Novel bacterial enzymes for plant biomass degradation discovered by meta-omics approach
Marutha Muthu, Mukil

IMPORTANT NOTE: You are advised to consult the publisher’s version (publisher’s PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2017

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
General discussion, conclusion and future prospective
1. Introduction
The objective of this study was to examine existing wheat straw (WS) degradation strategies on the basis of traditional methods coupled to those developed on the basis of metagenomics approaches. In order to do so, a first focus needed to be placed on WS quality, as this was thought to strongly determine the type of degradation strategy that might be optimal. Moreover, variation in WS quality was deemed to be a key factor. In the following section, I will discuss the relevance of WS quality and variability therein, with the implications for its degradability. Then, I discuss my achievements with respect to physical pretreatment of WS, metagenomics and the possibilities of the newly found enzymes.

2. Wheat straw – Quality, structure, composition and implications for degradation
Wheat straw is produced in billions of tons every year throughout the world. It has good utilization potential but has not been fully explored. Raw wheat straw (RWS) includes cellulose (30-50%), hemicellulose (15-33%) and lignin (5-26%) fractions, next to other materials (such as pectin, sugars, acids and proteins) [1]. Here, I opted for the use of WS harvested from a Dutch field as the target material for degradation improvement.

2.1. Variation in chemical composition
WS has generally a high cellulose content, similar to rice straw (RS) and poplar. The hemicellulose content of WS is approximately similar to that of rice straw but 1.45 times higher than that of poplar [2]. As is obvious from the ranges given in the foregoing, a huge variation exists in the chemical composition of WS, as driven by various factors such as cultivation conditions, cultivar type, climate, soil type and location, as well as soil treatment (Table 1). For instance, WS from tropical summer wheat is rich in cellulose and hemicellulose moieties (37-50% and 15-33%, respectively) that are often harvested (at well-matured conditions) between summer and autumn. On the other hand, winter WS of different wheat cultivars showed variation in chemical composition even when grown under the same climatic condition but at different geographical locations (around Denmark). In the latter WS, lignin and ash content was different between pretreated and unpretreated straw (Table 1; [3]). Moreover, it contained high (19-27%) lignin contents, which provides structural strength and mechanical support to the plant stem and leaf needed for often windy winter conditions. In comparison, summer WS had a low level (5.6%) of lignin [4]. In addition, the physical parts of wheat plants (internodes, leaf-sheath, leaf-blade, nodes and grains) also showed large variation in the WS chemical composition [5]. Here, I opted for the use of a local WS, which was predicted to have high (hemi)cellulose contents, but also nearly 22% of lignin. Such raised lignin contents are often difficult to remove from the (hemi)cellulose matrix, impairing the degradation processes. Thus, to efficiently
degrade the (hemi)celluloses in our local WS, special treatment strategy had to be designed. Overall, in the light of the huge variation and complexity of the composition of WS, we are far from a thorough understanding of what would constitute generally optimal WS degradation strategies.

