General introduction and scope of this thesis

Diego Javier Jiménez and Jan Dirk van Elsas
1. Plant biomass polysaccharides

Plant biomass is composed of cellulose (polymer of glucose), hemicellulose (heterogeneous polymer of pentose and hexose sugars), lignin (highly complex aromatic heteropolymer) and pectin. Together, the polysaccharides and lignin (in jargon: lignocellulose) provide high complexity and rigidity to the plant cell wall (Figure 1A). The lignocellulosic biomass consists of a defensive inner structure which contributes to the hydrolytic stability and structural robustness of the plant cell walls and its resistance to microbial degradation. In addition, the presence of cross-links between cellulose and hemicellulose with lignin via ester and ether linkages provides to the plant biomass recalcitrance as well as resistance to herbivores and pathogens (Cragg *et al.*, 2015). It is important to notice that the energy in plant biomass is stored in the form of cellulose and hemicellulose. These two moieties are the most abundant organic polysaccharides found in nature.

Cellulose, the major constituent of plant cell walls, consists of β-1,4-linked D-glucose units that form linear polymeric chains of about 8,000-12,000 glucose units. In its crystalline form, cellulose consists of chains that are packed together by hydrogen bonds to form highly insoluble structures. In addition, to the crystalline structure, cellulose also contains amorphous regions (Himmel *et al.*, 2007) (Figure 1B). Three main polysaccharides constitute the hemicellulose fraction of most plants, i.e. xylan (arabinoxylan), xyloglucan and galacto(gluco)mannan. These fractions are usually classified according to the main sugar residues present in the backbone. Xylan (arabinoxylan) is composed of β-1,4-linked D-xylose units, which may be substituted by different side groups, such as D-galactose, L-arabinose, glucuronic acid, acetyl, feruloyl and p-coumaroyl residues. Xyloglucan consists of a β-1,4-linked D-glucose backbone together with D-xylose side linkages. In addition, L-arabinose and D-galactose residues can be attached to the xylose residues, and L-fucose can be attached to galactose residues. Galacto(gluco)mannan consists of a backbone of β-1,4-linked D-mannose (mannans) and D-glucose (glucomannans) residues with D-galactose side chains (de Souza 2013) (Figure 1C). In hemicellulose, the diversity of side groups attached to the main backbone confers high structural complexity and variability. Thus, the high degree of substitutions in the hemicellulose requires the action of various enzymes that are able to release the main sugars (xylose, arabinose and glucose). In general terms, the cellulose and hemicellulose moieties provide the most important “sweet fuel” for microbial metabolism in environments with heterotrophic nutrition.

2. Microbial degradation of plant biomass

The degradation of plant biomass is pivotal for life on Earth, because it is responsible for large portions of the carbon flux in the biosphere. In addition, plant-derived carbonaceous compounds play a pivotal role in the lifestyle of many microorganisms. The breakdown of lignocellulose and release of simple sugars for microbial metabolism drives large parts of the microbiomes of terrestrial and aquatic environments. Thus, saprotrophic nutrition (associated with heterotrophic microorganisms), encompasses the living off dead organic matter, which in the aforementioned ecosystems is mainly composed by plant biomass (de Souza 2013). In nature, large numbers of very diverse microorganisms (bacteria and fungi) are...
known to be capable of degrading plant lignocellulosic material. The complex structure of the plant biomass and the wide range of environmental conditions under which it is consumed reflect the high diversity of microorganisms and enzymes involved in this degradation process (Cragg et al., 2015). Along evolution, microorganisms have developed very fine tuned mechanisms that allow to obtain energy and carbon from the plant biomass. A key facet of the energy generation mechanisms from plant polysaccharides is the first step, i.e. the primary attack on the lignocellulose. In this first process step, the production and secretion of so-called carbohydrate-active enzymes is primordial. Collectively, these enzymes are defined as “degraders of the plant cell wall”, thus including all types of hydrolysis/bond-cleavage reactions. Ultimately, the process releases the monosaccharides (singular sugars) that can be captured, imported and used as substrates for energy generation and anabolism in the microbial cell. An efficient attack on the plant biomass is known to involve the synergistic action of several enzymes, under which the glycoside hydrolases (GHs), polysaccharide lyases (PLs), carbohydrate esterases (CEs) and lytic polysaccharide monooxygenases (LPMOs), to mention the most relevant enzyme types (Zhao et al., 2013; Cragg et al., 2015).

**The enzymatic machineries**

Regarding its macromolecular organization, the lignocellulolytic machinery differs significantly between aerobic and anaerobic microorganisms. In anaerobic bacteria (e.g. *Clostridium*, *Acetivibrio*, *Bacteroides* and *Ruminococcus*), the machinery (cellulase/hemicellulase apparatus) is commonly assembled as a large multi-enzyme complex, called the cellulosome (Doi et al., 2003). Basically, the catalytic components of such cellulosomes include a structure named dockerin, which consists of non-catalytic modules that bind to cohesin modules, located in a large non-catalytic protein that is acting as a scaffold. The protein-protein interaction between dockerins and cohesins allows the integration of the hydrolytic enzymes into the complex. It has been demonstrated that scaffolding is also responsible for the anchoring of the whole complex onto crystalline cellulose, through a non-catalytic carbohydrate-binding module (CBM). The spectrum of enzymes associated with cellulosomes currently includes numerous endo- and exo-acting GHs, several of which appear to be processive enzymes that cleave the cellulose chain in a sequential manner. Notably, hemicellulases and other polysaccharide-degrading enzymes, such as chitinases and PLs, can be associated with cellulosomes (Bayer et al., 2004).

