Novel bacterial enzymes for plant biomass degradation discovered by meta-omics approach

Marutha Muthu, Mukil
CHAPTER 1

General Introduction

Mukil Maruthamuthu
Introduction

Prokaryotes constitute the first organisms on Earth; they have dominated the biosphere, playing central roles in biological systems [1]. Estimates have consistently suggested that the immense majority of prokaryotic microorganisms in natural environments is difficult to cultivate [2]. “Uncultivable” indicates that current bench methods are unable to grow such microorganisms in the laboratory. This implies a lack of knowledge of their biological traits and mechanisms that allow growth. Currently, there are ample opportunities – using direct molecular techniques - to assess the molecular traits that reside in such cryptic organisms, allowing us to gain access to previously hidden metabolic diversity. Thus, by applying such novel methods, new natural products can be obtained and the factors that contribute to both ecological balance and health revealed. The genetic information (DNA) present in ecosystems is responsible for programming the physical characteristics of the organisms living therein. Differences in DNA sequences between organisms create diversity. Such diversity manifests itself as biological diversity through the structure, organization, regulation and expression of DNA. There are ample data that indicate that diversity is extremely high in the Prokaryotes [3]. In contrast, the diversity in the Eukarya is much lower, although still considerable and concentrated in the microbial members. The biochemical properties carried in this genetic diversity are thought to support the stability (resilience) of natural systems. To unlock the genetic diversity contained within microbiomes, including both “cultivable” and “uncultivable” microorganisms, the use of metagenomics is very useful [4]. This approach can reveal how the members of microbial communities function and relate to their niches, enabling inferences about processes and their rates in the systems studied.

Importance and potential of biological catalysts

In biotechnological processes, there is an increasing demand for biological catalysts (enzymes) with novel properties, to be used in all aspects of the biochemical reactions that are pursued (such as increased reaction rate, faster kinetics, better specificity and being non-hazardous to nature) [5]. Biocatalyst are widely used in the detergent, textile, cosmetics, leather, pharmaceutical, biotechnology, chemical, food processing and brewing industries, to improve the productivity and efficiency of industrial processes [6]. In addition, the environmental burden of industrial processes needs to be reduced and - in this process - enzymes are also important [7]. Recent research has explored the potential of enzymes (known or unknown ones) to mediate new catalytic functions that are difficult to achieve through chemical catalysis [5]. Moreover, improvement of existing enzymes is receiving great interest [8]. In particular, in silico protein design tools are playing vital roles in efforts that aim to foster the characterization of biocatalysts [9]. Thus, enzymes are considered that have specific properties, e.g. binding to target positions of substrates with better affinity.
and specificity, allowing optimized catalytic activity. For example, the efficiency of production of sugar monomers from lignocellulosic (plant biomass) material poses quite a challenge in the light of the demand by Industry [10]. Thus, carbohydrate-active enzymes (CAE; cellulases, lipases/esterases and xylanases) are required. There is a huge diversity of CAE, which have been grouped into more than 276 families, among which glycosidases (135), transferases (101), lyases (24) and esterases (16) (CAZy: Carbohydrate-Active Enzyme, [11]. Only a small part of this diversity has been unlocked so far. Thus, accessing the enzyme diversity in natural or derived systems will allow to explore the wide catalytic potential of enzymes as related to relevant reaction contexts.

From microbial ecology to biotechnology
Recent assessments estimate that about $10^{30}$ bacterial cells inhabit Earth, with $2.6 \times 10^{29}$ cells in terrestrial and $1.2 \times 10^{29}$ in marine systems [12]. Only part of this huge biomass is accessible by cultivation; studies in marine systems indicated extremely low levels, i.e. 0.001–0.1%, of culturability and those in terrestrial systems about 1% culturability [13]. Given this general lack of culturability of extant microorganisms, the use of metagenomics is indicated to make progress with explorations for biotechnology. Moreover, such an approach will also foster fundamental ecological studies on microbial behavior and survival. Thus, if evidence for the growth of particular microorganisms in particular conditions in an ecosystem can be found, this will enable the discovery of novel enzymes that are required for growth and survival under those conditions [14]. For example, recent studies have shown that the addition of chitin to soil enhances the bacterial abundance along with increased chitinase activities, allowing the unlocking of novel chitinases [15]. Thus, an interplay exists between our attempts to understand the functioning of ecosystems and to foster biotechnology, in terms of the discovery of novel enzymes or antibiotics. Historically, the discovery of streptomycin and other antibiotics sprang from very basic studies of the taxonomy and ecology of actinomycetes in soil, as conducted by soil microbial ecologists [16]. Similarly, turbomycin, one of the first antibiotics discovered by a metagenomics approach, was identified in a soil-derived library [17], in a clone that had hemolytic activity. Thus, to optimize or maximize biotechnological applications, it is essential that both basic biology assessments and utility screens are pursued in metagenomics [6].