Table 1: Variation in WS chemical composition

<table>
<thead>
<tr>
<th>No</th>
<th>Cellulose</th>
<th>Hemicellulose</th>
<th>Lignin</th>
<th>Ash</th>
<th>Others</th>
<th>Country/Type</th>
<th>Sampling season</th>
<th>Plant parts</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>45.5</td>
<td>31.3</td>
<td>22.2</td>
<td>ND</td>
<td>1</td>
<td>The Netherlands</td>
<td>Winter</td>
<td>Straw</td>
<td>[43]</td>
</tr>
<tr>
<td>2</td>
<td>31.6</td>
<td>16.2</td>
<td>18.5</td>
<td>17.2</td>
<td>16.5</td>
<td>China</td>
<td>Winter</td>
<td>Straw</td>
<td>[44]</td>
</tr>
<tr>
<td>3</td>
<td>41.3</td>
<td>28.2</td>
<td>21.4</td>
<td>ND</td>
<td>9.1</td>
<td>Denmark</td>
<td>Winter</td>
<td>Straw</td>
<td>[45]</td>
</tr>
<tr>
<td>4</td>
<td>41.6</td>
<td>18.7</td>
<td>22.2</td>
<td>ND</td>
<td>17.5</td>
<td>Germany</td>
<td>Winter</td>
<td>Straw</td>
<td>[46]</td>
</tr>
<tr>
<td>5</td>
<td>31.2</td>
<td>21.7</td>
<td>23.4</td>
<td>8.1</td>
<td>15.6</td>
<td>China</td>
<td>Summer</td>
<td>Straw</td>
<td>[47]</td>
</tr>
<tr>
<td>6</td>
<td>37.1</td>
<td>23.5</td>
<td>15.8</td>
<td>8</td>
<td>15.6</td>
<td>Estonia</td>
<td>Summer</td>
<td>Straw</td>
<td>[48]</td>
</tr>
<tr>
<td>7</td>
<td>40.7</td>
<td>25.2</td>
<td>21.7</td>
<td>7.6</td>
<td>4.8</td>
<td>China</td>
<td>Summer</td>
<td>Stem</td>
<td>[49]</td>
</tr>
<tr>
<td>8</td>
<td>35.1</td>
<td>27.5</td>
<td>14.2</td>
<td>11.3</td>
<td>11.9</td>
<td>China</td>
<td>Summer</td>
<td>Leaf</td>
<td>[49]</td>
</tr>
<tr>
<td>9</td>
<td>43.3</td>
<td>31.7</td>
<td>17.3</td>
<td>ND</td>
<td>7.7</td>
<td>China</td>
<td>Summer</td>
<td>Straw</td>
<td>[50]</td>
</tr>
<tr>
<td>10</td>
<td>30.6</td>
<td>32.3</td>
<td>16.8</td>
<td>5</td>
<td>15.3</td>
<td>Portugal</td>
<td>Summer</td>
<td>Straw</td>
<td>[51]</td>
</tr>
<tr>
<td>11</td>
<td>43.1</td>
<td>19.2</td>
<td>25</td>
<td>1.9</td>
<td>10.8</td>
<td>China</td>
<td>Summer</td>
<td>Straw</td>
<td>[52]</td>
</tr>
<tr>
<td>12</td>
<td>35.2</td>
<td>22.2</td>
<td>22.1</td>
<td>7.4</td>
<td>13.1</td>
<td>Spain</td>
<td>Summer</td>
<td>Straw</td>
<td>[53]</td>
</tr>
<tr>
<td>13</td>
<td>30.8</td>
<td>15.9</td>
<td>5.6</td>
<td>12.5</td>
<td>15.2</td>
<td>Mexico</td>
<td>Summer</td>
<td>Straw</td>
<td>[4]</td>
</tr>
<tr>
<td>14</td>
<td>35.4</td>
<td>21.6</td>
<td>22.6</td>
<td>4.4</td>
<td>16</td>
<td>Cary, NC</td>
<td>Summer</td>
<td>Straw</td>
<td>[54]</td>
</tr>
<tr>
<td>15</td>
<td>36.8</td>
<td>31.6</td>
<td>20.8</td>
<td>ND</td>
<td>9.1</td>
<td>South Korea</td>
<td>Summer</td>
<td>Straw</td>
<td>[54]</td>
</tr>
<tr>
<td>16</td>
<td>36.4</td>
<td>33.5</td>
<td>11.2</td>
<td>12.5</td>
<td>14.6</td>
<td>China</td>
<td>Summer</td>
<td>Straw</td>
<td>[56]</td>
</tr>
<tr>
<td>17</td>
<td>32.4</td>
<td>20.9</td>
<td>17.7</td>
<td>ND</td>
<td>29</td>
<td>Boulder, CO</td>
<td>Summer</td>
<td>Straw</td>
<td>[57]</td>
</tr>
<tr>
<td>18</td>
<td>39.3</td>
<td>20.7</td>
<td>16.2</td>
<td>ND</td>
<td>23.8</td>
<td>France</td>
<td>Winter</td>
<td>Straw</td>
<td>[58]</td>
</tr>
<tr>
<td>19</td>
<td>35.9</td>
<td>23.9</td>
<td>19.3</td>
<td>4.1</td>
<td>16.8</td>
<td>Denmark</td>
<td>Summer</td>
<td>Straw</td>
<td>[59]</td>
</tr>
<tr>
<td>20</td>
<td>32.6</td>
<td>24.2</td>
<td>26.5</td>
<td>4.6</td>
<td>12.1</td>
<td>Sweden</td>
<td>Winter</td>
<td>Straw</td>
<td>[60]</td>
</tr>
<tr>
<td>21</td>
<td>37.1</td>
<td>27.6</td>
<td>20.9</td>
<td>3.7</td>
<td>10.7</td>
<td>cv: Ambition, Denmark</td>
<td>WWS</td>
<td>Straw</td>
<td>[3]</td>
</tr>
<tr>
<td>22</td>
<td>35.6</td>
<td>27.6</td>
<td>19.9</td>
<td>4.7</td>
<td>12.2</td>
<td>cv: Hereford, Denmark</td>
<td>WWS</td>
<td>Straw</td>
<td>[3]</td>
</tr>
<tr>
<td>23</td>
<td>36.4</td>
<td>27.2</td>
<td>19.9</td>
<td>5.2</td>
<td>11.3</td>
<td>cv: Skalmeje, Denmark</td>
<td>WWS</td>
<td>Straw</td>
<td>[3]</td>
</tr>
<tr>
<td>24</td>
<td>36.6</td>
<td>28.4</td>
<td>19.3</td>
<td>4.2</td>
<td>11.5</td>
<td>cv: Smuggler, Denmark</td>
<td>WWS</td>
<td>Straw</td>
<td>[3]</td>
</tr>
<tr>
<td>25</td>
<td>35.9</td>
<td>27</td>
<td>20.1</td>
<td>4.6</td>
<td>12.4</td>
<td>cv: Frument, Denmark</td>
<td>WWS</td>
<td>Straw</td>
<td>[3]</td>
</tr>
</tbody>
</table>