In aerobic microorganisms, generally the degradation of plant biomass occurs by the secretion of freely-diffusible single catalytic proteins. However, some single enzymes might contain more than one active site, for instance a multi-modular enzyme with GHs and CBM domains. The CBM could potentiante the activity of the appended catalytic module (GHs), and it has been proposed that this is likely due to an increasing proximity of the catalytic module towards the target substrate. Aerobic cellulose-degradation activity levels are determined by (secreted) endoglucanases (EG), which randomly act on the internal (amorphous) sites in the cellulose chain, next to exoglucanases or cellobiohydrolases (CBH), which progressively act on the reducing or non-reducing ends of cellulose chains. Thus, either cellobiose or glucose is released and β-glucosidases will act, which
hydrolyze soluble cellodextrins and cellobiose to glucose (Himmel et al., 2010) (Figure 1B). The aforementioned enzymes are produced by a wide range of bacteria and fungi (e.g. Aspergillus, Trichoderma, Bacillus, Cellulomonas and Streptomyces) (Doi, 2008).

The complexity of the hemicellulose fraction requires the concerted action of endo-enzymes cleaving the main chain internally, exo-enzymes that release monomeric sugars and accessory enzymes that cleave the side chains of the polymers or associated oligosaccharides. This leads to the release of various mono- and disaccharides depending on the hemicellulose type. The complete degradation of hemicellulose is only achieved after the release of all substitutions present on the main backbone. In most hemicelluloses, xylan is a major structural component. Xylan can be degraded through the action of β-1,4-endoxylanase, which cleaves the backbone into small oligosaccharides, and β-1,4-xylosidase, which cleaves the oligosaccharides into xylose. The class of hemicellulases includes a variety of enzymes with different activities. Included in this class are xylanases, arabinofuranosidases, glucuronidases, mannases, mannosidases, fucosidases, galactosidases, furfural acid esterases and acetyl xylan esterases (de Souza, 2013) (Figure 1C). As xylose and arabinose are the most common sugars in the hemicellulose fraction of cereal crop agro-residues (e.g. wheat straw), their complete release from the hemicellulose is performed by synergic action of xylanases and arabinofuranosidases (van den Brink and de Vries, 2011). Notably, some hemicellulases could have dual activity (e.g. β-xylosidase/α-arabinosidase) (Shi et al., 2013).

In addition to the polysaccharides, the plant structural composition entails the key structural component lignin. Lignin is a complex aromatic heteropolymer, composed of phenylpropanoid aryl-C3 units, linked together via a variety of ether and C–C bonds. In the lignin bioconversion, laccases and peroxidases are important. Laccases-like enzymes are multi-copper-containing enzymes capable of catalyzing the oxidation of a wide range of (phenolic and non-phenolic) lignin aromatic compounds. In addition, these enzymes can be involved in the attack on the linkages connecting the lignin and hemicellulose moieties in plant cell walls (Singh et al., 2011). The bioconversion of lignin is widely known to be carried out by fungi, but it is relatively unknown in bacteria. However, recent research suggests that this process can be carried out by different bacterial genera (e.g. Streptomyces, Pseudomonas, Acinetobacter, Aeromonas, Rhodococcus, Bacillus, Caulobacter, Brevundimonas, Enterobacter and Klebsiella) (Dubé et al., 2008; Wang et al., 2013; Wu et al., 2010; Ahmad et al., 2011; Pan et al., 2011; DeAngelis et al., 2011a; DeAngelis et al. 2011b; Woo et al., 2014). Thus, the bioconversion of lignin, by fungi and bacteria, and the consequent generation of energy is performed by different catabolic pathways, such as those via protocatechuic acid, β-aryl ether, biphenyl or diarylpropane (Bugg et al., 2011; Pollegioni et al., 2015).

The role of lytic polysaccharide monooxygenases (LPMOs)

In the foregoing section, we have argued that the primary attack on plant biomass largely occurs via enzyme complexes that include enzymes with defined cleaving activity on the substrates that make up lignocellulose. However, the recent discovery of the so-called lytic polysaccharide monooxygenases (LPMOs) (Hervé et
al., 2010; Forsberg et al., 2011; Horn et al., 2012) has revolutionized our view as to how lignocellulose is degraded. LPMOs are metalloenzymes (copper-dependent) that attack cellulose (or other polymers, see later) using a mechanism involving molecular oxygen and an electron donor (Vaaje-Kolstad et al., 2010) (Figure 1B). Recently, cellobiose dehydrogenase, which is able to incite an electron flow, has been proposed as a key electron source for such reactions occurring in nature (Horn et al., 2012). Other studies that aimed to characterize LPMO activity have led to the discovery that such enzymes can be active on a plethora of polymers, such as cellulose (Forsberg et al., 2011), chitin (Forsberg et al., 2014), celldextrins (Isaksen et al., 2014) and hemicellulose (Agger et al., 2014). More specifically, LPMOs have been described as enzymes that catalyze the oxidative cleavage of (1-4)-linked glycosidic bonds of crystalline polysaccharide surfaces. Genes encoding LPMOs have been discovered in diverse bacteria, such as Cellvibrio, Serratia and Thermobifida (Millward-Sadler et al., 1995; Vaaje-Kolstad et al., 2005; Moser et al., 2008). These enzymes were originally classified as family GH61 and CBM33 in the CAZY database (www.cazy.org). However, they have recently been reclassified, and now belong to families AA10 (formerly CBM33) and AA9 (formerly GH61) (Levasseur et al., 2013). In family AA10, high sequence diversities are observed between such enzymes within the phylum Proteobacteria (e.g. Pseudomonas, Vibrio, Shevanella, Erwinia and Stenotrophomonas), Firmicutes (e.g. Bacillus, Lactobacillus, Listeria and Lactococcus) and Actinobacteria (e.g. Streptomyces, Micromonaspora, Salinispora and Actinosynnema).

