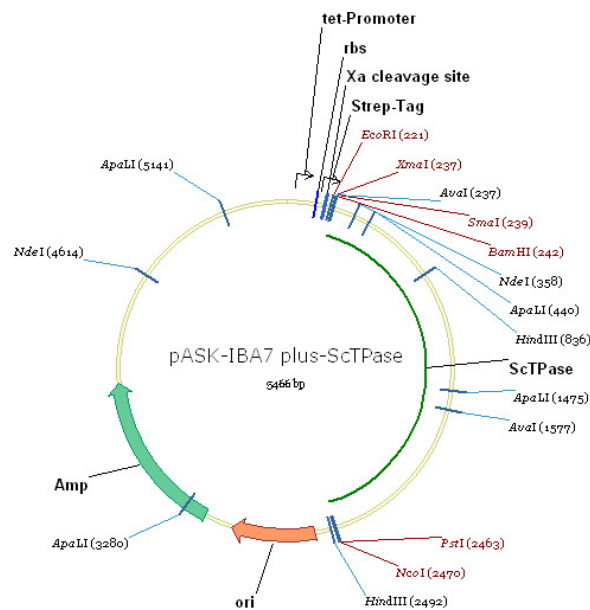


Figure X. Vector map of pASK-IBA7plus with *ScTPase* WT.

Originally, the recombinant *ScTPase* wild type (WT) was cloned into a pQE30 vector. Purification of the full-length protein via His-tag was not successful due to the enzyme's inability to sustain activity in elevated imidazole concentrations upon elution. Thus, the *ScTPase* was purified from crude extracts using anion exchange chromatography with a subsequent desalting step and a final hydrophobic interaction chromatography step as described previously (Goedl *et al.* 2006). To simplify purification and to obtain highly pure *ScTPase*, we decided to subclone the gene into a commercial pASK-IBA7plus vector (see Figure X). This expression vector provided the protein with an N-terminal Strep-

tag for convenient affinity purification with a Strep-Tactin matrix and elution with desthiobiotin (Schmidt and Skerra 2007). Strep-tag purification is a gentle method yielding highly pure protein in a one-step purification.

Unfortunately, no crystal structure of *ScTPase* is available. Crystallization of GTs is generally a troublesome task due to difficulties in sufficient expression levels, purity, stability and crystallization of these enzymes (Breton *et al.* 2006). Anyway, as GTs of the same family all take on the same fold – in case of the GT4 family this is the GT-B fold – a rational approach based on homology modeling seems feasible. Our rational approach presented here includes extensive sequence comparisons, secondary structure predictions, and homology modeling, which provided the basis to identify targets for site-directed mutagenesis. After construction and expression of the mutants, we screened them for the target activity.

2.2.2. Results and discussion

Construction, expression and purification of the Strep-tagged ScTPase

We determined the kinetic parameters for the Strep-tagged WTII *ScTPase* and compared them with the previously published values for the fungal and recombinant His-tagged *ScTPase* (Table ii). As can be deduced from the negligible differences, these enzymes are comparable in their kinetic behavior.

Table ii. Kinetic parameters of fungal, His-tagged, and Strep-tagged *ScTPase*. The parameters for the fungal and the His-tagged enzyme are from Goedl *et al.* (2006).

			fungal <i>ScTPase</i>	His-tagged <i>ScTPase</i>	Strep-tagged <i>ScTPase</i>
Phosphorolysis	direction	K_m phosphate [mM]	1.4	0.8	0.85
		K_m trehalose [mM]	71	53	62.2
		k_{cat} [s^{-1}]	13.3	16.4	14.2
		k_{cat}/K_m phosphate [$mM^{-1}s^{-1}$]	9.5	21	16.7
		k_{cat}/K_m trehalose [$mM^{-1}s^{-1}$]	0.19	0.31	0.23
Synthesis	direction	K_m glucose [mM]	35	46	67.9
		K_m G1P [mM]	5.8	2.2	1.81
		k_{cat} [s^{-1}]	9.5	14.1	3.05
		k_{cat}/K_m glucose [$mM^{-1}s^{-1}$]	0.27	0.31	0.045
		k_{cat}/K_m G1P [$mM^{-1}s^{-1}$]	1.6	6.4	1.35

Sequence analysis and homology modeling of the ScTPase gene

The GT4 member *PhTreT* catalyzes trehalose synthesis from UDP-glucose and glucose. It is also able to use GDP-glucose and adenosine 5'-diphospho- (ADP) glucose as glycosyl donors. Although *PhTreT* generally has a greater sequence identity with GT4 trehalose synthases and phosphorylases, its C-terminal donor binding domain shows high identity to that of trehalose 6-phosphate synthases (Ryu *et al.* 2005). In addition to that, the donor binding site residues are strictly conserved with the residues

described in a trehalose 6-phosphate synthase from the GT20 family whose crystal structure is available (Gibson *et al.* 2002, Ryu *et al.* 2005). The GT20 family also consists of retaining GT-B fold GTs. Primary sequence alignment with several GT4 trehalose phosphorylases and trehalose synthases revealed that all these residues which were described as essential for NDP-donor sugar binding are also strictly conserved in these enzymes with the only exception of one single leucine which was substituted by a valine or – in one case – an isoleucine in the phosphorylases. This substitution is very unlikely to be a determinant for NDP-preference due to the similar nature of these amino acids and the fact that only the backbone of the leucine 365 is involved in interactions with the NDP moiety (see Figure XI). This was observed in the crystal structure of the *E. coli* trehalose 6-phosphate synthases (OtsA) (Gibson *et al.* 2002).

Table iii. GTs used for sequence comparison. ScTPase was aligned with the following GTs which were chosen either because of similarity in the sequence or the activity. G1P – glucose 1-phosphate; n.c. – not classified; Gluc – glucose; UDP-GlcNAc – UDP-*N*-acetyl-D-glucosamine; Donors in italics indicate a putative donor.

Name	Organism	Acronym	Family	Donor	pdb code
trehalose phosphorylase	<i>Grifola frondosa</i>	<i>GfTPase</i>	GT4	G1P	
hypothetical protein	<i>Coprinopsis cinerea</i>	<i>CcTPase</i>	n.c.	<i>G1P</i>	
trehalose phosphorylase	<i>Pleurotus pulmonarius</i>	<i>PpTPase</i>	GT4	<i>G1P</i>	
trehalose phosphorylase	<i>Lentinus sajor-caju</i>	<i>PsTPase</i>	GT4	G1P	
trehalose phosphorylase	<i>Schizophyllum commune</i>	<i>ScTPase</i>	GT4	G1P	
trehalose synthase	<i>Pyrococcus horikoshii</i>	<i>PhTreT</i>	GT4	UDP-Gluc	2X6Q, 2X6R, 2XA1, 2XA2, 2XA9, 2XMP
trehalose synthase	<i>Pyrococcus furiosus</i>	<i>PfTs</i>	GT4	<i>UDP-Gluc</i>	
putative trehalose synthase	<i>Staphylothermus marinus</i>	<i>SmTs</i>	GT4	<i>UDP-Gluc</i>	
trehalose synthase	<i>Pyrobaculum aerophilum</i>	<i>PaTs</i>	GT4	<i>UDP-Gluc</i>	
1L-myo-inositol-1P- α - <i>N</i> -acetylglucosaminyltransferase	<i>Corynebacterium glutamicum</i>	MshA	GT 4	UDP-GlcNAc	3C48, 3C4Q, 3C4V
sucrose phosphate synthase	<i>Halotheothrix orenii</i>	SpsA	GT4	UDP-Gluc	2R60, 2R66, 2R68
glycogen synthase	<i>Pyrococcus abyssi</i>	<i>PaGlgA</i>	GT5	UDP/ADP-Gluc	2BFW, 2BIS
glycogensynthase 1	<i>Agrobacterium tumefaciens</i>	GlgA	GT 5	ADP-Gluc	1RZU, 1RZV
L-glycero-D-manno-heptose-II α -1,3-glucosyltransferase I	<i>Escherichia coli</i>	WaaG	GT 4	UDP-Gluc	2IV7, 2IW1
phosphatidylinositol mannosyltransferase	<i>Mycobacterium smegmatis</i>	PimA	GT 4	GDP-Mannose	2GEJ, 2GEK
eurekanate-attachment enzyme	<i>Streptomyces viridochromogenes</i>	AviGT4	GT4	<i>unknown</i>	2IUY, 2IV3
trehalose 6-phosphate synthase	<i>Escherichia coli</i>	OtsA	GT20	UDP-Gluc	1GZ5, 1UQT, 1UQU

Interestingly, in a study on the phosphate subsite of ScTPase, Arg507 and Lys512 were identified as part of a conserved motif (Arg507-XXXX-Lys512) pivotal for binding recognition of the phosphate group (Goedl and Nidetzky 2008). Overall, this shows how closely related GT4 phosphorylases and synthases are and that ScTPase – in principle – features all the residues which have been identified as key for NDP-binding. The EX₇E-motif (right column in Figure XIA) is a conserved donor-binding motif among GTs featuring an asp or Glu at the first position in case of the

families GT4 and GT20 (Rosen *et al.* 2004). It is involved in contacts with the glycosyl moiety of the donor as well as with the leaving group. The second Glu often establishes interactions with the ribose of NDP-donors (Troutman and Imperiali 2009). In OtsA, the backbone of the Leu interacts with the second phosphate of the NDP-donor.

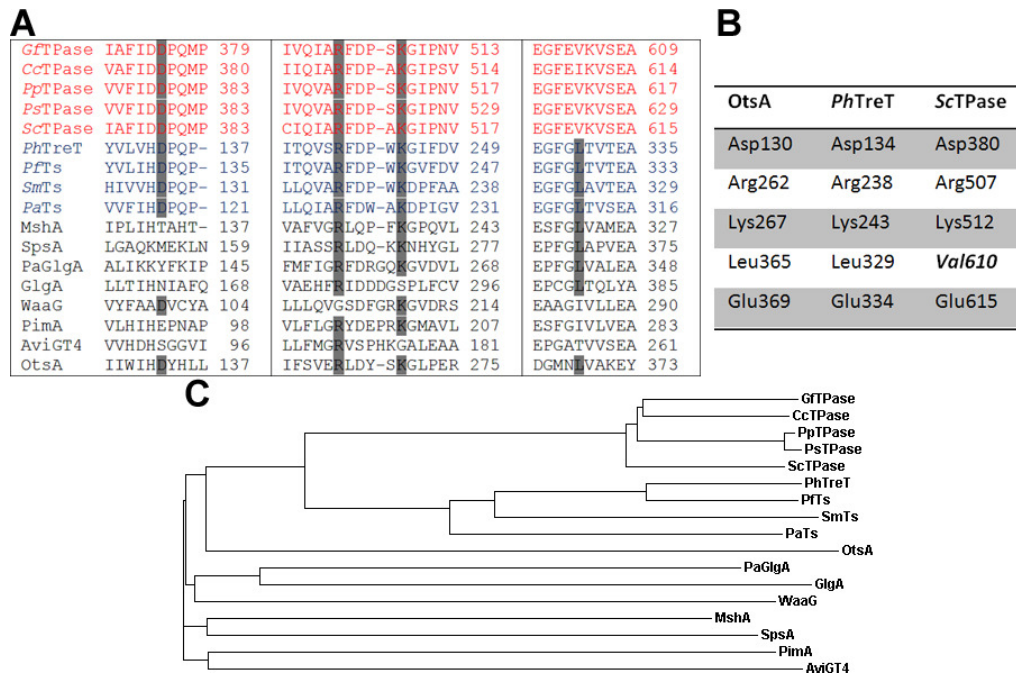


Figure XI. A: Sequence alignment of (putative) donor-binding regions in selected trehalose phosphorylases, trehalose synthases, and trehalose 6-phosphate synthases. For explanation of the acronyms see Table iii. Red letters – trehalose phosphorylases; blue letters: trehalose synthases; black letters: other transferases employing NDP-donors. B: Residues in *PhTreT* and *ScTPase* which correspond to the UDP-binding residues found in the crystal structure of *OtsA*. C: Phylogenetic tree of the compared enzymes.

We employed primary sequence alignment and secondary structure prediction with various family members before choosing the structure most appropriate as a basis for the *ScTPase* model (see Table iii). Automated secondary structure prediction and alignment with *psipred* and *pGenTHREADER*¹¹ offered the 1L-myo-inositol-1P- α -*N*-acetylglucoaminyltransferase from *Corynebacterium glutamicum* (*MshA*) as the best candidate to model the N-terminal *ScTPase* shortcut (aa 277–737) with a score of 128 (Figure XII) (Bryson *et al.* 2005, Jones 1999, Lobley *et al.* 2009). The N-terminal part was not predicted due to low sequence identity and the fact that it is cleaved off and not pivotal for *ScTPase* activity. *MshA* shares only 17.8 % of sequence identity with *ScTPase* but uses an UDP-glycosyl donor.

In the next step, we generated a model based on this secondary structure alignment using *Modeller 9v4*¹² (Eswar *et al.* 2006). The result of the modeling is shown in Figure XIII. Although the GT-B fold is conserved among the family and the two Rossmann-fold domains and the interdomain

¹¹ <http://bioinf.cs.ucl.ac.uk/psipred/>, last visited 07/28/2011

¹² <http://salilab.org/modeller/>, last visited 07/28/2011

cleft are nicely modeled, many large loops are visible. These loops are probably a result of the low overall sequence identity making the secondary structure prediction and alignment unreliable. Many of the loopy regions have most likely not been modeled adequately. However, the model provides us with a preliminary look on what *ScTPase* might look like and the highly conserved substrate binding and active site regions are very likely aligned to each other hence allowing to draw rational conclusions.