cv: Cultivar; WWS: winter wheat straw

2.2. Structural interconnectedness of WS chemical moieties

In WS, cellulose is coupled to hemicellulose by hydrogen bonds, whereas hemicellulose is attached to lignin via a covalent bond. The most abundant homopolymer sugar is a linear chain of several hundred to ten thousand β-(1→4) linked D-glucose units. The cellulose chains group together to form microfibrils and these microfibrils together form cellulose fibers. Cellulose fibers have different morphologies based on the variation in fibril structure [6]. The next major moiety, hemicellulose, is a heteropolymer consisting mainly of xylans, xyloglucans, glucomannans and mannans. Xylose is the main sugar unit for the first two polymers, whereas mannose is the main backbone sugar for the latter two. The xylose backbone is joined by β-1, 4 glycosidic bonds, often branched with other sugars such as arabinose, glucose, galactose and fructose in β-1 or β-6 positions. Hemicellulose covers the microfibrils of cellulose, and therefore a significant amount of hemicellulose has to be digested for efficient cellulose digestibility. A third polymer, pectin, forms a dynamic structure of the wheat cell wall located between the cellulose microfibrils. It is composed of hetero-polysaccharides such as galacturonic acid residues substituted by methoxyl esters. Pectin is soluble and can be easily digested by enzymes. Finally, the WS lignin moiety is a
multifaceted, hard, insoluble and heterogeneous mixture of polymers linked via a variety of C-C and C-O ether linkages that protects the aforementioned polysaccharides from microbial degradation. In order to obtain the desired end products for bio-refineries, the removal or ‘loosening’ of lignin is an important key step.

2.3. A combination of physical and enzymatic methods is needed for wheat straw degradation

The variation in WS chemical composition combined with the structural interconnectedness highly influences the degradation process. Therefore, selecting a suitable WS degradation method is key to efficiently degrade any new WS. Most currently used WS degradation methods involve a first (physical) pretreatment followed by enzymatic hydrolysis. Table 2 extensively lists and compares the efficiencies of different WS degradation methods, as described in the literature.

Table 2: Wheat straw: different pre-treatments

<table>
<thead>
<tr>
<th>No</th>
<th>Pre-treatment (PT) methods</th>
<th>PT Conditions</th>
<th>Enzyme hydrolysis (EH)</th>
<th>EH Condition</th>
<th>Glucose &amp; Xylose yield (%)</th>
<th>EH rate (%)</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Steam explosion (SE)</td>
<td>210°C and 2.5 min and (4% (v/v) H2SO4, 120°C and 30 min)</td>
<td>NS50013, NS50010, endoxylanase (XlnC)</td>
<td>50°C, 24 h</td>
<td>28.6 &amp; 14.7 g/L</td>
<td>67% / 82%</td>
<td>[8]</td>
</tr>
<tr>
<td>2</td>
<td>PHP (phosphoric acid plus hydrogen peroxide)</td>
<td>70.2 % H3PO4 (w/w), 40 °C, 2.0 h</td>
<td>Celluclast 1.5L, Novozyme 188</td>
<td>50°C, 72 h</td>
<td>946.2 mg/g</td>
<td>95%</td>
<td>[44]</td>
</tr>
<tr>
<td>3</td>
<td>Hydrothermal pretreatment</td>
<td>180 °C, 20 min</td>
<td>Cellic CTec2, Cellic HTec2</td>
<td>50°C, 96 h</td>
<td>24.9 &amp; 7.4g</td>
<td>72%</td>
<td>[54]</td>
</tr>
<tr>
<td>4</td>
<td>Organosolv</td>
<td>0 % (v/v) aqueous EtOH, 210 °C, 0 min</td>
<td>Accellerase 1500</td>
<td>50°C, 72 h</td>
<td>814.9 mg/g</td>
<td>82%</td>
<td>[61]</td>
</tr>
<tr>
<td>5</td>
<td>Dilute alkaline</td>
<td>2 % (w/v) NaOH, 121 °C, 30 min</td>
<td>NS50013, NS50010, NS50030</td>
<td>50°C, 72 h</td>
<td>775 mg/g</td>
<td>78%</td>
<td>[62]</td>
</tr>
<tr>
<td>6</td>
<td>Wet explosion</td>
<td>35 % (v/v) H2O2, 180–185 °C, 15 min</td>
<td>Celluclast 1.5L, Novozyme 188</td>
<td>50°C, 72 h</td>
<td>741.6 mg/g</td>
<td>76%</td>
<td>[63]</td>
</tr>
<tr>
<td>7</td>
<td>Steam explosion</td>
<td>Soaked for 18 h at 45 °C in 0.9 % (w/v) H2SO4, or water, 180 °C, 10 min</td>
<td>Celluclast 1.5L, Novozyme 188</td>
<td>50°C, 72 h</td>
<td>761.6 mg/g</td>
<td>76%</td>
<td>[64]</td>
</tr>
</tbody>
</table>