3. The ecology behind the microbial degradation of plant biomass

The role of bacteria and fungi

In the degradation of plant biomass, bacteria and fungi occupy diverse niches that depend on the availability and accessibility of the carbon and energy sources (e.g. cellulose), spatiotemporal and physicochemical conditions and physical barriers. Moreover, aspects of both competition and cooperation are key issues that govern the degradative process (Boer et al., 2005). In this respect, fungi have developed very intricate strategies that enable them to be at an advantage in the degradation process. Key in this is the ability to get close access to the lignocellulose via hyphal extension, enzyme secretion and pH modulation. Such fungi have large arsenals of enzymes that operate best at acid pH and deconstruct cellulose and lignin. However, the bacteria can also occupy different niches locally, and this relates to their plasticity and versatility to grow under varied conditions, with production of diverse enzymes involved in (hemi)cellulose deconstruction (Berlemont and Martiny 2015). It is generally assumed that in largely aerobic soils, the fungi constitute the first lignocellulose attackers, while the resulting sugars are consumed by both fungi and bacteria, which incites a competition for these carbon sources (Boer et al., 2005). An interwoven microbial kingdom succession may ensuing, with some microorganisms (e.g. bacteria) acting on the soluble easy to deconstruct part of the lignocellulose, and other (e.g. fungi) on the highly recalcitrant compounds. Thus, particular enzymatic reactions (e.g. those involved in lignin bioconversion) may occur at the end of the process. As a consequence, the chemical and spatial heterogeneity of the lignocellulose changes over time. This
process theoretically results in the ongoing dynamic formation of novel niches with a potential increase of microbial diversity, or the creation of more stable conditions with a concomitant decrease (Voříšková and Baldrian 2013).

**Figure. 1.** Lignocellulose structure A) and microbial enzymatic machineries involved in plant polysaccharide deconstruction. B) Enzymes involved in the degradation of cellulose and C) hemicellulose. CBH: cellobiohydrolases; EG: endoglucanases; CBM: carbohydrate-binding module; AA10: lytic polysaccharide monooxygenase (LPMOs); $e^*$ donor electron.

**Social interactions**

In plant biomass bioconversion processes, microbial social interactions are important. Cooperation is likely to play an important role, where multiple individuals engage in a common task for mutual benefit (in this case release of the “public goods”, the sugars). The defining characteristic of the “public goods” is that consumption of it by one species does not actually or potentially reduce the amount available to be consumed by another species. In this biodegradative process, division of labor is eminent, where two or more classes of individuals engage in specialized tasks, involving complementary and thus synergistic behavior (in this case, the synergic action of different secreted GHs and LPMOs). The cooperation can be “egalitarian”, in that all individuals contribute to the common good, and obtain their gains in a more or less equal fashion. Alternatively, as addressed in the above, it can involve division of labor, where individuals engage in different tasks from which they might obtain different rewards, directly or via benefits to kin (Crespi 2001). Thus cooperation, resulting in an optimization of “public goods” to
the benefit of the whole, in still natural terrestrial or aquatic systems, is spatially- and temporally-explicit in terms of the successional events that drive the process. Clearly, consortia of microorganisms (bacteria and fungi) produce diverse and complementary enzymes for the breakdown of plant biomass in spatially- and temporally-explicit social interactions, resulting in overall process rates that are called “natural” from an anthropogenic viewpoint.

In mixed (shaken) systems, the situation is clearly different. In such systems (e.g. liquid enrichment cultures), conditions are hypothesized to more strongly drive the selection in competition of specific microorganisms that are “fittest” and thus take care of the production of “public goods” (sugars released) for the benefit of the whole. It is open to research to what extent such systems shift the driving forces in the microbiomes involved from a “cooperative, social” form towards a “competitive” form. However, it is evident that social “cheaters” will arise that profit from the “public goods” made available by the biodegradative action of the “social” community members. Overall, a balance must be struck, as this will otherwise lead to a decline of the fitness of the whole (see below).

4. Ecological behavior, selection and characterization of lignocellulolytic microbial consortia

Selection of biodegradative microbial consortia under specific conditions (e.g. using plant biomass as sole carbon source in aerobic liquid cultures) is hypothesized to strongly decrease the diversity (from thousand to hundreds species), as a result of a lowering of the number of niches due to the more uniform environment that is created and the removal of spatial barriers (Figure 2). In the selection, the most competitive and versatile microorganisms (bacteria and fungi) are expected to survive (“survival of the fittest”) and the niches they can occupy are related to the complexity of the plant biomass (in this case, different types and amounts of sugars, next to other compounds). The spatiotemporally determined substrate availability and physicochemical conditions are relevant factors that are at the basis of the competitive processes (Boer et al., 2005). Thus, the analysis of the microbial consortia that emerge from the selective process is useful to improve our understanding of microbial social behavior in respect to lignocellulose input. For instance, to unveil the interactions, successional patterns (diversity and abundance of particular organismal types on temporal scales) and enzyme prevalences. In addition, unraveling such interactions may aid in our comprehension of the productivity (lignocellulose degradation, releasing monomers) as related to microbial consortium composition and enzymatic profile.

Ecological scenarios

In general terms, several ecological scenarios can occur in lignocellulose-degrading microbial consortia (bacteria and fungi). For instance, competition (resulting in competitive exclusion), cooperation (mutualism), division of labor (synergism) or commensalism may play different roles (Crespi 2001; Faust and Raes 2012). These scenarios can be also integrated, for instance, there may be cases of synergism-competition and synergism-commensalism (Figure 2). Additionally, the interactions will apply to the bacterial and/or fungal communities or across these.
Thus, bacterial and fungal species can compete for the sugar resources released in lignocellulose degradation. Due to competitive exclusion, two species that occupy the same niches will exclude each other (-/-) (Gause 1934). In mutualism, bacterial and fungal species can degrade lignocellulose moieties (e.g. lignin by a fungus -or a specific bacterial species- and then degradation of (hemi)cellulose by bacteria) and then consume the released sugars without competition, as their niches are different (+/+).

Moreover, in synergism, bacterial and fungal species produce different types of enzymes that are complementary to each other, and so complete the degradation (e.g. cellulases from a fungal species and hemicellulases from bacteria, or EG from species 1, xylanase from species 2 and arabinofuranosidase from species 3). In such synergism, the species can either compete for the resulting sugars (-/-) or not at all (+/+), depending of the formation of novel niches. For instance, one species can take advantage of the sugars with the subsequent decreasing of “the survival of the fittest” (commensalism) (+/-). However, co-existence is maintained if the species each occupy another niche (e.g. another type of carbon source); this could be common on plant biomass enriched cultures due to the high complexity of the substrate and the spatiotemporal substrate availability (Figure 2).

**Selection of efficient lignocellulose-degrading microbial consortia in practice**

The use of microbial consortia instead of single strains has been proposed as a powerful approach to degrade plant biomass and enhance the discovery of the key enzymes involved in this process (DeAngelis *et al.*, 2010). This is related to the diverse range of enzymes that are required for the complete deconstruction of the
lignocellulosic matter. Different methods useful to enrich functional microbial consortia have been reported. Of these, the most useful for the production of lignocellulolytic microbial consortia is the dilution-to-stimulation approach (Lee et al., 2013) (Figure 3). In this approach, an environmental community is grown on a lignocellulosic substrate in liquid batch cultures (oxic or anoxic) that are diluted in sequential transfers. Thus, microorganisms with the highest capacities to thrive on the plant biomass (as sole energy and carbon source), in the presence of inorganic sources of nitrogen, sulfur and phosphorus, are selected. Along the sequential transfers, a reshaping of the microbial communities is achieved, in which the diversity is predicted to decrease, consequently increasing the prevalence of specific degrader organisms. A state of relative stability may be reached when the microbial communities start to reveal similar structural profiles across the transfers (Figure 3). This enrichment approach follows the “habitat biasing” concept, which was recently proposed by Ekkers et al. (2012). In the experimental design, parameters such as the inoculum source, the plant biomass substrate, incubation time and temperature, pH, redox potential and electron acceptors, oxygen availability, dilution factor, the presence of toxic compounds and composition of the mineral medium (e.g. nitrogen, sulfur iron and phosphorus sources and vitamins) are highly relevant. Clearly, these parameters drive the composition, richness and lignocellulolytic capacities of the resulting microbial consortia (Brossi et al., 2015; Korenblum et al., 2016; DeAngelis et al., 2012). We here posit that stochastic factors right at the onset of the selection process may also drive the microbial diversity and consortial composition in these enrichment processes.

In the dilution-to-stimulation approach, the aim is to produce an efficient microbial consortium that consists of productive collaborating partner organisms. However, particular members of the selected consortium that become prevalent may be mere scavengers (social cheaters) of the “public goods”, i.e. the monosaccharides that are produced by the degradative consortium members. The presence of these cheaters poses a challenge to the fitness of the “productive” consortium members, and this is bound to the (carrying) limits of the system. Moreover, plant biomass is known to not only contain (recalcitrant) polymers, but also (easily degradable) small soluble substrates. The latter may also increase the proliferation of opportunist microorganisms that cannot deconstruct the aforementioned recalcitrant polysaccharides. To remove such small molecules, washes (twofold) of the plant biomass, with water and ethanol, have been proposed (Gladden et al., 2011a). In other studies, simplified cellulosic substrates have been used (e.g. carboxymethyl cellulose or filter paper) (Eichorst et al., 2013; Zhou et al., 2014). However, it has been suggested that, due to the reduced complexity of the substrate, the diversity of enzymes (e.g. hemicellulases) required to degrade the plant biomass may be lost in the cellulolytic microbial consortia.
Characterization of lignocellulolytic microbial consortia

Along enrichment by the dilution-to-stimulation approach, the degree of degradation of plant biomass is a key issue (e.g. measured by substrate weight loss, Fourier transformed infrared spectroscopy- FTIR- or by two-dimensional nuclear magnetic resonance spectroscopy- 2D-NMR-). In addition, CO₂ production, release of reducing sugars (e.g. measured by 3,5-dinitrosalicilic acid method or by high-performance liquid chromatography- HPLC) and other metabolites are key parameters to be evaluated. Finally, the rate of growth (e.g. measured by microscopy or quantitative PCR) and the levels of different enzymatic activities (e.g. measured by chromogenic substrates) are important. To monitor development of the microbial consortia, several molecular tools (e.g. denaturing gradient gel electrophoresis - DGGE) and culture techniques (plating and isolation) have been applied. However, bacterial 16S rRNA and fungal ITS1 amplicon sequencing constitute more advanced and powerful techniques to evaluate the diversity, structural composition and stability of the microbial consortia along the sequential transfers. Moreover, meta-omics approaches (metagenomics, metatranscriptomics and metaproteomics) can unveil the lignocellulolytic machinery of the resulting microbial consortia. Most currently available microbial lignocellulolytic consortia have been bred from soil, compost and/or sediment using either pretreated or untreated agricultural residues (e.g. wheat straw, corn stover and switchgrass) as carbon and energy sources (Table 1). Additionally, microbial consortia have been cultivated under several conditions (e.g. aerobic-mesophilic, aerobic-thermophilic or anaerobic-thermophilic, at alkaline or neutral pH’s) (Table 1). In the studies of Table 1, the microbial consortia have been characterized by different approaches. Notably, one consortium (*) was characterized by metabolic profiling, enzymatic activity evaluation, plant biomass saccharification tests, perturbation analysis (response of the microbial consortium to perturbation with a variety of substrates) and expression of synthetic cellulases retrieved from the metagenomic data.
Table 1. Recently reported plant biomass-degradin microbial consortia

<table>
<thead>
<tr>
<th>Microbial sources</th>
<th>Plant biomass substrates</th>
<th>Culture conditions</th>
<th>pH</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soil</td>
<td>Wheat straw</td>
<td>Aerobic-mesophilic</td>
<td>7.2</td>
<td>Brossi et al., 2015</td>
</tr>
<tr>
<td></td>
<td>Corn stover</td>
<td>Aerobic-mesophilic</td>
<td>9.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Switchgrass</td>
<td>Aerobic-mesophilic</td>
<td>7.2</td>
<td></td>
</tr>
<tr>
<td>Soil</td>
<td>Rice straw</td>
<td>Aerobic-mesophilic</td>
<td>7.0</td>
<td>Sharma and Malik 2015</td>
</tr>
<tr>
<td>Soil</td>
<td>Switchgrass</td>
<td>Anaerobic-mesophilic</td>
<td>5.5</td>
<td>DeAngelis et al., 2012</td>
</tr>
<tr>
<td>Soil</td>
<td>Wheat straw</td>
<td>Aerobic-mesophilic</td>
<td>ND</td>
<td>Wang et al., 2011</td>
</tr>
<tr>
<td>Sediment</td>
<td>Wheat straw</td>
<td>Aerobic-mesophilic</td>
<td>7.2</td>
<td>Cortes-Tolalpa et al., 2016</td>
</tr>
<tr>
<td>Sediment</td>
<td>Switchgrass</td>
<td>Anaerobic-mesophilic</td>
<td>7.2</td>
<td>Korenblum et al., 2016</td>
</tr>
<tr>
<td>Sediment</td>
<td>Wheat straw</td>
<td>Anaerobic-mesophilic</td>
<td>9.0</td>
<td>Porsch et al., 2015</td>
</tr>
<tr>
<td>Cow rumen</td>
<td>Wheat straw</td>
<td>Anaerobic-mesophilic</td>
<td>6.1</td>
<td>Lazuka et al., 2015</td>
</tr>
<tr>
<td>Cow rumen</td>
<td>Rice straw</td>
<td>Aerobic-thermophilic</td>
<td>8.0</td>
<td>Wongwilaivalin et al., 2013</td>
</tr>
<tr>
<td>Bagasse pile</td>
<td>Activated sludge</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compost</td>
<td>Switchgrass</td>
<td>Aerobic-thermophilic</td>
<td>6.5</td>
<td>Gladden et al., 2011a; Gladden et al., 2011b; Park et al., 2012; D’haeseleer et al., 2013; Gladden et al., 2014 *</td>
</tr>
<tr>
<td>Compost</td>
<td>Rice straw</td>
<td>Aerobic-thermophilic</td>
<td>6.5</td>
<td>Gladden et al., 2011a; Gladden et al., 2011b; Park et al., 2012; D’haeseleer et al., 2013; Gladden et al., 2014 *</td>
</tr>
<tr>
<td>Compost</td>
<td>Switchgrass</td>
<td>Aerobic-thermophilic</td>
<td>6.5</td>
<td>Gladden et al., 2011a; Gladden et al., 2011b; Park et al., 2012; D’haeseleer et al., 2013; Gladden et al., 2014 *</td>
</tr>
<tr>
<td>Compost</td>
<td>Corn stover</td>
<td>Aerobic-thermophilic</td>
<td>ND</td>
<td>Reddy et al., 2011; Yu et al., 2015</td>
</tr>
<tr>
<td>Compost</td>
<td>Switchgrass</td>
<td>Aerobic-thermophilic</td>
<td>6.5</td>
<td>Eichorst et al., 2014</td>
</tr>
<tr>
<td>Compost</td>
<td>Switchgrass</td>
<td>Aerobic-thermophilic</td>
<td>ND</td>
<td>Allgaier et al., 2010; Dougherty et al., 2012</td>
</tr>
<tr>
<td>Compost</td>
<td>Rice straw</td>
<td>Aerobic-mesophilic</td>
<td>ND</td>
<td>Wang et al., 2016</td>
</tr>
</tbody>
</table>

ND: not determined, * Microbial consortia well characterized

5. Metagenomics and its application to explore genes/enzymes involved in plant biomass breakdown

Metagenomics approaches seek to understand microbiology at the aggregate level, transcending the individual organisms that make up communities. Thus, metagenomics places a focus on genes and on how these influence the collective functions in a microbiome (Handelsman et al., 2007). Metagenomics-based analyses have become valuable tools for bioprospecting. In addition, they can provide useful information on the metabolic potential and ecology of microbial communities. Functional metagenomics studies performed so far can be divided into three categories depending on the environment that is accessed (Steele et al., 2008): 1) substrate-biased (e.g. amended environments or enrichment cultures); 2) highly diverse (e.g. forest soils) and 3) extreme (e.g. hot springs).
In order to explore environmental habitats for the enzymes involved in lignocellulose deconstruction, a very promising approach is offered by the metagenomic analysis of lignocellulolytic consortia (Reddy et al., 2013). This approach has been coined “targeted metagenomics” (Suenaga et al., 2012). It promises an easy way to highlight functional traits and monitor the abundance of key genes involved in the biodegradative process. The metabolic potential and functional profile of such lignocellulolytic microbial consortia can be unveiled by high-throughput sequencing (Figure 4), and a comprehensive picture of carbohydrate-active enzymes (e.g. GHs) can be obtained using the CAZy database (Cantarel et al., 2009). An alternative strategy is the identification of GHs catalytic domains using hidden Markov models on the basis of the Pfam database (Li et al., 2009). Such analyses have been useful to understand the relative abundance of certain GHs in environments like the gut microbiomes of herbivorous insects and soil microbiomes (Suen et al., 2010; Cardenas et al., 2015), lignocellulosic biomasses during degradation (Ventorino et al., 2015) and different lignocellulolytic microbial consortia (Allgaier et al., 2010; D’haeseleer et al., 2013; Wongwilaiwalin et al., 2013; DeAngelis et al., 2013a).

Target enzymes can be accessed using functional metagenomics approaches, i.e. high-throughput sequencing followed by a search for specific and complete genes. The latter can then be further custom-synthesized and codon-optimized, after which expression can be achieved in a suitable host. This process is known as “synthetic metagenomics”. It has been successfully applied in two different lignocellulolytic microbial consortia (Dougherty et al., 2012; Gladden et al., 2014). Alternatively, a metagenomic library can be prepared from the lignocellulolytic microbial consortia and then screened for genes encoding enzymes involved in plant biomass degradation (Montella et al., 2015). Upon successful detection, the metagenomic fragment can be sequenced and the target genes expressed in order to evaluate the ensuing enzymatic activities (Figure 4). Searches for (hemi)cellulases using metagenomic library screenings have already been made in rumen, soil, compost and termite gut microbiomes (Del Pozo et al., 2012; Nacke et al., 2012; Jeong et al., 2012; Bastien et al., 2013). However, only few studies have focused on the exploration of cellulolytic enrichment cultures (Beloqui et al., 2010; Mori et al., 2014).
Figure 4. Functional metagenomic approaches for lignocellulolytic microbial consortia. Asterisk indicates the evaluation of the metabolic profile by currently available tools for the taxonomic and functional classification of metagenomic datasets (e.g. Huson et al., 2007; Meyer et al., 2008; Macdonald et al., 2012).

6. Metasecretomics – an approach to unveil the dynamics of secreted enzymes in microbial consortia

Recently, metasecretomics approaches have emerged as powerful tools that allow to disentangle the diversity of enzymes released from microbial consortial cells after growth on plant biomass. Metasecretomics is defined as the analysis of the total surface-bound and secreted proteins that make up the “secretome” of a microbial community. Metasecretomics allows to quantify the enzymatic potential and lignocellulolytic capacities of specific microbial consortium. In particular, it can be useful to analyze the diversity of secreted proteins along a temporal scale, in order to describe the network of enzymes over time and understand the flux of secreted enzymes involved in lignocellulose degradation. Moreover, metasecretomics can help us to understand the pattern of secreted proteins depending on the plant biomass. In addition, assessment of the taxonomic affiliation of the secreted proteins will enable us to identify the metabolically-active microorganism, thus correlating specific functions with the taxa that are truly involved in plant biomass degradation.

Unfortunately, the approach has so far been underexplored on plant biomass-degrading microbial consortia. Basically, two main methods can be used:1) detection of extracellular enzymatic activities and the correlation of such activities with the detected proteins (which are then labeled “probably active”), and 2) identification and quantification of extracellular proteins by two-dimensional gel electrophoresis or liquid chromatography–tandem mass spectrometry (Adav et al., 2012a,b; D’haeseeleer et al., 2013). Additionally, proteins that have secretion signal peptides, found in metagenomics data, could be used as proxies to determine the metasecretome profiles of the microbial consortia under study. In the first method, different substrates can be used to quantify the enzymatic activities, for instance
pNP-glycosides (Gladden et al., 2011a) or chromogenic polysaccharide hydrogels (Kračun et al., 2015). The enzymatic activities detected will lead one to track the presence of active proteins. For instance, if the metasecretome reveals activity on para-nitrophenyl-beta-D-xylanopyranoside, this indicates the presence of active beta-xylosidases. In the second method, the extracellular protein fractions can be tested for enzymatic activity (in gel) using zymograms (Vandooren et al., 2013). The latter can be performed with co-polymerized carboxymethyl cellulose or xylan from beechwood (Gladden et al., 2011b), and also by using fluorogenic substrates, like 4-methyl umbelliferyl-beta-D-xylopyranoside. Such approaches are quite novel and need to be explored in depth, as they can give clues as to the process-relevant enzymes and activities, also providing information useful for the design of enzyme cocktails for future biotechnological applications.

7. Production of bio-based compounds from plant biomass

Beyond the relevance for natural cycling processes, the degradation of lignocellulose has great interest in terms of biotechnological applications. Thus, the understanding of the ecology and enzymology of this process, and how these define efficiency, is indispensable for biorefining. Lignocellulosic matter constitutes a carrier of energy as well as (reduced) carbon, and hence it is of both ecological (as the “sweet fuel” for heterotrophic microbiomes) and biotechnological relevance. Industries around the world have explored the possibilities for the production of bioethanol, biodiesel, biomethane and monomers for bioplastics from plant biomass. In particular the production of building blocks (e.g. 5-furan dicarboxylic, lactic and itaconic acids) for plastics has raised interest (Koopman et al., 2010; Vennestrøm et al., 2011; Ramos et al., 2015). The development is driven by the high demand and cost of fossil fuel, and also by its environmental and techno-scientific importance. Key plant biomass to be converted into valuable products includes woody substrates, agricultural residues (e.g. sugarcane bagasse, wheat straw and corn stover) and dedicated energy crops (e.g. switchgrass) (Chandel and Singh, 2011). As an example, the process that converts plant biomass into bioethanol includes three steps: 1) depolymerisation (via thermochemical and enzymatic routes) of the polysaccharides into fermentable monosaccharides, 2) fermentation to yield ethanol and 3) ethanol recovery (Zheng et al., 2009; Limayem and Ricke, 2012) (Figure 5). In the first step, which is of most interest to this thesis, there are several issues of efficiency that need attention, i.e. the choice of the proper plant biomass, the cost of conversion (in which thermochemical conversion is expensive) and the toxicity of particular byproducts, next to the low yields that are generally provided by current commercial enzymes (see below).

Plant biomass pretreatment and byproducts

In the biotechnological exploration of plant biomass, pretreatment is required to make the lignocellulose amenable to further treatments, such as enzymatic digestion. These pretreatments use various techniques, including dilute acid hydrolysis, steam explosion, ammonia-mediated fiber explosion, alkaline hydrolysis, torrefaction and the use of imidazolium-based ion liquids (Kumar et al., 2009, Chew and Doshi, 2011; Tumuluru et al., 2011; van der Pol et al., 2014; Socha
et al., 2014). Unfortunately, most of these processes generate byproducts that can inhibit subsequent enzymatic attack and fermentation steps. Thus, in lignocellulosic hydrolysates, toxic compounds such as furaldehydes (furfural and 5-hydroxymethylfurfural), acids (e.g. acetic, formic and levulinic acid) and phenolic compounds (e.g. vanillin, syringaldehyde and coniferyl aldehyde) occur commonly (Palmqvist and Hahn-Hagerdal 2000; Parawira and Tekere 2011) (Figure 5). From an economic point of view, and depending on the substrate, alkaline hydrolysis (using sodium hydroxide at high temperature) and steam explosion methods appear to have the largest potential for the biorefining. Specifically, alkaline pretreatment is based on saponification of intermolecular ester bonds cross-linking xylan and lignin, whereas the basis of steam explosion methods is an explosive decompression at ~200°C (van der Pol et al., 2014).

![Figure 5. Schematic representation of bioethanol production. Asterisk means pectin and other compounds (e.g. ash). C5: pentose sugar; C6: hexose sugar.](image)

**Commercial enzymatic cocktails for plant biomass saccharification**

Beyond understanding the action of single enzymes, the ability to understand how cocktails of enzymes work together synergistically will be crucial to harness the paradigms of nature into industrial applications. In addition, information as to how microbial communities adjust their enzyme cocktails to different substrates will be valuable for many biotechnological applications (Cragg et al., 2015). The leading industrial source of cellulase cocktails is *Trichoderma reesei*. Several strains exist and their secretomes (total secreted proteins fractions) have been widely used to develop commercial cocktails for plant biomass hydrolysis (e.g. Celluclast 1.5L, CTecc2 and CTecc3 from Novozymes). However, *T. reesei* secretomes are dominated by CBH and EG, whereas only low quantities of xylanases, arabinofuranosidases, LPMOs and β-glucosidases are produced. Hence, addition of exogenous enzymes to these cocktails will improve the hydrolytic efficiency.
Hemicellulolytic enzymes such as endoxylanases, xylosidases, arabinofuranosidases, mannosidases and fucosidases may be indispensable in the formulation of efficient enzyme cocktails. For instance, Poidevin et al. (2014) showed that a T. reesei cocktail significantly improved lignocellulose saccharification efficiency when supplemented with secretomes of the fungus Podospora anserine. Berlin et al. (2007) studied the effect of supplementing a commercial T. reesei cocktail with xylanases and pectinases, observing enhanced saccharification by the supplemented cocktail on pretreated plant biomass. Additionally, Harris et al. (2010) showed that the co-expression of a highly active LPMO protein in a commercial T. reesei cellulase resulted in a 2-fold reduction of the total enzyme load required to reach 91% cellulose conversion of lignocellulosic biomass. Notably, LPMOs can act synergistically with GHs (cellulases and hemicellulases) to improve saccharification yields. Since their discovery, LPMOs have been included in all modern commercial cellulase cocktails (e.g. Cellic CTec2) (Cannella et al., 2012; Cannella and Jørgensen 2014). However, LPMO activity depends on the presence of molecular oxygen, which is thus an important factor in industrial saccharification and fermentation processes (Müller et al., 2015).
SCOPE AND AIMS OF THIS THESIS

Our understanding of the dynamic functioning and key interactions in plant biomass-degrading microbial consortia is still in its infancy. This lack of comprehension is all the more pitiful in the light of the great demand for such microbial consortia and their secreted enzymes in biotechnology. The characterization of plant biomass-degrading microbial consortia by metagenomics and metasecretomics-based approaches appears to be very powerful, as it will spur the understanding of the dynamics, interactions, successions, ecological behavior, metabolic profile and the lignocellulolytic machineries involved. In addition, the approach will allow to pinpoint which kind of genes/enzymes and microorganism are most relevant for the degradation of plant biomass.

Hypotheses and research questions:

1. The selection of microbial consortia in lignocellulose-driven sequential-batch enrichments narrows the richness yet fosters efficient biodegradation.
2. Although the selection of lignocellulolytic microbial consortia, from highly diverse source microbiomes, is stochastic, similar consortia develop in similarly-treated parallel cultures.
3. Lignocellulolytic microbial consortia selected in liquid media will contain bacterial next to fungal constituents. In both groups, organisms with different roles will come up, allowing the synergistic deconstruction of lignocellulose.
4. Specific members of the Bacteroidetes phylum will become highly relevant in lignocellulolytic microbial consortia, given their high genomic potential to deconstruct structural polysaccharides.
5. Highly complex enzyme mixtures are required for the efficient lignocellulose degradation, and hence complex microbial consortia are selected as the most efficient (fittest) microbial units upon which selection acts.
6. The lignocellulolytic microbial consortia and their carbohydrate-active gene and secreted enzyme profiles are strongly directed by the composition of the substrates used for their growth.

From the aforementioned hypotheses, the following research questions are distilled:

- To what extent are plant biomass-degrading microbial consortia, with respect to diversity, composition, metabolic profile and lignocellulolytic machinery, influenced by the inoculum source, plant biomass type and the conditions of liquid cultures, in particular temperature and oxygen availability?
- What is the composition of bacterial versus fungal communities, and the relative abundance of key members of these, in soil-derived lignocellulolytic consortia?
- Do directed enrichments in sequential-batch cultures with plant biomass allow to obtain higher frequencies of genes that are potentially associated with lignocellulose degradation, like hemicellulases?
- What types of enzymes, present in the soil-derived microbial consortia, are key to the degradation of plant biomass?
This thesis aims to increase our understanding of how soil microbial communities respond to a carbon source input (plant residues, in particular wheat straw) at neutral pH and under oxic and mesophilic conditions, which kinds of interactions take place in lignocellulose-enriched consortia, what the key degradative consortium members are (bacteria and fungi) and what functional roles they have. Moreover, questions are posed with respect to what kinds of proteins (proxies for processes) are enriched compared with the inoculant soil microbiome and what key enzymes (as judged by the glycosyl hydrolase – GH - families) are selected for the deconstruction of lignocellulose. The study develops along a metagenomics-metasecretomics-based approach and characterizes, in terms of taxonomy and functional make-up, different plant biomass (wheat straw)-degrading microbial consortia. The information retrieved, with respect to the ecology and lignocellulolytic machineries of the wheat straw-degrading microbial consortia, is used for the design of efficient microbial consortia or enzymes cocktails for biotechnological applications.

The specific aims of this thesis are:

1) Produce deliberately biased (enriched) microbial consortia using forest soil samples as inocula on lignocellulosic agricultural residues (in particular wheat straw).
2) Evaluate the stability, composition and diversity of the resultant microbial consortia in comparison to the inoculum source.
3) Disentangle the lignocellulolytic capacities present in the resultant microbial consortia with respect to metabolic potential, the carbohydrate–active gene profiles and the secreted protein fractions.
4) Identification of enzymes that are key in the degradation of (hemi)cellulosic compounds by screening of metagenomic libraries derived from the degrader microbial consortia.
5) Develop strategies for the design of efficient enzyme cocktails for biorefining purposes.
CHAPTER 1

THESIS OUTLINE

Chapter 1 introduces the topic of the thesis, focusing on the ecology and enzymology behind the microbial degradation of plant biomass. In addition, it highlights the potential social interactions of lignocellulolytic microbial consortia and the methods that are required to characterize these communities. Finally, in a general way, the importance of plant biomass degradation for biotechnological applications is explained.

Chapter 2 addresses the production, by dilution-to-stimulation, and characterization, by 16S rRNA gene/ ITS1 qPCR and PCR-DGGE and isolation of key bacteria and fungi, of two microbial consortia that degrade either pretreated (TWS) or untreated wheat straw (RWS). In addition, this chapter reports a novel, easy and fast method to detect oxidoreductase activity in the presence of 5-hydroxymethylfurfural.

Chapter 3 expands the data of chapter 2, by focusing on the “metataxonomic” analysis of the two microbial consortia, coined RWS and TWS. In addition, the chapter shows the successional microbial diversity, stability and community composition of the two consortia. Moreover, the metagenomic predictor PICRUSt was used to indicate the key genes involved in lignocellulose degradation on the basis of bacterial 16S rRNA sequences. Moreover, specific (hemi)cellulolytic enzymatic activities in the metasecretomes of the resulting and stable microbial consortia are reported.

Chapter 4 focuses on the metagenomic evaluation, in a time course, of the RWS and TWS lignocellulolytic microbial consortia. This chapter also provides insight in the metabolic potential and lignocellulolytic machinery of these consortia. In addition, the overrepresentation of carbohydrate transporters, polysaccharide sensing genes and GHs are presented and examined. Finally, predictive descriptions of partial pathways related to plant biomass deconstruction are provided.

Chapter 5 presents a functional screening for hemicellulases in the metagenomic libraries generated from the RWS and TWS microbial consortia. In this chapter, the functional screening is expanded by using a multi-substrate approach with six chromogenic compounds. High hit rates were obtained with this method, as compared with recently-published methods. In addition, clones containing genes for putative novel thermoalkaliphilic enzymes (one xylanase, one galactosidase and one glucosidase) were identified.

Chapter 6 analyzed the proteins that are secreted (metasecretome) by a wheat straw-degrading microbial consortium (RWS) bred on either wheat straw, xylose or xylan. Liquid chromatography–tandem mass spectrometry was used as the tool to analyze the proteins in the supernatants of each enriched consortium. Taxonomic and functional affiliation of the secreted proteins enabled us to identify the metabolically-active microorganisms, helping to correlate specific functions with taxa that are predicted to be involved in plant biomass deconstruction. Diverse
types of hemicellulases were identified in the RWS metasecretome, giving clues with respect to the design of enzyme cocktails for future improvements in biorefining.

Chapter 7 focuses on three plant biomass-degrading microbial consortia that were trained to degrade once-used wheat straw (WS1-M), switchgrass (SG-M) and corn stover (CS-M). These microbial consortia were characterized by metagenomics (taxonomic and carbohydrate-active enzyme profiles) and metasecretomics approaches. A key finding was the dominance of lytic polysaccharide monooxygenases as well as of enzymes involved in xylan/arabinoxylan degradation. This chapter also reports on the analysis of enzymatic activities in the metasecretome produced from each microbial consortium by using a new generation of chromogenic substrates. Based on these approaches, we found that the microbial consortia have the potential to degrade specific and complex plant-derived polysaccharides.

Chapter 8 summarizes and discusses the main findings of this thesis against recent relevant literature. Theoretical considerations, perspectives, concluding and final remarks with respect to future work are exposed in the light of ecological studies and biotechnological applications.