Metagenomics
Metagenomic gene discovery studies need fine tuning to allow the retrieval of enzymes with potential useful applications [18]. Here, the approaches, in terms of the organisms being accessed and mined, are complex [19]. First, suitable microbiome DNA needs to be produced. Second, suitable genetic fragments need to be cloned into suitable hosts and then genes should be expressed to allow the
detection of useful functions (Figure 1). In effect, many metagenomics approaches
 can be seen as being analogous to the classical genome library construction and
 screening method, with the difference that the ‘genome’ cloned is not from a
 single organism, but rather from the entire microbial community present in an
 environmental sample. It is, thus, a reflection of the ‘community genome’.
 ‘Genome’ coverage - in this case - is an ephemeral notion, since different
 community members will be present in different numbers in the sample, and their
 genomes will be extracted with different efficiencies. Hence, genes of different
 organisms will be present in very different, and often completely serendipitous or
 idiosyncratic prevalences in the DNA used to construct metagenomic libraries
 [20].

 Different strategies can be taken in metagenomics, according to the
 primary goal. Small (plasmid-based) or large (fosmid-, cosmid- and/or bacterial
 artificial chromosome (BAC)-based) insert libraries can be constructed for
 archiving and sequence homology screening purposes (Figure 1). Some genes in
 such libraries may be expressed in the cloning host and may thus be found in
 activity screens, however there will be a bias towards genes from organisms related
 to the host [21]. The reason may be that there is greater biochemical compatibility
 among closely-related that across distantly related organisms. Moreover, selection
 of the cloning vector is often guided by the need for insertion of the adequate size
 of DNA inserts. Here, as inserts between 20 to 40 Kb were required, fosmids were
 chosen, as these vectors work efficiently and accommodate these insert sizes [22].

 Metagenomics: method, importance and challenges
 Two basic approaches have been used for the recovery of novel biocatalysts from
 different environmental sources by metagenomics. In both cases, shotgun libraries
 are produced in suitable vectors. Then, the difference between the two approaches
 are determined by the library screening method applied: function- and sequence-
 based screening of metagenomic libraries (Figure 1).

 Sequence-based metagenomics - In sequence-based metagenomics screening,
 libraries are screening by sequencing, and comparisons of sequences with
databases will give information about the presumed types of genes that are
 present. Coupled with sequencing and exploration of DNA from the sample
 source, this method will allow us to infer potential function, next to assemble
 operons, genomes and genes (Figure 1) [23]. Moreover, this application
 encompasses the design of DNA probes and/or primers that are derived from
 already-known genes or protein families. In this way, only novel variants of
 identified functional group of proteins can be screened [24]. In contrast, function-
 based metagenomics is based on the screening for a specific function or activity as
 the basis for clone selection (Figure 1) [23]. Function-based screening plays a key
 role in identifying new products for biotechnological applications. Bottlenecks are
the often-low expression of genes in metagenomic libraries, the screening for specific genes without a clear phenotype and (substrate) induced genes. However, function-based screening has been successfully used to recognize novel enzymes (For example: agarases [25], amylases [26], esterase and lipases [25], glycosyl hydrolase [27]) and other products [28]. It is the only approach that enables to identify genes that effectively encode novel functions [4,18].

**Figure 1.** Flow chart of our work

**Plant biomass**

Plant secondary cell walls are formed by specific plant cell type. In those cells with structural support, secondary cell walls are dynamic in different aspects, e.g. water transport [29]. The secondary cell wall is composed of the main polymers cellulose, hemicellulose, lignin and proteins (Figure 2). Building the complex cell wall structure encompasses a harmonized execution of different biosynthetic pathways (Figure 2). These include the biosynthesis and inter-conversion of precursor sugars, the polymerization of precursors into large polymers, the transport of precursors or polymers to the cell wall and the final assembly of component polymers into the cell wall structure (Figure 2). The secondary structure that is formed is often rigid and highly stable and does not easily decay. Upon decay, however, the major polymers used for energy production are, next to proteins, cellulose and hemicellulose. The latter are considered to be the most energy-rich polysaccharides. Evolution has brought these biopolymers together in very prominent architectures, foremost to boundless biodiversity from the molecular to
the macroscopic level. They are hierarchically organized (Figure 2); cellulose is embedded in a complex matrix composed of hemicellulose and lignin.