PT: Pre-treatment; EH: Enzyme hydrolysis;

2.3.1. Pretreatment

A major hindrance in WS degradation is due to the problem of ‘recalcitrance’, i.e. the natural resistance of the plant cell wall to microbial and enzymatic deconstruction. To overcome this recalcitrance, pre-treatment is highly required, in order to i) disrupt the cell wall structure, ii) remove the lignin moieties and iii) improve the access of the enzymes to the sugar polymers. Pre-treatment specifically breaks down the macroscopic rigidity of WS and decreases the physical barriers to mass transport. Various pre-treatment methods have been concocted, which are biological, physical (mechanical comminution, extrusion), chemical (alkali, acid, ozonolysis, organosol, ionic liquids) and physico-chemical (steam explosion, liquid hot water, ammonia fiber explosion, wet oxidation, microwave pretreatment, ultrasound, CO2 explosion) in nature. An extensive overview of different studies revealed up to five methods to be promising for WS (Table 3). The optimal pre-treatment method has to elevate the sugar yields of the subsequent enzymatic hydrolysis. Also, it has to minimize the degradation level of
sugars as well as the formation of inhibitory compounds (5-HMF and furfural). I summarized the findings in Table 2. From the collective data, and placing a focus on WS, I distill that steam explosion can be superior in many aspects (Table 3). Moreover, it is performed with no chemicals and is economic and fairly energy-efficient [7]. The in-house steam explosion method that we developed was found to clearly disrupt many of the cell walls of the WS material and enhanced the break-down of cellulose fibres. This was evident when released sugars were quantified, as the steam-treated WS yielded 6 times more sugar than untreated WS (1.2 mg/ml versus 0.2 mg/ml) (Chapter 5). At the same time, our in-house method still revealed reduced levels of total sugars in the liquid fraction (1.2 g/L) when compared to a previous study (5.8 g/L) [8]. This suggested that major fractions of steam exploded WS material remained in the solid fraction, which was to be subjected to enzymatic hydrolysis treatment. However, a well-known problem is the generation of inhibitory compounds during WS steam explosion that affect any subsequent enzymatic treatment. Therefore, there is an increasing demand for new enzymes to effectively degrade the (hemi)cellulose moieties as well as tolerant to inhibitory compounds.

<table>
<thead>
<tr>
<th>Table 3: Different pre-treatments and their effect on lignocellulosic biomass (wheat straw)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE</td>
</tr>
<tr>
<td>Increases accessible surface area</td>
</tr>
<tr>
<td>Cellulose decrystallization</td>
</tr>
<tr>
<td>Hemicellulose solubilization</td>
</tr>
<tr>
<td>Lignin removal</td>
</tr>
<tr>
<td>Generation of toxic compounds</td>
</tr>
<tr>
<td>Lignin structure alteration</td>
</tr>
</tbody>
</table>

H: High; M: Medium; L: Low; n.d: not determined; -: not detected; SE: Stem explosion; AFEX: ammonia fibre expansion.