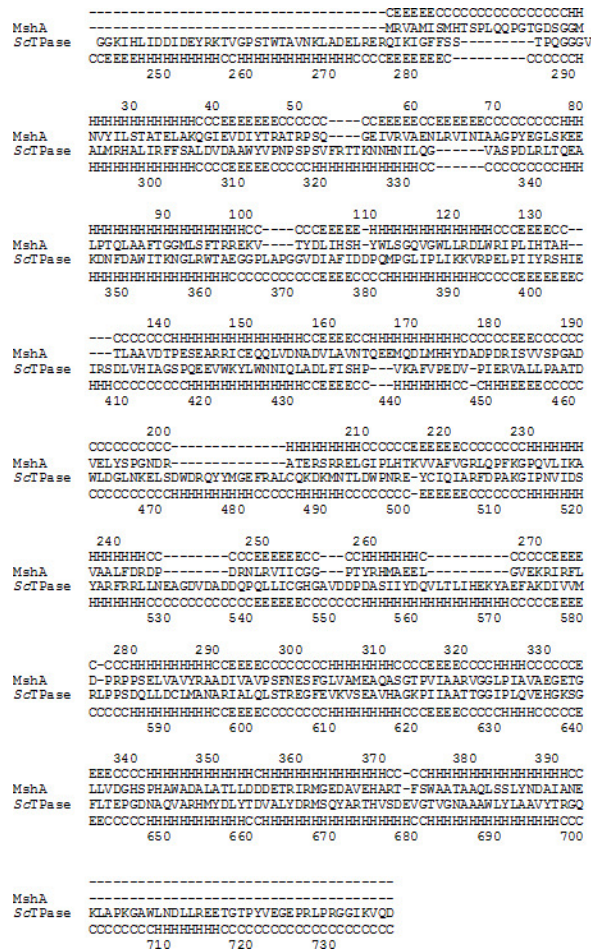


Figure XII. Pspired secondary structure prediction and alignment. C denotes coil, E denotes a β -sheet, H denotes a helix.

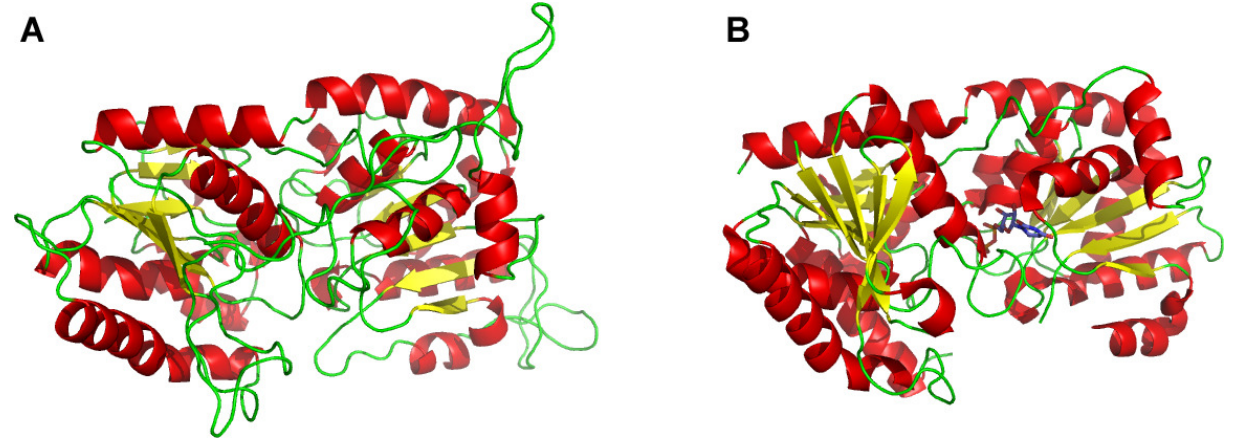


Figure XIII. Model of *ScTPase* (A) based on the crystal structure of *MshA* chain B (B, pdb code 3C4Q).

GfTPase	IYDQVLQLIH-AKYKEYAPDIVVMRCPSPDQLLNTM ³¹ ANAKFALQLSTR-EGFEVKVSEA 609
CcTPase	IYDQTMQLIASAPYNEYANDIVVMRLPSPDQLLNALM ³¹ NSRIALQLSIR-EGFEIKVSEA 614
PpTPase	IYDQIMALVNSDPYKEYAHDIVVMRLPSPDELLNANM ³¹ ANSRIALQLSTR-EGFEVKVSEA 617
PsTPase	IYDQIMALVNSDPYKEYAHDIVVMRLPSPDELLNANM ³¹ ANSRIALQLSTR-EGFEVKVSEA 629
ScTPase	IYDQVLTLIH-EKYAEFAKDIVVMRLPSPDQLLDCM ³¹ ANARIALQLSTR-EGFEVKVSEA 615
PhGT	YFEKTLRKIG----EDYDVKVLTNLIGVHAREVNAF ³¹ QRASDVILQMSIR-EGFG ³¹ GLTVTEA 335
PfTs	YFEKTLRKIG----EDYDIKVLTNLTGVHAREVNAF ³¹ QRASDVILQMSIR-EGFG ³¹ GLTVTEA 333
SmTs	YYEKTLYAG----MDEDIFILTNLKNVGTLEVNAF ³¹ QRAATVVLQMSRR-EGFGLAVTEA 329
PaTs	VYKKTVEAAG----NDRDIHLLM-LPPDSHIEVNAF ³¹ QRAATVVMQKSIR-EGFG ³¹ GLTVSEA 316

Figure XIV. Partial sequence alignment of trehalose phosphorylases and trehalose synthases. Red: conserved methionine/glutamine; blue: EX₇E-motif.

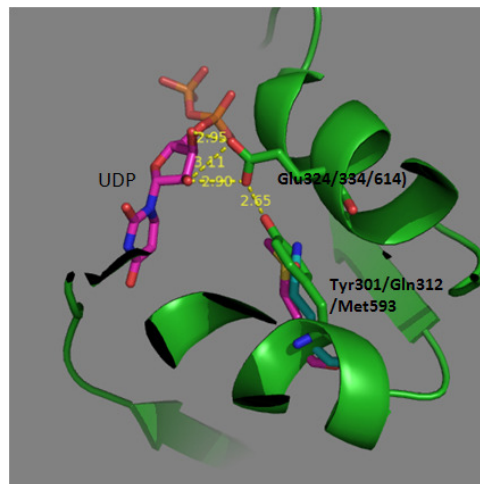


Figure XV. Donor binding site overlay of the regions of interest of MshA, PhTreT and ScTPase. The UDP is bound to the crystal structure of MshA and models of PhTreT and ScTPase were aligned to the MshA structure.

As suggested by our model (Figure XV) the conserved glutamine of the trehalose synthases – which is substituted by a tyrosine in case of MshA – could play a key role in positioning of the Glu324 (614 in ScTPase numbering). This glutamate is part of the EX₇E-motif which is a known conserved motif in various glycosyltransferases (Absmanner *et al.* 2010). This motif is highly involved in interactions with the donor substrate via hydrogen bonding (Figure XV) but also pivotal for catalysis in many GTs. Mutational studies on the EX₇E-motif of other GT4 enzymes showed that the first glutamate is involved in catalysis and that its substitution, e.g. by alanine, leaves the mutant catalytically inactive (Abdian *et al.* 2000, Woo *et al.* 2010). In the PhTreT crystal structure, the first glutamate (E326) also interacts with the two phosphate moieties of the UDP-glucose (Woo *et al.* 2010). There is evidence from docking studies, mutational studies, and from crystal structure suggesting that the second glutamate of the EX₇E-motif is involved in hydrogen bonding with the ribose of the NDP-sugar in various synthases (Abdian *et al.* 2000, Muniz *et al.* 2004, Woo *et al.* 2010). Interestingly, the EX₇E-motif is also highly conserved in the GT4 family phosphorylases which rather employ G1P as donor substrate (Figure XIV). This raises the question, what is the role of the second glutamate of the EX₇E-motif in phosphorylases? And, how can this highly conserved motif participate in binding of two different donor substrates?

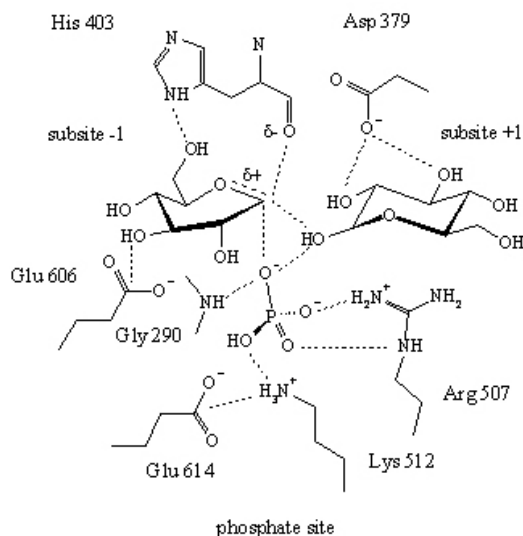


Figure XVI. The hypothetical catalytic center of *ScTPase* (image courtesy of C. Luley-Goedl).

In order to elucidate the role of the second glutamate (E614) of the EX₇E-motif in *ScTPase* (E₆₀₆X₇E₆₁₄) we mutated this position to alanine, glutamine, and lysine. The results of the kinetic characterization of phosphorolysis showed that the presence of the positively charged, long side chain of lysine affected activity and affinity towards phosphate most. The E614A and the E614Q mutant showed significantly reduced phosphorolysis activity and an increase in K_m by approximately 10-fold (Table iv). It is conceivable that the glutamine of E614Q possibly interferes with the correct positioning of the phosphate group in the transition state. Generally, the impact of the E614 mutations on k_{cat} and the K_m of phosphate suggest that this residue is important for catalysis. However, in phosphorylases, the second glutamate probably might rather play a role in correct positioning of the phosphate group in the transition state. This is in contrast to synthases where the second glutamate of the EX₇E-motif maintains hydrogen bonds to the ribose moiety of an NDP or NDP-sugar.

Table iv. E614 mutants of *ScTPase*: phosphorolysis of trehalose. Kinetic parameters were determined at constant, saturating trehalose concentrations with varied phosphate concentrations. Table courtesy of C. Luley-Goedl.

	k_{cat} [s ⁻¹]	Fold decrease	K_m Pi	Fold increase
<i>ScTPase</i> WT	16.4	1	0.8	1
E614A	0.62	26	6.3	8
E614Q	0.21	78	8.0	10
E614K	$1.7 \cdot 10^{-3}$	9650	23.5	30

In general, the phosphate site of *ScTPase* has been studied well using mutational investigations (Goedl and Nidetzky 2008). Other residues important for phosphate binding, such as the highly conserved Arg₅₀₇X₄Lys₅₁₂-motif, have been identified and again, these residues are found in other synthases, such as *PhTreT* (Arg₂₃₉X₄Ly_{S244}). This, together with the double role of the EX₇E-motif, leads to the assumption that the surrounding of the donor substrate pocket determines the

specificity for a sugar phosphate or an NDP-sugar as a donor in GT4 GTs – and probably in retaining GT-B fold GTs in general. Apart from a general requirement of the NDP moiety for additional space in the binding pocket, residues involved in interactions with the donor subsite residues might play a role. Therefore, we focused on residues surrounding the donor binding site in our rational approach. Based on the results of the homology modeling and primary sequence alignment (Figures XII-XV), we identified the methionine at position 593 as a potential target: the methionine at this position is highly conserved in trehalose phosphorylases but substituted by a glutamine in trehalose synthases (Figure XIV). The first mutation suggested is hence M593Q. The second mutant carries a modified synthase-like EX₇E-motif. We decided on introducing the entire EX₇E-motif of *PhTreT*, which uses UDP-glucose as a donor. It differs in four amino acids in this motif compared to *ScTPase*.

In addition to the modeling, we also used the adaptive Poisson-Boltzmann server (APBS version 1.0.0)¹³ to calculate the electrostatic potential on the surface of the *ScTPase* model (Baker *et al.* 2001). The result of this evaluation is shown in Figure XVII. Interestingly, the surface seems predominantly negatively charged but with a strikingly opposite charge exactly in the cleft where the substrates are supposed to bind. It is conceivable that this feature might direct the substrate right into the active site.

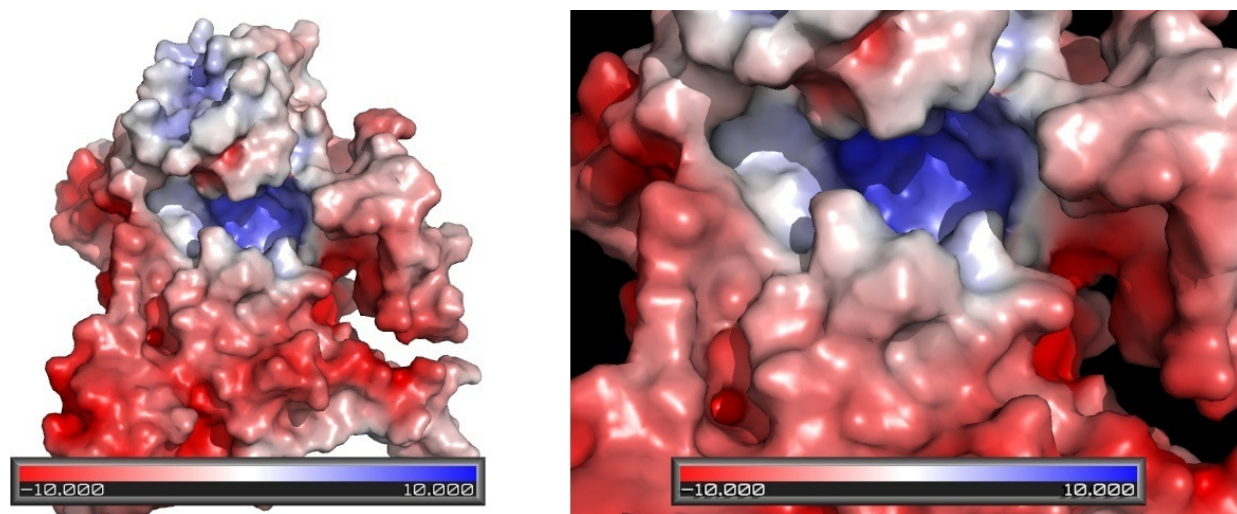


Figure XVII. Electrostatic potential of the *ScTPase* model as evaluated with APBS.

Characterization of the M593Q and the EX₇E-motif mutants

We initially assayed the purified mutants for *ScTPase* WT activity using the continuous coupled activity assay. The activity of *ScTPase* M593Q was 9 mU/mg in the crude extract and 250 mU/mg in the purified fractions. For the quadruple mutant, we barely found activity in the crude extract (<0.1 mU/mg) as well as in the purified protein (4 mU/mg). By contrast, *ScTPase* WT activity was published with 0.7 U/mg in the crude extract and 6.3 U/mg in the pure preparation (Goedl *et al.* 2006). Both mutants hence exhibited significantly reduced WT activity.