**Figure 2: Structure of lignocellulose**

**Classification of lignocellulosic biomass** - The presence of polysaccharides as well as the proportion of the different sugars of the hemicelluloses differs between different plants (Table 1). In contrast, **cellulose** is considered as an abundant homo-polysaccharide of the plant cell wall across the board. It is composed of chains with high numbers of D-glucose linked together by β-1, 4-glucosidic bonds (Figure 3). The linear structure of the cellulose chain (formed with hydrogen bonds) leads to crystalline fibrils. In these, the hydrogen bonds of the cellulose make the structure flexible, insoluble in most solvents and quite resistant to microbial degradation [30,31]. The surface of cellulose is hydrophobic, which results in the formation of a repelled layer of water that may hinder diffusion of enzymes and degradation products near the cellulose surface [32].

**Hemicelluloses** are complex heterogeneous polysaccharides composed of D-glucose, D-galactose, D-mannose, D-xylose, L-arabinose, D-glucuronic acid and 4-O-methyl-D-glucuronic acid (Figure 3). Hemicelluloses have a degree of polymerization below 200, and side chains can be acetylated [33]. They are classified according to the main sugar of the backbone of the polymer, e.g. xylan (β-1,4-linked xylose) or mannan (β-1,4-linked mannose) [34,35]. Plants belonging to the grass family (Poaceae), e.g. rice, wheat, oat and switch grass, have hemicelluloses that are composed of mainly glucurono arabinoxylans [36]. Due to differences in hemicellulose composition, agricultural waste products like wheat straw and corn stover, as well as hardwood materials, are rich in the pentose sugar xylose, whereas softwoods are rich in the hexose sugar mannose (Table 1) [34,35,37–40].

**Lignin** is a complex network formed by polymerized phenyl propane units. It constitutes the most abundant non-polysaccharide fraction in lignocellulose. The
three monomers in lignin are p-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol that are joined through alkyl–aryl, alkyl–alkyl and aryl–aryl ether bonds. Lignin embeds the cellulose, thereby offering protection against microbial and chemical degradation. Furthermore, it is able to form covalent bonds with some hemicelluloses, e.g. benzyl ester bonds with the carboxyl group of 4-O-methyl-D-glucuronic acid in xylan. More stable ether bonds, also known as lignin carbohydrate complexes (LCC), can be formed between lignin and arabinose or galactose side groups in xylans and mannans [41]. In general, herbaceous plants, such as grasses, have the lowest content of lignin, whereas softwoods have the highest lignin content (Table 1).

Figure 3: Structural components of hemicellulose

Pretreatment of lignocellulosic materials
The saccharification of (hemi)cellulose from lignocellulose is hindered by different factors of physico-chemical or structural nature. A mechanical or physical (pre)treatment may aid in the opening up of the structures, so as to obtain the polysaccharide fibers (Figure 1). Several pretreatment techniques are currently playing roles in this process, including ammonia fiber explosion, steam explosion (SE), chemical and biological treatments. Steam explosion appears as a method of choice (Garrote et al., 1999), as it uses no chemicals, yields high amounts of hemicelluloses with low degraded byproducts, does not cause equipment corrosion, is user-friendly (no acids) and allows the disruption of the solid residues
from bundles to individual fibers. However, it may produce inhibitory compounds such as furfural and hydroxyl methyl furfural (HMF). In effect, steam-explosion was an effective process for increasing methane yield from wheat straw (upto 30%), as compared to untreated wheat straw [43]. Although (energy) expensive, SE has great potential for the industrial scale [44]. This process has to meet several requirements: (1) increase the production of sugars or the enzyme convertibility, (2) reduce the loss of sugars and avoid the degradation of carbohydrate, (3) avoid the formation of inhibitors for further processes, and (4) be scalable to industrial level. However, in practice it is very difficult to consider all issues. Clearly, loss of sugars and cost effectiveness are points of concern in scaling up. In fact, the combination of steam explosion and enzyme hydrolysis has been recommended as a standard technology that best meet current industrial demands [45]. During steam explosion, the lignocellulosic material pore size can be increased, however, the method does not affect the microfibrils [46]. Thus, enzymes can have easy access to the microfibrils to increase hydrolyses process.