2.3.2. Enzymatic hydrolysis

Following steam explosion, enzymes are required to efficiently decompose the loosened up WS moieties into their primary components. Most likely, enzymes in the classes cellulases, hemicellulases, putative accessory enzymes and even laccases/ligninases are essential. Cellulases involve endo-1, 4-β-glucanases (EG) that cleave internal bonds and exo-1,4-β-glucanases that cleave two to four units from the ends of cellulose strands. Furthermore, cellobiohydrolase cleaves the disaccharide cellobiose into two glucose moieties. As discussed above, given its intimate linkage to cellulose, enhanced attack on hemicellulloses may offer major improvements in total degradation success. Various enzymes may be needed for this. Thus, de-branching enzymes like β-xylosidase, arabinofuranosidase, feruloyl esterase, acetylxylan esterase, and a-glucuronidase may be required to degrade side chains. When the side chains are degraded, the main xylan backbone is unlocked.
for xylanases to act. The unlocked xylan will be hydrolyzed to xylobiose moieties that can be attacked by β-xylosidase, thus releasing xylose monomers. It also cleaves the end of the xylo-oligosaccharide (xylan backbone) to release more xylose [9]. Finally, the lignin moiety appears to pose the main obstacle to degradation due to its aromatic nature and highly branched polymer network. The key issue here is the optimal loosening of the lignin from the other moieties, allowing enzymatic access to the latter. Several commercial enzyme cocktails (Celluclast 1.5L, CTec2, Spezyme CP, Novozymes 188) are currently available for the enzymatic hydrolysis of WS. These commercial enzymes are complex mixtures of approximately 80 to 200 proteins [10], with the specific function of each individual protein in the complex mixture being unclear [11]. In general, such enzymes are mined from microorganisms such as bacteria and fungi. For example, Celluclast 1.5L has been produced from the biodegradative fungus Trichoderma reesei. It contains five different endoglucanases (EG; Cel7B, Cel5A, Cel12A, Cel61A, and Cel45A, respectively), two cellobiohydrolases (CBH; CBHI (Cel7A) and CBHII (Cel6A)) and several xylanases [12]. In our work, steam-treated WS with commercial enzyme cocktails (Celluclast 1.5L) yielded relatively increased amount of glucose by ~180-200 % than treated WS alone [13]. Therefore, Celluclast 1.5L cocktail is best in recovering almost 80% of glucose (cellulose moiety) but less efficient in extracting (hemicellulose moiety) xylose (56%) [14]. Thus, more opportunities exist with respect to the exploration of efficient hemicellulases by metagenomics approach.

3. Metagenomics to get improved WS degrading enzymes
Metagenomics involves the direct extraction of DNA from microbiomes from environmental samples, often producing a metagenomics library. Such DNA encompasses genes that may be impossible to get at by cultivation-based methods. The metagenomic DNA thus obtained is sequenced and analyzed for genes of interest. Thus, selected genes for, for instance, biocatalysts can be singled out from the metagenomic DNA and cloned to obtain ‘clean’ production systems. Subsequently, the new biocatalytic enzyme activity will be tested. The overall method is defined as ‘functional or function-directed metagenomics’.

3.1. Functional screening of metagenomic libraries for identification of genes encoding novel enzymes
Metagenomics applied to WS degrader consortia has great potential to uncover genes for (hemi)cellulolytic enzymes [15]. The fed-batch set up with RWS used by us was supposed to select bacterial communities that enhance the prevalence of genes that encode key target enzymes [16] (Figure 1). Here, fosmid-based libraries were produced from 10-fold selected WS degrading microbial consortia. Our functional screening relied on the ready expression of the gene of interest detected by activity screening with specific substrate. Thus, activity-based screening was
successfully used as described by Ekkers et al (2012) [17] and van Elsas et al (2008) [18]. The screening method primarily focussed on the degradation of a specific substrate (mimicking bonds in WS moieties) which resulted in a colour change of a substrate, (Figure 2) finally allowing us to pre-select (hemi)cellulolytic clones.

Figure 1: Strategies used in this thesis: Screening enzymes to impact in wheat straw degradation

3.1.1. Substrate specificity - To address glycosidase activities, more than 15 distinct chromogenic and fluorimetric substrates were employed; varying levels of success were found [19]. For example, bromo-4-chloro-3-indolyl-β-D-galactopyranoside, the common substrate for β-galactosidases, produced the lowest positive hit rate of 1:700,000 [20] while azurine hydroxyethyl cellulose (for endo-celluloses) yielded high hit rates of 1 per 108 clones screened [21]. In Chapter 2, we give full details of the screens for genes encoding (hemi)cellulolytic enzymes in a fosmid library. Briefly, we adopted two screening strategies:

1) High-throughput screening using mixed chromogenic substrates, and,
2) Detection of enzyme activity in cell lysates (crude extracts).

The innovative approach consisted of an integration of a maximum of six substrates into the screening medium, thus enhancing hit rates as compared to the use of single substrates. This methodology yielded a total of 71 positive hits (fosmid clones), being 50 from RWS (untreated wheat straw) and 21 from TWS (heat-treated wheat straw). This corresponds to, respectively, 1 hit per 440-
screened clones (RWS), and 1 hit per 1,047 screened clones (TWS). To date, most similar screens revealed consistently low incidence rates with respect to lignocellulolytic enzymes [22]. Such hit rates are dependent on multiple contributing factors, including the enzyme activity being sought (Chapter 3), the metagenomics source and host organism, the DNA extraction method and the choice of cloning vector (Chapter 3 to 5). In addition, the sensitivities of the substrates and the screening conditions (such as pH, Tm, stability and addition of ions) also play vital roles in the approach (Chapter 2).