¹³ <http://www.poissonboltzmann.org/apbs>, last visited 07/28/2011

Then we aimed at determining synthase activity of the mutants. To this end we measured uridine diphosphate (UDP) formation from UDP-glucose and glucose discontinuously. We also determined the degree of hydrolysis of UDP-glucose concomitantly. Unfortunately, we did not detect UDP release above the level of UDP-glucose hydrolysis. Even though the synthesis reaction is of primary interest in characterizing the mutants, we also assayed degradation of trehalose in presence and absence of UDP. Thus we were able to determine a significant degree of error hydrolysis in the mutants which hydrolyzed trehalose in absence of UDP. This so-called “error hydrolysis” was significantly improved as compared to the WT and the observed k_{cat} is very likely solely due to hydrolysis, not transfer (Table v).

Table v. Determination of synthase activity in the mutants. n.d. – not detectable; WT values are from Goedl *et al.* (2006).

		M593Q	E609G/V610L/ K611T/S613T	WT
Synthesis direction	k_{cat} [s^{-1}]	$\sim 1 \times 10^{-5}$	n.d.	n.d.
	k_{cat} hydrolysis [s^{-1}]	$\sim 1 \times 10^{-5}$	n.d.	n.d.
Degradation direction	k_{cat} [s^{-1}]	0.11	0.09	n.d.
	k_{cat} hydrolysis [s^{-1}]	0.05	0.07	$\sim 3 \times 10^{-4}$

PhTreT/ScTPase Chimera

The construct harboring the synthetic gene of the chimera is shown in Figure XVIII. The chimera was transformed into electrocompetent *E. coli* JM109 cells and expressed as described for the Strep-tagged WTII. However, the chimera protein formed excessive inclusion bodies and no protein was detected in the soluble fraction after purification via Strep-Tactin columns. Also, we did not detect ScTPase WT activity using the continuous coupled assay neither in the crude extract nor in the purified fraction.

2.2.3. Conclusion and outlook

We performed extensive sequence comparison, alignment, secondary structure predictions and modeling with ScTPase to provide a basis for rational engineering of a phosphorylase to a synthase. The close evolutionary relationship of ScTPase to synthases in the same family – especially to the trehalose synthase – PhTreT – becomes obvious in the phylogenetic tree (Figure XI). PhTreT shares a 28 % sequence identity with ScTPase but employs a different donor substrate. Unfortunately, the rational design of three different mutants – ScTPase M593Q, E609G/V610L/K611T/S613T and the chimera – did not lead to the desired switch in donor specificity. This indicates that probably a greater structural change is necessary to achieve this. Interestingly, both mutants showed a significantly reduced WT-activity meaning that the exchanged amino acids play an important role in ScTPase activity. Nevertheless, the bioinformatic as well as the mutational studies identified how important this region is for donor binding and that it is obviously crucial for activity. Comparison with other GT4 synthases suggest that mutations in the C-terminal donor binding regions should lead to success. Due

using T4 DNA ligase following the protocol provided by the manufacturer. Next, we transformed the now 5466 bp vector into electrocompetent *E. coli* JM109 cells. Eventually, we sequenced the cloned insert and the plasmid.

We cultivated the cells harboring the pASK-IBA7plus *ScTPase* WT (see Figure X) to express Strep-tagged *ScTPase* WT (“WTII”) in lysogeny broth (LB) supplemented with 115 µg/mL ampicillin. We allowed the pre-culture to grow overnight at 37°C and 130 rpm and inoculated the main cultures to a start OD₆₀₀ of 0.01. Then, we let the main cultures grow to an OD₆₀₀ of 0.8 at 37°C and 130 rpm. At this point we cooled the cultures to 18 °C and induced expression by the addition of anhydrotetracycline (AHT) at a concentration of 2µg/mL. The cells were allowed to express at 18°C for 18 hours. Finally, we harvested the cells in a centrifuge at 5000 rpm and 4°C for 30 minutes, washed them with potassium phosphate buffer (50 mM, pH 7.5), centrifuged them again and deep froze the pellets at -80°C until further use.

In order to purify the Strep-tagged *ScTPase*, we first of all disrupted the cells using an Aminco French Press. We removed the cell debris by ultracentrifugation at 30000 rpm and 4°C for 1 hour. The supernatant from this step is the crude cell extract which we filtrated through a cellulose acetate filter with a pore width of 1.2 µM before applying it on the column. It is crucial that the pH value of the crude cell extract is set to a value above 7.5 to achieve good separation on the Strep-Tactin column. We applied the filtrated crude extract (pH 7.6) on the gravity flow column with the Strep-Tactin sepharose matrix. Next, we washed with five column volumes of washing buffer (50 mM KH₂PO₄/K₂HPO₄ pH 8.0, 76 mM NaCl) and finally eluted the protein of interest with the elution buffer containing 50 mM KH₂PO₄/K₂HPO₄ pH 8.0, 76 mM NaCl and 2.5 mM desthiobiotin. The purified protein was then desalted on a NAP-25 column (GE Healthcare) according to the manufacturer’s protocol and concentrated using Vivaspinn columns (Vivascience) with a molecular weight cut off (MWCO) of 10 kDa. Aliquots of purified Strep-tagged *ScTPase* in 50 mM KH₂PO₄/K₂HPO₄ pH 7.0 were stored at -20°C until further use.

Activity of the purified protein was determined with the continuous coupled assay as described elsewhere (Eis and Nidetzky 1999). To determine the kinetic parameters of the Strep-tagged WTII we followed the reactions in synthesis direction by measuring the phosphate released and the degradation reaction by measuring the amount of glucose 1-phosphate formed as described earlier (Eis and Nidetzky 1999, Eis *et al.* 2001, Goedl *et al.* 2006).

Sequence alignments and homology modeling

We performed primary sequence alignments with ClustalW (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>, last visited 07/28/2011) (Thompson *et al.* 2002). For secondary sequence prediction we employed psipred and for secondary structure alignment pGenTHREADER, both located at <http://bioinf.cs.ucl.ac.uk/psipred/> (last visited 07/28/2011) (Bryson *et al.* 2005, Jones 1999, Lobley *et al.* 2009). Homology modeling was performed with Modeller 9v4

(<http://salilab.org/modeller>, last visited 07/28/2011) (Eswar *et al.* 2006). We evaluated the electrostatic surface potential of the *Sc*TPase model with the adaptive Poisson-Boltzmann server (APBS version 1.0.0, <http://www.poissonboltzmann.org/apbs>, last visited 07/28/2011) (Baker *et al.* 2001).

Site-directed mutagenesis, expression and purification of the mutants

The generation, expression and purification of the E614 mutants of the pQE30-*Sc*TPase WT was accomplished as described previously (Goedl *et al.* 2006). We introduced the M593Q mutation into the pQE30-*Sc*TPase WT template using a two-stage PCR protocol (Wang and Malcolm 1999). At first, two extension reactions were run separately for four cycles. They contained the forward and the reverse oligonucleotide primer, respectively. Then, the reactions were combined and the amplification was continued for another 18 cycles. The primer in forward direction was 5'-CGACTGCCTCCAGGCCAACGC-3' and in reverse direction 5'-GCGTTGGCCTGGAGGCAGTCG-3'. After the PCR amplification, we digested residual WT template with DpnI and eventually transformed the circular gapped plasmids in *E. coli* JM109 cells via electroporation. We isolated plasmid DNA of pQE30-*Sc*TPase M593Q from positive clones growing on LB-Amp plates and subjected them to dideoxy sequencing to confirm the mutations. For the motif quadruple mutation (E609G/V610L/K611T/S613T) we employed two consecutive PCR approaches. First, we inserted the E609G/V610L mutation using two-stage PCR as described above with the degenerative forward primer 5'-CGAGGGCTTCGGGTTAAAGGTCTCC-3' and the complementary degenerative reverse primer 5'-GGAGACCTTTAACC CGAAGCCCTCG-3'. Introduction of the E609G/V610L mutation was confirmed through dideoxy sequencing. We then used the isolated E609G/V610L plasmid as a template for a so-called "fusion PCR" in the second round to introduce the K611T/S613T mutation into the E609G/V610L plasmid. We employed a fusion PCR protocol similar to one published (Szewczyk *et al.* 2006). The rationale here was that two separate PCR reactions are set up – one for amplification with the forward degenerative primer which carries the desired mutation with a flanking primer, and similar to the other one for amplification with the reverse degenerative primer – which also carries the desired mutation – with a flanking primer. The flanking primers are selected as such that they are at a position in the gene having an appropriate restriction site. After amplification of the two separate gene fragments, these fragments share a common sequence at their 3' or 5' terminus, respectively: the former degenerative primer. This is now utilized in a second PCR amplification round to fuse these two fragments together. In this reaction, the two fragments are allowed to fuse at the common sequence and the generated single strand overhang at the 3' or 5' prime is then filled by the PCR polymerase in the first couple of cycles. Then, the same non-degenerative primers as before are added and the reaction is allowed to proceed for another couple of cycles now amplifying the fused product. In case of the quadruple mutant, we employed the following non-degenerative primers carrying the restriction sites: forward primer 1 ("DraIII fwd") 5'-AAGAATGGCCTTCGCTGGACTGC-3' and reverse primer 4 ("NotI rev") 5'-

TCCTCGCGCAGTAGGTCGTTTCAGC-3'. The degenerate primers to introduce the mutations were primer 2 ("K611T/S613T fwd") 5'-CGGGTTAACGGTCACCGAAGCCGTC-3' and primer 3 ("K611T/S613T rev") 5'-GACGGCTTCGGTGACCGTTAACCCG-3'. According to the protocol described above, we first set up two separate reactions: reaction A with primers 1 and 2 yielding a 324 base product and reaction B with primers 3 and 4 yielding a 786 base product. Then, we isolated these products via preparative gel electrophoresis on an agarose gel and started a new PCR reaction with them – the "fusion" reaction. We allowed 8 cycles for fusion of the two fragments to a larger, 1085 base fragment. Then we added primers 1 and 4 thus initiating amplification of the fusion product (20 cycles). We then isolated the product via preparative gel electrophoresis on an agarose gel and digested the termini using DraIII and NotI restriction enzymes according to the manufacturer's protocol. Similarly, the *ScTPase* pQE30 WT plasmid was digested with these two restriction enzymes and the resulting approximately 4000 base fragment was isolated via preparative gel electrophoresis on an agarose gel. Finally, the digested fusion PCR product harboring the mutation was ligated into the 4000 base fragment of the WT using a T4 DNA ligase. The resulting circular plasmid was transformed into electrocompetent *E. coli* JM109 cells and positive clones were selected on LB-amp agar plates. Eventually, the correct sequence of the quadruple mutant (E609G/V610L/K611T/S613T) was confirmed by dideoxy sequencing. The PCR conditions and reagents employed can be found in Table vi. Both, the M593Q and the quadruple motif-mutant were then expressed and purified as described previously (Goedl *et al.* 2006).

Characterization of the mutants

We performed assays of the E614 mutants in phosphorolysis direction as described previously at saturating concentration of trehalose (400 mM) and varied phosphate concentrations (Goedl and Nidetzky 2008). We initially assayed the purified M593Q and EX₇E-motif mutants for *ScTPase* WT activity using the continuous coupled assay as described for the WT (Eis and Nidetzky 1999). To test for synthase activity, we performed a discontinuous assay in absence of phosphate in 50 mM MES buffer (pH 6.6) with 100 mM glucose and 25 mM UDP-glucose as substrates. We incubated the samples at 30°C and 300 rpm for 17 hours (M593Q). Using a discontinuous coupled pyruvate kinase/lactic dehydrogenase (PK/LDH) assay, we aimed at detecting the UDP formed during the reaction. The rationale of this UDP-assay is that in a first reaction step the PK converts UDP to uridine triphosphate (UTP) using phosphoenol pyruvate (PEP). The second product formed, pyruvate, is then reduced to lactate by LDH that concomitantly oxidizes NADH to NAD⁺. The UDP-assay mixture contained 420 µL of buffer (50 mM HEPES, 10 mM MgCl₂, 50 mM KCl, 1 mM EDTA), 10 µL NADH (10 mg/mL), 10 µL PK/LDH (197.2 U/mL PK and 246 U/mL LDH in reaction buffer) and 50 µL sample. The absorption at 340 nm was measured for approximately 15 minutes at 30°C until a stable value ("A1") was reached. Then, we added 10 µL of a 50 mM PEP solution and again observed the absorption at 340 nm at 30°C for approximately 15 minutes and read the absorption value "A2".

The difference between A1 and A2 (ΔE) was then used to calculate the amount of UDP formed which is equimolar to the amount of NADH oxidized to NAD^+ by the PK/LDH system. To calculate the amount of UDP (NAD^+) formed we used Lambert Beer's law and the extinction coefficient of NADH at 340 nm of $6.22 \text{ mM}^{-1}\text{cm}^{-1}$. We also had appropriate blanks and controls to determine possible hydrolysis of the UDP-glucose. The hydrolysis blank did not contain glucose to test for UDP-glucose hydrolysis by the mutant.