Screening microbiomes for glycosyl hydrolases
The enzymes that break glycosidic bonds between sugar residues or a sugar and a non-sugar moiety within carbohydrates or oligosaccharide are known as glycoside hydrolases. The glycoside hydrolase family enzymes are scattered throughout the prokaryotic, eukaryotic and archaeal domains [47]; they have remarkable functional diversity. Till now, 135 glycoside hydrolase (GH) families were identified. The classification of GH enzymes based only on substrate specificities is inappropriate, since the same protein fold often harbors several types of specificities. Therefore, one key study proposed a better classification pattern based on the consequence of protein folding, as dictated by the amino acid sequence [47]. The 135 different GH families provide insight into the comparative structural features within a family, their evolutionary relationships with other family members and their mechanisms of action. Through functional metagenomics, several carbohydrate-active enzymes (glycosyl hydrolase family) have been identified using p-nitrophenyl tagged substrates. For example, p-nitrophenyl-β-D-glucopyranoside and p-nitrophenyl-α-L-arabinopyranoside were used to screen for (hemi)cellulose hydrolytic enzymes in *E. coli* clones harboring metagenomic fosmid libraries derived from (hemi)cellulose-depleting microbiomes of a fresh cast of earthworms. This study revealed two glycosyl hydrolases that had no resemblance (63% and 75% of identity with hypothetical proteins) to any identified glycosyl hydrolase family enzymes; two novel families (G03-3 and G04-9) of β-galactosidases/α-arabinopyranosidases were thus defined [48]. The purified G03-3 and G04-9 enzymes were tested against pNP-β-D-galactoside (pNPGal) and pNP-Ara. Remarkably, these two enzymes showed approximately 90- and 2-fold higher catalytic efficiencies with β-galactoside than with α-arabinopyranoside. *In-silico* structural alignment and 3D modelling were performed, and these methods
predicted (based on the crystal structure) similarity to the UDP-glucose epimerase of *Thermus thermophilus*. Thus the analyses suggested that the two enzymes belong to two new families of functional GHs with β-galactosidase (EC 3.2.1.23) and α-arabino pyranosidase activities.

<table>
<thead>
<tr>
<th>Table 1: Presence of polysaccharides in lignocellulosic materials</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Samples</strong></td>
</tr>
<tr>
<td>Grasses</td>
</tr>
<tr>
<td>Wheat straw</td>
</tr>
<tr>
<td>Rice straw</td>
</tr>
<tr>
<td>Corn stover</td>
</tr>
<tr>
<td>Hardwood</td>
</tr>
<tr>
<td>Birch</td>
</tr>
<tr>
<td>Willow</td>
</tr>
<tr>
<td>Softwood</td>
</tr>
<tr>
<td>Spruce</td>
</tr>
<tr>
<td>Pine</td>
</tr>
</tbody>
</table>

**Importance of enzymes for hydrolysis of lignocellulose**

Efficient hydrolysis of lignocellulose requires a number of efficient enzymes that work on the cellulose, hemicellulose and lignin moieties of the substrate. Concerning the required enzymes for cellulose, cellulolytic enzymes are divided into three major classes; i) cellobiohydrolases (CBH) (EC 3.2.1.91), which break the cellulose chain into cellobiose units from the ends; ii) endo-1,4-β-D-glucanases (EG) (EC 3.2.1.4), which hydrolyse internal β-1,4-glucosidic bonds; iii) 1,4-β-D-glucosidases (EC 3.2.1.21), which hydrolyse cellobiose to glucose and also cleave glucose units from cello oligosaccharides. All these enzymes work synergistically to hydrolyse cellulose by creating new accessible sites for each other, removing obstacles and relieving product inhibition [49]. The degradation of hemicellulose requires several tasks from the hemicellulases (Figure 3). The main structural part in hemicellulose is xylan, which is composed of d-xylopyranosyl units linked by β-1,4-glycosidic bonds. The hemicellulase system includes as enzymes, among others, endo-1,4-β-D-xylanases (EC 3.2.1.8), which hydrolyse internal bonds in the xylan chain; 1,4-β-D-xylanases (EC 3.2.1.37), which attack xylo oligosaccharides from the non-reducing end and liberate (EC 3.2.1.22), α-D-arabinofuranosidases (EC 3.2.1.55), α-glucuronidases (EC 3.2.1.139), acetyl xylan esterases (EC 3.1.1.72) and feruloyl and p-cumaric acid esterases (EC 3.1.1.73) [50,51]. Although several of these glycosyl hydrolase family enzymes have been characterized and crystallized to better understand the structural functions, there is a paucity of knowledge about the full diversity across these (Table 2). While enzymes from organisms in natural systems exert habitually selective catalytic activities, these are often not appropriate for the burdens of industrial processes in which greater xylose; endo-1,4-β-D-mannanases (EC 3.2.1.78), which cleave internal bonds in
Table 2: Important (hemi)cellulases compiled from CAZy* and PDB* databases