![Figure 2: Activity screening of fosmid clones using chromogenic substrate](image)

**3.2. Targeted synthetic metagenomics for enzyme production**

Genes cloned directly from metagenomes may not be optimally suited for expression in the cloning host *E. coli* due to codon usage incompatibilities [23]. Thus, the codon usage of a gene of interest may have to be adjusted. However, codon swapping has possible disadvantages that are related to a fast depletion of the precise cognate tRNAs in the host organism. An exact “smart mix” of major and minor codons appears to be necessary, which may be gene- and host-dependent. However, it is often unclear to what extent a novel protocol can optimize gene expression. We applied synthetic gene technology using a preconceived optimal codon mix for the expression of gene XylM1989 in *E. coli*. In chapter 4, I report this novel bi-functional hemicellulolytic enzyme produced in *E. coli*. The homology-modelled structure of XylM1989 displayed an architecture that was predicted to have a catalytic triad and a substrate-binding site, which are responsible for the two activities shown for the enzyme (β-xylosidase/α-arabinosidase). This type of constellation of active sites resulting in two different catalytic activities has been described for enzymes similar to XylM1989 [24].
bi-functional enzyme is thus predicted to work on the hemicellulose moiety of the WS, specifically attacking side chains [25]. It is likely that such bi-functional enzymes will find their way into innovative WS degraded strategies.

4. Selection and biochemical characterization of enzymes

I here described the identification of several genes for novel proteins with catalytic function. A good biocatalyst must meet particular criteria to be suitable for biotechnological industry applications. In particular, the pH characteristics, Tm, stability, kinetics, response to additives and general condition tolerance have to be known and optimal [26]. In this thesis, I mined seven selected fosmids that were predicted to encode diverse (hemi)cellulases (chapter 2). From these, eight glycosyl hydrolase family genes were selected. Four of the predicted enzymes (P1, P2, P5, and P6) and one bi-functional hemicellulase XylM1989 (Chapter 4) showed significant hemicellulolytic activities (Chapter 3), and were thus taken on for further exploratory work.

The enzymes P1 and P2 worked efficiently under thermo-alkaliphilic conditions. Such thermo-alkaliphilic tolerance (without any additives) recently improved industrial needs [27] and so these enzymes are highly recommended, for instance for pulp bleaching processes. In this respect, enzymes P1 and P2 showed high homologies with protein (1) a β-galactosidase (74%, Enterobacter hormaechei) and protein (2) a β-xylosidase (84%, Enterobacter morti); they were identified as responsible for the thermo-alkaliphilic glycosidase activities of extracts of cultures of E. coli carrying the respective fosmids (chapter 2). P1 showed elevated enzymatic activity with increased Tm, pH and was highly stable when compared with previously known β-galactosidase (BglA); it released high amounts of p-nitrophenol (58.7 U/mg) [28]. I foresee that enzymes like P1 and P2 may play vital roles in the food industry, hydrolysing for instance lactose. P2 showed an interesting bi-functional (β-xylosidase/α-arabinosidase) activity (pH 9.0; Tm 50°C), which was found to be higher than that of other β-xylosidase enzymes (Rubgx1 and XynB5) [29,30].

4.1. Unexpected enzymatic activity in proteins P5 and P6, initially predicted to be a diguanylate cyclase and an aquaporin

Proteins P5 and P6 were anticipated as a diguanylate cyclase (79% with Enterobacter cloacae) and an aquaporin (60%, Hyphomonas neptunium), respectively. Remarkably, both showed similar activities on pNP-α-D-glucopyranoside at pH 4.0 – 6.0 under the same conditions. They were also highly stable for 120 min. Maximal activity was revealed with a release of 107.63 U/mg of p-nitrophenol at pH 10.0. This enhanced activity under alkaline conditions might relate to the excess OH ions altering the enzyme shape or active site. Thus alkaline conditions might alter the pattern of substrate binding, in consistency with the “induced-fit” theory of enzyme action. The theory states that the active site of an enzyme does not
necessarily match the shape complementary to that of substrate; rather, when the substrate binds to the enzyme, the active site of the enzyme assumes its shape.