Table vi. PCR conditions.

mutant	PCR reaction mixture	Temperature cycling	
M593Q <i>Two stage PCR</i>	1 X pfU buffer	95°C	1 min
	3% (v/v) DMSO	95°C	50 s
	0.2 mM dNTPs	63°C	30 s
	primer	72°C	14 min
	template (WT) pfU polymerase (3 U)	72°C	5 min
E609G/V610L <i>Two stage PCR</i>	1 X Phusion GC buffer	98°C	30 s
	0.2 mM dNTPs	98°C	10 s
	primer	60°C	15 s
	template (WT)	72°C	1.5 min
	Phusion polymerase (1 U)	72°C	5 min
E609G/V610L/ K611T/S613T <i>1st PCR – amplification of the 324* and the 786 base fragment</i>	1 X Phusion GC buffer	98°C	30 s
	4% (v/v) DMSO	98°C	10 s
	0.2 mM dNTPs	65°C	20 s
	primer	72°C	8s*/14s
	template (E609G/V610L) Phusion polymerase (3 U)	72°C	5 min
E609G/V610L/ K611T/S613T <i>2nd PCR – fusion of fragments and amplification of 1085 base fragment</i>	1 X Phusion HF buffer	98°C	30 s
	0.2 mM dNTPs	98°C	10 s
	(primer 1 & 4 added after the first 8 cycles)	65°C	20 s
	template (324 and 786 base fragments, equimolar)	72°C	14 s
	Phusion polymerase (1 U)	98°C	10 s
		65°C	15 s
		72°C	20 s
	72°C	5 min	

In degradation direction we performed an assay using 200 mM trehalose and 50 mM UDP in 50 mM MES buffer (pH 6.6). We incubated the samples for 17 hours (M593Q) at 30°C and 300 rpm. To determine transfer on UDP, we measured formation of glucose and glucose 1-phosphate discontinuously. The latter one needs to be determined, as small amounts of phosphate might be present in any UDP preparation. This would lead to G1P formation by the WT *Sc*TPase activity which needs to be subtracted from the glucose formed. Again, appropriate blanks and controls to check for increased trehalose hydrolysis were performed. The control for increased trehalose hydrolysis did not contain UDP. We determined the glucose formed using a discontinuous hexokinase (HK) assay: glucose and adenosine triphosphate (ATP) are converted to glucose 6-phosphate (G6P) and ADP by

HK. The G6P-dehydrogenase (G6PDH) subsequently oxidizes G6P to D-gluconolactone 6-phosphate¹⁴ and at the same time reduces NAD⁺ to NADH. To this end we first added 25 μ L of the sample to 500 μ L of HK solution 1 (70 mM NaH₂PO₄ pH 7.7, 4 mM MgSO₄, 1.6 mM ATP, 1.6 mM NAD⁺) and determined the absorption at 340 nm at 30°C after a stable value (“A1”) was reached. Then, we added 5 μ L of HK solution 2 (100 U/mL HK, 180 U/mL G6PDH, 100 mM MgSO₄, 50 % glycerol) and again observed the absorption at 340 nm at 30°C for approximately 30 minutes and read the absorption value “A2”. Similar to the UDP assay described above, the difference between A1 and A2 (ΔE) was then used to calculate the amount of glucose formed which is equimolar to the amount of NAD⁺ reduced to NADH by the G6PDH. To calculate the amount of glucose (NADH) formed we used Lambert Beer’s law and the extinction coefficient of NADH at 340 nm of 6.22 mM⁻¹cm⁻¹. For measuring the amount of G1P produced by the *ScTPase* mutants, we employed a discontinuous coupled G1P assay in which a phosphoglucomutase (PGM) converts G1P to G6P. Next, G6PDH oxidizes G6P to D-gluconolactone 6-phosphate and at the same time reduces NAD⁺ to NADH. For the measurement, 40 μ L sample were added to the G1P assay mixture which contained 400 μ L buffer (50 mM Tris-HCl pH 7.7, 10 mM MgCl₂, 10 mM EDTA, 10 μ M glucose 1,6-bisphosphate), 10 μ L NAD⁺ (30 mM), and 2 μ L G6PDH (958.5 U/mL). A1 was determined as described for the HK assay. Then, we added 2 μ L of the PGM (846 U/mL) and measured A2, again as described for the HK assay. Also, the amount of G1P formed was calculated as described for the HK assay.

PhTreT/ScTPase Chimera

As inserting the EX₇E-motif of *PhTreT* did not switch the donor substrate specificity of *ScTPase*, we anticipated that inserting the entire C-terminal region of *PhTreT* responsible for the donor binding would do so. As this region is significantly more “crowded” in *ScTPase*, we expected that this chimeric protein would offer more space for the bulky UDP to bind. Therefore, we constructed a chimera of both, substituting aa 525-601 of *ScTPase* by aa 247-333 of *PhTreT* as shown in Figure XVIII. This construct was ordered from Genscript (Piscataway, NJ, USA) subcloned in the pASK-IBA7plus vector.

¹⁴ In aqueous solution, D-gluconolactone 6-phosphate rapidly autohydrolyzes to D-gluconate 6-phosphate.

2.3. Evolutionary approach: High-throughput assays for the directed evolution of glycosyltransferases

in preparation for ChemBioChem

FULL PAPERS

High-throughput assays for glycosyltransferases

High-throughput assays for the directed evolution of glycosyltransferases

Patricia Bubner,^[a] Tibor Czabany,^[a] Christiane Luley-Goedl,^[a] and Bernd Nidetzky*^[a]

Glycorandomization requires promiscuous enzymes. Directed evolution is widely employed to enhance nature's toolbox with glycosyltransferases (GTs) featuring novel and promiscuous activities. The evolutionary alteration of GT specificity is the method of choice if detailed structural information is not available, which is often the case for GTs. However, the bottleneck in many directed evolution approaches of GTs is the lack of an efficient and non-ambiguous

high-throughput screening (HTS). Besides, cost and availability of substrates such as NDP sugars are an important factor as well. Here, we present a novel screening method for GTs and compare it to previously published screening assays in terms of sensitivity, selectivity, high-throughput suitability, applicability to screening in crude cell extracts and unambiguousness.

Introduction

Complex carbohydrates play a pivotal role in manifold biological processes and molecules – such as in cell signaling, cell growth, immune defense, inflammation, glycoproteins, glycolipids, and polysaccharides^[1]. Moreover, many carbohydrate-based natural products are possible potent novel drugs^[2]. Chemical synthesis of complex carbohydrates is extremely difficult and usually involves multiple steps with disappointing yields^[2b]. The directed and targeted assembly of intricate carbohydrates and the capability of selectively adding or changing sugar residues attached to natural products is a highly desired goal in glycorandomization, which aims at the diversification of sugar-containing compounds^[3].

For this purpose, nature's toolbox offers glycosyltransferases (GTs) which are highly specific enzymes catalyzing the transfer of a glucosyl moiety from an activated donor to an acceptor. Their specificity, however, is a hurdle in order to harness nature's wisdom in directed assembly of intricate carbohydrates. Therefore, expanding the substrate promiscuity of GTs is of high interest. Unfortunately, crystallographic data of

GTs is rare due to general difficulties in sufficient expression levels, purity, stability and crystallization of these enzymes^[4]. Rational approaches – i.e. based on homology models with known structures of related GTs – towards increased promiscuity are therefore often unsuccessful, labor intense and time consuming. Directed evolution is a way to circumvent these issues, however, due to a lack of suitable high or medium-throughput screening assays, increasing the promiscuity of GTs is still a arduous process^[1, 2b].

A high-throughput screening (HTS) assay for GT screening needs to meet some general requirements: i) good sensitivity and signal-to-noise ratio even in crude cell extracts, ii) signal stability over the time span needed, iii) easy detectability

[a] DI Patricia Bubner, Dr. Tibor Czabany, Dr. Christiane Luley-Goedl, Prof. Dr. Bernd Nidetzky
Institute of Biotechnology and Biochemical Engineering
Graz University of Technology
Petersgasse 12, A-8051 Graz, Austria
Fax: (+)43-316-873-434
E-mail: bernd.nidetzky@tugraz.at

and reproducibility, iv) efficiency, performance, easy handling, and rapidness, v) reasonable costs, vi) possibility to screen large libraries; and vii) general applicability on GTs using various substrates. A tailored assay is required for every screening task and well-thought-out design is crucial to avoid failure, i.e. finding lots of false positives or false negatives. „You get what you screen for“ was postulated in 1999 and became an axiom in HTS assay development^[5]. A couple of years later, while aiming at evolving a family GT42 sialyltransferase towards novel acceptors using fluorescence-activated cell sorting (FACS), this postulate was proven true. Aharoni et al. used bodipy-labeled acceptor sugar derivatives in their screening – and ended up with a highly prevalent mutation (F91Y) which was identified as a dye-binding site. The F91Y mutation generated a hydrophobic pocket for bodipy-binding thus significantly increasing this mutant's k_{cat}/K_m values towards various bodipy-labeled acceptors^[1, 2b].

In the last decade, HTS has gained general importance due to significant advances and reduced cost in the required robotics. In enhancing GT promiscuity, an effective HTS assay is the bottleneck. Nucleoside monophosphate- (NMP) and nucleoside diphosphate- (NDP) sugar dependent GTs – so-called Leloir enzymes – are the major group of GTs and of special interest due to their broad availability and capabilities^[6]. A general hindrance in HTS assays for these GTs is the cost and availability of substrates such as NDP-sugars. In recent years, a handful of assays has been published, some of which have been proposed to be suited as HTS assays, some of which have already been validated.

Direct measurement of enzymatic activity would be highly desirable. However, it is not possible to determine GT activity directly using fluorescence or absorbance as no change in these parameters occurs during cleavage or formation of the glycosidic bond. Hence, assaying glycosyl transfer requires coupling in some way. As mentioned above, coupling of the acceptor to fluorescent molecules bears the risk of evolving a dye-binding site in the enzyme. This might be circumvented by the use of two different fluorescent labels as recently suggested^[1, 7]. However, FACS-based screenings require laborious preparations and special laboratory equipment and as the labeled substrates must be uptaken by cells, they are probably not generally suitable. In directed evolution of the GT-B-fold olandomycin GT (OleD), Williams *et al.* also employed a fluorescence-based assay. They capitalized on the fact that OleD is capable of using 4-methylumbelliferone as an acceptor. The transfer of a glucosyl moiety to the C7 hydroxyl group of this fluorescent dye quenches fluorescence. Using this assay the authors identified a hit with 30-fold improved activity directly from

the crude cell extract^[9]. As the great majority of GTs does not adopt fluorescent dyes as acceptors, a screening with general applicability is required.

Here, we compare four different HTS assays under screening conditions regarding their suitability and applicability as a general HTS for Leloir GTs: i) a pH-shift assay, ii) a commercially available immuno-assay (Transcreener[®]), iii) a phosphatase-coupled assay, and iv) a dehydrogenase-coupled assay. In i) and iii) a change in absorbance is measured while ii) and iv) are fluorescence-based assays. The principle of the assays is outlined in Figure 1. To probe these assays, we used a sucrose synthase (SuSy) from soybean. SuSy catalyzes the synthesis of sucrose from UDP-glucose and fructose. In contrast to most synthases, the reaction equilibrium is on the side of the UDP-glucose thus favoring the degradation reaction.

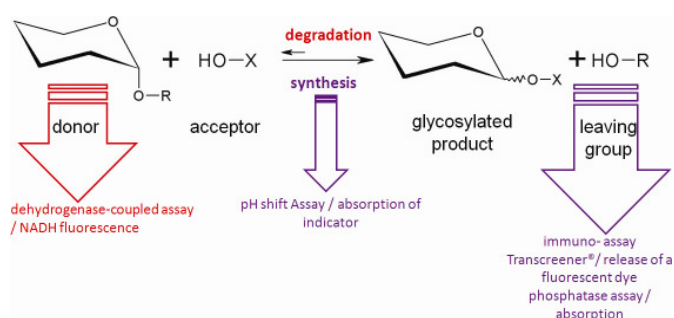


Figure 1. Targets for a general HTS assay of glycosyltransferases and assays compared in this paper.

In 2004, Deng and Chen published a pH-sensitive assay for an UDP-galactose dependent β -1,4-galactosyltransferase and proposed that it is applicable to other glycosyltransferases^[8]. A couple of years later, Persson and Palcic probed the applicability of a similar pH-indicator based glycosyltransferase assay for HTS with an α -1,3-galactosyltransferase and an α -1,3-*N*-acetylgalactosaminyl-transferase, both using the respective NDP-donor substrate^[9]. In 96-well assay format, they determined the change in absorbance of a pH indicator, bromothymol blue, during the reaction. They suggested that the decrease in absorbance and the according color change of the indicator result from release of a proton during glycosyl transfer^[9]. Park *et al.* reported in their study how they employed a similar pH-shift assay in an *E. coli* whole cell system^[2a].

Direct immunodetection with antibodies, which are able to specifically detect a certain antigen, is a widely employed method that can also be a useful method in directed evolution of GTs. For example, Hancock *et al.* published an ELISA-based assay for the directed evolution of a glycosphingolipid-synthesizing enzyme^[10]. However, this assay is very specific and requires immobilization of the acceptors. For a general screening

of NDP-donor dependent GTs, we here tested the commercial Transcreener® ADP2 FI assay (BellBrook Labs, Madison, WI, USA). This immuno-assay can selectively detect NDPs in a mixture of nucleoside triphosphates (NTPs). Originally, this assay was designed to follow the activity of kinases by monitoring formation of ADP from ATP. This assay uses antibodies which are more selective for ADP than for ATP by a factor of >100. At the same time, these antibodies feature a similar affinity towards other NDPs such as GDP and UDP. The use of this method as a GT assay requires the antibodies to be able to distinguish between an NDP and an NDP-sugar. For determining a signal upon antibody binding, the antibody is coupled to Alexa Fluor®594, a well established fluorescent dye. NDPs generated during the reaction displace the dye which thereupon becomes unquenched and thus, the fluorescence intensity increases^[11].

The phosphatase-coupled assay capitalizes on the activity of phosphatases which cleave inorganic phosphate off the leaving groups of the GT reaction (NMP or NDP)^[6]. The thus released phosphate can subsequently be detected using a colorimetric assay such as with malachite green (620 nm)^[6] or based on formation and reduction of a phosphomolybdate complex (850 nm)^[12].

The dehydrogenase-coupled assay described here also makes use of a coupled highly specific enzymatic activity. In contrast to the other assays described above, this screening assay is based on the degradation reaction of GTs, the transfer of a glucosyl moiety on an NDP (see Figure 1). Subsequently, a specific dehydrogenase is employed to oxidize the produced NDP-sugar to an NDP-glycuronic acid. Concomitantly, the NDP-dehydrogenase reduces NAD⁺ to NADH which is detected via its intrinsic fluorescence (see Figure 2).

We probed the HTS applicability of these four assays in a comprehensive study. We tested them regarding their suitability for screening from crude extract, the stability of the signal, their sensitivity and suitability as a general GT screening. The pH-sensitive assay has been reported to work for galactosyltransferases in crude extract; we here test it with a glucosyltransferase. The Transcreener®ADP2 FI assay has not been proposed but not evaluated as an assay for glycosyltransferases yet, and therefore also not as a GT HTS assay in crude extract, before. Also, the phosphatase assay has not been tested under true screening conditions concerning the overall handling or direct assaying from crude extract. Finally, we here present a novel dehydrogenase-coupled assay developed in our laboratory to avoid screening for mutants with increased donor-substrate hydrolysis (“error hydrolysis”) and compare it to the other assays under screening conditions.

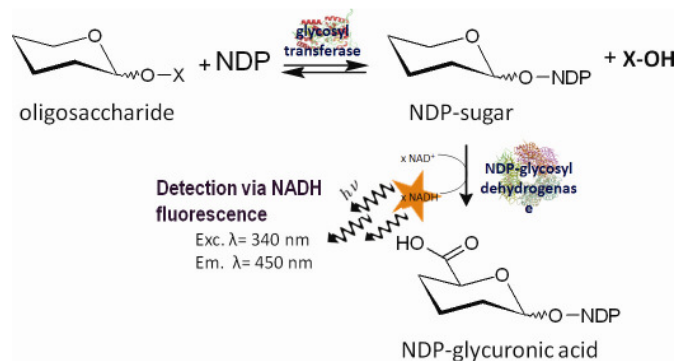


Figure 2. Reaction scheme of the dehydrogenase-coupled assay.

Results

pH shift assay

We employed phenol red as an indicator in a weakly buffered system, which allows to determine the slightest changes in pH value. Preliminary to testing the reaction of the SuSy in the assay setting, we carried out extensive testing to determine the best setting. To this end, we tested the two different indicators that have been used in the crude cell extract screenings in previous studies – bromothymol blue^[9] and phenol red^[8] – in two different buffer systems at two starting pH values: TES (pKa=7.6) and MOPS (pKa=7.3), each at a concentration of 2.5 mM and at pH 7.3 and 7.5. Phenol red has a pKa of 7.4 and changes its color from red to yellow in a pH range of 8.4–6.8. This change in color is also detectable as a change in absorbance at 557 nm. The pKa of bromothymol blue is 7.1 and the color change occurs in a pH range of 7.6–6.0 from blue to yellow, detectable photometrically at 615 nm. The pH range of the indicators is compatible to the pH range in which most GTs show high activity. The buffers were chosen according to their pKa values which are supposed to be close to the one of the indicator and their known general compatibility with GTs. We titrated both systems with HCl to evaluate the performance of the indicators in the respective system. In Figure 3 we plotted the absorption decrease against the proton concentration. We eventually decided to use phenol red because in our test setting it was easier to detect visually, showed a better signal stability, a higher slope in absorption difference, and superior sensitivity at very small pH changes (>0.3 mM H⁺). For the buffer system, we decided on using TES at pH 7.3 as the buffers performed quite similar and the GTs used in our laboratory are stable in this buffer.

Next, we tested the assay in crude *E. coli* cell extract and with SuSy. The mixture employed for cell disruption contained 2.5 mM TES (pH 7.3). In order to start the assay, we added 10 mU of pure SuSy and 2 μL of crude *E. coli* extract or buffer to 100 μL of a substrate solution in a 96-well plate. The

substrate solution contained 0.5 mM UDP-glucose, 40 mM fructose, 0.01 mM phenol red and 2.5 mM TES (pH 7.3). We also measured appropriate substrate and enzyme blanks. After starting, we measured absorption at 557 nm at discrete points in time for the next 24 hours and also checked for visible changes. We did not observe a detectable pH shift for the SuSy in buffer. Wells with crude extract showed a great amount of precipitate formation within the first hour and a concomitant color change to yellow, which was also occurring in the wells where no SuSy had been added.

After stopping the SuSy reaction with the Saheki solution, the color development was observed visually. Positive wells could be distinguished visually when pure SuSy in buffer was used. After 15 minutes, absorption spectra were recorded (500-850 nm) of each well. A small amount of precipitate formed in samples with crude extract did not bias the measurement significantly. With crude extract present, the blank value is very high and prevents identification of positive hits. Therefore, we diluted the assay solution by a factor of 10 in pure water after incubation and measured the absorption again. Thus, we were able to reproducibly detect "hits" with an activity of >1 mU/well (Figure 5). However, only if the activity of the GT is high enough, a significant difference can be seen visually.

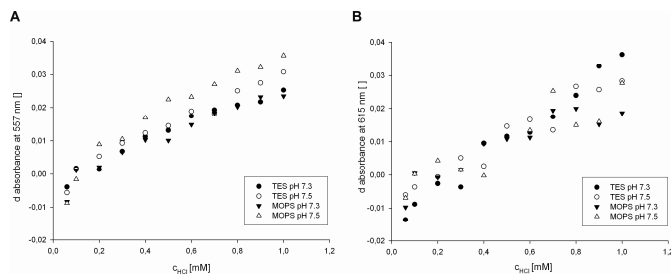


Figure 3. Testing of phenol red (A) and bromothymol blue (B) with increasing proton concentration.

Transcreener[®] assay

To test whether the assay is capable of distinguishing UDP from UDP-glucose and whether UDP-glucose is degraded in the crude cell extract, we set up a primary mock reaction plate. This 96-well plate contained UDP in varying concentrations with and without the addition of crude extract as well as samples with UDP-glucose, again with and without crude extract. With the fluorescence signal highly amplified, we were able to detect a signal of as little as 0.1 μM UDP in buffer (see Figure 4). The crude extract added significantly to the signal, however, we were still able to safely distinguish UDP concentrations of 0.01 mM UDP from the background. Moreover, the assay proved to be capable of distinguishing between UDP and UDP-glucose, which gave no signal at all. Besides, UDP-glucose did not get degraded to detectable amounts of UDP in the crude extract during the time span observed.

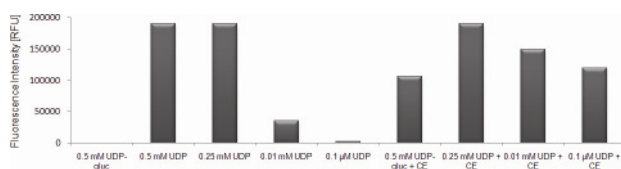


Figure 4. Transcreener[®] evaluation. UDP-glic denotes UDP-glucose. CE denotes crude extract.

Phosphatase assay

We used the non-specific shrimp alkaline phosphatase (SAP) for dephosphorylation of the UDP produced during SuSy reaction.

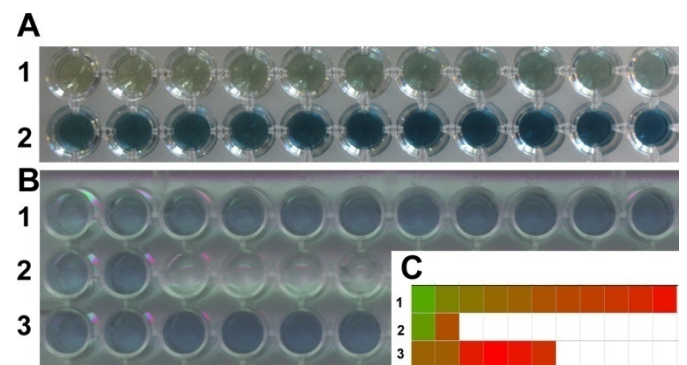


Figure 5 Phosphatase assay. A: phosphate calibration (0–1 mM) without (1) and with (2) crude extract (CE, 2 μL/well). B: (1) increasing phosphate concentrations (0–1 mM) from left to right after 10 X dilution; (2) left well: blank with CE, right well: blank with CE and SAP, 10 X dilution; (3) spiking experiments with 0, 0.1, 1, 5, 10, and 15 mU SuSy/well (from left to right), 10 X dilution. C: absorption intensity of the samples in B determined at 850 nm (green: low absorption, red: high absorption).

Dehydrogenase-coupled assay

By determining the fluorescence intensity of different NADH solutions in buffer and with crude extract added, we learned that the fluorescence signal is slightly decreasing over time but – over all – relatively stable even after 24 hours. We determined the detection limit of NADH in buffer and in the crude extract by measuring the intrinsic fluorescence of NADH concentrations from 0.0001 mM – 10 mM. We were able to distinguish an amount as low as 0.05 mM from the background, which corresponds to an enzymatic activity of approximately 0.08 mU/well, when using human UDP-glucose dehydrogenase (hUGDH) as the coupled enzyme. hUGDH catalyzes the oxidation of UDP-glucose to UDP-glucuronic acid^[13] and amplifies the signal in this assay by a factor of two due to reducing two NAD⁺ molecules per UDP-glucose oxidized.

Discussion

In our study we tested four different assays under screening conditions and evaluated them regarding their signal stability, easy detectability, handling, and their ability to detect low amounts of GT activity in crude cell extract (sensitivity). Easy detectability of the signal, such as a visual color change, absorbance or fluorescence, is necessary to quickly identify a hit among several thousand clones screened. Employing the visible shift of a pH indicator upon glycosyl transfer in a weakly buffered system, the pH-shift assay is supposed to be a simple, color-based assay with the advantage of visual detection of positive hits. The underlying rationale is that a proton should be released upon transfer of the glycosyl moiety from the NDP-sugar to the acceptor. A previous study determined the detection limit of this method with 5–10 mU of GT activity^[9]. Corresponding to the proton release the indicator used, e.g. bromothymol blue or phenol red, shows a visible color change. However, in contrast to previously published studies reporting the success of this method, no pH-shift was observed using SuSy, probably because no proton is released during SuSy reaction. Hence, this assay is not suited as a general GT assay. Besides, screening in crude cell extracts quickly led to precipitation. This did not only cause turbidity thus biasing absorbance measurements, but also a visible drop in pH. Generally, if used for GTs, where a measurable pH-shift occurs during reaction, the pH-shift assay is a cheap method offering continuous incubation and visual identification of hits. Due to the sensitivity of the weakly buffered system extensive testing is required prior to the screening in order to avoid ambiguous results or false positives.

The Transcreener[®] assay was originally designed as a general ADP detection method in presence of ATP. Besides, the antibodies used in this immuno-assay are promiscuous enough to generally discriminate NDPs from NTPs. In our study we show that it is also capable of distinguishing between an NDP and an NDP-sugar and, moreover, to do so in crude cell extracts: the release of UDP upon transfer of glucose from UDP-glucose on fructose by SuSy was successfully determined in *E. coli* crude extract. Under the conditions tested, this commercial immuno-assay showed clear and reliable results, a long signal stability (24 hours), and was very sensitive with a limit of detection in crude cell extracts of <0.015 mU. It offers the screening for activity with various NDP-sugars without the need for extensive testing. False positives can only occur due to increased donor substrate hydrolysis (“error hydrolysis”) releasing NDP without formation of the transfer product. Detection via fluorescence is additionally advantageous as it is not biased by possible precipitations in the crude extract. For HTS of large libraries, the assay is probably too

costly. Besides, due to the discontinuous nature of the assay and the several steps of liquid handling required, it is quite time-consuming.

The phosphatase-coupled assay proved to be a suitable assay under screening conditions if significant expression of the target protein can be achieved. Wu et al. neither probed their assay in a 96-well format nor in crude extract. Although they used two specific phosphatases, the recombinant ectonucleoside triphosphate diphosphohydrolase (ENTPD) CD39L3 and the 5'-nucleotidase CD73 in their assay, they also suggested the use of a non-specific phosphatase for general GT screenings. Here, we present our results with a non-specific alkaline phosphatase, which costs approximately 2000 times less per unit than the specific phosphatases. For inorganic phosphate detection we employed the Saheki-method in contrast to Wu et al., who used the malachite green phosphate detection method^[6]. Generally, the phosphatase-coupled assay is well suited for the screening of purified enzymes. The signal is stable and visual detection is possible with either colorimetric phosphate assay if the activity of the target enzyme is high enough. In *E. coli* crude cell extract, however, the sensitivity of the assay is drastically reduced probably due to side-reactions of the phosphatase with abundant phosphorylated cellular components such as ATP, DNA, RNA and proteins. This generates a high background which might mask a low-level positive signal from a barely active or scarcely expressed “hit”. Just as in the Transcreener[®] assay, this is a discontinuous assay aiming at detection of the released NDP, which could lead to false positive detection due to increased error hydrolysis by a mutant enzyme.

Persson and Palcic stated that an assay based on NADH detection is not suited for screening in crude cell extracts due to competing NADH oxidase activity^[9]. However, in the setting where we tested the dehydrogenase-coupled assay in crude extracts of three different widely employed *E. coli* strains (JM109, TOP10, BL21DE3), we always found a stable NADH fluorescence signal, even when low substrate concentrations were present. The limit of detection in crude extracts was as low as 0.08 mU of GT activity. The principle of the dehydrogenase-coupled assay is outlined in Figure 2: a nascent NDP-sugar is converted by an NDP-glucosyl dehydrogenase, which concomitantly reduces NAD⁺ to NADH. In our case, the hUGDH employed produces to equivalents of NADH per UDP-glucose converted thus multiplying the signal by a factor of two and increasing sensitivity. We were successful in monitoring the conversion of sucrose and UDP to UDP-glucose and fructose by SuSy in crude extracts using hUGDH as coupled enzyme. The overall signal decreased slightly with time over the observed time

course of 24 hours but was in all cases still well distinguishable from a stable blank. The blank value of crude extracts was of significant intensity but we were able to detect GT activity with a low detection limit. Therefore, the dehydrogenase-coupled assay proved to be a stable and continuous assay suitable for HTS of GTs from crude extracts. This method is specific for the donor sugar of the GT reaction thus offering the advantage of not producing false positive results due to increased donor hydrolysis. False positives resulting from increased hydrolysis are an issue in GT screening in general indeed and many GTs are known to hydrolyze the donor sugar into NDP and sugar parts to some extent in synthesis direction. During directed evolution, increase of the enzyme active site accessibility to water could easily result in an increased donor substrate hydrolysis – this might be even more probable than finding enhanced promiscuity in the first place. If using an assay sensitive to NDP formation, this would surely yield false positives. This cannot occur in the dehydrogenase-coupled assay because a signal is only produced upon successful transfer of the glycosyl group on the NDP. The general low activity of synthases in degradation direction should be overcome by the removal of the donor-sugar from the reaction mixture by the dehydrogenase thus speeding up the reaction. Additionally, NADH detection using fluorescence is very sensitive even from crude extracts, hence, even low GT activities can be observed. Its strict substrate limitation due to the exquisite specificity of the coupled dehydrogenase, however, limits its application to UDP-glucosyl transferases. To circumvent the substrate limitation, increasing the substrate promiscuity of the coupled enzyme would, of course, be a possibility. Use of a different dehydrogenase activity – or a cocktail of activities – according to the donor substrate of the GT activity to be evolved is another possibility.

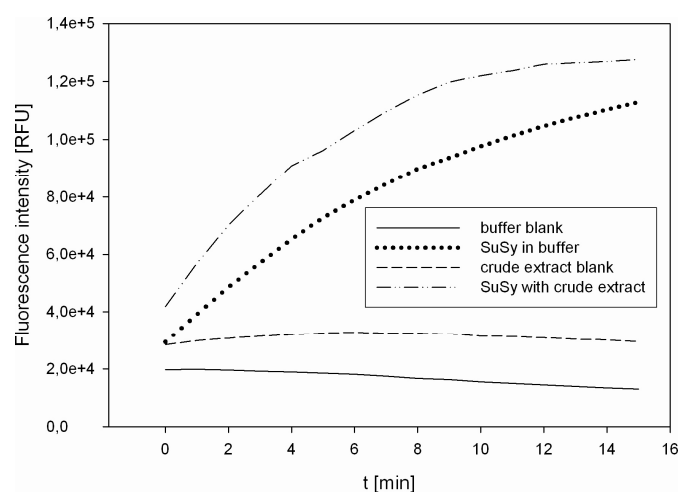


Figure 6. Continuous detection of SuSy activity with the dehydrogenase-coupled assay.

Conclusions

We probed four very different assays for their suitability under screening conditions and as general HTS assay for GTs. The first one tested was a pH-shift assay that was expected to be suited as a general GT assay according to previous reports^[8-9]. This assay relies on the release of a proton during reaction, which might not be the case for all GTs. In our assay setting, for instance, the SuSy exhibited no detectable pH-shift upon reaction. We therefore conclude that the pH-shift might be a result of the reaction mechanism and/or a pKa effect of the reaction products. Secondly, we used a commercially available immuno-assay (Transcreener[®]) developed to assay kinase activity^[11]. This assay had neither been tested as a HTS assay for GTs before nor had it been probed in crude cell extracts. It performed well in detecting UDP, even in crude cell extracts, with a remarkably high sensitivity and was able to distinguish between UDP and UDP-glucose. Its backdraws are that it is a discontinuous and quite costly method. The third assay tested was a phosphatase-coupled assay adapted from a protocol published recently^[6]. This assay relies on the dephosphorylation of the leaving group – NDP – by a phosphatase and subsequent detection of the released phosphate by a colorimetric assay. Again, this is a discontinuous assay which requires several individual operation steps. This of course affects the efficiency and HTS applicability of this assay. Finally, we developed a novel GT HTS assay which – in contrast to the three others tested – screens in degradation direction detecting the donor substrate. This prevents false positives which can result from mutants with increased error hydrolysis. The assay is based on a dehydrogenase-coupled reaction: only if transfer to the NDP occurs, the resulting NDP-sugar will be oxidized by a specific NDP-sugar dehydrogenase. During the reaction, NAD⁺ is reduced and formation of NADH is followed by an increase in fluorescence, which proved to be a sensitive method even in crude cell extracts. Contrary to previous reports^[9], NADH oxidases present in *E. coli* crude extracts did not pose a problem during time course of the continuous reaction. In fact, the signal was stable for at least 24 hours, which is advantageous if screening for very low activities. We believe that using a promiscuous enzyme in the coupled HTS presented here would make this assay a well-suited, highly sensitive general GT screening assay.

For a successful HTS, the qualitative pre-screen needs to be sufficiently sensitive and reliable to detect positives without showing abundant false positives, which would make a quantitative re-screen labor-intensive and probably disappointing – if not impossible. Insensitivity and unstable signal, on the other

hand, would lead to false negatives. Developing a general HTS assay for GTs suffers from the intrinsic restrictions of the reaction to screen: glycosidic bond formation and reaction products cannot be measured directly. To determine the main product of synthesis direction, some sort of labeling, e.g. with fluorescent labels, would be required. However, in a library screen, this might result in specificity for the bulky label rather than the product^[1]. If basing the screening on detection of the leaving group, NDP for example, it is likely to find a lot of false positives due to increased error-hydrolysis of the NDP-sugar donor. These issues can only be circumvented by screening in degradation direction, as transfer of the glycosyl moiety on the NDP must occur for a positive signal. Unfortunately, the dehydrogenase-coupled assay presented here suffers from a strict substrate specificity so far restricting the screening to UDP-glucose. The advantages are, however, striking: it is a highly sensitive, reliable and stable continuous assay.

A robust screening is able to identify hits with desired properties without bias by false positives or false negatives. After the tedious and labor-intensive process of mutagenesis, cloning, transformation, and expression to generate a library containing a couple of thousand mutant GTs, the screening itself should require as few operation steps as possible. Additional liquid handling multiplies the material cost as well as it is time-consuming and a possible source of errors. Therefore, a screening assay suitable for the HTS of large libraries needs to be of reasonable cost and it should involve as few handling steps as possible: keep it small and simple!

Experimental Section

Materials: All materials were purchased of the highest purity available from Carl Roth (Karlsruhe, Germany) or Sigma (St. Louis, MO, USA) unless stated otherwise. SAP (1 U/ μ L) was obtained from Fermentas (St. Leon-Rot, Germany). The Transcreener[®] ADP² FI assay was from BellBrook Labs (Madison, WI, USA). The 384- and 96-well plates used for cultivation and absorption measurements as well as the black 96-well half-area plates used for fluorescence measurement were purchased from Greiner Bio One (Kremsmünster, Austria). We used a BMG Labtech Omega plate reader for 96-well absorbance and fluorescence intensity measurements. The recombinant Strep-tagged hUGDH was produced and purified to apparent homogeneity as described previously^[13].

Preparation of crude *E. coli* cell extracts: *E. coli* TOP10 and BL21(DE3) Gold cells harboring a pASK-IBA7plus plasmid (IBA GmbH, Göttingen, Germany) providing resistance against ampicillin were cultivated in lysogeny broth (LB) supplemented with 115 μ g/mL ampicillin. The pre-culture was grown for 16 hours at 30 °C in 384-well

plates (50 μ L LB per well) and the main culture in 96-well plates (200 μ L LB per well) at 37 °C for 24 hours. After harvesting of the cells at 4000 rpm (4 °C, 30 min), we prepared crude *E. coli* cell extract by cell disruption with Bug Buster[®] 10X Protein Extraction reagent (Novagen, Merck, Darmstadt Germany) according to the manufacturer's protocol. To this end, the 10X concentrate was diluted in the buffer indicated for the respective assay which was supplemented with 1 mM dithiothreitol (DTT) and 2 μ L/mL Lysonase[™] (Novagen, Merck, Darmstadt Germany). After cell disruption, we centrifuged the plates for 30 minutes at 4000 rpm and 4 °C. The cleared supernatant thus obtained is referred to as crude cell extract.

Activity assay for sucrose synthase: Activity of SuSy was checked in 0.1 mM UDP, 300 mM sucrose, 10 mM NAD⁺, 20 mU/mL hUGDH, 50 mM HEPES-KOH (pH 7.5). The change in absorption of NADH was followed at 340 nm.

pH-shift assay: Before probing SuSy activity with this assay, we tested the assay setup with several buffer/indicator systems for this assay. Solutions with either one of the two indicators bromothymol blue or phenol red were prepared at four conditions: TES and MOPS, each at a concentration of 2.5 mM and at pH 7.3 and 7.5. In a 96-well plate, we titrated the solution with increasing amounts of HCl and measured the decrease in absorbance at 557 nm (phenol red) or 615 nm (bromothymol blue).

The buffer system used for cell disruption contained 2.5 mM TES (pH 7.3). In order to start the assay, we added 10 mU of SuSy and 2 μ L of crude *E. coli* extract or buffer to 100 μ L of a substrate solution in a 96-well plate. The substrate solution contained 0.5 mM UDP-glucose, 40 mM fructose, 0.01 mM phenol red and 2.5 mM TES (pH 7.3). We also measured appropriate substrate and enzyme blanks. After starting, we measured absorption at 557 nm at discrete points in time for the next 24 hours and also checked for visible changes.

Transcreener[®] assay: We prepared the crude extract here using the 10X Bug Buster[®] solution with 200 mM HEPES buffer (pH 8.0). The following samples (total volume of 50 μ L per well) were prepared in triplicates for this primary mock reaction plate (all in 200 mM HEPES, pH 8.0): UDP in concentrations ranging from 1 nM to 0.5 mM UDP with and without crude extract (2 μ L), UDP-glucose at 0.5 mM with and without crude extract; and appropriate blanks. The primary mock reaction plate was then incubated for 15 minutes at 30 °C and afterwards fluorescence was measured using the Transcreener[®] assay using an excitation wavelength λ_{Ex} of 590 nm and an emission wavelength λ_{Em} of 617 nm. The Transcreener[®] assay setup was prepared according to the manufacturer's manual with the recommended controls.

Phosphatase assay: For cell disruption, the 10X Bug Buster[®] solution was used in a 50 mM HEPES buffer (pH 8.0). We tested crude cell extract of SuSy (0.29 U/mg) expressed in *E. coli* BL21DE3

gold as well as cell extract of empty *E. coli* TOP10 and supplemented it with varying amounts of purified SuSy (8.2 U/mg) to test the sensitivity of the assay. Each well contained 10 μ L crude extract, 1 μ L SAP, 50 mM HEPES-KOH buffer (pH 8.0), 10 mM $MgCl_2$, 1 mM UDP-glucose, and 100 mM fructose in a total volume of 50 μ L. After mixing we sealed the plate with a transparent film and incubated it for 1 hour at 30°C. Concomitantly, appropriate controls and the activity of the pure enzyme were measured in the same way. Finally, the colorimetric detection of the released phosphate was performed according to the method described by Saheki et al.^[12]. We added 190 μ L of the Saheki solution to the reaction and measured the absorption at 850 nm. Then, we diluted the samples by a factor of 10 with pure water due to the high blank value of the crude extract. Deionized water was used for the preparation of all solutions to avoid phosphate contamination from water.

Dehydrogenase coupled assay: For cell disruption, the 10X Bug Buster[®] solution was used in a 50 mM Tris-HCl buffer (pH 7.5). We prepared a black 96-well half area plate containing 50 μ L of the substrate solution (250 mM sucrose, 0.5 mM UDP, 10 mM NAD^+ , 50 mM Tris-HCl, pH 7.5 and 6 mU hUGDH per well). To some wells we added 2 μ L of crude *E. coli* extract. Then, we added 5 mU SuSy to start the reaction. Appropriate controls and blanks were also measured. After starting the reaction, the fluorescence intensity ($\lambda_{Ex}=340$ nm, $\lambda_{Em}=430-450$ nm) was measured continuously for 14 hours to check signal stability. In addition to that we determined a NADH standard curve (0.0001 mM – 10 mM) with and without addition of crude extract (2 μ L/well).

Acknowledgements

We thank Sigrid Egger for preparation of the hUGDH and helpful discussion, Alexander Gutmann and Linda Bungaruang for the preparation of the SuSy from Soybean, and Karin Longus and Jacqueline Harg for hands-on help in assay testing.

Keywords: carbohydrates · high-throughput screening · glycosyltransferase · genetic engineering · assay development

- [1] R. Kittl, S. G. Withers, *Carbohydr Res* **2010**, *345*, 1272.
[2] a)S. H. Park, H. Y. Park, J. K. Sohng, H. C. Lee, K. Liou, Y. J. Yoon, B. G. Kim, *Biotechnol Bioeng* **2009**, *102*, 988; b)A. Aharoni, K. Thieme, C. P. Chiu, S. Buchini, L. L. Lairson, H. Chen, N. C. Strynadka, W. W. Wakarchuk, S. G. Withers, *Nat Methods* **2006**, *3*, 609.
[3] G. J. Williams, C. Zhang, J. S. Thorson, *Nat Chem Biol* **2007**, *3*, 657.
[4] C. Breton, L. Snajdrova, C. Jeanneau, J. Koca, A. Imberty, *Glycobiology* **2006**, *16*, 29R.
[5] C. Schmidt-Dannert, F. H. Arnold, *Trends Biotechnol* **1999**, *17*, 135.
[6] Z. L. Wu, C. M. Ethen, B. Prather, M. Machacek, W. Jiang, *Glycobiology* **2011**, *21*, 727.
[7] G. Yang, J. R. Rich, M. Gilbert, W. W. Wakarchuk, Y. Feng, S. G. Withers, *J Am Chem Soc* **2010**, *132*, 10570.
[8] C. Deng, R. R. Chen, *Anal Biochem* **2004**, *330*, 219.
[9] M. Persson, M. M. Palcic, *Anal Biochem* **2008**, *378*, 1.
[10] S. M. Hancock, J. R. Rich, M. E. Caines, N. C. Strynadka, S. G. Withers, *Nat Chem Biol* **2009**, *5*, 508.
[11] K. M. Kleman-Leyer, T. A. Klink, A. L. Kopp, T. A. Westermeyer, M. D. Koeff, B. R. Larson, T. J. Worzella, C. A. Pinchard, S. A. van de Kar, G. J. Zaman, J. J. Hornberg, R. G. Lowery, *Assay Drug Dev Technol* **2009**, *7*, 56.
[12] S. Saheki, A. Takeda, T. Shimazu, *Anal Biochem* **1985**, *148*, 277.
[13] S. Egger, A. Chaikuad, K. L. Kavanagh, U. Oppermann, B. Nidetzky, *J Biol Chem* **2011**, *286*, 23877.

2.3.1. Appendix to: “High-throughput assays for the directed evolution of glycosyltransferases”

2.3.1.1. Rationale of the evolutionary approach

The C-terminal region of *ScTPase* is responsible for the binding of the donor. Based on our previous rational design studies and comparison to the synthase *PhTreT* (see 2.2.), we expect that random mutagenesis within a 454 bp region of the *ScTPase* gene generates *ScTPase* mutants capable of using UDP-glucose as a donor sugar. Figure XIX shows a partial primary sequence alignment of the donor binding region of *PhTreT* and *ScTPase*. In the synthase, this region has fewer residues and probably offers more space for binding of the UDP-moiety. The regions also show a high degree of conservation between the two enzymes with, however, differences whose effect is elusive. Thus, directed evolution of this donor binding region should provide us with information about the role of these residues in donor binding in synthases and phosphorylases. The targeted 454 bp stretch corresponds to possible changes in 152 of the 748 amino acids of *ScTPase* (D472–A625). We intend to achieve an average mutation rate of 5 base changes within the 454 bp frame. As a second step, we plan to screen a library of 10000 clones using the dehydrogenase-coupled assay as described in 2.3 to identify hits exhibiting the desired activity.

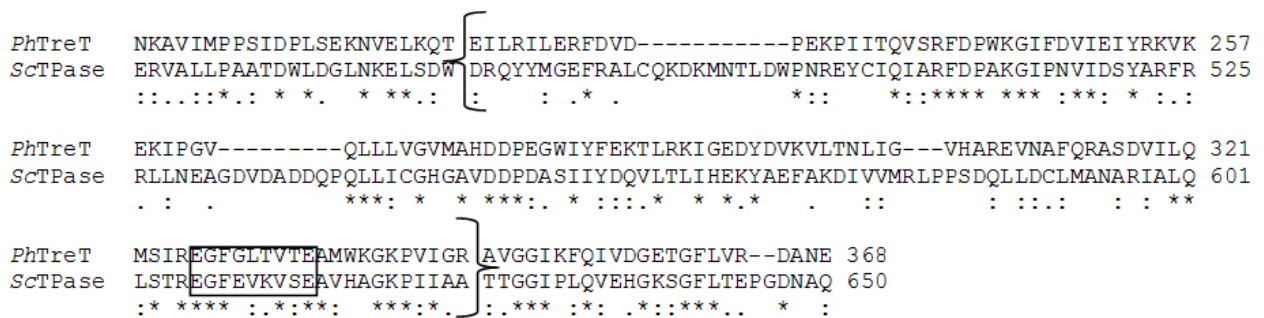


Figure XIX. Partial sequence alignment of the donor binding regions of *PhTreT* and *ScTPase*. The brackets indicate the target region for directed evolution; the rectangle marks the EX₇E-motif.

2.3.1.2. Materials and methods

Materials

We used the GeneMorphII random mutagenesis kit (Stratagene) to introduce the desired mutation rate into the target region of the *ScTPase* WT gene. Phusion DNA polymerase was purchased from Finnzymes (Thermo Electron). DpnI and the plasmid miniprep kit was from Fermentas. The oligonucleotide primers were from Invitrogen. Sterile 384- and 96-well flat bottom plates and black 96-well half-area plates were obtained from Greiner Bio One (Kremsmünster, Austria). Bug Buster[®] 10X protein extraction reagent and Lysonase[™] were from Novagen (Merck, Darmstadt, Germany). We used a BMG Labtech Omega plate reader for 96-well absorbance and fluorescence intensity measurements.

Library generation and subcloning

Firstly, we generated a library of the 454 bp stretch carrying random mutations using the GeneMorphII kit and the following primers: forward primer – 5'-CCTCAACAAGGAGCTCAGCGACTGG-3';

reverse primer -5'-CGGGATGCCCGCCGGTGGTGG-3'. Including the flanking primers, the amplified PCR product had a total length of 499 bp – Table vii shows the PCR conditions. Subsequently, we purified the PCR product using preparative agarose gel electrophoresis. Then, we subcloned the library into the pASK-IBA7plus vector employing an overlap extension PCR cloning method (Bryksin and Matsumura 2010). To this end, we used the gene library as a “megaprimer” in a PCR reaction with a high fidelity polymerase (see Table vii and Figure XX). We added DpnI to digest residual WT template according to the manufacturer’s recommendations. We then transformed the subcloned library into electrocompetent *E. coli* JM109 cells and plated aliquots on LB plates supplemented with ampicillin (115 µg/mL). Typically we had approximately 2000 clones per µL of transformed PCR reaction. We isolated plasmids of 10 single colonies with a minipreparation kit and sent them for sequencing to verify the mutation rate of 5 ± 2 mutations in the mutated ScTPase gene stretch.

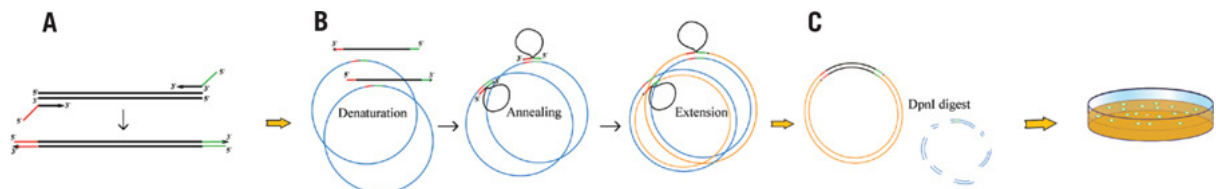


Figure XX. Overlap extension PCR cloning (Bryksin and Matsumura 2010).

Table vii. PCR conditions for library generation.

PCR	PCR reaction mixture	Temperature cycling	
ScTPase mutant library (499 bp)	1 X Mutazyme II reaction buffer	95°C 2 min	
	0.2 mM dNTPs	95°C 30 s	
	0.5 mM forward primer	64°C 30 s	35x
	0.5 mM reverse primer	72°C 1 min	
	template (ScTPase WTII)	72°C 10 min	
Mutazyme II DNA polymerase (2.5 U)			
Overlap extension PCR cloning	1 X Phusion HF buffer	100°C 2 min	
	0.25 mM dNTPs	98°C 30 s	18x
	~75 ng Megaprimer	60°C 30 s	
	~40 ng template (ScTPase WTII)	68°C 7 min 50s	
	Phusion polymerase (0.2 U)	68°C 10 min	

Transformation, picking and cultivation of mutants

We transformed the library into electrocompetent *E. coli* JM109 cells and plated aliquots on LB plates supplemented with ampicillin (115 µg/mL). The cells were allowed to grow overnight (~16 hours) at 37°C. Subsequently, 10000 single colonies were picked using a picking robot (Genetix QPix). The robot transferred the single colonies to 384-well plates with 50 µL LB-ampicillin medium (115 µg/mL ampicillin). We left 16 wells per plate empty as controls. These pre-culture plates were allowed to grow at 37°C for 36 hours without shaking in controlled humidity to avoid evaporation. The OD₆₀₀ was controlled routinely. We inoculated the main culture plates (96-well plates with 200 µL LB-ampicillin medium per well) using a sterile 96-pin stamp. This stamp transfers an approximate volume

of 5 μL . We supplemented the residual pre-culture volume with sterile glycerol (60%, 20 $\mu\text{L}/\text{well}$) and deep froze them at -70°C . We incubated the main cultures at 37°C without shaking for 21 hours in controlled humidity until the OD_{600} reached approximately 0.6–0.8. At this point we induced protein expression by addition of AHT¹⁵. Upon induction we allowed for protein expression at 18°C without shaking for 24 hours. Finally, we harvested the cells by centrifugation at 4000 rpm and 4°C for 15 minutes, disposed of the supernatant and deep froze the pellets at -20°C .

Cell disruption and library screening

For the screening we first of all disrupted the cells using Bug Buster[®] 10X protein extraction reagent according to the manufacturer's protocol. To this end we diluted the 10X concentrate in 50 mM MES buffer (pH 6.6) supplemented with 1 mM dithiothreitol (DTT) and 2 $\mu\text{L}/\text{mL}$ Lysonase[™]. Using a pipetting station (Tomtec Quadra Tower) we added 40 $\mu\text{L}/\text{well}$ of the 1X Bug Buster[®] mixture to the frozen pellets and resuspended twenty times. Then, we incubated the plates at ambient temperature for 30 minutes at 750 rpm on a plate incubator and eventually added 40 μL of 50 mM MES buffer (pH 6.6) to each well. We centrifuged the plates for 30 minutes at 4000 rpm and 4°C . The cleared supernatant thus obtained is referred to as crude cell extract.

For the screening for synthase activity in the *Sc*TPase library we prepared a screening mixture which would have the following concentrations upon addition of the crude extract: 250 mM trehalose, 0.5 mM UDP, 10 mM NAD^+ , 50 mM MES pH 6.6 and 2 mU hUGDH/well. Again, using the pipetting station, we put 30 $\mu\text{L}/\text{well}$ of this screening mixture in black half-area 96-well plates, transferred 20 μL of the fresh crude cell extracts in each well and mixed by resuspension. We then measured the NADH fluorescence intensity continuously for 24 hours ($\lambda_{\text{Ex}}=340\text{ nm}$, $\lambda_{\text{Em}}=430\text{--}450\text{ nm}$). We also had appropriate controls containing UDP-glucose or crude extracts of “empty” cells. In order to probe the quality of the cell disruption, we randomly picked a plate and checked for WT activity using the continuous coupled assay as described for the *Sc*TPase elsewhere (Eis and Nidetzky 1999).

2.3.1.3. Results and outlook

After having screened 10000 clones we identified several possible (signal intensity approximately double of the blank) and one definitely positive hit (signal intensity approximately five times of the blank) whose identification is pending so far. Next, we plan to re-screen potentially positive clones and to identify and characterize the positive hit. Generally, the dehydrogenase-coupled assay was well suited for the screening in crude extracts.

¹⁵ At induction, we added 10 μL of the following mix per well of the main cultures: LB-medium with 2.3 mg/mL ampicillin and 4 $\mu\text{g}/\text{mL}$ AHT.

2.4. Conclusion and outlook

In the quest for the “synthase switch” in *ScTPase* we went from rational design to directed evolution of a GT4 phosphorylase. Despite of promising hints and results gained through primary and secondary sequence alignments, homology modeling and site-directed mutagenesis, the determinants of the different donor substrate specificity in phosphorylases and synthases remained elusive. Generally, rational design was hampered due to the lack of a crystal structure of *ScTPase*. In the course of these rational investigations we identified a part of the C-terminal region of *ScTPase* as a promising target for directed evolution. Unfortunately, there is no reliable general assay available for GTs. As a stable and unambiguous HTS-assay is crucial for a successful evolutionary approach, we tested four different assays in terms of stability, unambiguousness, HTS-suitability and general applicability: a pH-shift assay, an immuno-assay (Transcreener[®]), a phosphatase-coupled assay, and a dehydrogenase-coupled assay. In spite of its specificity which limits this assay to UDP-glucose dependent transferases, the dehydrogenase-coupled assay was suited best for our task. Firstly, because evolutionary related transferases sharing high sequence identity with *ScTPase* – such as *PhTreT* – use UDP-glucose as donor-sugar; secondly, because this assay targets the NDP-sugar produced in the degradation direction of the reaction and hence increased hydrolysis cannot cause false positives; and finally, because long-term continuous measurement is possible even in crude extracts which enables for detection of mutants with very low activity. Using this assay, we screened a library of 10000 clones. We found a couple of possible hits which need to be verified in a re-screening due to their low activity. We also identified one positive hit whose sequence still needs to be identified and characterized.

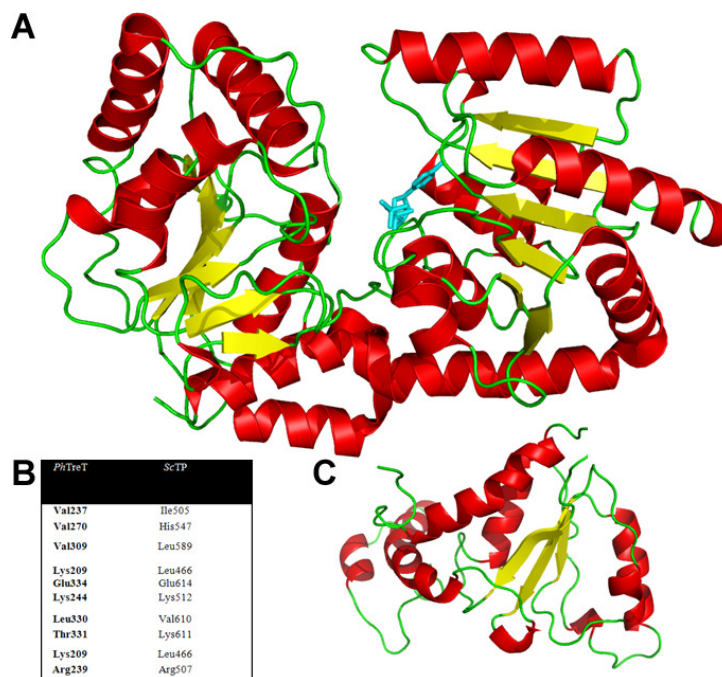


Figure XXI. Novel model of *ScTPase* based on the *PhTreT* crystal structure. A: *ScTPase* C-terminus model (~65 kDa) with UDP. B: residues which have been identified to be responsible for UDP-binding in *PhTreT* and the respective residues in *ScTPase*. C: *ScTPase* N-terminus model (~16 kDa).

Generally, rational design is based on knowledge of the structure of the target enzyme. There is no structural information available for *ScTPase*, however, due to the high degree of structural conservation of the GT-B-fold within the GT4 family modeling should give good results even with low sequence identity members such as MshA (approximately 18 % sequence identity). Recently, the structure of *PhTreT* was solved at 2.3 Å resolution. *PhTreT* shares 28% sequence identity with *ScTPase*, as well as some conserved regions and motifs (Figure XVIII and Figure XIX). We calculated a new model of the 65 kDa part of *ScTPase* based on the *PhTreT* structure using Yasara¹⁶ (Krieger *et al.* 2009). Upon storage at 4°C, *ScTPase* spontaneously splits in a shorter and completely inactive N-terminal piece (approximately 16 kDa) and an approximately 65 kDa part which is fully active (Eis *et al.* 2001, Goedl *et al.* 2006). We performed a p-blast search¹⁷ with the short N-terminal part and interestingly, it appears to have the closest similarity with an ATPase domain of paraplegin¹⁸, a human mitochondrial protease (Karlberg *et al.* 2009) and also modeled this part of *ScTPase*. However, the exact function of the N-terminal part is not clear yet.

We anticipate that our combined rational and evolutionary approach, comprising homology modeling, site-directed and random mutagenesis, will lead us to new insights in what discriminates a synthase from a phosphorylase.

¹⁶ <http://www.yasara.org/homologymodeling.htm>, last visited 07/28/2011

¹⁷ <http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>, last visited 07/28/2011

¹⁸ pdb: 2QZ4

2.5. Abbreviations

aa	amino acid
ADP-glucose	adenosine 5'-diphospho- glucose
AHT	anhydrotetracycline
ATP	adenosine triphosphate
AviGT4	eurekanate-attachment enzyme from <i>Streptomyces viridochromogenes</i>
bp	base pair
CAZymes	carbohydrate-active enzymes
CcTPase	putative trehalose phosphorylase from <i>Coprinopsis cinerea</i>
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
ENTP	ectonucleoside triphosphate diphosphohydrolase
FACS	fluorescence-activated cell sorting
G1P	α -D-glucose 1-phosphate
G6P	α -D-glucose 6-phosphate
G6PDH	G6P-dehydrogenase
GDP	guanosine diphosphate
GfTPase	trehalose phosphorylase from <i>Grifola frondosa</i>
GlgA	glycogensynthase 1 from <i>Agrobacterium tumefaciens</i>
Gluc	α -D-glucose
GT	glycosyltransferase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HK	hexokinase
HTS	high-throughput screening
hUGDH	human UDP-glucose dehydrogenase
LB	lysogeny broth
LDH	lactic dehydrogenase
MES	2-(<i>N</i> -morpholino)ethanesulfonic acid
MOPS	3-(<i>N</i> -morpholino)propanesulfonic acid
MshA	1L-myo-inositol-1P- α - <i>N</i> -acetylglucoaminytransferase from <i>Corynebacterium glutamicum</i>
MWCO	molecular weight cut off
NAD⁺/NADH	nicotinamide adenine dinucleotide (oxidized/reduced form)
NDP	nucleoside diphosphate
NMP	nucleoside monophosphate
NTP	nucleoside triphosphate
OleD	olandomycin glycosyltransferase
OtsA	trehalose 6-phosphate synthase from <i>Escherichia coli</i>
PaGlgA	glycogen synthase from <i>Pyrococcus abyssi</i>
PaTs	trehalose synthase from <i>Pyrobaculum aerophilum</i>
PCR	polymerase chain reaction
PEP	phosphoenol pyruvate
PfTs	trehalose synthase from <i>Pyrococcus furiosus</i>
PGM	phosphoglucomutase
PhTreT	trehalose synthase <i>Pyrococcus horikoshii</i>
PimA	phosphatidylinositol mannosyltransferase from <i>Mycobacterium smegmatis</i>
PK	pyruvate kinase
PpTPase	trehalose phosphorylase from <i>Pleurotus pulmonarius</i>
PsTPase	trehalose phosphorylase from <i>Lentinus sajor-caju</i>
RNA	ribonucleic acid

SAP	shrimp alkaline phosphatase
ScTPase	trehalose phosphorylase from <i>Schizophyllum commune</i>
SmTs	putative trehalose synthase from <i>Staphylothermus marinus</i>
SpsA	sucrose phosphate synthase from <i>Halothermothrix orenii</i>
SuSy	sucrose synthase
TES	2-[[1,3-dihydroxy-2-(hydroxymethyl)propan-2-yl]amino]ethanesulfonic acid
Tris	tris(hydroxymethyl)aminomethane
UDP	uridine diphosphate
UDP-GlcNAc	uridine 5'-diphospho-N-acetyl-D-glucosamine
UDP-glucose	uridine 5'-diphospho-glucose
UTP	uridine triphosphate
WaaG	L-glycero-D-manno-heptose-II α -1,3-glycosyltransferase I from <i>Escherichia coli</i>
WT	wild type

2.6. References

- Abdian, P. L., Lellouch, A. C., Gautier, C., Ielpi, L. & Geremia, R. A. Identification of essential amino acids in the bacterial alpha -mannosyltransferase aceA. *The Journal of biological chemistry* **275**, 40568-40575, (2000).
- Absmanner, B., Schmeiser, V., Kampf, M. & Lehle, L. Biochemical characterization, membrane association and identification of amino acids essential for the function of Alg11 from *Saccharomyces cerevisiae*, an alpha1,2-mannosyltransferase catalysing two sequential glycosylation steps in the formation of the lipid-linked core oligosaccharide. *The Biochemical journal* **426**, 205-217, (2010).
- Aharoni, A., Thieme, K., Chiu, C. P., Buchini, S., Lairson, L. L., Chen, H., Strynadka, N. C., Wakarchuk, W. W. & Withers, S. G. High-throughput screening methodology for the directed evolution of glycosyltransferases. *Nat Methods* **3**, 609-614, (2006).
- Argüelles, J. C. Physiological roles of trehalose in bacteria and yeasts: a comparative analysis. *Arch Microbiol* **174**, 217-224, (2000).
- Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* **98**, 10037-10041, (2001).
- Breton, C., Snajdrova, L., Jeanneau, C., Koca, J. & Imberty, A. Structures and mechanisms of glycosyltransferases. *Glycobiology* **16**, 29R-37R, (2006).
- Bryksin, A. V. & Matsumura, I. Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. *Biotechniques* **48**, 463-465, (2010).
- Bryson, K., McGuffin, L. J., Marsden, R. L., Ward, J. J., Sodhi, J. S. & Jones, D. T. Protein structure prediction servers at University College London. *Nucleic Acids Res* **33**, W36-38, (2005).
- Chang, C. W. Predictable enzymatic glycosylation. *Chem Biol* **16**, 579-580, (2009).
- Davies, G. & Henrissat, B. Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853-859, (1995).
- Davies, G. J. Sweet secrets of synthesis. *Nat Struct Biol* **8**, 98-100, (2001).
- Davies, G. J., Gloster, T. M. & Henrissat, B. Recent structural insights into the expanding world of carbohydrate-active enzymes. *Curr Opin Struct Biol* **15**, 637-645, (2005).
- Eis, C. & Nidetzky, B. Characterization of trehalose phosphorylase from *Schizophyllum commune*. *The Biochemical journal* **341** (Pt 2), 385-393, (1999).
- Eis, C., Watkins, M., Prohaska, T. & Nidetzky, B. Fungal trehalose phosphorylase: kinetic mechanism, pH-dependence of the reaction and some structural properties of the enzyme from *Schizophyllum commune*. *The Biochemical journal* **356**, 757-767, (2001).
- Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., Pieper, U. & Sali, A. Comparative protein structure modeling using Modeller. *Curr Protoc Bioinformatics* **Chapter 5**, Unit 5 6, (2006).

- Gibson, R. P., Turkenburg, J. P., Charnock, S. J., Lloyd, R. & Davies, G. J. Insights into trehalose synthesis provided by the structure of the retaining glucosyltransferase OtsA. *Chem Biol* **9**, 1337-1346, (2002).
- Goedl, C., Griessler, R., Schwarz, A. & Nidetzky, B. Structure-function relationships for *Schizophyllum commune* trehalose phosphorylase and their implications for the catalytic mechanism of family GT-4 glycosyltransferases. *The Biochemical journal* **397**, 491-500, (2006).
- Goedl, C. & Nidetzky, B. The phosphate site of trehalose phosphorylase from *Schizophyllum commune* probed by site-directed mutagenesis and chemical rescue studies. *FEBS J* **275**, 903-913, (2008).
- Goedl, C., Schwarz, A., Mueller, M., Brecker, L. & Nidetzky, B. Mechanistic differences among retaining disaccharide phosphorylases: insights from kinetic analysis of active site mutants of sucrose phosphorylase and alpha,alpha-trehalose phosphorylase. *Carbohydrate research* **343**, 2032-2040, (2008).
- Jones, D. T. Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* **292**, 195-202, (1999).
- Karlberg, T., van den Berg, S., Hammarstrom, M., Sagemark, J., Johansson, I., Holmberg-Schiavone, L. & Schuler, H. Crystal structure of the ATPase domain of the human AAA+ protein paraplegin/SPG7. *PLoS One* **4**, e6975, (2009).
- Kittl, R. & Withers, S. G. New approaches to enzymatic glycoside synthesis through directed evolution. *Carbohydrate research* **345**, 1272-1279, (2010).
- Krieger, E., Joo, K., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D. & Karplus, K. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins* **77 Suppl 9**, 114-122, (2009).
- Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem* **77**, 521-555, (2008).
- Lairson, L. L. & Withers, S. G. Mechanistic analogies amongst carbohydrate modifying enzymes. *Chem Commun (Camb)*, 2243-2248, (2004).
- Lobley, A., Sadowski, M. I. & Jones, D. T. pGenTHREADER and pDomTHREADER: new methods for improved protein fold recognition and superfamily discrimination. *Bioinformatics* **25**, 1761-1767, (2009).
- Muniz, J. R., Alves, C. A., de Pieri, C., Beltramini, L. M., Selistre-de-Araujo, H. S., Vettore, A. L., da Silva, F. R., Arruda, P., Garratt, R. C., Oliva, G. & Souza, D. H. Overexpression, purification, biochemical characterization, and molecular modeling of recombinant GDP-mannosyltransferase (GumH) from *Xylella fastidiosa*. *Biochem Biophys Res Commun* **315**, 485-492, (2004).
- Park, S. H., Park, H. Y., Sohng, J. K., Lee, H. C., Liou, K., Yoon, Y. J. & Kim, B. G. Expanding substrate specificity of GT-B fold glycosyltransferase via domain swapping and high-throughput screening. *Biotechnol Bioeng* **102**, 988-994, (2009).
- Rosen, M. L., Edman, M., Sjoström, M. & Wieslander, A. Recognition of fold and sugar linkage for glycosyltransferases by multivariate sequence analysis. *The Journal of biological chemistry* **279**, 38683-38692, (2004).
- Ryu, S. I., Park, C. S., Cha, J., Woo, E. J. & Lee, S. B. A novel trehalose-synthesizing glycosyltransferase from *Pyrococcus horikoshii*: molecular cloning and characterization. *Biochem Biophys Res Commun* **329**, 429-436, (2005).
- Sawangwan, T., Goedl, C. & Nidetzky, B. Glucosylglycerol and glucosylglycerate as enzyme stabilizers. *Biotechnol J* **5**, 187-191, (2010).
- Schmidt, T. G. & Skerra, A. The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat Protoc* **2**, 1528-1535, (2007).
- Szewczyk, E., Nayak, T., Oakley, C. E., Edgerton, H., Xiong, Y., Taheri-Talesh, N., Osmani, S. A. & Oakley, B. R. Fusion PCR and gene targeting in *Aspergillus nidulans*. *Nat Protoc* **1**, 3111-3120, (2006).
- Thompson, J. D., Gibson, T. J. & Higgins, D. G. Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics* **Chapter 2**, Unit 2 3, (2002).

- Troutman, J. M. & Imperiali, B. *Campylobacter jejuni* PglH is a single active site processive polymerase that utilizes product inhibition to limit sequential glycosyl transfer reactions. *Biochemistry* **48**, 2807-2816, (2009).
- Wang, W. & Malcolm, B. A. Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis. *Biotechniques* **26**, 680-682, (1999).
- Woo, E. J., Ryu, S. I., Song, H. N., Jung, T. Y., Yeon, S. M., Lee, H. A., Park, B. C., Park, K. H. & Lee, S. B. Structural insights on the new mechanism of trehalose synthesis by trehalose synthase TreT from *Pyrococcus horikoshii*. *J Mol Biol* **404**, 247-259, (2010).

2.7. Additional acknowledgements

Christine Winkler provided helpful information for the library generation, the transformation, and the cultivation of the library. Gustavo Arruda Bezerra helped in p-Blast searches and modeling with Yasara.

Appendix

List of publications

Papers

- Bubner, P., Dohr, J., Plank, H., Mayrhofer, C., and Nidetzky, B. (2011). Cellulases dig deep: *in situ* observation of the mesoscopic structural dynamics of enzymatic cellulose degradation. *J. Biol. Chem.*, under review.
- Bubner, P., Plank, H., Nidetzky, B. (2011). Visualizing cellulase action. Review in preparation for *Biotech. Bioeng.*
- Bubner, P., Luley-Goedl, C., Egger, S., Nidetzky, B. (2011). High-throughput assays for the directed evolution of glycosyltransferases. Full paper in preparation for *ChemBioChem*.

Oral Presentations

- Bubner, P., Luley-Goedl, C., Nidetzky, B. (2008) Determinants of donor substrate specificity in glycosyltransferases. Privatissimum Angewandte Enzymologie/Biotechnologie (Graz, Austria). May 2008.
- Bubner, P., Luley-Goedl, C., Nidetzky, B. (2009) Glycorandomization – strategies to promote substrate promiscuity in glycosyltransferases. Dissertantinnenseminar 1 (Graz, Austria). October 2008.
- Bubner, P., Luley-Goedl, C., Nidetzky, B. (2009) Structure-function relationships in *Schizophyllum commune* Trehalose Phosphorylase. Dissertantinnenseminar 2 (Graz, Austria). February 2009.
- Bubner, P., Luley-Goedl, C., Nidetzky, B. (2009) From rational design to directed evolution: trying to trick *Schizophyllum commune* Trehalose Phosphorylase. Privatissimum Angewandte Enzymologie/Biotechnologie (Graz, Austria). October 2009.
- Bubner, P., Luley-Goedl, C., Egger, S., Nidetzky, B. (2010) High-throughput screening of glycosyltransferases: opportunities and pitfalls. Privatissimum Angewandte Enzymologie/Biotechnologie (Graz, Austria). May 2010.
- Bubner, P., Luley-Goedl, C., Egger, S., Nidetzky, B. (2010) Novel high-throughput screening methodologies to advance directed evolution of glycosyltransferases. Privatissimum Angewandte Enzymologie / Biotechnologie (Graz, Austria). January 2010.
- Bubner, P., Dohr, J., Plank, H., Mayrhofer, C., Nidetzky, B. (2011) Cellulase action, unraveled! AFM-based investigations deliver novel insights into the degradation of cellulose. Privatissimum Angewandte Enzymologie / Biotechnologie (Graz, Austria). January 2011.
- Bubner, P., Luley-Goedl, C., Egger, S., Nidetzky, B. (2011) Novel high-throughput screening methodologies to advance directed evolution of glycosyltransferases. 15. Österreichischer Kohlenhydratworkshop (Graz, Austria). February 2011.

Poster Presentations

- Bubner, P., Luley-Goedl, C., Egger, S., Nidetzky, B. (2010) Novel high-throughput screening methodologies to advance directed evolution of glycosyltransferases. Pacificchem (Honolulu, HI, USA). December 2010.
- Bubner, P., Luley-Goedl, C., Sawangwan, T., Schwarz, A., Mueller, M., Nidetzky, B. (2010) Tapping the potential of a biocatalyst: Industrial scale synthesis of natural products using *Leuconostoc mesenteroides* sucrose phosphorylase. Pacificchem (Honolulu, HI, USA). December 2010.
- Bubner, P., Luley-Goedl, C., Egger, S., Nidetzky, B. Novel high-throughput screening methodologies to advance directed evolution of glycosyltransferases. Doc Day, NAWI Graz Doctoral School Molecular Biosciences and Biotechnology (Graz, Austria). February 2011.