<table>
<thead>
<tr>
<th>Enzyme function</th>
<th>EC Number</th>
<th>GH Family</th>
<th>Number of PDB entries</th>
<th>Number of Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Bacteria Eukaryota Archaea</td>
<td>of Genes</td>
</tr>
<tr>
<td>Cellulases</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cellobiohydrolases (CBH)</td>
<td>EC 3.2.1.91</td>
<td>5, 6, 9</td>
<td>30</td>
<td>28</td>
</tr>
<tr>
<td>endo-1,4-β-D-glucanases (EG)</td>
<td>EC 3.2.1.4</td>
<td>5, 6, 7, 8, 9, 10, 12, 26, 44, 45, 48, 51, 74, 124</td>
<td>188</td>
<td>51</td>
</tr>
<tr>
<td>1,4-β-D-glucosidases</td>
<td>EC 3.2.1.21</td>
<td>1, 2, 3, 5, 9, 30, 116</td>
<td>63</td>
<td>85</td>
</tr>
<tr>
<td>Hemicellulases</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>endo-1,4-β-D-xylanases</td>
<td>EC 3.2.1.8</td>
<td>3, 5, 8, 9, 10, 11, 12, 16, 26, 30, 43, 44, 51, 62, 98.</td>
<td>241</td>
<td>81</td>
</tr>
<tr>
<td>1,4-β-D-xylidosidases</td>
<td>EC 3.2.1.37</td>
<td>1, 3, 5, 30, 39, 43, 51, 52, 54, 116, 120</td>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td>endo-1,4-β-D-mannanases</td>
<td>EC 3.2.1.78</td>
<td>5, 9, 26, 44, 113, 134</td>
<td>33</td>
<td>17</td>
</tr>
<tr>
<td>1,4-β-D-mannosidases</td>
<td>EC 3.2.1.25</td>
<td>1, 2, 5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>α-D-galactosidases</td>
<td>EC 3.2.1.22</td>
<td>4, 27, 31, 36, 57, 97, 110</td>
<td>22</td>
<td>24</td>
</tr>
<tr>
<td>α-D-arabinofuranosidases</td>
<td>EC 3.2.1.55</td>
<td>2, 3, 10, 43, 51, 54, 62</td>
<td>37</td>
<td>19</td>
</tr>
<tr>
<td>α-glucuronidases</td>
<td>EC 3.2.1.139</td>
<td>4, 67</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>acetyl xylan esterases</td>
<td>EC 3.1.1.72</td>
<td>5, 11</td>
<td>31</td>
<td>4</td>
</tr>
<tr>
<td>feruloyl and p-cumaric acid esterases</td>
<td>EC 3.1.1.73</td>
<td>10, 78</td>
<td>ND</td>
<td>9</td>
</tr>
</tbody>
</table>

* Carbohydrate-Active enzymes (CAZy); Protein data bank (PDB)
ND: Not deteriminer; EC: Enzyme Commission number; -: absent from database
enzymes; α-D-galactosidases stability and productivity are required [52]. Recent developments and industrial applications of a variety of biocatalysts including hydrolases, reductases, transaminase, oxidases and others, have been reviewed in detail [53]. To boost degradative processes, enzyme mixtures are of increasing importance in bio-refinery industries. Ideally, the function of each enzyme in such a mix should be known, next to its interactivity with other enzymes [54]. In particular, enzyme characteristics such as stability, degradation, tolerance to inhibitors, allosteric hindrance and mainly competition and/or collaboration between enzymes, needs to be understood. Moreover, the effects of factors like temperature and pH need attention, as these are important in enzyme reactions and may affect the nature of protein structure, such as folding and posttranslational modification (PTMs) [55]. Some commercial enzyme cocktails are already in use in lignocellulose degradation processes for the production of biofuels. At present, Celluclast, CelliC® CTec2 and CelliC® CTec3 are available as rather efficient tailor-made cellulolytic enzyme cocktails [56,57]. However, Celluclast mostly exhibits cellulase activity, and low hemicellulolytic activity. Thus hemicellulolytic enzymes added to Celluclast may promote the enzyme hydrolysis activity to obtain more reducing sugars lignocellulosic feedstocks.

Overall aim of this thesis

An improved understanding of how natural microbiome address the problem of degrading recalcitrant plant biomass will provide information that will ultimately permit expansion of bio-refining processes. Thus fuels, plastics and value-added chemicals can be more efficiently produced from biomass. I here propose to exploit the power of metagenomics to access the wealth of microbial genetic diversity in natural microbiomes. A thorough analysis of these will enable effective enzyme cocktails to be produced and modulated for enhanced lignocellulose biodegradation.

Hypotheses and research questions

1. Microbial consortia developed from soil on the proper lignocellulose substrate will harbor genes for novel carbohydrate-active enzymes. Functional metagenomics will enable to analyze the presence of glycosyl hydrolase family genes such as (hemi)cellulases.
2. Specific novel glycosyl hydrolase family genes on fosmids can be successfully expressed, enabling the discovery of novel catalytic protein space.
3. From consortial metagenomes, predicted GH genes can be codon-optimized and expressed. Such targeted synthetic metagenomics is useful for the selection and characterization of GHs with novel properties.
4. The commercial enzyme mix Celluclast can be strongly improved in efficiency by the addition of novel enzymes produced by targeted metagenomics.
Research questions

- What are the key carbohydrate-active genes that play roles in lignocellulosic microbial consortia? Can functional metagenomics be used to find novel glycoside hydrolase family enzymes?
- What are the characteristics of novel identified enzymes revealed by metagenomics?
- What are the key enzymes that enhance the catalytic activity of commercial enzyme cocktails? Does plant biomass pre-treatment influence the enzyme catalytic activity? To what extent it can be useful to the bio-refinery industries?

Outline of the thesis

Chapter 1 introduces the topic of this thesis. It describes the importance of microbiomes and the molecular functions they carry in plant biomass degradation. Moreover, it touches on key methods that allow to discover the novel biocatalysts beneficial to the biotechnological applications.

Chapter 2 gives an overview of function-based metagenomics library screening for the identification of (hemi) cellulolytic enzymes from lignocellulolytic microbial consortia. In this chapter, I, together with co-authors, show the impact of using multi-substrate methodology (based on six chromogenic substrates) to enhance the metagenomics hit rate. Here, we pinpoint clones where three putative thermoalkaliphilic enzymes were covered (a xylanase, a galactosidase, and a glucosidase).

Chapter 3 develops the data of chapter 2, with respect to the molecular characterizations of the selected genes. Here, eight glycosyl hydrolase (GHase) family candidate genes were selected using molecular cloning, gene expression and purification studies in Escherichia coli. Four of the enzymes showed significant activities on pNP-β-D-galactopyranoside, pNP-β-D-xylopyranoside, pNP-α-L-arabinopyranoside and/or pNP-α-D-glucopyranoside. Two of the enzymes showed thermo-tolerance and were found to encode a ‘new family’ of α-glucosidase activity.

Chapter 4 enhances the strategy of targeted and synthetic metagenomics, enabling the production of glycosyl hydrolases for lignocellulose degradation. Here, we address the practical difficulties and limitations in codon optimization, gene expression and protein purification. Specifically, we characterized a novel dual-function hemicellulolytic enzyme with high furan aldehyde tolerance and optimal catalytic activity at slightly alkaline pH. The features of this enzyme will be useful in the efficient degradation of pretreated plant biomass.
Chapter 5 describes the power of enzymatic hydrolysis in plant biomass degradation. However, practical limitations are huge while using single enzymes. In order to enhance the efficiency of hydrolysis in plant biomass pretreatment, we developed a steam explosion procedure that helps to expose the microfibrils to enzymes. Remarkably, four novel (hemi)cellulases were identified that efficiently enhance the activity of commercially available cellulase enzyme cocktails.

Chapter 6 highlights the overall findings of this thesis, with a focus on the methodologies and novelties. Furthermore, the prospects for lignocellulosolytic enzymes, enzyme cocktails, pretreatments and metagenomic library screening are examined. In addition, avenues to future work are discussed to improve the biotechnological applications.

References


58. Wiselogel A, Tyson J JD. Biomass feedstock resources and composition, in handbook on