4.1.1 Newly identified protein P5
Remarkably, protein P5, predicted to have diguanylate cyclase activity, showed α-glucosidase activity. This in spite of the fact that no such activity was detected from its mother fosmid clone 10BT (chapter 2). In the presence of multiple substrates, this fosmid clone had shown some activity, presumably with βNP-xyl. The expression of the α-glucosidase encoding gene may have been suppressed at the (fosmid) genomic level, but might be expressed from the inducible promoter of the used vector expression system (chapter 3). Also, protein P5 was found to clearly belong to the “GGDEF” domain protein family [31]. This domain was first identified in a response regulator involved in cell differentiation of C. crescentus [32]. It was also observed in Salmonella enterica enzymes involved in cellulose biosynthesis and biofilm formation [33]. For more than 20 years, all GGDEF domain family enzymes have been classified as diguanylate cyclases and/or phosphodiesterases [34]. The former enzyme produces cyclic di-GMP (cdiG), a messenger that regulates the bacterial lifestyle transition from a motile to a sessile, biofilm-forming state [35]. Most bacterial genomes have genes that encode a range of GGDEF domain proteins [33], allowing functional diversity across them. As yet, the function of most GGDEF domain proteins has not been experimentally demonstrated [36]. Surprisingly, the GGDEF domain family protein P5 revealed 100% α-glucosidase activity in the context of (hemi)cellulose degradation. Detailed structural studies should be performed that address the dynamics of such glycoside hydrolase activities at varying pH values in WS degradation processes. Moreover, protein P5 was found to have a motif “GDSL” by CDART prediction.

4.1.2 Newly identified protein P6
Remarkably, protein P6 (presumptively identified as an aquaporin) was identified to have two different conserved motifs, denoted NPA and GDSL. The NPA motif is a key structural feature of proteins that plays a crucial role in water transport across membranes. Surprisingly, the GDSL hydrolase family domain (encompassing esterases/lipases) of protein P6 had no shared sequence homology with any of the glycosyl hydrolases. In general, aquaporins are partially-hydrophobic integral transmembrane proteins involved in a wide variety of biological processes including photosynthesis, respiration, signal transduction, molecular transport (water), and – finally – also catalysis [37]. In fact, membrane proteins often cause experimental problems, as they are unstable and difficult to express in aqueous solution. However, P6 was successfully expressed and was soluble, with good stability (95% of activity/120 min/ 50°C); this stands in contrast to the predicted properties of aquaporins. Therefore, our findings suggest that P6 might not be a typical trans-membrane protein but belong to a new family
General discussion of glycosyl hydrolases with α-glucosidase activity. A closer scrutiny of the available information revealed that this GDSL family subgroup was categorized as a so-called SGNH hydrolase. Thus, our protein P6 might belong to this subgroup. The SGNH family of hydrolytic enzymes involves a wide range of catalytic functions such as lipase, protease, carbohydrate esterase, (thio)esterase, aryl esterase and acyltransferase activities [38]. The newly-identified carbohydrate esterase family 3 (CE3) gene axe2 product (acetyl xylan esterase) has a GDSL motif, which removes acetyl groups from acetylated xylan [39]. Although untested, such activity might likely enable protein P6 to hydrolyze substituents on the xylan backbone, aiding in the degradation of hemicellulose [40].

4.2. Enzymes showed high tolerance to inhibitors (5-HMF and furfural)
The tolerance to inhibitory compounds, such as 5-HMF and furfural, that are often generated during WS steam explosion, is not often studied [41]. According to a previous study, the amount of 5-HMF and furfural released from WS during steam explosion was 0.2 and 0.7%, respectively [8]. Here, tested the activities of the selected enzymes against different concentration of 5-HMF and furfural. Interestingly, enzymes P1, P5, P6 and XylM1989 all showed 40 to 60% tolerance to 0.5% (w/v) 5-HMF and furfural. Unfortunately, enzyme P2 showed 100% inhibition. To conclude, three of the novel enzymes showed approximately 50% tolerance. Therefore, we speculate that these are superior accessory enzymes to include in enzyme cocktails for LCB degradation processes.

5. Application of novel enzymes to improve the LCB degradation activity of existing (commercial) enzyme preparations
In recent studies, activities of commercial cocktails have been enhanced by supplementation with accessory enzymes. For example, supplementing AFCel (A. fumigatus cellulase) with Spezyme CP on corn straw consistently increased the sugar yield when compared with Spezyme CP alone [42]. Celluclast 1.5L is a so-called “efficient enzyme cocktail” that is widely used for cellulose degradation. Here, we speculate that admixture to it of our novel accessory enzymes enhances WS hydrolytic activities.

5.1 Do the new enzymes act as a complement for commercial cocktails to break (hemi)cellulose moieties?
We evaluated the synergistic activity of Celluclast 1.5L and our hemicellulases P1, P2, P5, and P6 on SE WS (in different combinations) by measuring the final sugar yield (chapter 5). Noticeably, the addition of enzymes P1+P5 and/or P1+P6 to the Celluclast 1.5L produced significantly raised amounts of xylose (303%), arabinose (343%) and glucose (13%) when compared to Celluclast 1.5L alone. This clearly showed that the added enzymes, with Celluclast 1.5L, worked extensively on the hemicellulose moieties, increasing the yield of xylose and arabinose. We
speculate that the action of the Celluclast 1.5L/enzyme mixture on the hemicellulose moiety might be on the xylan, the main chain of $\beta$-1-4 link of D-xylose. Thus, xylose and arabinose were primarily released from the xyloglucan backbone (Chapter 1, Figure 3). Also, the effects of enzyme P1 ($\beta$-galactosidase) might have been on the side chains of xylose-glucose (of xyloglucan), thus increasing the release of xylose and glucose. Later, enzymes P5 and P6, which are $\alpha$-glucosidases, might have cleaved $1\rightarrow4$ linked $\alpha$-glucose residues to release $\alpha$-glucose moieties. To conclude, in a presumably joint and spatially-explicit action, our (hemi)cellulases were – together - effective in enhancing the activity on SE WS of the commercial Celluclast 1.5L.

6. Outlook and recommendations for further work

This study applied metagenomics to lignocellulolytic microbial communities bred (from soil) on WS, in order to produce a suite of novel enzymes with biotechnological potential. Clearly, several new (hemi)cellulases were enriched, selected, produced and characterized from the WS degrading microbial consortia. When microbial populations undergo enrichments on wheat straw as the single driving carbon and energy source, the (spatial) conditions in the system are modulated so as to favour those organisms that function well under wet and well-mixed conditions. With the deliberate choice of this strategy, the organisms and enzymes that come to prevail may actually be less relevant in their natural soil habitat. Clearly, the approach in my thesis was tuned to biotechnological applications via increasing the prevalence of promising organisms that might serve as the best potential sources of LCB degrading enzymes. However, finding the best enzymes is still fraught with uncertainties that derive from choices in the technologies used. For instance, in the construction of metagenomic library, *Escherichia coli* is a commonly used, but not always optimal, host. Transfer of cloned DNA from *E. coli* to alternative hosts systems offers a promising perspective in future efforts to enhance frequencies. However, this was beyond the scope of the current study. In future studies one might adopt broad-host-range plasmid systems for cloning and subsequent spread across different hosts, in order to maximize the potential of successful expression and detection of the genes being targeted.

6.1. Improvement of enzyme cocktails

As discussed above, a key prerequisite for most plant biomass saccharification processes is LCB pre-treatment. Here, we opted for SE pretreatment of WS. I found that sugars released during SE entered the liquid fractions of the substrate. In future endeavors, this released sugar has to be captured and included in the preparation, by adjusting the steam explosion process. Though our newly identified P5 and P6 enzymes worked effectively on the SE WS hemicellulose moieties, we still ignore their exact action on the substrate. Hence, I posit here that
crystallization structure analysis is important to understand the catalytic mechanism involved in the reaction. Optimal characterization of the enzyme characteristics and its mode of action is always mandatory in order to sustain our understanding of efficient enzyme cocktails for complex substrates. Therefore, the required dosage of our enzyme has to be understood. It will then need to be standardized to supplement the commercial cocktails. Furthermore, the degree of enzyme loading is important in industrial processes. Less enzyme loading at a similar degradation rate practically implies that these enzymes are active and effective in WS degradation. In this thesis, we adopted only one commercial enzyme cocktail (Celluclast 1.5L) to test our enzymes. For future work, I suggest to evaluate different commercial cocktails with our enzymes (with different dosage), in order to improve the global efficiency of cocktails for (hemi)cellulose hydrolysis. The Dutch (summer) wheat used was found to have high lignin content (22.2%), whereas for maximal degradation of hemicellulose, lignin is mainly a blocking module. Though the steam explosion applied was able to alter the lignin structure, the amount of lignin loosened up or removed is still unknown. Therefore, a robust evaluation of lignin removal might fine tune the enzyme dosage that has to be applied for effective degradability. Moreover, the generation of enzyme inhibitors during steam explosion is a hindrance to subsequent enzymatic hydrolysis approaches. Therefore, measuring the concentration of inhibitors provides valuable information regarding the types of enzymes required, and their dosage. The enzyme biochemical characteristics such as stability, tolerance to inhibitors and recovery from the conversion reaction, are important parameters for industrial purposes. In addition, other ‘conditions’ such as medium ionic strength, pH, temperature and buffer strength, also play important roles in LCB conversion, and therefore the effects of these factors have to be characterized.

7. References:


