

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/22379> holds various files of this Leiden University dissertation

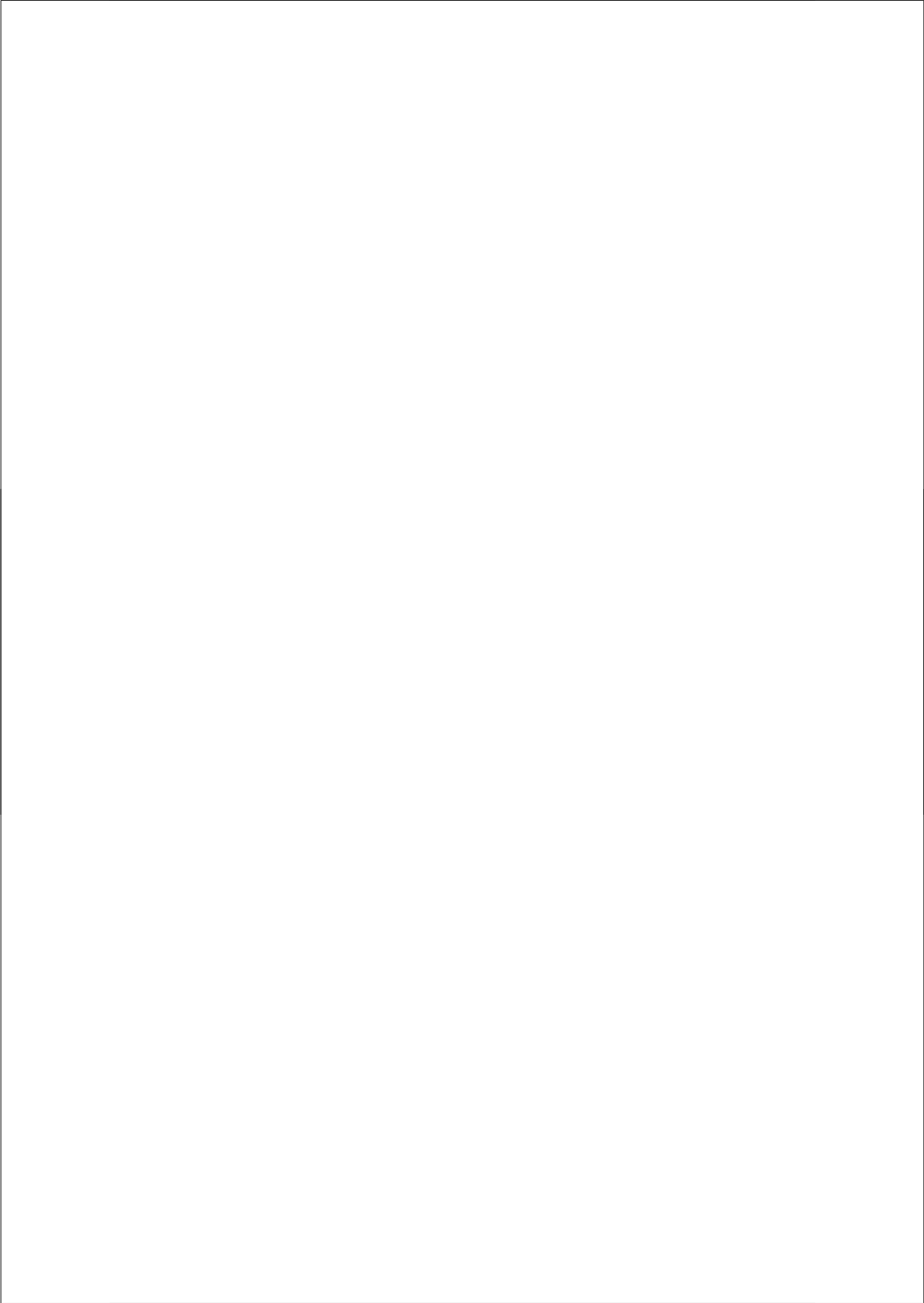
Author: Zha, Ying

Title: Unraveling substrate dynamics and identifying inhibitors in hydrolysates of lignocellulosic biomass by exometabolomics

Issue Date: 2013-11-27

**Unraveling Substrate Dynamics and Identifying Inhibitors
in Hydrolysates of Lignocellulosic Biomass
by Exometabolomics**

Ying Zha



Unraveling Substrate Dynamics and
Identifying Inhibitors in
Hydrolysates of Lignocellulosic Biomass by
Exometabolomics

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 27 november 2013
klokke 11:15 uur

door

Ying Zha

geboren te Yinchuan
in 1983

PROMOTIE COMMISSIE

Promotor: Prof. dr. P.J. Punt

Overige leden: Prof. dr. ir. J.J. Heijnen (Technische Universiteit Delft)
Dr. L. Karaffa (University of Debrecen)
Prof. dr. A.K. Smilde (Universiteit van Amsterdam)
Prof. dr. C. J. ten Cate
Prof. dr. T. Hankemeier
Prof. dr. C.A.M.J.J. van den Hondel

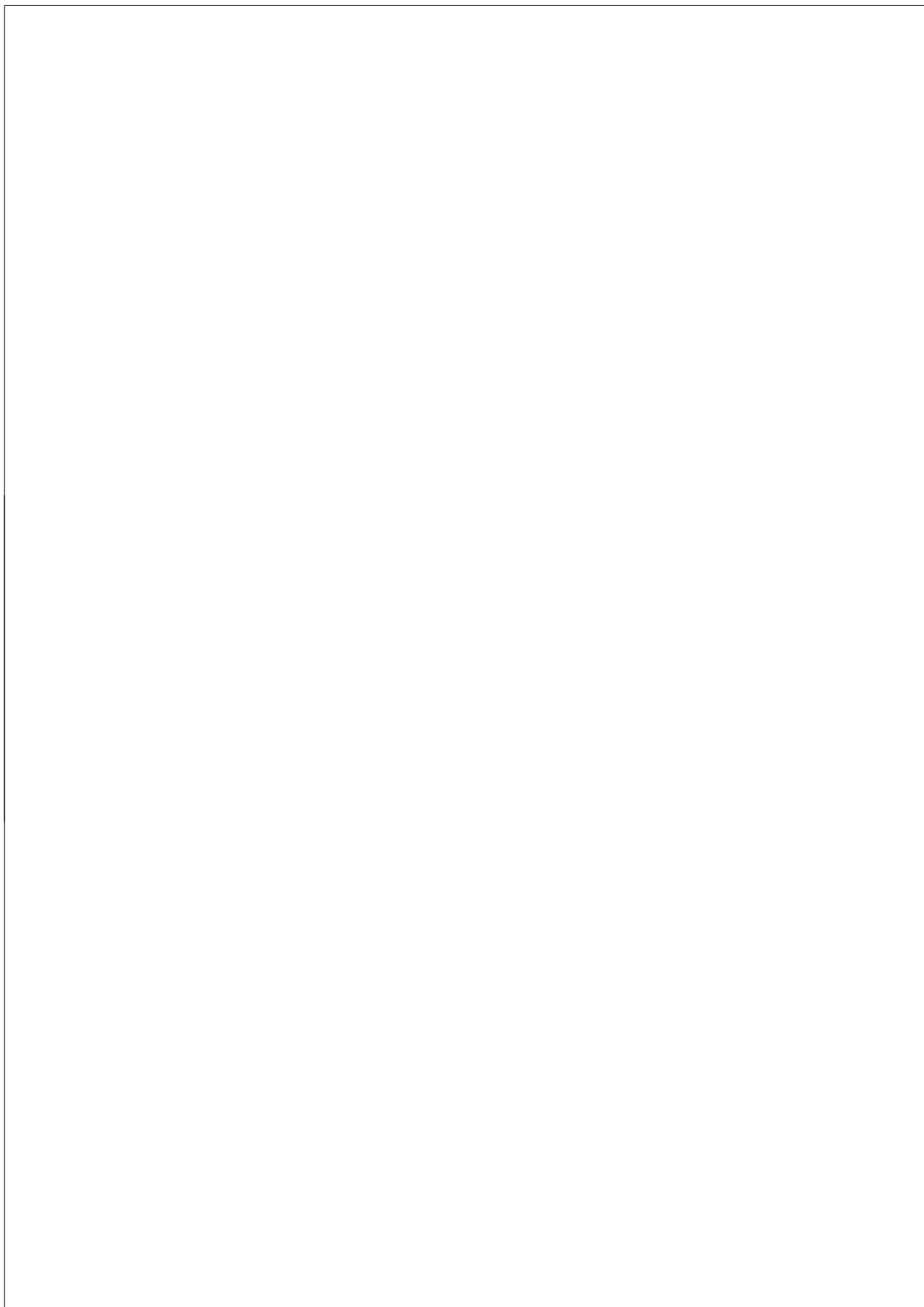
The work described in this thesis was financially supported by the Netherlands Metabolomics Centre (NMC) and carried out at TNO Microbiology & Systems Biology.

Printed by: *Ipskamp Drukkers*

ISBN: 978-94-6191-893-2

CONTENTS

	Outline	1
Chapter 1	Exometabolomics approaches in studying the application of lignocellulosic biomass as fermentation feedstock	3
Chapter 2	Preparation and evaluation of lignocellulosic biomass hydrolysates for growth by ethanologenic yeasts	33
Chapter 3	Analysis of oligosaccharides in lignocellulosic biomass hydrolysates by high-performance anion-exchange chromatography coupled with mass spectrometry (HPAEC-MS)	49
Chapter 4	Inhibitory compounds in lignocellulosic biomass hydrolysates during hydrolysate fermentation processes	73
Chapter 5	Identifying inhibitory compounds in lignocellulosic biomass hydrolysates using an exometabolomics approach	99
Chapter 6	<i>Pichia anomala</i> 29X: A resistant strain for lignocellulosic biomass hydrolysate fermentation	129
	Summary	147
	Samenvatting	149
	List of publications	153
	Curriculum vitae	155



OUTLINE

Lignocellulosic biomass is the 2nd generation feedstock for biofuel production through fermentation processes. The material has a rigid structure, which needs to be broken down by a pretreatment procedure to expose cellulose for hydrolysis. The hydrolysis products, so called biomass hydrolysates, contain next to the sugar monomers, toxic compounds released and formed during the pretreatment process. These compounds inhibit the growth of the fermenting host(s). To improve the fermentability of biomass hydrolysates, identification of these inhibitory compounds is of great importance. **Chapter 1** of this thesis reviews the approaches and techniques that have been used to study the inhibitors in various biomass hydrolysates, and introduces a non-targeted methodology to systematically identify biomass hydrolysate inhibitors: the exometabolomics approach. To identify hydrolysate inhibitors through an exometabolomics approach, a wide range of biomass hydrolysates needs to be prepared. **Chapter 2** describes the detailed procedures of four pretreatment methods and the overall fermentability of the generated hydrolysates. The hydrolysis efficiency of the carbohydrate polymers in pretreated biomass was analyzed by using high-performance anion-exchange chromatography coupled with mass spectrometry (HPAEC-MS), and the results of this analysis are presented in **Chapter 3**.

The last three chapters of the thesis focus on identifying inhibitory compounds in lignocellulosic biomass hydrolysates and studying their effects on fermenting yeasts during fermentation processes. **Chapter 4** examines the fermentability of a series of biomass hydrolysates, in relation to the presence and dynamics of a target group of inhibitors in these hydrolysates. **Chapter 5** reports the detailed experimental procedure and results of the actual exometabolomics approach introduced in Chapter 1. The research question, identification of inhibitors in biomass hydrolysates, was answered by statistically correlating the fermentability of different biomass hydrolysates with their biochemical compositions. Finally, in search for potential ethanologenic host organisms resistant to biomass hydrolysate inhibitors, a *Pichia anomala* strain was isolated. In **Chapter 6**, the properties and fermentation performance of this strain in biomass hydrolysates were tested. Through further research and possibly genetic modifications, the strain has the potential to become a suitable yeast for fermenting lignocellulosic biomass hydrolysates.

CHAPTER 1

Exometabolomics Approaches in Studying the Application of Lignocellulosic Biomass as Fermentation Feedstock

Ying Zha and Peter J. Punt

Metabolites **2013**, 3, 119-143

Abstract

Lignocellulosic biomass is the future feedstock for the production of biofuel and bio-based chemicals. The pretreatment-hydrolysis product of biomass, so-called hydrolysate, contains not only fermentable sugars, but also compounds that inhibit its fermentability by microbes. To reduce the toxicity of hydrolysates as fermentation media, knowledge of the identity of inhibitors and their dynamics in hydrolysates need to be obtained. In the past decade, various studies have applied targeted metabolomics approaches to examine the composition of biomass hydrolysates. In these studies, analytical methods like HPLC, RP-HPLC, CE, GC-MS and LC-MS/MS were used to detect and quantify small carboxylic acids, furans and phenols. Through applying targeted metabolomics approaches, inhibitors were identified in hydrolysates and their dynamics in fermentation processes were monitored. However, to reveal the overall composition of different hydrolysates and to investigate its influence on hydrolysate fermentation performance, a non-targeted metabolomics study needs to be conducted. In this review, a non-targeted and generic metabolomics approach is introduced to explore inhibitor identification in biomass hydrolysates, and other similar metabolomics questions.

Introduction

In the last decade, more and more attention has been paid to using lignocellulosic biomass as feedstock for bulk chemical production with biotechnology processes [1,2]. This biomass, including for example wheat straw, corn stover and bagasse, consists mainly of agricultural residues, which is renewable and not competitive with world food supply [3,4]. If microorganisms could use such biomass efficiently as fermentation feedstock, production processes would be less expensive and more environmentally friendly.

Lignocellulosic biomass is mainly composed of cellulose, hemicellulose and lignin (Figure 1). Cellulose is a polysaccharide consisting of D-glucose, and it forms the backbone structure of lignocellulose; hemicellulose is composed of a matrix of different polysaccharides, such as xylan, arabinoxylan and xyloglucan; in addition, lignin is a complex aromatic polymer, functioning as the supportive structure of lignocellulose [5,6]. Due to the rigid structure of lignocellulosic biomass, very few microorganisms can use the biomass directly for growth and production. Therefore, prior to feeding the biomass into fermentors, a pretreatment-hydrolysis step is carried out to break down the structure of lignocellulosic biomass and hydrolyze the exposed polysaccharides into monomers [7,8]. The conditions under which feedstock is pretreated are quite harsh, involving high temperature, high pressure and an acidic/alkaline environment [9-11]. Pretreatment not only results in the disruption of the lignocellulose structure but also in the formation and release of compounds, which could negatively influence the fermentation processes. Therefore, when biomass hydrolysates (hydrolysis products of lignocellulosic biomass) are used as fermentation media, their fermentability is reduced compared to synthetic media with pure sugar monomers as carbon source (Figure 1) [12,13].

To identify and ultimately reduce the effects of inhibitory compounds on the fermentation processes, insight into biomass hydrolysate composition and its relationship with fermentation performance is required. One way to obtain this insight is through a so-called metabolomics approach. That is, by studying the relationship between (the change of) metabolite levels and performance of the biological system [14,15].

Metabolomics is a functional genomics approach aimed at studying the diversity of biological systems by analyzing intra- and extra-cellular metabolites. Compared to genomics, transcriptomics and proteomics, metabolomics reflects most directly the physiological status of a biological system, as metabolites links most closely to the phenotype of an organism

[15,16]. In the last two decades, a diverse range of techniques that can detect and quantify metabolites with various properties have been developed. Metabolomics has been applied in the areas of pharmacy, food and nutrition, plant research and biotechnology [17,18]. Metabolomics studies include detecting metabolite level change caused by genetic modification and/or altered environmental conditions [19,20], finding bio-makers that improve the performance of a biological system [21], and sample classification [22].

Metabolites are small organic compounds participating as intermediates or products in metabolic pathways. Metabolites that are secreted into fermentation media are defined as exo-metabolites together constituting the so-called exo-metabolome. As the chemical properties of different metabolites are diverse, usually several different analytical techniques are required to conduct a metabolomics study [16,23,24].

Different metabolomics approaches may be adopted, such as metabolite target analysis, metabolite profiling, metabolomics and metabolic fingerprinting [14]. With metabolite target analysis and metabolite profiling, a selection of metabolites is made based on previous research and expert knowledge, and for the most part a single analytical technique is chosen for measuring this group of compounds. These approaches allow a simple sample analysis process and avoid dealing with complex data-sets. However, though widely applied, these approaches are often biased, neglecting the metabolites that are not in the selection. This can artificially amplify effects of selected compounds on the performance of the biological system, losing information like synergetic effect with compounds not selected for analysis [25]. When it is not known which metabolites are of importance in the research question, a non-targeted metabolomics approach becomes essential, since the approach does not involve compound pre-selection.

Metabolomics approaches, mostly targeted, have been used to study the composition of lignocellulosic biomass hydrolysates, in relation to their performance as fermentation media. The “exo-metabolites” in such metabolic footprinting studies are components of biomass hydrolysates [26]. These exometabolomics studies help to identify compounds that inhibit the growth of fermenting microbes, reveal the dynamics of some inhibitory compounds in detoxification and fermentation processes, and provide evidence to optimize pretreatment conditions. To further investigate the overall composition of different types of biomass hydrolysates, and study potential inhibitors in these hydrolysates unbiased, a non-targeted exometabolomics approach should also be adopted.

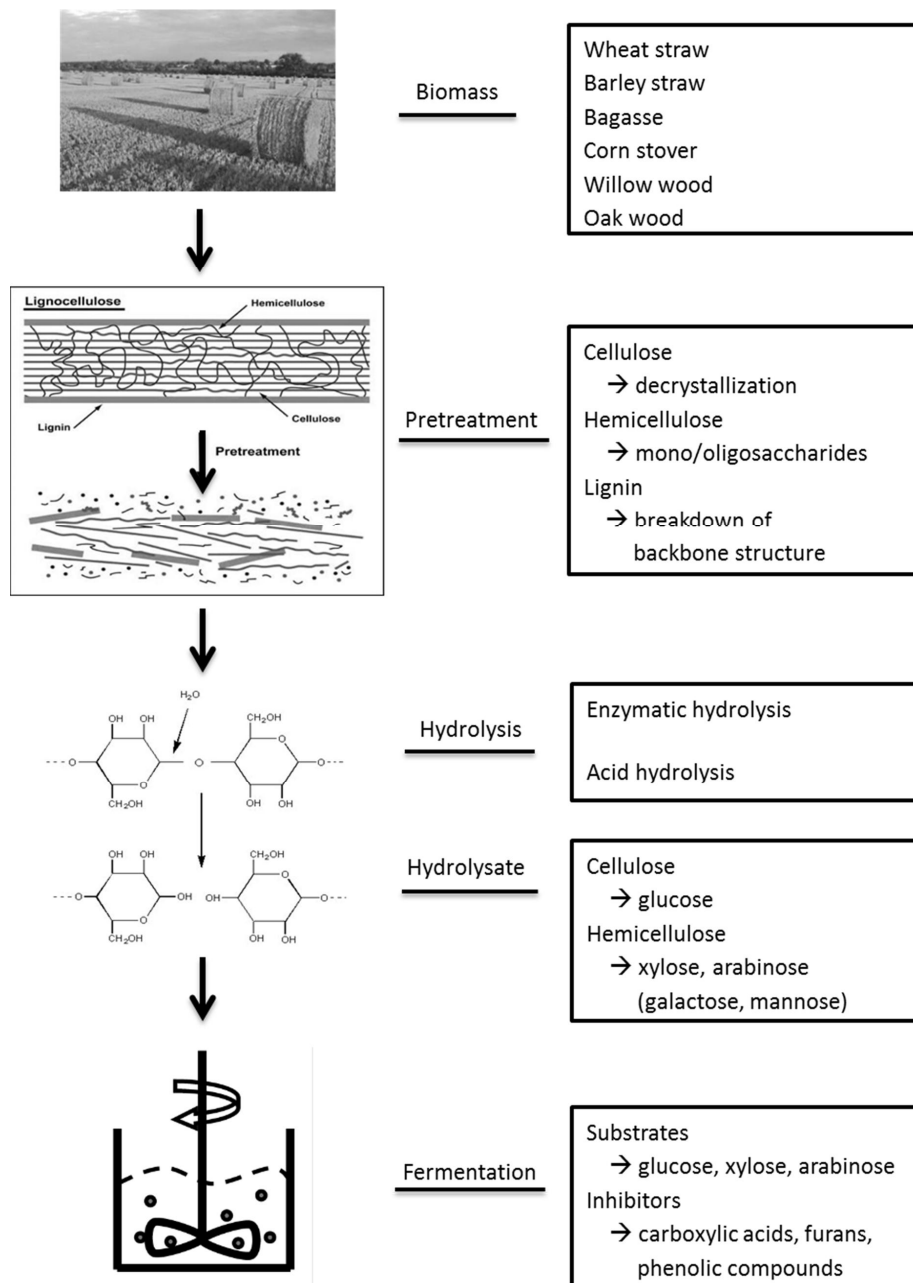


Figure 1 Schematic workflow for the preparation of lignocellulosic biomass hydrolysates and their use in microbial fermentation.

In this review, we present several targeted exometabolomics approaches with which the composition of lignocellulosic biomass hydrolysates was studied. The analytical methods used for analyzing the non-sugar compounds in biomass hydrolysates are summarized. The use of targeted approaches in improving pretreatment conditions and fermentation performance of hydrolysates is illustrated. Furthermore, a non-targeted and generic exometabolomics approach is introduced. The approach is applied to identify inhibitors in different types of biomass hydrolysates unbiased and to study their dynamics in fermentation processes.

General approach of metabolomics studies

In general, the goal of a metabolomics study is to address biological questions by measuring relevant metabolites in a biological system. The measured metabolites are used to reveal their relationship with the performance of the biological system through statistical means. A flowchart illustrating the general metabolomics approach is shown in Figure 2A.

The first step is to define a research question that clearly describes the aim of the study. The question should be informative and specific, pointing out both the analytical targets and the biological system of the study [27]. When the research question is clear and specific, it can be translated into a statistical question, based on which experimental design is carried out and tentative statistical methods are chosen.

Based on the defined research question, an estimation of the amount of metabolites to be measured can be made. The number of metabolites to be measured relates not only to the property of the biological system, but also to the coverage of the analytical methods used. For instance, when both GC-MS and LC-MS were used to analyze the metabolome of *Escherichia coli*, the detection of between 250 and 500 metabolites was estimated [16]. Based on the number of metabolites to be analyzed, the number of different experiments can also be determined. The larger the number of metabolites, the more experiments should be carried out to acquire a reliable answer to the research question [28–30].

Knowing the required number of different experiments to be conducted, experimental design can be performed. The requirements of the designed experiments are that they (1) closely relate to the research question, (2) reflect real-life situations, and (3) result in a range of well-spread measurable phenotypes [31,32]. To ensure the success of the experimental design, information about the operability and repeatability of the experiments needs to be obtained

beforehand. Preferably, more experiments than desired are initially conducted, so that, when certain experiments do not meet the requirements mentioned above, they can be discarded from the design.

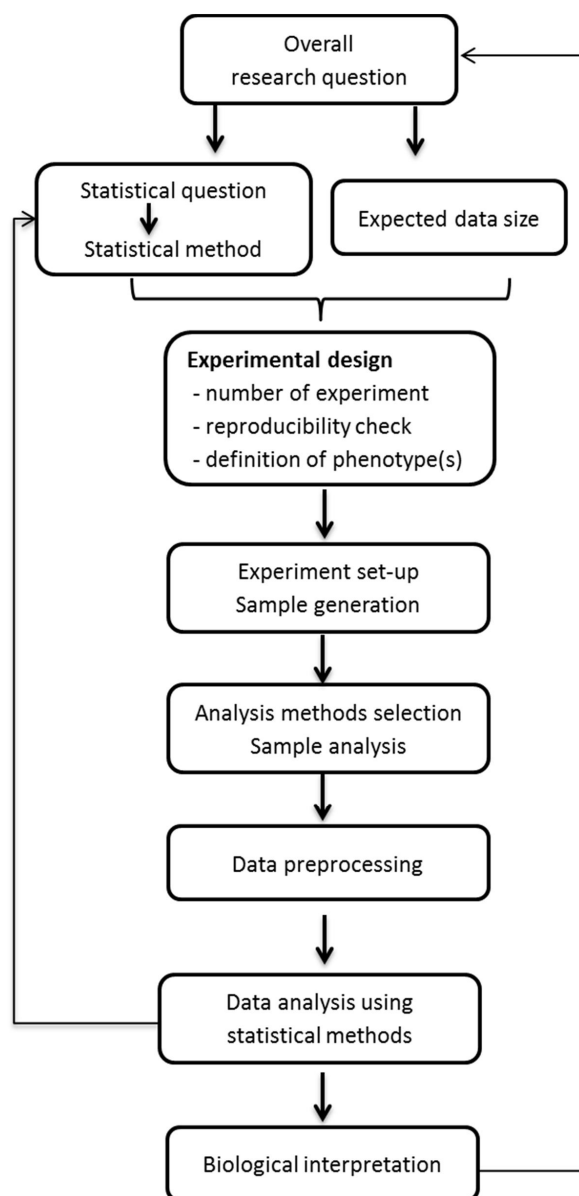


Figure 2A: Overview of a general metabolomics workflow.

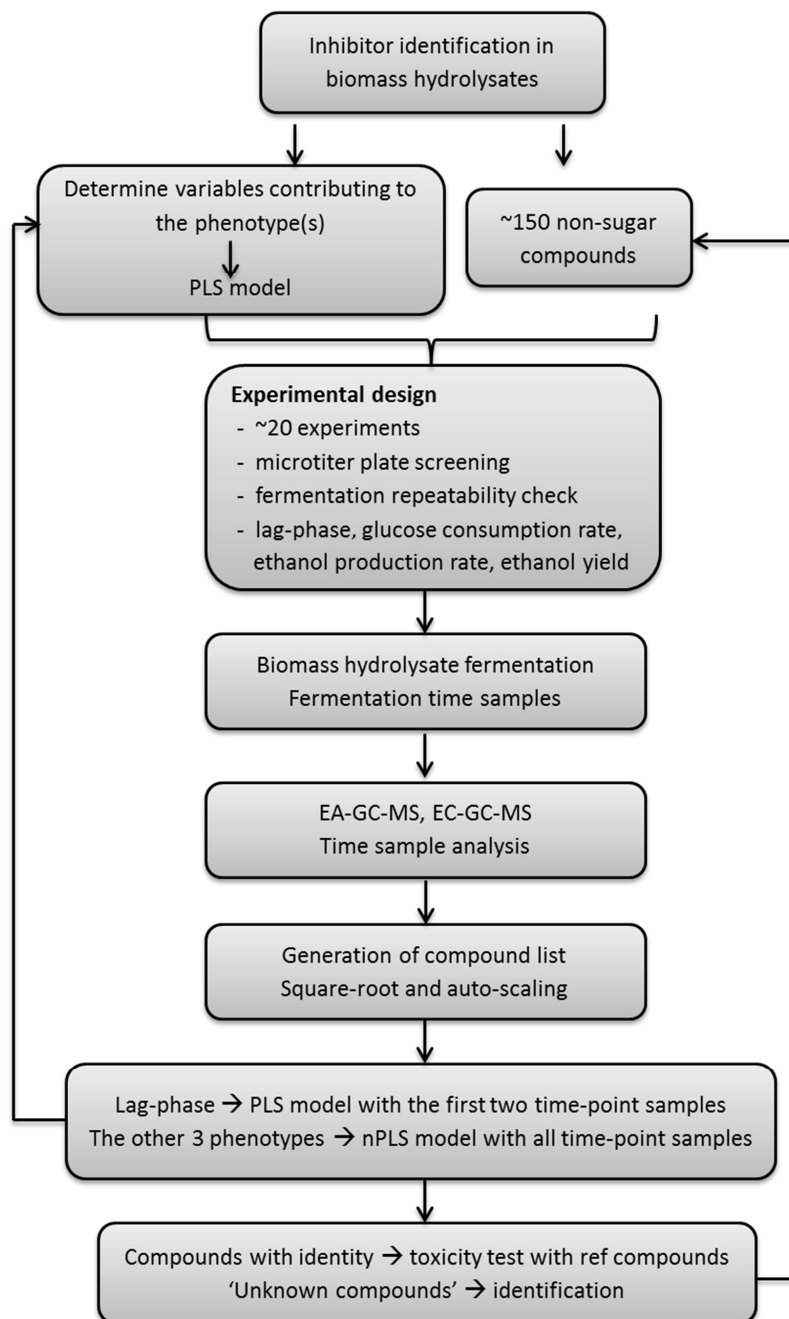


Figure 2B: Non-targeted metabolomics workflow used for studying the use of lignocellulosic biomass hydrolysate as fermentation medium, in particular by identifying inhibitory compounds.

Another key point in experimental design is defining the phenotype(s) to characterize the performance of the different experiments. Depending on the selected phenotype, experimental set-up and sampling strategies will be determined. Phenotypes are parameters defined to describe the research question in a quantitative manner. There is no universal phenotype definition, since the focus of each study is different, and often more than one phenotype is needed to fully represent the research question. The importance of defining proper phenotypes and its influence in answering the research question are illustrated by Braaksma *et al.* [19]. In her study on enzyme production by the filamentous fungus *Aspergillus niger*, six different phenotypes were defined to be able to fully address the research question. In addition, different metabolite target groups were found to be correlating to different phenotypes. Therefore, defining a series of phenotypes that address different aspects of the research question is strongly recommended.

As soon as phenotypes are defined, experimental process and sampling can be set up to obtain parameters needed to calculate the phenotypes. For practical reasons, it is preferred to set up as simple an experimental process and sampling method as possible, given that all necessary parameters can be acquired. One should bear in mind that the sampling method is also determined by the biological system and the sample analysis techniques of the study. The sample work-up of extracellular metabolomics (exometabolomics) is much simpler than intracellular metabolomics, which needs sample quenching, metabolites extraction and biomass correction [25,26]. When GC-MS is used to analyze samples, extra sample preparation steps, like derivatization, are often required, which is generally not required for LC-MS and NMR analysis [23].

In targeted metabolomics, analytical methods are chosen based on the properties of the pre-selected compounds. In non-targeted metabolomics, analytical techniques also need to be chosen, as it is not possible to use all available methods for sample analysis. Besides, it is more informative to focus on metabolite classes that are relevant to the aim of the study. Therefore, analytical methods in non-targeted metabolomics should still be selected based on the research question and known properties of the biological system. For instance, when it is known that volatile compounds may be important to the research question, methods allowing the analysis of these compounds, such as solid phase microextraction (SPME), should be used [33,34]; and as the focus of the study is on carbohydrates, methods like high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) or LC-MS should be selected [35,36].

In non-targeted metabolomics, sample analysis results in a list of detected compounds, both known and unknown, and their relative quantities, presented as peak areas in chromatograms. The analysis results of all samples in the experimental design form a data-set, which will be studied statistically. Before the data-set is analyzed statistically, it needs to be preprocessed. Generally, data preprocessing involves the following aspects, (1) peak area correction with internal standards, (2) data-set normalization, and (3) data-set transformation. Peak area correction is conducted to minimize the influence of sample matrix, an effect caused by the overall composition of the sample; normalization is carried out to reduce the redundancy of the data-set; and transformation is performed to increase the useful information content the data-set carries. There are multiple ways to preprocess a data-set, and the methods chosen are specific to the analytical technique used and the statistical model selected. Detailed discussions on data preprocessing are given by Roessner *et al.* [37] and van den Berg *et al.* [38].

To find the relationship between the preprocessed data-set and the defined phenotypes in non-targeted metabolomics study, multivariate data analysis (MVDA) tools are applied. The most commonly used tools are principal component analysis (PCA), partial least square (PLS), and discrimination/classification methods. PCA model points out variables (metabolites) that contribute the most to the data-set structure [39]; PLS model seeks metabolites that are most responsible for a certain phenotype [40]; discrimination/classification methods determine if a sample belongs to a specific group [28]. Based on the research question, one or several of the MVDA tools are selected to analyze the preprocessed data-set. Two other factors to be considered when conducting MVDA are (1) fusing of the data-sets generated by different analytical methods and its influence on the model building results, and (2) methods for model validation. Simply using MVDA tools for analyzing metabolomics data-sets without checking the validity of the models can produce misleading or even wrong results. Rubingh *et al.* addressed the complexity of the real-life metabolomics data. Several model validation methods were provided to attain more reliable and comprehensive data analysis results [29]. Compared to non-targeted metabolomics, the compound list in a targeted approach is very short. Since the compounds are pre-selected, their absolute concentrations can be determined with reference compounds. This simplifies or even omits data preprocessing, and makes data analysis straightforward and simple.

The last step in a metabolomics study is to translate the statistical analysis results into the biological context to answer the research question. Some analytical results speak for themselves, like the ones in discrimination/classification studies [41], while others are complex, especially those involving metabolites identification [42]. There are several tools

that assist the biological interpretation, which are illustrated by van der Werf *et al.* [25]. Additionally, it should be noted that non-targeted metabolomics analysis might suggest compounds that seem to be 'incorrect' based on expert knowledge. They are either not previously found in any similar biological systems, or known to function in an unrelated biological process. Such compounds should also be taken into account for future research, since they may play a role in further understanding the biological system studied.

Targeted approach: Applying targeted metabolomics approaches to study the sugar and lignin degradation products in lignocellulosic biomass hydrolysates

Most of the targeted approaches start with analyzing the structure of lignocellulosic biomass, which reveals several main degradation products in biomass hydrolysates, the pretreatment-hydrolysis product of lignocellulose. As shown in Figure 1, cellulose, hemicellulose and lignin are the three main components of lignocellulosic biomass. Cellulose is the linear polymer of β -1,4-linked D-glucose residues, hemicellulose is a heteropolymer mainly containing xylan, arabinoxylan and xyloglucan, when hydrolyzed generating xylose, mannose, galactose, arabinose and glucose [43]. Lignin is a complex macromolecule composed of phenylpropane units, which are the dehydrogenation products of *para*-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol [13]. The degradation products of the sugar monomers of cellulose and hemicellulose, and lignin are generally categorized into small carboxylic acids, furans and phenolic (aromatic) compounds [12,44]. Formic, acetic and levulinic acid are the most common small carboxylic acids, while furfural and 5-hydroxy-methylfurfural (HMF) are the representatives of furans [45]. Comparatively, the diversity of phenolic (aromatic) compounds in biomass hydrolysates is much greater [46–49]. In this section, the analytical methods used to detect and quantify these three categories of compounds are presented. Furthermore, application of targeted metabolomics approaches on identifying inhibitors in biomass hydrolysates and improving hydrolysate preparation methods is reviewed.

Analytical methods for studying hydrolysate composition

As many of the targeted studies referred to in this review are focused on specific classes of compounds, analytical methods used to detect and quantify these are discussed separately. However, general aspects of these analytical tools are often not specific for the compound classes.

Exometabolomics approaches in studying lignocellulosic biomass

Table 1 Analytical methods used for detecting compounds in lignocellulosic biomass hydrolysates.

analytical method	extraction / derivatization	detected compounds	identification	quantification	ref
HPLC	no	formic, acetic acid, levulinic acid, lactic acid, glycolic acid, malic acid, citric acid, succinic acid, oxalic acid	no	yes	[49-51, 55,58-62]
		furfural, HMF furfuryl alcohol, 2,5-bis-hydroxymethylfuran			[49,55,58, 60,61,63,64]
RP-HPLC	precipitation-filtration, MTBE / no	formic acid, lactic acid, acetic acid, levulinic acid, furfural, HMF, phenolic compounds	partial	yes	[52]
	MTBE / no	gallic acid, furfural, HMF, protocatechuic acid, vanillin, coniferyl alcohol, syringaldehyde, sinapic acid	partial /GC-MS		[51]
	no	reference phenolic compounds	GC-MS		[50]
CE	no	formic acid, acetic acid, levulinic acid, glycolic acid, lactic acid, furfural, HMF	no	yes	[53,63,65]
GC-MS	solvent / no	acetic acid, furfural, acetamide	yes/ partial	yes	[66]
	MTBE / silylation	gallic acid, HMF, vanillin, protocatechuic acid, syringaldehyde		no	[51]
	DCM / EC-derivatization	levulinic acid, furfural, furfuryl alcohol, 2-furanmethanol acetate, HMF, phenolic compounds		yes	[67]
	SPE / silylation	phenolic compounds phenolic compounds			[60,61]
	EA / silylation	furfural, HMF, furfuryl alcohol, 2-furoic acid, phenolic compounds		yes	[55]
		phenolic compounds		no	[49]
		phenolic compounds		yes	[57,59]
	no / silylation	lignin derived monomer and dimers			[65]
LC-MS/MS	precipitation-filtration, MTBE / no	aliphatic acids, furans, phenolic compounds	yes	yes	[54,68]

MTBE: methyl tertiary butyl ether; DCM: dichloromethane; SPE: solid phase extraction; EC: ethylchloroformate; EA: ethylacetate.

Small carboxylic acids and furans

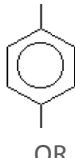
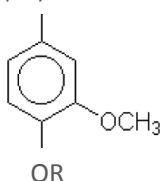
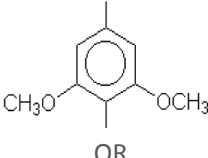
Several methods have been extensively used to detect and quantify small carboxylic acids and furans in biomass hydrolysates, among which are High-Performance Liquid Chromatography (HPLC), and Capillary Electrophoresis (CE) (Table 1).

HPLC is the most standard method for quantifying monomer sugars, simple small carboxylic acids, furfural and HMF, though the analytical system and column used may vary. The method requires little sample work-up and detects a limited range of target compounds, which are quantified by making calibration curves using external standards. RP-HPLC is a variation of HPLC that detects a much larger group of compounds with identification possibility only when followed up by GC-MS [50,51]. RP-HPLC assigns identity to detected compounds mainly by comparing their retention time to and/or spiking samples with reference compounds. In the identification process, no compound structural analysis is involved and the availability of reference compounds is a necessity. Therefore, the identification conducted by RP-HPLC requires prior knowledge [52].

It can also be seen from Table 1 that an extraction step using methyl tertiary butyl ether (MTBE) is often used before analyzing hydrolysate samples with RP-HPLC in combination with detection based on refractive index (RI). This is because hydrolysate samples normally contain high concentrations of sugars, like glucose. These huge sugar peaks appear in RP-HPLC chromatograms interfere with the RI detection of target compounds, like furans. Therefore, to minimize the disturbance, sugars are removed by extracting hydrolysates with organic solvent before conducting analysis. This applies also to GC-MS method, which requires an extraction step before the derivatization step in sample preparation (Table 1). Besides small carboxylic acids and furans, phenolic (aromatic) compounds can also be studied by RP-HPLC. This will be discussed in the next section.

CE is yet another method for analyzing the described compounds in hydrolysates. Compared to RP-HPLC, the targets of CE are more specific, mainly small organic acids. Like in HPLC, little sample work-up is needed for CE, and the method cannot be used for identification of novel compounds. When analyzing hydrolysate samples, it is preferred to measure both carboxylic acids and furans with one analytical method. Since HPLC is capable of detecting both acids and furans, the method is often chosen above CE. Recently, it was shown that CE can also separate saccharides and furans in hydrolysate samples, and the quantification results of CE on furfural and HMF are highly comparable to HPLC [53]. Therefore, CE has the potential to become a routine analytical method for measuring hydrolysate samples.

Table 2 Phenolic (aromatic) compounds detected in the studies listed in Table 1.

<i>p</i> -hydroxyphenyl residue (H)	detected in more than one study*	detected in one study	hydrolysate	ref
polymer 	phenol 4-hydroxybenzaldehyde 4-hydroxybenzoic acid salicylic acid (2-hydroxybenzoic acid) 3,4-dihydroxybenzaldehyde benzoic acid catechol (1,2-dihydroxybenzene) <i>p</i> -coumaric acid (4-hydroxycinnamic acid) piceol (4-hydroxyacetophenone)	hydroquinone	spruce-dilute acid	[60]
		4-methoxyphenol <i>p</i> -coumaryl alcohol phloretic acid (3-(4-hydroxyphenyl) propionic acid)	wheat straw-steam explosion	[55]
		<i>o</i> -cresol (2-methylphenol) gentisic acid (2,5-dihydroxybenzoic acid) protocatechuic acid (3,4-dihydroxybenzoic acid)	willow-acid steam	[57]
		caffeic acid (3,4-dihydroxy cinnamic acid)	corn stover-dilute acid or ammonia fiber expansion	[54]
Guaiacyl residue (G)				
polymer 	guaiaacol vanillin vanillic acid homovanillic acid ferulic acid (4-hydroxy-3-methoxycinnamic acid) 3-hydroxy-4-methoxycinnamic acid coniferyl aldehyde dihydroconiferyl alcohol acetovanillone (acetoguaiacone) G-CH2COCH3	G-CHOHCOCH3 G-COCOCH3 G-CH2COCH2OH G-COCHOCH3	spruce-dilute acid	[60]
		vanillyl alcohol G-CH2CH2COOH G-CHCHCHO	wheat straw-steam explosion	[55]
		<i>trans</i> -isoeugenol	willow-acid steam	[57]
Syringyl residue (S)				
polymer 	syringaldehyde syringic acid	acetosyringone	wheat straw-alkaline wet oxidation	[61]
		syringol S-CHCHCHO	wheat straw-steam explosion	[55]
Other structures		biphenyl-type dimer diarylpropane-type dimer pinoresinol-type dimer	Japanese beech-hot compressed water	[65]
		4-hydroxycoumarin <i>o</i> -toluic acid <i>p</i> -toluic acid	corn stover-dilute acid or ammonia fiber expansion	[54]

* The compounds listed in this column appeared in two or more studies listed in the "ref" column and the following three references: [49] [52] [67]. The hydrolysates used in these three studies were corn stover-dilute acid, yellow poplar organosolv, and bagasse and oak hydrolysates.

Phenolic (aromatic) compounds

As addressed before, phenolic (aromatic) compounds are mostly the degradation products of lignin, and due to the complexity of lignin structure, the chemical structure of this group of compounds in biomass hydrolysates is very diverse. The potential phenolic compounds in hydrolysates derived from the three basic lignin building blocks, namely *para*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) residues, are summarized by Klinke *et al.* [13] (Table 2). It was estimated that about 60 different phenolic compounds could be found in various hydrolysates, including compounds with unknown structures.

To detect, identify and quantify these phenolic compounds in hydrolysates, several different methods have been applied, including RP-HPLC, Gas Chromatography–Mass Spectrometry (GC-MS) and Liquid Chromatography–Mass Spectrometry² (LC-MS/MS), see Table 1. A common characteristics of the three techniques is that they all possess the possibility of (partial) identification, which is essential for studying a diverse group of lignin degradation products with many ‘unknowns’.

Compound identification with RP-HPLC and LC-MS/MS is mainly done by first constructing chromatograms with a relative large group of reference compounds. The generated chromatograms are then compared with the peaks in the sample chromatogram [52,54]. By comparing the retention time in LC and/or extract mass information provided by MS, identities can be assigned to peaks in hydrolysatesamples. Since for each identified compound, its reference is already available, quantification can be directly carried out by generating calibration curves.

In contrast to LC methods, GC-MS appears to be more open, as reference compounds are only involved in a later stage of the analysis. The initial identification with GC-MS is often conducted by comparing compound fragment profiles with a mass spectral library [48,55,56]. In some cases, reference compounds are used to confirm the identity of characterized peaks [48]. Even when identity is not assigned, an indication of the category the compound belongs to can be given [57]. Therefore, GC-MS seems to be a preferred method for studying phenolic (aromatic) compounds in biomass hydrolysates. The method has been adopted in multiple studies, resulting in the identification and quantification of a variety of phenolic (aromatic) compounds, see Table 2. It can be seen that most identified phenolic compounds fell into the categories of the three lignin building blocks, primarily aldehyde and acid forms. More derivatives of *para*-hydroxyphenyl residues (H) were found than guaiacyl derivatives (G) than syringyl derivatives (S). Phenolics dimers and non-phenolic aromatic compound, namely toluic acid, were also detected. These analysis results confirmed that there is a

diverse group of phenolic compounds in biomass hydrolysate, indicating that the phenolic compound composition in different hydrolysates vary.

Application of targeted approaches in studying biomass hydrolysates

The detection and quantification of the degradation products of sugars and lignin in lignocellulosic biomass hydrolysates not only revealed the presence and level of such compounds, but also provided information to (1) test the toxicity of these compounds towards microbes, (2) study the formation conditions of these compounds, (3) trace their dynamics in a detoxification treatment or during a fermentation process. The applications of targeted metabolomics approaches on studying these aspects of biomass hydrolysates are discussed below.

Inhibitor identification

Biomass hydrolysates generated from different pretreatment methods exhibit inhibitory effects when used as fermentation media. Some elongate lag-phase, some reduce growth rate, some lower product yield, while others abolish growth completely [12,44]. The inhibitory effects are the results of compounds present in hydrolysates, which are formed or released during the pretreatment process. These inhibitory compounds are mostly sugar and lignin degradation products, which can be different in each hydrolysate. To improve the fermentability of biomass hydrolysates, identifying these compounds is crucial. The identification has been carried out by using targeted metabolomics approach.

Most studies start with selecting a group of compounds that are potentially inhibitory in biomass hydrolysates. The selection was made based on expert knowledge as well as previous research results. For instance, in the study of Chen *et al.* [52], aliphatic acids, phenols, aromatic acids and aromatic aldehydes were selected as they were reported as major degradation products in biomass hydrolysates [13]. According to the chemical properties of the selected compounds, analytical methods were established to measure and, in some cases, quantify these compounds. Both RP-HPLC and GC-MS have been used in such studies, and pure reference compounds were used for both identification and quantification purposes [50,52,59]. In some studies, the presence of the selected compounds in the actual hydrolysate was checked [52,58], while in other studies, their inhibitory effects towards one or several microbes were tested by spiking with various concentrations [50,69].

In some other studies, the pre-selection of potential inhibitors was not conducted, hydrolysates were typically analyzed with GC-MS, and the mass spectra of the resulting peaks were used for compound characterization [49]. The characterization was either done by comparing the mass spectra of the detected peaks to a mass spectral library [48,55,56], or comparing them to a series of reference compounds [51,59]. When a mass spectral library is used, a large group of compounds can be characterized [55]. However, instead of exploring the inhibitory effect of each detected compound, the authors decided to focus on vanillin and furfural based on previous research results. This makes such a study targeted from this point on. Compared to approaches using reference compounds, the benefit of directly analyzing hydrolysates with GC-MS is that as soon as the compound is characterized, its presence in the hydrolysate is also confirmed. The concentration of the characterized compound can be determined with its reference compound, and its toxicity can be tested according to its concentration present in the hydrolysate [51,67].

Pretreatment condition optimization

It is known that the inhibitory compounds in biomass hydrolysates are mainly formed during pretreatment process, which is in most cases operated under harsh conditions (Figure 1) [9–11]. The fermentability of a specific hydrolysate is, to a great extent, determined by its pretreatment [11]. Thus, studying the relationship between biomass pretreatment and its resulting hydrolysate composition provides valuable information for selecting appropriate pretreatment conditions.

A targeted metabolomics approach has been used to study the influence of pretreatment conditions on fermentable sugars and inhibitors formation of a specific pretreatment method [62,63]. The approach started with designing experiments by varying specific pretreatment conditions, such as temperature and residence time, both individually and together. All different pretreatment conditions were quantitatively represented by a series of combined severity factors (CS), and under each CS, a pretreatment experiment was carried out. Samples were taken from the resulting hydrolysates of different CS for analysis. The fermentable sugars and inhibitors to be analyzed were pre-selected based on expert knowledge, which in turn determined the analytical methods. As the inhibitors selected in these studies were small carboxylic acids, furfural and HMF, HPLC and CE were used to quantify these compounds in the hydrolysate samples (Table 1). Based on the analysis results, the authors evaluated the influence of CS on the formation of fermentable sugars, as well as on the release of the selected inhibitors, which provided criteria for choosing the optimal pretreatment conditions.

A similar approach has been applied by Klinke *et al.* to not only determine the optimal pretreatment conditions, but also study the correlation between pretreatment conditions and the degradation products [61]. In such a study, a much larger range of potential inhibitory compounds were selected, which included not only carboxylic acids and furans, but also phenolic compounds. Hydrolysates, prepared at different pretreatment conditions, were analyzed with GC-MS for their phenolic contents. The identification of the phenols was conducted by comparing their MS spectra with a mass spectral library, and standards were used to verify the identity and quantify these compounds in hydrolysates. The correlation between pretreatment conditions and the detected degradation products was studied statistically, using principal component analysis (PCA), revealing the influence of each single pretreatment condition on the formation of degradation products.

Monitoring compound dynamics during detoxification and fermentation

To reduce the toxicity of biomass hydrolysates as fermentation media, detoxification methods have been developed to remove inhibitors in hydrolysates [44,70]. The effects of detoxification were improved fermentability and increased product yield [71–73]. To study beyond the effect of hydrolysate detoxification, the composition change in terms of (potential) inhibitory compounds in hydrolysates needs to be monitored during the detoxification. Such studies were conducted using targeted metabolomics approaches. The most straightforward way of studying a detoxification process was by using the already identified inhibitors as monitoring targets. These inhibitors mainly include small carboxylic acids, furfural and HMF. Typically, the concentration of these compounds was determined before and after the detoxification process, using HPLC [72,74,75]. The targets of each detoxification method can be different, as far as monitored compounds were considered. For instance, it was discovered that the chemical detoxification by overliming was specifically effective to furans [74,76]. In the study of Martinez *et al.* [76], besides the selected inhibitors, the authors also looked at the unknown peaks in the HPLC chromatogram. Among those unknown peaks, three decreased after overliming, indicating that more compounds could be involved in resulting the detoxification effect of this specific method.

When the detoxification targets are neither small carboxylic acids nor furans, a different targeted metabolomics approach than the one discussed above should be applied. In the case of enzymatic detoxification using laccase, phenols were assumed to be the detoxification targets, as laccase is a phenol oxidase. This assumption was verified by Larsson *et al.* [60] through quantifying small carboxylic acids, furans and total phenols in spruce hydrolysate. To study the detoxification effect of laccase on individual phenolic compounds, both HPLC and GC-MS were adopted [56,57,77]. When HPLC was used, a pre-

selection of phenolic compounds was made based on the reported toxicity of these compounds, and their detectability by HPLC [77]. When hydrolysates were analyzed with GC-MS, the compound pre-selection was not done. The phenols detected by GC-MS were characterized either by comparing to a mass spectral library [56] or using reference compounds [57]. The advantage of using GC-MS is that the relative quantity of some unidentified compounds can also be determined to check if they were (partially) removed from the hydrolysate after detoxification.

Similar to detoxification, it was observed that during a fermentation process, the hydrolysate toxicity reduces. This is because the fermenting microbe can transform inhibitors to their less toxic form [45,78]. Targeted metabolomics approach also contributed to study the chemical conversion of these compounds. In such studies, the identified inhibitors were taken out of the context of hydrolysates and added into synthetic medium for growth testing. The conversions of these compounds were predicted based on expert knowledge, and analytical methods were selected accordingly. The conversion of furfural and HMF were monitored by analyzing their alcohol forms during fermentation processes with HPLC [64,79,80]. In addition, the conversions of vanillin and coniferyl aldehyde were investigated with RP-HPLC and GC-MS [50]. To examine these conversions in hydrolysates, GC and GC-MS were used to monitor different forms of furan and phenolic compounds, namely aldehydes, alcohols, ketones, and acids [67,81]. Similar trends of conversion from aldehyde to alcohol and acid form were observed in hydrolysates, though their quantitative relationships were not as straightforward as those in synthetic medium. These results suggested that aldehydes are more likely to be the inhibitory forms of furans and phenols in biomass hydrolysates.

By monitoring the dynamics of above mentioned compounds during detoxification and the fermentation process, it was shown that all three groups of proposed inhibitors could negatively influence hydrolysate fermentability. Especially for phenolic compounds, their toxicity was confirmed both in the laccase study and by their conversions during fermentation processes. Phenolic compounds have much greater diversity in hydrolysates compared to small carboxylic acids and furans. The overall composition of phenolic compounds was hardly studied in relation to their toxicity in biomass hydrolysates. It seems that besides the identified phenols, more of this kind of compounds are present in hydrolysates exhibiting inhibitory effects [57,59,63]. To investigate these unknown inhibitors, a non-targeted metabolomics approach needs to be carried out.

Non-targeted approach

Research case: applying metabolomics approach to study inhibitors and their dynamics in lignocellulosic biomass hydrolysates as fermentation media

As discussed in the previous section, when used as fermentation media, hydrolysates show toxicity towards fermenting microbes, due to the degradation products of (hemi-) cellulose and lignin. The toxicity varies with different types of hydrolysates, and is mainly determined by the pretreatment hydrolysis method used, but is also influenced by the biomass type [11,78].

Targeted metabolomics has been used to study the toxicity of biomass hydrolysates in fermentation processes by analyzing the composition of (hemi-) cellulose and lignin degradation products. However, it is believed that besides the identified inhibitors, there are still other non-sugar compounds and their derivatives present in biomass hydrolysates that may show toxicity or influence the toxicity of other compounds by synergistic or antagonistic effects. This is because the identified inhibitors alone do not fully explain the toxicity of biomass hydrolysates [57,59,63]. To explore the identity of these unknown compounds, the composition of biomass hydrolysates needs to be studied in a non-targeted manner, alongside the dynamics of these compounds and their effects during fermentation processes. The metabolomics approach introduced in section 2 is adopted to carry out such a study (Figure 2B). In the following sections, the steps of this study are described in more detail.

Define research question

The aim of the study was to identify compounds that (negatively) influence the hydrolysate fermentability through analyzing the composition of different hydrolysates. The corresponding research question was to identify inhibitors in biomass hydrolysates relevant for ethanolic fermentation of *S. cerevisiae*. This question can be differentiated into an experimental and a statistical research question. The experimental question was to determine which non-sugar compounds in hydrolysates are responsible for the hydrolysate toxicity towards microbes in a fermentation process. The statistical question was to determine which of the variables contribute the most to the fermentation performance phenotype(s) (Figure 2B). The variables are the detectable non-sugar compounds in hydrolysates, while the phenotypes were defined to quantitatively describe the fermentation processes.

Experimental design

In the next step, experiments were designed to answer the research question. The statistical question was first considered before any wet-lab experiments were designed. The three aspects of the statistical question were (1) selecting statistical model(s), (2) estimating the number of detectable metabolites, and (3) determining how many experiments to be carried out. In our particular case, partial least square (PLS) model was selected, as it provides, as described in section 2, those variables that most closely relate to the phenotypes. To estimate the number of detectable non-sugar compounds present in hydrolysates, the results of previous hydrolysate composition studies were used. As summarized in Tables 1 and 2, there were in total about 10 small carboxylic acids, 5 furans and 60 phenolic compounds identified. We assumed that a similar number of inhibitory compounds in hydrolysates were not yet detected, giving 150 compounds in total. This number was used to determine the number of experiments to be carried out in the experimental design. The non-sugar compounds in biomass hydrolysates are mainly (hemi-) cellulose and lignin degradation products. The formation of these compounds are interrelated, for instance, formic acid is partially formed from HMF, and furfuryl alcohol is the conversion product of furfural [45,79]. As lignocellulosic biomass is consisted of a relatively small number of building-blocks (Figure 1), it was assumed that the above-mentioned 150 detectable compounds present in hydrolysates, represent only 15–20 groups of compounds formed completely independently. For regression models like PLS, the number of experiments is preferred to be larger than the independent variables in the system. Therefore, approximately 20 experiments were to be carried out.

Knowing that about 20 different experiments were to be conducted, the wet-lab experiments were designed by resolving the following four aspects: (1) generating different experiments, (2) checking the diversity and reproducibility of these experiments, (3) setting up experimental and sampling procedures, and (4) defining phenotypes.

Different experiments were acquired by conducting batch fermentation with different biomass hydrolysates. These hydrolysates were prepared with various biomass types and different pretreatment-hydrolysis methods [82–84]. To obtain about 20 experiments, six biomass types and four pretreatment-hydrolysis methods were selected. The six biomass types were wheat straw, barley straw, corn stover, bagasse, willow wood and oak wood. They represented the most widely used biomass in the category of agriculture residue, sugar industry by-product, and wood [51,62,82,85–87]. Straw is the main agriculture residue in Europe, while corn stover is mostly produced in North and South America. Of the four

pretreatment-hydrolysis methods, three used enzymatic hydrolysis, and their pretreatment methods included acid, alkali and oxidative treatment. The fourth method used high concentration of sulfuric acid for both pretreatment and hydrolysis [11].

The hydrolysates were first prepared in small volume, *i.e.*, 50 ml, to check their diversity in fermentability by conducting a screen experiment on microtiter plates. This screen experiment confirmed that, as far as growth rate was considered, there was significant diversity among these hydrolysates [11]. Moreover, two hydrolysates were used to examine the reproducibility of batch fermentation. As shown in Figure 3, of both hydrolysates, the fermentation process was clearly presented by both duplicates. Through these pre-experiments, a good basis was formed for the full-scale experiment.

The full-scale experiment was carried out by fermenting all these hydrolysates individually. These fermentations had a fixed set-up and the same inoculum, so that the difference in fermentation performance was only caused by different hydrolysates [67]. For each fermentation, samples were taken during the whole fermentation process. With these samples biomass formation, glucose and ethanol concentration were measured. These measurements were used to visualize the fermentation process and calculate phenotypes.

Phenotypes are the quantitative description of a fermentation process. In this study, four phenotypes were defined, which were lag-phase, glucose consumption rate, ethanol production rate and ethanol yield (Figure 2B). Lag-phase was a phenotype expressed in hours, which was used to describe the time window before growth starts. Glucose consumption rate and ethanol production rate expressed how quick the microbe grows and how fast the product is produced. Ethanol yield indicated the production efficiency. Each of these phenotypes tackled a different aspect of the fermentation, and together described the whole fermentation process. It should be noted that more phenotypes could be defined, such as growth rate and productivity. However, since the fermentation aspects these phenotypes describe directly relate to one of the four phenotypes defined above, there was little value to include them.

Sample selection and analysis

To analyze the hydrolysate composition during a fermentation process, samples representing the fermentation process were selected. The fermentation process was divided into three different phases based on the phenotypes, namely lag phase, growth phase and stationary phase, see Figure 3. Based on these phases, samples were selected for analysis:

three at the beginning of each phase, one at the mid-point of growth phase, and one at the end of stationary phase (Figure 3). In our particular case, these five samples represented the whole fermentation process.

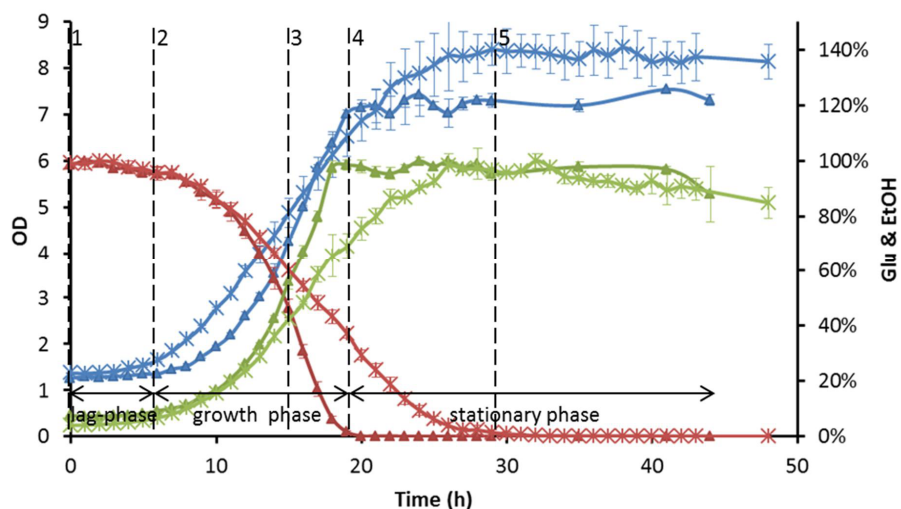


Figure 3 Duplicate fermentation results of the following two hydrolysates: wheat straw-mild alkaline (triangle) and wheat straw-dilute acid (star). Blue: OD, red: glucose percentage, green: ethanol percentage. The three fermentation phases and the five selected sample points are illustrated with wheat straw-mild alkaline fermentation (triangle).

It was decided to analyze the selected samples with two GC-MS methods for their non-sugar composition, as GC-MS is capable of detecting a broad range of compounds, including both knowns and unknowns. As the compounds of interest in this study are potential inhibitors in biomass hydrolysates, it is important to remove sugars from the fermentation samples. This is mainly because sugars were present in large quantity in those samples, which severely interferes with the detection of non-sugar compounds [49,59]. For this purpose, two sample work-up methods were used, namely, ethyl acetate extraction and ethylchloroformate derivatization.

Ethyl acetate extraction GC-MS (EA-GC-MS) was adopted from the method described by Heer *et al.* [55]. In this method, the hydrolysate samples were extracted with ethyl acetate (EA), compounds that are apolar, e.g. with aromatic rings, dissolved in EA, while polar compounds, like sugars, remained in the water phase. In this way, sugars were removed from the hydrolysate samples, and the extracted compounds were concentrated. As the

solubility of different compounds varies in EA, recovery was a main issue in EA-GC-MS method. Therefore, before analyzing hydrolysates, the recovery of EA method was assessed with several reference compounds. This allowed the estimation of compound recovery in fermentation samples.

Ethylchloroformate derivatization GC-MS (EC-GC-MS) was adapted to the use for the analysis of biomass hydrolysate samples [67]. The method converts acids to their ethyl ester form through derivatization with ethylchloroformate and extracts the derivatized sample with dichloromethane (DCM). This allowed the detection of carboxylic acids, amino acids, aromatic compounds and furans. The main issue of EC-GC-MS method was the diverse reactivity of different compounds with ethylchloroformate. This may result in detecting compounds present at high concentration with low recovery. The compounds detected by EC-GC-MS partly overlapped but also complemented the ones measured with EA-GC-MS.

From these analyses, 'compound lists' will be generated for each method. The initial peak characterization will be done by comparing the mass spectra of these peaks with the mass spectral reference libraries available in our lab.

Data analysis and interpretation

The statistical question of our study is, to determine which variables contribute the most to the selected phenotypes. To answer this question, the two data-sets generated in the previous step will be analyzed by building statistical models. The model to be used is partial least square (PLS), which provides compounds that most closely relate to the four defined phenotypes. To conduct model building, the following aspects need to be carefully studied (1) data preprocessing, (2) model input, and (3) model validation method.

Based on the property of the acquired data-sets, square-root transformation and autoscaling will be conducted to preprocess the data. These two methods are to reduce the heteroscedasticity and to amplify the variation in the data-sets, respectively [38].

To model lag-phase, a data-set containing the first two time-point samples (Figure 3) can be used as model input. This is because lag-phase ends at the second sampling point, and it is assumed that after growth starts, the hydrolysate composition has no influence on lag-phase anymore. To model the other three phenotypes, all five time-point samples are to be used, since the influence of any of the five sampling points on these phenotypes cannot be excluded.

One way to validate the models is to check their ability of predicting the phenotypes of a new data-set. A so-called double cross validation method is preferred to validate the PLS models in this study, as it evaluates the model quality in a more strict manner [28,88]. The modeling results will provide, for each phenotype, a set of compounds that contribute the most to that specific phenotype. The next step will then be to evaluate if these compounds are actually inhibitory to the fermenting microbe. The experimental evaluation of the toxicity of known compounds can be relatively simple. However, to evaluate the toxicity of 'unknown' compounds, further compound identification is required.

Conclusions

This review illustrated the application of exometabolomics approaches, both targeted and non-targeted, in studying lignocellulosic biomass hydrolysates as fermentation media. Through analyzing the composition of hydrolysates, targeted exometabolomics has been applied to identify inhibitory compounds, improve hydrolysate preparation method, and monitor compound dynamics during detoxification and fermentation process. To further reveal the overall non-sugar composition of various hydrolysates and identify fermentation inhibitors in an unbiased manner, a non-targeted approach was introduced. Its application was demonstrated in our research to identify inhibitors in biomass hydrolysates relevant for ethanolic fermentation of *S. cerevisiae*, emphasizing the essential role of experimental design, phenotype definition, selection of both analytical methods and statistical models in the non-targeted metabolomics approach.

Acknowledgment

This project was (co) financed by the Netherlands Metabolomics Centre (NMC) which is part of the Netherlands Genomics Initiative / Netherlands Organization for Scientific Research.

References

1. Sun Y, Cheng J: Hydrolysis of lignocellulosic materials for ethanol production: A review. *Bioresource Technol.* **2002**, 83, 1–11.
2. Zhang YH: Reviving the carbohydrate economy via multi-product lignocellulose biorefineries. *J. Ind. Microbiol. Biot.* **2008**, 35, 367–375.
3. Hahn-Hägerdal B, Galbe M, Gorwa-Grauslund MF, Lidén G, Zacchi G: Bio-ethanol – the fuel of tomorrow from the residues of today. *Trends Biotechnol.* **2006**, 24, 549–556.
4. Metzger JO, Hüttermann A: Sustainable global energy supply based on lignocellulosic biomass from afforestation of degraded areas. *Naturwissenschaften* **2009**, 96, 279–288.
5. Carpita NC, Gibeaut DM: Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J.* **1993**, 3, 1–30.
6. Sarkar P, Bosneaga E, Auer M: Plant cell walls throughout evolution: Towards a molecular understanding of their design principles. *J. Exp. Bot.* **2009**, 60, 3615–3635.
7. Mosier N, Wyman C, Dale B, Elander R, Lee YY, Holtzapple M, Ladisch M: Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technol.* **2005**, 96, 673–686.
8. Agbor VB, Cicek N, Sparling R, Berlin A, Levin DB: Biomass pretreatment: Fundamentals toward application. *Biotechnol. Adv.* **2011**, 29, 675–685.
9. Hendriks ATWM, Zeeman G: Pretreatments to enhance the digestibility of lignocellulosic biomass. *Bioresource Technol.* **2008**, 100, 10–18.
10. Alvira P, Tomás-Pejó E, Ballesteros M, Negro MJ: Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review. *Bioresource Technol.* **2010**, 101, 4851–4861.
11. Zha Y, Slomp R, Groenestijn J, Punt PJ: Preparation and Evaluation of Lignocellulosic Biomass Hydrolysates for Growth by Ethanologenic Yeasts. In *Microbial Metabolic Engineering: Methods and Protocols*, Qiong C Ed., Humana Press: New York, USA, 2012, pp. 245–259.
12. Palmqvist E, Hahn-Hägerdal B: Fermentation of lignocellulosic hydrolysates. II: Inhibitors and mechanisms of inhibition. *Bioresource Technol.* **2000**, 74, 25–33.
13. Klinke HB, Thomsen AB, Ahring BK: Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass. *Appl. Microbiol. Biot.* **2004**, 66, 10–26.
14. Fiehn O: Metabolomics—the link between genotypes and phenotypes. *Plant Mol. Biol.* **2002**, 48, 155–171.
15. Roessner U, Bowne J: What is metabolomics all about? *Biotechniques* **2009**, 46, 363–365.
16. van der Werf MJ, Overkamp KM, Muijlwijk B, Coulter L, Hankemeier T: Microbial metabolomics: Toward a platform with full metabolome coverage. *Anal. Biochem.* **2007**, 370, 17–25.
17. Fukusaki E, Kobayashi A: Plant metabolomics: potential for practical operation. *J. Biosci. Bioeng.* **2005**, 100, 347–354.
18. Nicolaou SA, Gaida SM, Papoutsakis ET: A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: From biofuels and chemicals, to biocatalysis and bioremediation. *Metab. Eng.* **2010**, 12, 307–331.
19. Braaksma M, Bijlsma S, Coulter L, Punt PJ, van der Werf MJ: Metabolomics as a tool for target identification in strain improvement: the influence of phenotype definition. *Microbiology* **2011**, 157, 147–159.

20. Ding M, Tian H, Cheng J, Yuan Y: Inoculum size-dependent interactive regulation of metabolism and stress response of *Saccharomyces cerevisiae* revealed by comparative metabolomics. *J. Biotechnol.* **2009**, 144, 279–286.
21. van Ravenzwaay B, Cunha GC, Leibold E, Looser R, Mellert W, Prokoudine A, Walk T, Wiemer J: The use of metabolomics for the discovery of new biomarkers of effect. *Toxicol. Lett.* **2007**, 172, 21–28.
22. Gu H, Pan Z, Xi B, Asiago V, Musselman B, Raftery D: Principal component directed partial least squares analysis for combining nuclear magnetic resonance and mass spectrometry data in metabolomics: Application to the detection of breast cancer. *Anal. Chim. Acta* **2011**, 686, 57–63.
23. Dunn WB, Ellis DI: Metabolomics: Current analytical platforms and methodologies. *TrAC Trends Anal. Chem.* **2005**, 24, 285–294.
24. Mashego M, Rumbold K, de Mey M, Vandamme E, Soetaert W, Heijnen J: Microbial metabolomics: past, present and future methodologies. *Biotechnol. Lett.* **2007**, 29, 1–16.
25. van der Werf MJ, Jellema RH, Hankemeier T: Microbial metabolomics: replacing trial-and-error by the unbiased selection and ranking of targets. *J. Ind. Microbiol. Biot.* **2005**, 32, 234–252.
26. Villas-Bôas SG, Noel S, Lane GA, Attwood G, Cookson A: Extracellular metabolomics: A metabolic footprinting approach to assess fiber degradation in complex media. *Anal. Biochem.* **2006**, 349, 297–305.
27. Bino RJ, Hall RD, Fiehn O, Kopka J, Saito K, Draper J, Nikolau BJ, Mendes P, Roessner-Tunali U, Beale MH, Trethewey RN, Lange BM, Wurtele ES, Sumner LW: Potential of metabolomics as a functional genomics tool. *Trends Plant Sci.* **2004**, 9, 418–425.
28. Westerhuis J, Hoefsloot H, Smit S, Vis D, Smilde A, van Velzen E, van Duijnoven J, van Dorsten F: Assessment of PLS-DA cross validation. *Metabolomics* **2008**, 4, 81–89.
29. Rubingh C, Bijlsma S, Derks E, Bobeldijk I, Verheij E, Kochhar S, Smilde A: Assessing the performance of statistical validation tools for megavariable metabolomics data. *Metabolomics* **2006**, 2, 53–61.
30. Verouden MPH, Westerhuis JA, van der Werf MJ, Smilde AK: Exploring the analysis of structured metabolomics data. *Chemometr. Intell. Lab.* **2009**, 98, 88–96.
31. Gibon Y, Rolin D: Aspects of Experimental Design for Plant Metabolomics Experiments and Guidelines for Growth of Plant Material. In *Plant Metabolomics: Methods and Protocols*, Hardy NW, Hall RD, Eds., Humana Press: New York, USA, 2012, pp. 13–30.
32. Kristal B, Shurubor Y, Paolucci U, Matson W: Methodological Issues and Experimental Design Considerations in Metabolic Profile-Based Classifications. In *Metabolome Analyses: Strategies for Systems Biology*, Vaidyanathan S, Harrigan GG, Goodacre R, Eds., Springer: New York, USA, 2005, pp. 173–194.
33. Qualley AV, Dudareva N: Metabolomics of Plant Volatiles. In *Plant Systems Biology*, Belostotsky DA, Ed., Humana Press: New York, USA, 2009, pp. 329–343.
34. Schauer N, Fernie AR: Plant metabolomics: towards biological function and mechanism. *Trends Plant Sci.* **2006**, 11, 508–516.
35. Lee YC: Carbohydrate analyses with high-performance anion-exchange chromatography. *J. Chromatogr. A* **1996**, 720, 137–149.
36. Bowman MJ, Dien BS, Hector RE, Sarath G, Cotta MA: Liquid chromatography–mass spectrometry investigation of enzyme-resistant xylooligosaccharide structures of switchgrass associated with ammonia pretreatment, enzymatic saccharification, and fermentation. *Bioresour. Technol.* **2012**, 110, 437–447.
37. Roessner U, Nahid A, Chapman B, Hunter A, Bellgard M: Metabolomics – The Combination of Analytical Biochemistry, Biology, and Informatics. In *Comprehensive Biotechnology*, 2nd Edition, Moo-Young M, Ed., Academic Press: Burlington, Canada, 2011, pp. 447–459.

38. van den Berg R, Hoefsloot HCJ, Westerhuis J, Smilde A, van der Werf MJ: Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* **2006**, 7, 142.
39. Daffertshofer A, Lamothe CJC, Meijer OG, Beek PJ: PCA in studying coordination and variability: a tutorial. *Clin. Biomech.* **2004**, 19, 415–428.
40. Krishnan A, Williams LJ, McIntosh AR, Abdi H: Partial Least Squares (PLS) methods for neuroimaging: A tutorial and review. *Neuroimage* **2011**, 56, 455–475.
41. Lutz U, Lutz RW, Lutz WK: Metabolic profiling of glucuronides in human urine by LCMS/MS and partial least-squares discriminant analysis for classification and prediction of gender. *Anal. Chem.* **2006**, 78, 4564–4571.
42. Thissen U, Coulter L, Overkamp KM, Jetten J, van der Werf BJC, van de Ven T, van der Werf MJ: A proper metabolomics strategy supports efficient food quality improvement: A case study on tomato sensory properties. *Food Qual. Prefer.* **2011**, 22, 499–506.
43. Gruben BS: General introduction. *PhD thesis: Novel transcriptional activators of Aspergillus involved in plant biomass utilization* **2012**, 3–49.
44. Taherzadeh MJ, Karimi K: Fermentation Inhibitors in Ethanol Processes and Different Strategies to Reduce Their Effects. In *Biofuels: Alternative Feedstocks and Conversion Processes*, Pandey A, Larroche C, Ricke SC, Dussap C-G, Gnansounou E, Eds., Academic Press, Waltham, Massachusetts, USA, 2011, pp. 287–311.
45. Almeida JRM, Modig T, Petersson A, Hähn-Hägerdal B, Lidén G, Gorwa-Grauslund MF: Increased tolerance and conversion of inhibitors in lignocellulosic hydrolysates by *Saccharomyces cerevisiae*. *J. Chem. Technol. Biot.* **2007**, 82, 340–349.
46. Ando S, Arai I, Kiyoto K, Hanai S: Identification of aromatic monomers in steam-exploded poplar and their influences on ethanol fermentation by *Saccharomyces cerevisiae*. *J. Ferment. Technol.* **1986**, 64, 567–570.
47. Nichols NN, Sharma LN, Mowery RA, Chambliss CK, van Walsum GP, Dien BS, Iten LB: Fungal metabolism of fermentation inhibitors present in corn stover dilute acid hydrolysate. *Enzyme Microb. Tech.* **2008**, 42, 624–630.
48. Raj A, Krishna Reddy MM, Chandra R: Identification of low molecular weight aromatic compounds by gas chromatography–mass spectrometry (GC–MS) from kraft lignin degradation by three *Bacillus* sp. *Int. Biodeter. Biodegr.* **2007**, 59, 292–296.
49. Koo BW, Park N, Jeong HS, Choi JW, Yeo H, Choi IG: Characterization of by-products from organosolv pretreatments of yellow poplar wood (*Liriodendron tulipifera*) in the presence of acid and alkali catalysts. *J. Ind. Eng. Chem.* **2011**, 17, 18–24.
50. Larsson S, Quintana-Sáinz A, Reimann A, Nilvebrant NO, Jönsson LJ: Influence of lignocellulose-derived aromatic compounds on oxygen-limited growth and ethanolic fermentation by *Saccharomyces cerevisiae*. *Appl. Biochem. Biotech.* **2000**, 84–86, 617–632.
51. Ranatunga TD, Jervis J, Helm RF, McMillan JD, Hatzis C: Identification of inhibitory components toxic toward *Zymomonas mobilis* CP4(pZB5) xylose fermentation. *Appl. Biochem. Biotech.* **1997**, 67, 185–198.
52. Chen SF, Mowery RA, Castleberry VA, van Walsum WG, Chambliss CK: High-performance liquid chromatography method for simultaneous determination of aliphatic acid, aromatic acid and neutral degradation products in biomass pretreatment hydrolysates. *J. Chromatogr. A* **2006**, 1104, 54–61.
53. Vaher M, Helmja K, Käsper A, Kurašin M, Väljamäe P, Kudrjašova M, Koel M, Kaljurand M: Capillary electrophoretic monitoring of hydrothermal pre-treatment and enzymatic hydrolysis of willow: Comparison with HPLC and NMR. *Catal. Today* **2012**, 196, 34–41.
54. Chundawat SPS, Vismeh R, Sharma LN, Humpala JF, da Costa Sousa L, Chambliss CK, Jones AD, Balan V, Dale BE: Multifaceted characterization of cell wall decomposition products formed during ammonia fiber expansion (AFEX) and dilute acid based pretreatments. *Bioresour. Technol.* **2010**, 101, 8429–8438.

55. Heer D, Sauer U: Identification of furfural as a key toxin in lignocellulosic hydrolysates and evolution of a tolerant yeast strain. *Microbial Biotechnology* **2008**, 1, 497–506.
56. Kolb M, Sieber V, Amann M, Faulstich M, Schieder D: Removal of monomer delignification products by laccase from *Trametes versicolor*. *Bioresource Technol.* **2012**, 104, 298–304.
57. Jönsson LJ, Palmqvist E, Nilvebrant N-O, Hahn-Hägerdal B: Detoxification of wood hydrolysates with laccase and peroxidase from the white-rot fungus *Trametes versicolor*. *Appl. Microbiol. Biotechnol.* **1998**, 49, 691–697.
58. Huang H, Guo X, Li D, Liu M, Wu J, Ren H: Identification of crucial yeast inhibitors in bio-ethanol and improvement of fermentation at high pH and high total solids. *Bioresource Technol.* **2011**, 102, 7486–7493.
59. Klinke HB, Olsson L, Thomsen AB, Ahring BK: Potential inhibitors from wet oxidation of wheat straw and their effect on ethanol production of *Saccharomyces cerevisiae*: Wet oxidation and fermentation by yeast. *Biotechnol. Bioeng.* **2003**, 81, 738–747.
60. Larsson S, Reimann A, Nilvebrant N-O, Jönsson LJ: Comparison of different methods for the detoxification of lignocellulose hydrolyzates of spruce. *Appl. Biochem. Biotech.* **1999**, 77, 91–103.
61. Klinke HB, Ahring BK, Schmidt AS, Thomsen AB: Characterization of degradation products from alkaline wet oxidation of wheat straw. *Bioresource Technol.* **2002**, 82, 15–26.
62. Panagiotopoulos IA, Bakker RR, de Vrije T, Koukios EG: Effect of pretreatment severity on the conversion of barley straw to fermentable substrates and the release of inhibitory compounds. *Bioresource Technol.* **2011**, 102, 11204–11211.
63. Larsson S, Palmqvist E, Hahn-Hägerdal B, Tengborg C, Stenberg K, Zacchi G, Nilvebrant N-O: The generation of fermentation inhibitors during dilute acid hydrolysis of softwood. *Enzyme Microb. Technol.* **1999**, 24, 151–159.
64. Liu ZL, Slininger PJ, Dien BS, Berhow MA, Kurtzman CP, Gorsich SW: Adaptive response of yeasts to furfural and 5-hydroxymethylfurfural and new chemical evidence for HMF conversion to 2,5-bis-hydroxymethylfuran. *J. Ind. Microbiol. Biot.* **2004**, 31, 345–352.
65. Lu X, Yamauchi K, Phaiboonsilpa N, Saka S: Two-step hydrolysis of Japanese beech as treated by semi-flow hot-compressed water. *J. Wood Sci.* **2009**, 55, 367–375.
66. Humpala JF, Chundawat SPS, Vismeh R, Jones AD, Balan V, Dale BE: Rapid quantification of major reaction products formed during thermochemical pretreatment of lignocellulosic biomass using GC–MS. *J. Chromatogr. B* **2011**, 879, 1018–1022.
67. Zha Y, Muilwijk B, Coulier L, Punt P: Inhibitory Compounds in Lignocellulosic Biomass Hydrolysates during Hydrolysate Fermentation Processes. *J. Bioprocess Biotechnol.* **2012**, 2, doi: 10.4172/2155–9821.1000112.
68. Sharma LN, Becker C, Chambliss KC: Analytical Characterization of Fermentation Inhibitors in Biomass Pretreatment Samples Using Liquid Chromatography, UV-Visible Spectroscopy, and Tandem Mass Spectrometry. In *Biofuels: Methods and Protocols*, Mielenz JR, Ed., Humana Press: New York, USA, 2009, pp. 125–143.
69. Delgenes JP, Moletta R, Navarro JM: Effects of lignocellulose degradation products on ethanol fermentations of glucose and xylose by *Saccharomyces cerevisiae*, *Zymomonas mobilis*, *Pichia stipitis*, and *Candida shehatae*. *Enzyme Microb. Technol.* **1996**, 19, 220–225.
70. Palmqvist E, Hahn-Hägerdal B: Fermentation of lignocellulosic hydrolysates. I: inhibition and detoxification. *Bioresource Technol.* **2000**, 74, 17–24.
71. Aliksson B, Cavka A, Jönsson LJ: Improving the fermentability of enzymatic hydrolysates of lignocellulose through chemical in-situ detoxification with reducing agents. *Bioresource Technol.* **2011**, 102, 1254–1263.
72. Huang X, Wang Y, Liu W, Bao J: Biological removal of inhibitors leads to the improved lipid production in the lipid fermentation of corn stover hydrolysate by *Trichosporon cutaneum*. *Bioresource Technol.* **2011**, 102, 9705–9709.

73. Miyafuji H, Danner H, Neureiter M, Thomasser C, Bvochora J, Szolar O, Braun R: Detoxification of wood hydrolysates with wood charcoal for increasing the fermentability of hydrolysates. *Enzyme Microb. Technol.* **2003**, 32, 396–400.
74. Millati R, Niklasson C, Taherzadeh MJ: Effect of pH, time and temperature of overliming on detoxification of dilute-acid hydrolyzates for fermentation by *Saccharomyces cerevisiae*. *Process Biochem.* **2002**, 38, 515–522.
75. Zhu J, Yong Q, Xu Y, Yu S: Detoxification of corn stover prehydrolyzate by trialkylamine extraction to improve the ethanol production with *Pichia stipitis* CBS 5776. *Bioresource Technol.* **2011**, 102, 1663–1668.
76. Martinez A, Rodriguez ME, York SW, Preston JF, Ingram LO: Effects of Ca(OH)₂ treatments (“overliming”) on the composition and toxicity of bagasse hemicellulose hydrolysates. *Biotechnol. Bioeng.* **2000**, 69, 526–536.
77. Moreno AD, Ibarra D, Fernández JL, Ballesteros M: Different laccase detoxification strategies for ethanol production from lignocellulosic biomass by the thermotolerant yeast *Kluyveromyces marxianus* CECT 10875. *Bioresource Technol.* **2012**, 106, 101–109.
78. Pienkos PT, Zhang M: Role of pretreatment and conditioning processes on toxicity of lignocellulosic biomass hydrolysates. *Cellulose* **2009**, 16, 743–762.
79. Taherzadeh MJ, Gustafsson L, Niklasson C, Lidén G: Conversion of furfural in aerobic and anaerobic batch fermentation of glucose by *Saccharomyces cerevisiae*. *J. Biosci. Bioeng.* **1999**, 87, 169–174.
80. Rumbold K, van Buijsen HJJ, Gray VM, van Groenestijn JW, Overkamp KM, Slomp RS, van der Werf MJ, Punt PJ: Microbial renewable feedstock utilization: A substrate-oriented approach. *Bioengineered Bugs* **2010**, 1, 359–366.
81. Thomsen MH, Thygesen A, Thomsen AB: Identification and characterization of fermentation inhibitors formed during hydrothermal treatment and following SSF of wheat straw. *Appl. Microbiol. Biotechnol.* **2009**, 83, 447–455.
82. Chen M, Zhao J, Xia L: Comparison of four different chemical pretreatments of corn stover for enhancing enzymatic digestibility. *Biomass Bioenergy* **2009**, 33, 1381–1385.
83. Laser M, Schulman D, Allen SG, Lichwa J, Antal MJ, Lynd LR: A comparison of liquid hot water and steam pretreatments of sugar cane bagasse for bioconversion to ethanol. *Bioresource Technol.* **2002**, 81, 33–44.
84. Rabelo SC, Filho RM, Costa AC: A comparison between lime and alkaline hydrogen peroxide pretreatments of sugarcane bagasse for ethanol production. *Appl. Biochem. Biotechnol.* **2008**, 148, 45–58.
85. Talebnia F, Karakashev D, Angelidaki I: Production of bioethanol from wheat straw: An overview on pretreatment, hydrolysis and fermentation. *Bioresource Technol.* **2010**, 101, 4744–4753.
86. Cardona CA, Quintero JA, Paz IC: Production of bioethanol from sugarcane bagasse: Status and perspectives. *Bioresource Technol.* **2010**, 101, 4754–4766.
87. Hashaiekh R, Fang Z, Butler IS, Hawari J, Kozinski JA: Hydrothermal dissolution of willow in hot compressed water as a model for biomass conversion. *Fuel* **2007**, 86, 1614–1622.
88. Smit S, van Breemen MJ, Hoefsloot HCJ, Smilde AK, Aerts JMFG, de Koster CG: Assessing the statistical validity of proteomics based biomarkers. *Anal. Chim. Acta* **2007**, 592, 210–217.

CHAPTER 2

Preparation and Evaluation of Lignocellulosic Biomass Hydrolysates for Growth by Ethanologenic Yeasts

Ying Zha, Ronald Slomp, Johan van Groenestijn, and Peter J. Punt

This chapter is published with minor modification in:
Microbial Metabolic Engineering: Methods and Protocols, 2012, 245-259

Abstract

Lignocellulosic biomass is a potential feedstock for bioethanol production. Biomass hydrolysates, prepared with a procedure including pretreatment and hydrolysis, are considered to be used as fermentation media for microorganisms, such as yeast. During the hydrolysate preparation procedure, toxic compounds are released or formed which may inhibit the growth of the microorganism and thus the product formation. To study the effects of these compounds on fermentation performance, the production of various hydrolysates with diverse inhibitory effects is of importance. A platform of methods that generates hydrolysates through four different ways and tests their inhibitory effects using Bioscreen C Analyzer growth tests is described here. The four methods, based on concentrated acid, dilute acid, mild alkaline and alkaline/oxidative conditions, were used to prepare hydrolysates from six different biomass sources. The resulting 24 hydrolysates showed great diversity on growth rate in Bioscreen C Analyzer growth tests. The approach allows the prediction of a specific hydrolysate's performance and helps to select biomass type and hydrolysate preparation method for a specific production strain, or vice versa.

Introduction

Lignocellulosic biomass, such as sugar cane bagasse, wheat straw and willow wood, are potential feedstocks for bioethanol production. They are inexpensive, abundant and not competitive with world food resources. Generally, lignocellulosic biomass is composed of cellulose, hemicellulose and lignin. To use lignocellulosic biomass as fermentation feed, a pretreatment and a hydrolysis step need to be carried out, which breaks the structure of the biomass and releases the sugar monomers, respectively. The resulting mixture after removing the lignin fraction, the lignocellulosic biomass hydrolysate, is then used as fermentation medium.

One of the drawbacks of the lignocellulosic biomass hydrolysate, when used as feedstock for bioethanol production, is its inhibitory effect on the growth of microorganisms [1,2]. The inhibitory effects are caused by the release of toxic compounds during the hydrolysate preparation process, since conditions like high temperature, high pressure, and acid/alkaline environment are used for the pretreatment. Growth inhibition mainly results in a longer lag-phase, a reduced growth rate and a lower ethanol productivity; and was found in both laboratory and industrial strains [3-5].

To restrict the inhibitory effect of the hydrolysate, identifying inhibitory compounds is a crucial step. In the last several decades, studies on hydrolysate inhibitory compounds have been conducted; the results showed that the potential inhibitors are aliphatic and aromatic acids, aldehydes and ketones from the degradation products of cellulose and hemicellulose, as well as phenolic compounds from lignin [6,7]. Among all the potential inhibitors, formic acid, acetic acid, furfural and hydroxymethylfurfural (HMF) were considered as the representatives, and were used to test the inhibition resistance of different microorganisms [8,9]. Heer *et al.* reported the identification of furfural as a key inhibitor in the wheat and barley straw hydrolysates, using ethylacetate extraction GC-MS [10].

Due to the complexity and the variability of the lignocellulosic biomass hydrolysate composition, and the requirement of advanced tools, such as LC-MS, GC-MS and solid-phase microextraction (SPME), for chemical analysis, the identification of inhibitory compounds has not yet been carried out systematically. In our research, we aim to use a non-targeted metabolomics approach to elucidate the substrate inhibitory compounds. The first requirement of this approach is a high-quality experimental design [11,12]. Such an experimental design should include a variety of hydrolysates that hold diverse inhibitory

effects on fermentation. It is expected that some inhibitory compounds are universal, while the others are more specific to a particular hydrolysate.

In our study, six types of biomass are selected and prepared with four different methods to test the inhibitory effects of the resulting 24 hydrolysates on different microorganisms. The procedures of the four hydrolysate preparation methods, namely, concentrated acid, dilute acid, mild alkaline and alkaline/oxidative method, are described here. In addition, we describe the procedure of the growth test using Bioscreen C Analyzer [13]. The results show that there is a great diversity of inhibitory effects among the 24 hydrolysates, which is contributed by both biomass type and hydrolysate preparation method.

The method provides a platform for testing lignocellulosic biomass hydrolysates on various microorganisms, especially bacteria and yeasts. The results obtained with this analysis platform will help to select biomass type and hydrolysate preparation method for a specific production strain, or vice versa.

Materials

Lignocellulosic biomass

Sugar cane bagasse is a kind gift from Zilor, Brazil. Corn stover is from the University of Cape Town, South Africa. Wheat straw, barley straw and willow wood are purchased from Oostwaardshoeve, the Netherlands. Oak sawdust is from ESCO, the Netherlands, a wood-flooring supplier. All biomass is pre-dried at 80°C for 5 h when received, and stored at room temperature. All biomass, except oak sawdust, is ground to pieces with average length of 3 mm.

Hydrolysate preparation

H₂SO₄, purity 95–97%, stored at room temperature. NaOH, purity ≥ 98%, stored at room temperature. Ca(OH)₂ purity ≥ 96%, stored at room temperature. Acetic acid, purity 100%, stored at room temperature. 30% H₂O₂, stored at 4°C. Penicillin:streptomycin (pen:strep) solution: 10000 unit penicillin and 10 mg streptomycin/ml. Used at 1:100 (v:v) dosage, stored at –20°C. Enzyme cocktail Accellerase 1500 (a gift from GENENCOR): endoglucanase activity 2200–2800 CMC U/g; beta-glucosidase activity 525–775 pNPG U/g.

Rotation is carried out on a two-deck modular roller apparatus (348971-C) from Wheaton Roller. All filtrations, unless otherwise stated, use Grade GF/D filter purchased from

Whatman. Concentration of H_2SO_4 is tested with HACH LANGE SO_4^{2-} testing kit (LCK353). Anion exchange is carried out with a MP-Cell from ElectroCell A/S (Tarm, Denmark), which includes two PVDF fluid distribution frames, four EPDM gaskets and two PVDF end frames. Anion exchange membranes are Neosepta AFN membranes from Eurodia Industrie SA, France. The ElectroCell equipment is assembled with stainless steel end plates and hardware, indicated as “EC” in Figure 1. Liquid in both donor and acceptor compartment are pumped in a circuit over a glass bottle, a pump and the cell and back to the bottle again. The liquid volume of the circuit is 0.45 l. The flow direction over the two cells was counter current at a rate of 0.5 l/min.

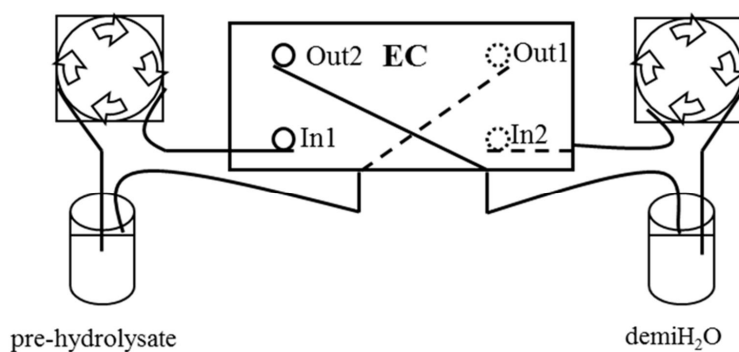


Figure 1 Ion exchange membrane facility for neutralization. EC: ElectroCell.

Yeast strains

Saccharomyces cerevisiae CEN.PK 113-7D is purchased from CBS Yeast Collection (CBS 8340). Ethanol Red yeast is a gift from Fermentis, France. Both strains are stored as 1 ml glycerol stocks at -80°C .

Bioscreen C Analyzer growth study

Bioscreen C Analyzer, LabSystems OY, Helsinki, Finland. In the Bioscreen C Analyzer, two honeycomb plates, each contains 100 wells, can be incubated at a preset temperature, and the growth of the microorganisms can be monitored by measuring optical density (OD 420–580) at a preset time interval, in our case, 15 min. The honeycomb plates can be shaken linearly at a preset speed. The Bioscreen C Analyzer is connected to a computer, which records all the measured optical density values.

Mineral medium (MM) is prepared according to the composition as described by van Hoek *et al.* [14]. Per liter: 5.0 g (NH₄)₂SO₄, 3.0 g KH₂PO₄, 0.5 g MgSO₄·7H₂O, 1 ml 1000x trace elements solution, 1 ml 1000x vitamin solution and 80 ml 250 g/l glucose solution.

Per liter 1000x trace element: 15.0 g EDTA, 4.5 g ZnSO₄·7H₂O, 0.3 g CoCl₂·6H₂O, 1.0 g MnCl₂·4H₂O, 0.3 g CuSO₄·5H₂O, 4.5 g CaCl₂·2H₂O, 3.0 g FeSO₄·7H₂O, 0.44 g Na₂MoO₄·2H₂O, 1.0 g H₃BO₄ and 0.1 g KI.

Per liter 1000x vitamin solution: 0.05 g biotin, 1.0 g calcium pantothenate, 1.0 g nicotinic acid, 25.0 g inositol, 1.0 g thiamine·HCl, 1.0 g pyridoxine·HCl, and 0.2 g *para*-aminobenzoic acid.

250 g/l glucose solution is sterilized by autoclaving at 110°C, 1000x vitamin solution is sterilized by filtering through 0.2 µm filter and 1000x trace element solution and the (NH₄)₂SO₄, KH₂PO₄, MgSO₄·7H₂O solution are sterilized by autoclaving at 121°C.

Methods

To investigate the inhibitory effects of the lignocellulosic biomass hydrolysates using exo-metabolomics, generating a diverse collection of hydrolysates is crucial [11,12]. Biomass type and hydrolysate preparation method are the sources for creating such diversity. In particular, the hydrolysate preparation method is a target of optimization. Therefore, establishing various hydrolysate preparation methods and applying them on different biomass types is of importance to allow a meaningful experimental design for process improvement.

The four hydrolysate preparation methods selected in our study are (1) concentrated acid, (2) dilute acid, (3) mild alkaline, and (4) alkaline/oxidative. Each method composes of a pretreatment step, which breaks down the structure of the biomass, and a hydrolysis step, which hydrolyzes the cellulose and hemicellulose into sugar monomers. Dilute acid, mild alkaline and alkaline/oxidative are relatively mild methods; they employ acidification, alkalization and oxidation for pretreatment, respectively. Dilute acid pretreatment takes place under high pressure for several minutes; alkaline/oxidative pretreatment is carried out at high temperature for many hours; while mild alkaline pretreatment undergoes high temperature for multiple days. All three methods apply enzymatic hydrolysis, which can be conducted in a fed-batch manner to reach high glucose concentration in the hydrolysate. Concentrated acid is more aggressive, it uses concentrated sulfuric acid for both pretreatment and hydrolysis. This method is more suitable for preparing hard wood hydrolysate than the three mild methods.

The detailed procedure of these four hydrolysate preparation methods are described in the following two sections, separating pretreatment and hydrolysis.

Pretreatment

To prepare 1 l lignocellulosic biomass hydrolysate, 500 g dried biomass is used for concentrated acid method, while 300 g is used for the other three methods. The pre-dried biomass contains 5–10% free water, indicating that the initial biomass amount before drying should be above 550 and 330 g. The drying takes place in an oven at 80°C, for at least 16 h.

Concentrated acid

500 g dried biomass is impregnated with 72% H₂SO₄ in a 5 l Duran laboratory glass bottle. The biomass dry matter to H₂SO₄ ratio is 1:2 (w:w), the concentration of 72% H₂SO₄ is 1155 g/l, so the amount of 72% H₂SO₄ added is 866 ml. The biomass is impregnated by rotating the bottle on the two-deck modular roller at room temperature for 24–48 h, rotating speed 10 rpm. The pretreated biomass is stored at 4°C before hydrolysis.

Dilute acid

300 g dried biomass is soaked in 2% (20.28 g/l) H₂SO₄, at a ratio of 8 g dried biomass/100 ml 2% H₂SO₄, in a 5 l polypropylene (PP) bucket, i.e. in total 3750 g 2% H₂SO₄ is used. The biomass is impregnated in an oven at 60°C for minimum 3 h, with the bucket lid closed. The impregnated biomass is cooled down to room temperature, and the free liquid in the wet biomass is partially removed by vacuum filtration. The wet biomass is transferred onto an aluminum dish, which is then autoclaved at 160°C for 3.5 min. After the wet biomass is cooled down to room temperature, its liquid content is determined by measuring the total wet biomass weight. If the liquid content in the biomass exceeds 1000 g, i.e. if the biomass dry matter content is below 23%, the wet biomass is air dried at room temperature (20°C) until the liquid content drops below 1000 g (see Note 1). The wet biomass is stored at 4°C before enzymatic hydrolysis.

Mild alkaline

A Ca(OH)₂ suspension is added to the dried biomass at 0.15 g Ca(OH)₂ /g biomass and at a 20% biomass dry matter content in a 5 l PP bucket: 300 g biomass is mixed with 48 g Ca(OH)₂ suspended in 1500 g demineralized water. The biomass is impregnated by placing the bucket in an oven at 80°C for 3–5 days, with the bucket lid closed. After the wet biomass is cooled down to room temperature, its liquid content is determined by measuring the total wet biomass weight. The wet biomass is air dried at room temperature (20°C) to reduce its

liquid content to below 1000 g, i.e. biomass dry matter content increases to 23% (see Note 1). Store the biomass at 4°C before enzymatic hydrolysis.

Alkaline/Oxidative

This pretreatment method is adapted from a method described by Zhao *et al.* [15]. 300 g dried biomass is mixed with 3.3% NaOH at a ratio of 1:3 (w:v) in a 5 l Duran laboratory glass bottle. The biomass is impregnated by placing the bottle in a water bath at 90°C for 90 min (see Note 2). The impregnated biomass is cooled down to room temperature and washed with demineralized water, using vacuum filtration, to remove the detached and dissolved lignin (see Note 3). Vacuum filtration is continued at the end of the washing to increase the dry matter content of the wet biomass.

Prepare peracetic acid (PAA) by adding acetic acid, 30% H₂O₂ and 96% H₂SO₄ at a ratio of 60:40:1 (v:v:v) together, and allowing to react at room temperature for 72–96 h before use. As soon as the PAA solution is ready for use, add it to the washed biomass at a ratio of 1:1 (w:w) with respect to the initial dried biomass amount, i.e. 300 g, in a 2 l Duran laboratory glass bottle. The mixture is impregnated in an oven at 70°C for 2 h, while the bottle is rotated at a speed of 10 rpm on the two-deck modular roller apparatus. After the impregnation, the biomass is washed with demineralized water to remove PAA. The washing is facilitated by vacuum filtration, and PAA is considered as removed when washing water pH reaches 3.5. The washed biomass is mixed with 3–4 l demineralized water, and the slurry pH is adjusted to 5.0 using 6 M NaOH. The free liquid in the mixture is then partially removed by vacuum filtration. The liquid content of the vacuum filtrated biomass is determined by measuring the total wet biomass weight. If liquid content of the wet biomass exceeds 1000 g, i.e. if the biomass dry matter content is below 23%, the wet biomass is air dried at room temperature (20°C) till the liquid content drops below 1000 g (see Note 1). Store the biomass at 4°C before enzymatic hydrolysis.

Hydrolysis

Two different hydrolysis methods are applied to the pretreated biomass, namely acid hydrolysis and enzymatic hydrolysis. The biomass pretreated with concentrated acid is hydrolyzed with acid, while the biomass pretreated with either dilute acid, mild alkaline, or peracetic acid method is hydrolyzed enzymatically.

Acid hydrolysis

At the end of the impregnation, 900 ml demineralized water is added to the impregnated biomass to reduce SO_4^{2-} concentration to around 42% (570 g/l). The slurry is then pre-hydrolyzed in the bottle in an incubator at 60°C for 8 h, with shaking speed 100 rpm (see Note 2). The pre-hydrolyzed slurry is cooled down to room temperature and vacuum filtrated to separate the liquid and the solid phase. The filter cake, which has around 50% biomass dry matter content, is collected and stored at 4°C, while the filtrate, the pre-hydrolysate, is neutralized by using anion exchange membranes to reduce SO_4^{2-} concentration to a required level (see Note 4). The anion exchange process is facilitated with ElectroCell (EC) equipment and two pumps, as shown in Figure 1. The donor compartment is filled with pre-hydrolysate, while the acceptor compartment contains 5 l demineralized water. In about 5 days, the SO_4^{2-} concentration of the pre-hydrolysate is reduced to the required level (see Note 4); during which the acceptor, demineralized water, is refreshed daily, and the SO_4^{2-} concentration is checked daily with HACH LANGE LCK 353 kit (see Note 5).

The neutralized pre-hydrolysate is then combined with the filter cake from the pre-hydrolysis in a 2 l Duran laboratory glass bottle. The slurry is hydrolyzed in a second round by placing the bottle in a water-bath at 95°C for 3 h (see Note 2). After the hydrolysis, the slurry is cooled down to room temperature and vacuum filtrated. The filter cake, mainly composed of lignin, is discarded. The filtrate, the hydrolysate, is again neutralized with the anion exchange membranes facilitated with EC equipment, as described before. In about 5 days, the SO_4^{2-} concentration is reduced to below 15 g/l (1.5%); during which the acceptor, demineralized water, is refreshed daily, and the SO_4^{2-} concentration is checked every other day (see Note 5). After anion exchange neutralization, the hydrolysate pH is adjusted to 5.0 using $\text{Ca}(\text{OH})_2$ (see Note 6). The hydrolysate is sterilized by using 0.2 μm filters (see Note 7) and stored at -20°C before use.

Enzymatic hydrolysis

The free liquid content of the wet biomass after pretreatment, W_{FL} , is determined by measuring the total wet biomass weight (see (1)).

$$W_{\text{FL}} (\text{g}) = W_{\text{WB}} - W_{\text{DB}} \quad (1)$$

Where, W_{WB} is the total wet biomass weight (g)

W_{DB} is the initial dried biomass weight (300 g)

$W_{\text{H}_2\text{O}}$ amount of demineralized water is added to a 2 l Duran laboratorial glass bottle (see

Note 8 and (2)).

$$W_{H_2O} (g) = 1200 g - W_{FL} \quad (2)$$

Wet biomass is added into the same bottle until a mixable slurry is formed (see Note 9). Record the amount of wet biomass added as W_{PWB} and calculate the added amount of initial dried biomass, W_{PDB} , using (3).

$$W_{PDB} (g) = \frac{W_{PWB}}{W_{WB}} \times 300 g \quad (3)$$

Enzyme Accellerase 1500 and pen:strep are then added into the slurry according to the mixing ratio, 0.5 ml Accellerase 1500/g initial dried biomass, and 1 ml pen:strep/100 ml slurry.

Estimate the H_2SO_4 or the $Ca(OH)_2$ content, $W_{PH_2SO_4}$ or $W_{PCa(OH)_2}$, in the slurry if dilute acid or mild alkaline pretreated biomass is used, with (4a) and (5a). Pre-neutralize the slurry by adding either $Ca(OH)_2$ powder or pure H_2SO_4 , $W_{ACa(OH)_2}$ or $W_{AH_2SO_4}$, according to (4b) and (5b).

$$W_{PH_2SO_4} (g) = (W_{PWB} - W_{PDB}) \times 0.0208 g/g \quad (4a)$$

$$W_{ACa(OH)_2} (g) = 95\% \times \frac{W_{PH_2SO_4}}{98 g/mol} \times 74 g/mol \quad (4b)$$

$$W_{PCa(OH)_2} (g) = W_{PDB} \times 0.15 g/g \quad (5a)$$

$$W_{AH_2SO_4} (g) = 105\% \times \frac{W_{PCa(OH)_2}}{74 g/mol} \times 98 g/mol \quad (5b)$$

Where, 0.0208 g/g is the concentration of 2% H_2SO_4

0.15 g/g is the adding ratio of $Ca(OH)_2$ in mild alkaline pretreatment

98 g/mol is the molar mass of H_2SO_4

74 g/mol is the molar mass of $Ca(OH)_2$

95 and 105% are rough compensation factors for reaching pH 5.0 instead of pH 7.0

The hydrolysis is conducted in an incubator at 50°C with shaking speed 150 rpm. After about 1 h, the slurry pH is measured with pH electrode and re-adjusted to 5.0 using either 6 M NaOH or 3 M H_2SO_4 . The hydrolysis is then continued under the same condition till the slurry is liquefied (see Note 10). The pH of the liquefied slurry is adjusted with either 6 M NaOH or 3 M H_2SO_4 to 5.0, before the next round of wet biomass addition. Step 3–7 are repeated until all the wet biomass is used, the hydrolysis slurry is then incubated at 50°C with shaking speed 150 rpm for another 48–72 h. After the enzymatic hydrolysis, the slurry is cooled down

to room temperature and vacuum filtrated or centrifuged to separate the hydrolysate from the remaining solids. The hydrolysate is stored at -20°C before use.

Bioscreen C Analyzer evaluation

The inhibitory effects of the generated hydrolysates are examined with growth tests, measuring growth rates. The growth tests were conducted in 100-well honeycomb plates in a Bioscreen C Analyzer [13], which monitors the growth curves by measuring optical density in a preset time interval.

Preculture preparation

The preculture of the growth test is conducted in a 500 ml shake flask with 100 ml MM. The preculture is inoculated with 1 ml glycerol stock and cultivated at 30°C for about 24 h with shaking speed 200 rpm. After the incubation, the pre-culture optical density (OD₆₀₀) is measured, and if it reaches approximately 3.5, 5 ml pre-culture is centrifuged at 15000 rpm, 4°C , for 20 min to separate the cells from the medium. After the centrifugation, the medium is discarded and the cells are resuspended in sterilized demineralized water to reach OD₆₀₀ 1.5–2.0, which is used as inoculum for Bioscreen C incubation.

Bioscreen C incubation

The growth test is conducted in triplicate in the 100-well honeycomb plates. The Bioscreen C Analyzer can cultivate two plates at the same time. To minimize medium evaporation, the wells that are on the outside border are filled with 400 μl demineralized water each, leaving 72 wells to be used for growth tests on one plate. Before filling the wells, Bioscreen C Analyzer is set at a temperature of 30°C , a detecting wavelength of 420–580 nm, a measuring interval of 15 min and no shaking (see Note 11) [13]. The experimental duration is set long enough to allow the growth in all wells to complete, e.g. 5 days. Fill each of the 72 wells with either 400 μl medium or hydrolysate according to the experimental design, and inoculate with 4 μl prepared inoculum. Place the inoculated plate(s) in the Bioscreen C Analyzer and start the experimental run. When all growth curves are complete, stop the experimental run and collect the measured optical density data from the computer. Calculate the $\ln(\text{OD}_{\text{measure}} - \text{OD}_{\text{blank}})$ value for each measurement point and plot these values against time (hours). Determine the slope of the linear part of the plot and use it as measure for growth rate.

Growth test results of 24 hydrolysates prepared from six biomass types

The hydrolysate preparation methods were applied to 6 different biomass, namely, sugar cane bagasse, corn stover, wheat straw, barley straw, oak sawdust and willow wood chips. The resulting 24 hydrolysates were used as media of the Bioscreen C Analyzer incubation to test their inhibitory effects on two model strains, *S. cerevisiae* CEN.PK 113-7D and Ethanol Red Yeast. The results are presented as growth rates in Figure 2 and 3. The reproducibility of the Bioscreen C Analyzer growth tests is high. The experiments were carried out in triplicate, of which the standard deviations of calculated growth rates were between 0.5 and 15%, with an average of 2.6%. An example of the Bioscreen C growth test results is shown in Figure 4. In general it can be seen that the optical density readings are stable for all growth curves. Occasional unstable readings are discarded for further growth rate calculations. The results show that the growth rate among different hydrolysates was diverse. This diversity was caused by both biomass type and hydrolysate preparation method, though the latter seems to be the major influencer, indicating that the forming or/and releasing of the inhibitory compounds tightly relates to the pretreatment method. As expected, the growth rates in all hydrolysates were lower than those in MM.

It can be seen from Figure 2 and 3 that both yeasts performed best in mild alkaline prepared hydrolysates, independent of the biomass type, suggesting that least inhibitory compounds were released or formed when mild alkaline was used as pretreatment method. However, the hydrolysis efficiency of the alkaline pretreated biomass was lower compare to the other 3 methods, especially when wood is used (results not shown). This makes mild alkaline a less attractive pretreatment method on industrial scale. On the contrary, the hydrolysates prepared with concentrated acid method had much stronger inhibitory effect on both yeasts. Particularly, when corn stover or oak sawdust was used, little growth was observed. Concentrated acid was a much severer method, which resulted in hydrolysates with high glucose concentration (80–95 g/l) but also strong inhibitory effects. This makes this method, because of its high efficiency, interesting for industrial use, valuable for studying the effects of inhibitory compounds in the hydrolysate.

As far as biomass type is concerned, the highest average growth rate among all four hydrolysate preparation methods was in wheat straw hydrolysates, considering concentrated acid and dilute acid methods particularly. The performance of wheat straw and barley straw hydrolysates were quite similar (see Figure 2 and 3). This observation is different from the result published by Almeida *et al.* [1], which indicated that barley straw hydrolysate is much more inhibitory than wheat straw hydrolysate. Bagasse and oak hydrolysates had the lowest average growth rate, especially in dilute acid prepared hydrolysates (see Figure 2 and 3).

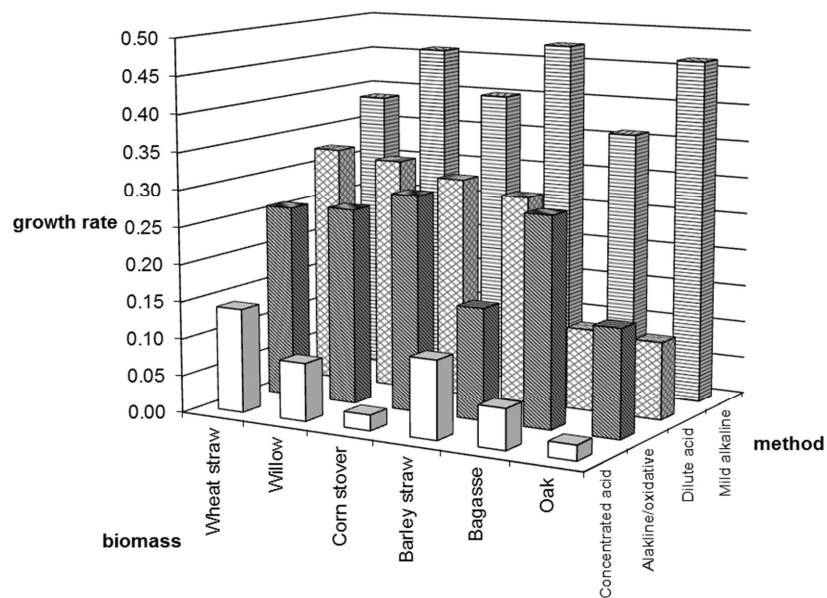


Figure 2 Growth rates of *S. cerevisiae* CEN.PK 113-7D in 24 different hydrolysates.

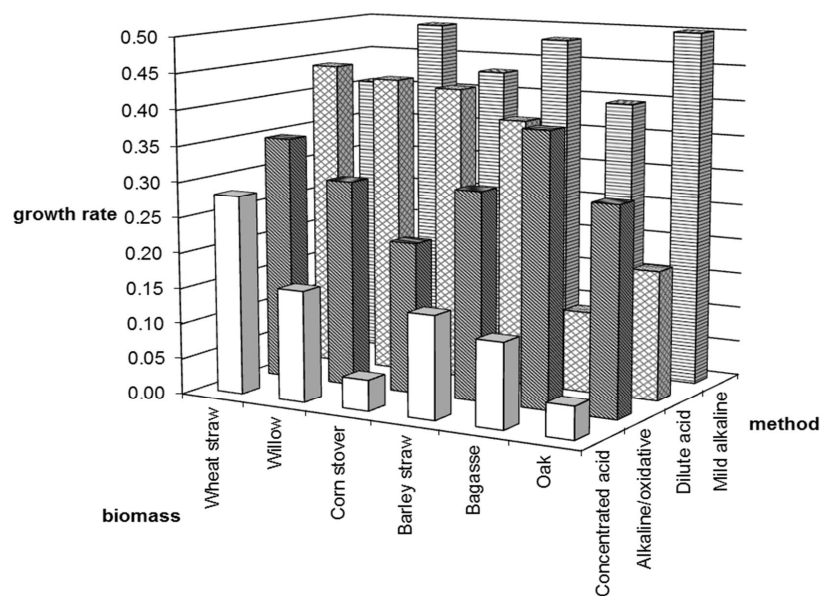


Figure 3 Growth rates of Ethanol Red Yeast in 24 different hydrolysates.

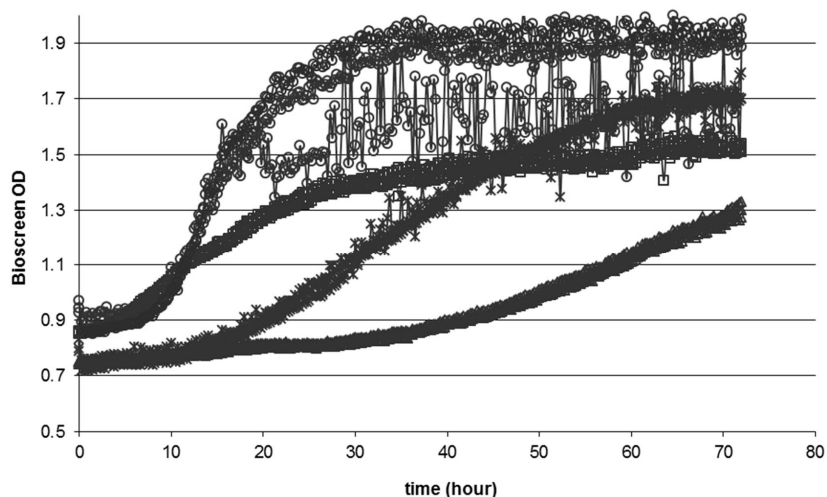


Figure 4 An example of Bioscreen C growth test results: the growth of *S. cerevisiae* CEN.PK 113-7D in hydrolysates prepared from bagasse. Circle, mild alkaline hydrolysate; star, dilute acid hydrolysate; square, alkaline/oxidative hydrolysate; triangle, concentrated acid hydrolysate.

The two microorganisms used in this experiment, *S. cerevisiae* CEN.PK 113-7D and Ethanol Red yeast, had similar performance in different hydrolysates, indicated by the comparable pattern of Figure 2 and 3. This may due to the close phenotype of the two microorganisms, both are ethanol production yeast strains. An overall comparison of the two strains shows that the average growth rate of Ethanol Red yeast is higher than CEN.PK, suggesting that the industrial strain is more resistant to the inhibitory compounds than the laboratorial one.

In conclusion, the method described here provides a platform for, on one hand, a quick and simple analysis of hydrolysate performance, which helps select the biomass type and the hydrolysate preparation method when a specific microorganism is used as production host. On the other hand, for screening of different microorganism strains in a particular hydrolysate to test their resistance on inhibitory compounds as well as other phenotypes, such as optimum growth pH and temperature.

Notes

1. To ensure the starting of enzymatic hydrolysis, a certain amount of water needs to be added to the pretreated biomass to form a mixable slurry. Limiting the free liquid content in the wet biomass to below 1000 ml allows the addition of minimum 200 ml demineralized water at the beginning of the hydrolysis.
2. To avoid pressure difference, don't completely close the bottle lid.
3. The washing process may cost 20–30 l demineralized water, depending on the biomass type. The detached and dissolved lignin is considered removed when the color of the washing water turns from dark black to light yellow.
4. The SO_4^{2-} concentration in the pre-hydrolysate should decrease to a specific level so that when it is mixed with the filter cake, the slurry has a SO_4^{2-} concentration of 150–200 g/l. This concentration range is optimum for the second round hydrolysis, which assists the releasing of the remaining sugars in the filter cake, while causes little sugar degradation in the hydrolysate.
5. SO_4^{2-} concentration testing kit, HACH LANGE LCK353, has a testing range of 150–900 mg/l. Therefore, the samples need to be diluted 100–1000 times before testing.
6. To avoid dilution, $\text{Ca}(\text{OH})_2$ is added as dry powder into the hydrolysate, so pH is only measured after complete $\text{Ca}(\text{OH})_2$ dissolution.
7. The filter sterilization is only suitable for small scale experiments. For industrial scale, contaminations may be eliminated by using antimicrobial compounds, such as antibiotics or antimicrobial plant (hop) extracts.
8. To make 1 l hydrolysate, there should be between 1300 and 1400 ml free liquid in the enzymatic hydrolysis slurry. Since, per 300 g dried biomass, in total 150 ml enzyme Accellerase 1500 will be added, 1200g - W_{FL} demineralized water is added here to make up the volume.
9. The slurry should be mixable to allow pH adjustment and to ensure the starting of enzymatic hydrolysis. For most biomass types, if 200 ml demineralized water was added firstly, the wet biomass adding amount is around 200 g.
10. It takes about 4 h to liquefy the slurry, which allows the proceeding to the next hydrolysis step.
11. Choosing no shaking is to avoid uneven cell distribution and to obtain steady optical density reading during the whole growth stage [13].

Acknowledgment

This project was co-financed by the Netherlands Metabolomics Centre (NMC) which is part of the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research.

References

1. Almeida JRM, Karhumaa K, Bengtsson O, Gorwa-Grauslund MF: Screening of *Saccharomyces cerevisiae* strains with respect to anaerobic growth in non-detoxified lignocellulose hydrolysate. *Bioresource Technol.* **2009**, 100, 3674–3677.
2. Modig T, Almeida JRM, Gorwa-Grauslund MF, Lidén G: Variability of the response of *Saccharomyces cerevisiae* strains to lignocellulose hydrolysate. *Biotechnol. Bioeng.* **2008**, 100, 423–429.
3. Tofighi A, Azin M, Mazaheri Assadi M, Assadirad MHA, Nejdassattari T, Fallahian MR: Inhibitory effect of high concentrations of furfural on industrial strain of *Saccharomyces cerevisiae*. *Int. J. Environ. Res.* **2010**, 4, 137–142.
4. Albers E, Larsson C: A comparison of stress tolerance in YPD and industrial lignocellulose-based medium among industrial and laboratory yeast strains. *J. Ind. Microbiol. Biot.* **2009**, 36, 1085–1091.
5. Martín C, Jönsson LJ: Comparison of the resistance of industrial and laboratory strains of *Saccharomyces* and *Zygosaccharomyces* to lignocellulose-derived fermentation inhibitors. *Enzyme Microb. Tech.* **2003**, 32, 386–395.
6. Klinker HB, Thomsen AB, Ahring BK: Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass. *Appl. Microbiol. Biot.* **2004**, 66, 10–26.
7. Palmqvist E, Hahn-Hägerdal B: Fermentation of lignocellulosic hydrolysates. I: inhibition and detoxification. *Bioresource Technol.* **2000**, 74, 17–24.
8. Almeida JRM, Modig T, Petersson A, Hahn-Hägerdal B, Lidén G, Gorwa-Grauslund MF: Increased tolerance and conversion of inhibitors in lignocellulosic hydrolysates by *Saccharomyces cerevisiae*. *J. Chem. Technol. Biot.* **2007**, 82, 340–349.
9. Nichols NN, Sharma LN, Mowery RA, Chambliss CK, van Walsum GP, Dien BS, Iten LB: Fungal metabolism of fermentation inhibitors present in corn stover dilute acid hydrolysate. *Enzyme Microb. Tech.* **2008**, 42, 624–630.
10. Heer D, Sauer U: Identification of furfural as a key toxin in lignocellulosic hydrolysates and evolution of a tolerant yeast strain. *Microb. Biotechnol.* **2008**, 1, 497–506.
11. Xia J, Psychogios N, Young N, Wishart DS: MetaboAnalyst: A web server for metabolomic data analysis and interpretation. *Nucleic Acids Res.* **2009**, 37, W652–W660.
12. Braaksma M, van den Berg RA, van der Werf MJ, Punt PJ: A Top-down Systems Biology Approach for The Identification of Targets for Fungal Strain and Process Development. In *Cellular and Molecular Biology of Filamentous Fungi*, Borkovich KA, Ebbole DJ, Ed., ASM Press: Washington, DC, 2010, pp. 25–35.
13. Franden MA, Pienkos PT, Zhang M: Development of a high-throughput method to evaluate the impact of inhibitory compounds from lignocellulosic hydrolysates on the growth of *Zymomonas mobilis*. *J. Biotechnol.* **2009**, 144, 259–267.
14. van Hoek P, van Dijken JP, Pronk JT: Effect of specific growth rate on fermentative capacity of baker's yeast. *Appl. Environ. Microbiol.* **1998**, 64, 4226–4233.
15. Zhao X, Peng F, Cheng K, Liu D: Enhancement of the enzymatic digestibility of sugarcane bagasse by alkali-peracetic acid pretreatment. *Enzyme Microb. Tech.* **2009**, 44, 17–23.

CHAPTER 3

Analysis of Oligosaccharides in Lignocellulosic Biomass Hydrolysates by High-Performance Anion-Exchange Chromatography Coupled with Mass Spectrometry (HPAEC–MS)

Ying Zha*, Leon Coulier*, Richard Bas, and Peter J. Punt

** These authors contributed equally to this study*

Bioresource Technology **2013**, 133, 221-231

Abstract

The carbohydrate composition of lignocellulosic biomass hydrolysates is highly complex. High performance anion exchange chromatography coupled with pulsed amperometric detection (HPAEC-PAD), a widely used method for carbohydrate analysis, provides limited chemical information on the detected peaks. To improve the detection and increase the chemical information of the carbohydrates, HPAEC was coupled with mass spectrometry (MS). Using a pooled hydrolysate sample, it was shown that HPAEC-MS can separate and detect many oligosaccharides in one experimental run based on retention time and mass. The method was validated on its linearity, reproducibility and response factors. The analysis of a group of different biomass hydrolysates revealed that remaining disaccharides was the bottleneck of the hydrolysis process. As an analytical tool, HPAEC-MS provides information for the improvement of hydrolysate pretreatment method and enzyme cocktail quality. Besides, the consumption ability of microbial host strains for various mono- and oligosaccharides in hydrolysates can be assessed.

Introduction

Lignocellulosic biomass is the largest feedstock form for 2nd generation biofuel production. Compared to fossil fuels, it is renewable, environmental friendly, while compared to starch and sugarcane, it is not competing with world food resources [1,2]. Generally, lignocellulosic biomass is composed of cellulose (35-50%), hemicellulose (30-35%) and lignin (10-25%), among which cellulose is the backbone structure, while hemicellulose and lignin are the binding networks [3]. To release the sugars in lignocellulosic biomass, a pretreatment step is required to break down the binding networks and expose cellulose. The exposed cellulose is then hydrolyzed by using either fungal enzyme cocktails or sulfuric acid [4]. The released monosaccharides, e.g. glucose and xylose, in the so-called hydrolysate, are the potential carbon sources for ethanol production through fermentation processes [5]. Due to the robust structure of lignocellulosic biomass and the inhibitory compounds formed during the pretreatment process, the hydrolysis efficiency on pretreated biomass is not satisfying [6,7]. This limits the increase of bioethanol yield. Therefore, to enhance hydrolysis efficiency of pretreated lignocellulosic biomass is of decisive importance for improving bioethanol yield.

In the last decade, efforts have been paid to improve the hydrolysis efficiency of fungal enzyme cocktails on lignocellulosic biomass. Genetic engineering of known cellulases production strains and screening for novel producers of lignocellulolytic enzymes are the two main paths [8,9]. The hydrolysis efficiency was mainly determined by measuring monosaccharides concentration in hydrolysates. However, more detailed knowledge on oligosaccharides composition in hydrolysates is crucial to further understand and improve the performance of fungal enzyme cocktails.

Knowing the general structure of lignocellulosic biomass, it can be predicted that biomass hydrolysates contain mono- and oligosaccharides composed of hexoses (C6), pentoses (C5) and the combination of C6 and C5 sugars. The C6 sugar units are mainly glucose, galactose and mannose, while the C5 sugar units are mostly xylose and arabinose [3,10]. These carbohydrates have also different degrees of polymerization (DP) and may carry sugar-acid residues, which makes the analysis of carbohydrate composition of biomass hydrolysates very challenging. High-performance anion-exchange chromatography coupled with pulsed amperometric detection (HPAEC-PAD) is a widely used method for analyzing carbohydrates, especially for analyzing different monosaccharides and disaccharides for which reference compounds are present [11,12]. Though HPAEC is a powerful tool for separating carbohydrates, the PAD detector provides only limited chemical information on the detected

peaks. Such chemical information is needed for complex samples, such as biomass hydrolysates. Mass spectrometry (MS), with or without coupling to chromatography, can be a very informative technique for analyzing oligosaccharides [13,14]. Being highly selective and sensitive, MS provides detailed chemical information on the peaks, such as degree of polymerization and building blocks. Therefore, the coupling of HPAEC with MS holds high potential for separating and characterizing complex mixture of carbohydrates. Several attempts have been made to couple HPAEC with MS [15-21]. And the performance of HPAEC-MS has been demonstrated for complex mixtures of galacto-oligosaccharides (GOS) [15], extracts of poplar leaves [22], vegetables [21], and bacterial cell hydrolysate [20]. Only one study with preliminary results could be found, in which HPAEC-MS was applied to biomass hydrolysates [19].

In this study, we demonstrate in more depth the performance of HPAEC-MS on a range of lignocellulosic biomass hydrolysates. HPAEC-MS could separate many of the carbohydrates and classify the peaks based on their degree of polymerization (DP) and building blocks, i.e. pentoses, hexoses, and aldonic acids, in one experimental run. The method was validated and the relative quantity of different classes of carbohydrates could be assessed using relative response factors. A series of biomass hydrolysates, generated from various biomass types and by different pretreatment-hydrolysis methods, together with their fermentation samples, were successfully analyzed with HPAEC-MS.

Materials and Methods

HPAEC-MS

Commercially available reference compounds of carbohydrates were purchased from either Sigma-Aldrich Chemie (Zwijndrecht, the Netherlands) or Megazyme International Ltd. (Wicklow, Ireland). Internal standards sucralose, ribose-¹³C₅ and glucose-d₇ were purchased from Sigma-Aldrich. All standard solutions were prepared in MilliQ water.

HPAEC-MS was performed on a Thermo Surveyor HPLC system (Thermo Electron Corporation, San Jose, CA) equipped with a Carbopac PA1 column (250 x 2.0 mm i.d., Dionex, Sunnyvale, CA) operated at 35°C. Elution was performed with a flow of 215 µl/min and the injection volume was 3 µl. The following eluents were used: 100 mM NaOH (A) and 100 mM NaOH + 500 mM NaOAc (B). The following gradient was used: 0-5 min, isocratic 100% A, 5-78 min, linear gradient from 100 to 74% A followed by a washing step with 100%

B for 6 min and re-equilibration for 10 min at 100% A. Prior to MS detection, an ASRS300-2 mm suppressor (Dionex) was used as an in-line desalter to convert the eluate into an MS-compatible solution. The membrane was continuously regenerated with acid generated by electrolysis of water. Milli-Q water was fed from an air-pressurized bottle into the regenerant chamber at a flow rate of 3.33 ml/min. A regenerant current of 150 mA was applied. Mass detection was carried out on a Thermo LTQ LT-1000 mass detector using electrospray ionization in the positive ion mode (ESI spray voltage, 3.5 kV; heated capillary temperature, 275°C; sheath gas, 30; auxiliary gas, 5; full scan range, m/z 125-1000; number of microscans, 3; maximum injection time, 200 ms.). The mass detector was tuned using a calibration solution of lactose. For all different classes of carbohydrates the main ion detection corresponded to $[M+Na]^+$.

HPAEC-MS/MS was performed on the same system and using identical conditions as described above. MSⁿ-experiments were performed on pre-selected peaks using wide band activation and based on dependent scan settings with collision energy of 30% (Xcalibur software V2.0, Thermo Electron Corporation).

Partial validation of HPAEC-MS

Partial validation of HPAEC-MS for oligosaccharides analysis was performed on three aspects, linearity, reproducibility and relative response factors. Linearity was determined by constructing calibration curves of reference compounds. The reference compounds were dissolved in MilliQ water with concentrations ranging from 1 to 100 µg/ml. From the constructed calibration curves, the coefficient of determination (R^2) was calculated (Table 1). Reproducibility was determined by preparing a pooled sample using 21 different hydrolysates (see section 2.3 and Table 4), followed by analyzing this sample in triplicate on three different days, and calculating the coefficient of variation (CV) of these nine HPAEC-MS runs (Table 2). Relative response factors were calculated by comparing the calibration curves of the reference compounds with glucose-d7 (Table 3).

Biomass hydrolysate samples preparation

Sugar cane bagasse is a kind gift from Zilor, Brazil, and corn stover is from the University of Cape Town, South Africa. Wheat straw, barley straw and willow wood are purchased from Oostwaardshoeve, the Netherlands, and oak sawdust is from ESCO, the Netherlands, a wood-flooring supplier. All biomass is pre-dried at 80°C for 5 hours when received, and stored at room temperature in air-tight bags. Prior to pretreatment, biomass (except oak saw

dust) was ground to pieces with average length of 3 mm, and was dried again at 80°C for minimum 16 hours.

Four pretreatment methods were used to prepare biomass hydrolysates, namely dilute acid (2% H₂SO₄), mild alkaline (3% Ca(OH)₂), alkaline/peracetic acid, and concentrated acid (72% H₂SO₄). The biomass pretreated with the first three methods was hydrolyzed with Accellerase 1500 (Genencor), while the concentrated acid pretreated biomass was hydrolyzed in acid (40% and 15% H₂SO₄). The detailed steps of these methods are described in Zha *et al.* [23]. Before HPAEC-MS analysis, all hydrolysates were filtrated through 0.22 µm filter, and internal standards were added (c=20 µg/ml). Each sample was then diluted ten times with MilliQ water, and 3 µl of the diluted sample was injected.

Table 1 Linearity of the HPAEC-MS analyzing results on carbohydrate standards, expressed as coefficient of determination (R²).

Compounds	R ² (n=8)
Glucose	0.9950
Xylose	0.9900
Galactose	0.9900
Arabinose	0.9870
Mannose	0.9990
Ribose	1.0000
Cellobiose	1.0000
Xylobiose	0.9890
Arabinobiose	0.9980
Mannobiose	0.9990
Cellotriose	0.9970
Xylotriose	0.9930
Mannotriose	0.9940
Xylotetraose	0.9970

Table 2 Reproducibility of the HPAEC-MS analyzing results on selected carbohydrates in pooled hydrolysate sample, expressed as coefficient of variation (CV). (GlcA: glucuronic acid, GalA: galacturonic acid)

Selected compounds	CV % (n=9)
DP1 C6	6
DP1 C5	5
DP2 C6	4
DP2 C5	5
DP2 C5C6	5
DP3 C6	6
DP3 C5	8
DP3 C5C6C6	9
DP3 C5C5C6	6
DP4 C5	9
Galactonic acid	13
Gluconic acid	16
GalA	7
Galactonic acid/ gluconic acid-C6	12
GlcA/GalA-C5	9

Table 3 Response factors of carbohydrates belong to several different classes in HPAEC-MS results.

Compounds	Classes	Peak areas	Response factor with regard to glucose-d7
Glucose	DP1 C6	6804392	1.0
Galactose (Gal)	DP1 C6	6897888	1.0
Arabinose (Ara)	DP1 C5	6321841	0.9
Mannobiose	DP2 C6	12161242	1.8
Cellobiose	DP2 C6	11111423	1.6
Xylobiose	DP2 C5	14217711	2.1
Arabinobiose	DP2 C5	10074466	1.5
3-Gal-Ara	DP2 C5C6	8614681	1.3
Xylotriose	DP3 C5	13503524	2.0
galactonic acid	C6 acid	3809735	0.6
gluconic acid	C6 acid	3809735	0.6
GalA (galacturonic acid)	C6 acid	2644657	0.4
GlcA (glucuronic acid)	C6 acid	1882777	0.3
Lactobionic acid	C6C6 acid	3911437	0.6

Bagasse hydrolysate fermentation samples preparation

Bagasse hydrolysates, prepared with diluted acid and concentrated acid method, were fermented in 2 l New Brunswick fermentors, with working volume of 1 l. The hydrolysates were stored at 4°C before being used as fermentation media. The detailed fermentation set-up and process are described in Zha *et al.* [24]. The strains used were *Saccharomyces cerevisiae* CEN.PK 113-7D (CBS 8340) and *Pichia anomala* TNO11-29 (CBS 132101), for which the loading was 0.1 g and 0.2 g cell dry weight per 1 l hydrolysate, respectively. During the whole batch fermentation process, temperature was kept at 30°C, pH at 5.0 by adding 2 M KOH or 1 M H₂SO₄, dissolved oxygen at 0 by flushing 0.5 l/min N₂ continuously. Fermentation samples were taken every 120 min during the fermentation process. The samples were directly cooled to 4°C, centrifuged to remove yeast cells, and stored at -20°C. For each fermentation, three samples were selected to be analyzed with HPAEC-MS. These three samples were taken immediately after inoculation, mid of growth phase, and in the stationary phase. Before HPAEC-MS analysis, the samples were treated as described above.

Results and Discussion

HPAEC-MS method development

The technical challenge of coupling HPAEC with MS is the desalting of the mobile phases prior to MS detection using a $\text{Na}^+ - \text{H}^+$ exchange suppressor [16-18]. For the analysis of biomass hydrolysates, HPAEC-MS has only been described once [19].

Here we describe an extensive study on the application of HPAEC-MS in analyzing a variety of different biomass hydrolysates (Table 4). Figure 1A shows the complex base peak HPAEC-MS chromatogram of a pooled hydrolysate sample, similar to PAD detection. The molecular mass information provided by mass spectrometry (MS) can be used to distinguish hexose (C5) and pentose (C6) sugars, as well as their degree of polymerization (DP) by extracting m/z related ions (Figure 1B). Thus MS helps to separate the coeluting peaks in the base peak chromatogram, and identify carbohydrates according to their DP and subunit types (hexose/pentose). This approach enables a deeper understanding of carbohydrates composition in biomass hydrolysates, and allows focusing on the compound of interests for further study.

The number of peaks in the chromatograms increases with the increase of DP until DP3 (Figure 1B), suggesting an increased complexity of the carbohydrates remaining in the hydrolysates. With DP4 and higher DP's (data not shown), the number of peaks and the peak areas decrease, indicating that the pooled hydrolysate sample mainly consisted of mono-, di- and trisaccharides. It should be noted that the HPAEC-MS method was tuned with lactose, a DP2 hexose, and thus the highest ionization efficiency is observed for disaccharides. Analysis of reference carbohydrates indeed showed a decrease in response with increase of DP (data not shown).

As shown in Figure 1B, HPAEC-MS is also able to detect uronic acids, such as glucuronic acid (GlcA) and galacturonic acid (GalA), aldonic acids, such as gluconic acid and galactonic acid, and their oligomers. These results demonstrate that both neutral and acidic mono- and oligosaccharides can be separated and detected, by extracting their characteristic m/z values, in a single experimental run with HPAEC-MS.

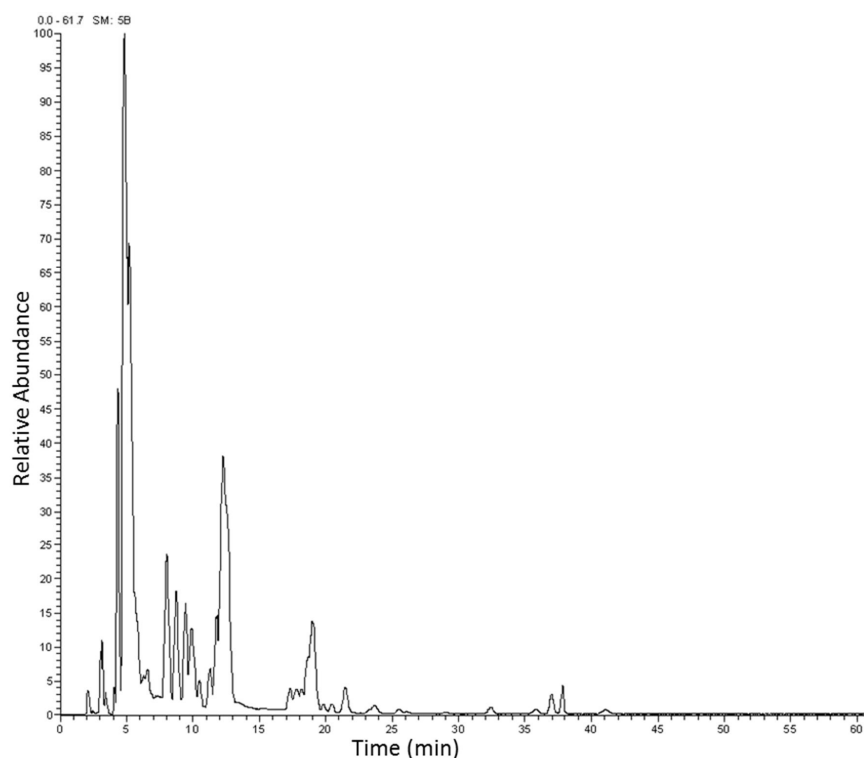
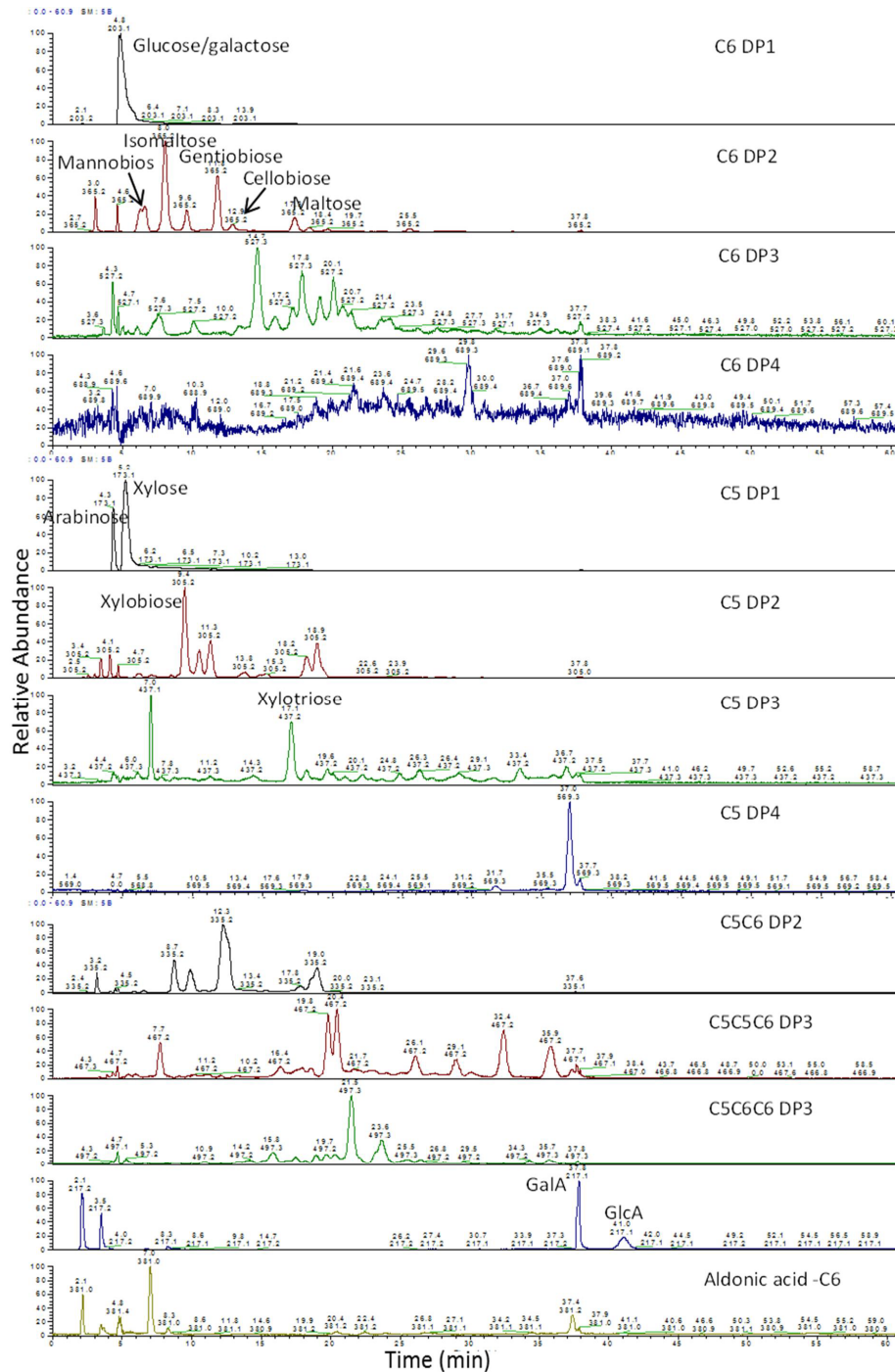


Figure 1A HPAEC-MS base peak chromatogram of the pooled hydrolysate sample.

Figure 1B (on the next page) HPAEC-MS extracted ion chromatograms ($[M+Na]^+$) of the pooled hydrolysate sample. C6: hexose, C5: pentose, GlcA: glucuronic acid, GalA: galacturonic acid, DP: degree of polymerization; extracted ion chromatograms corresponding to C6 DP1 (m/z 203.1), C6 DP2 (m/z 365.2), C6 DP3 (m/z 527.1), C6 DP4 (m/z 689.2), C5 DP1 (m/z 173.1), C5 DP2 (m/z 305.2), C5 DP3 (m/z 437.3), C5 DP4 (m/z 569.5), C5C6 DP2 (m/z 335.1), C5C5C6 DP3 (m/z 467.1), C5C6C6 DP3 (m/z 497.3), GalA/GlcA (217.1), and aldonic acid-C6 (381.0). Labeled peaks were identified by authentic standards.

Analysis of oligosaccharides in biomass hydrolysates with HPAEC-MS



Identification with HPAEC-MS/MS

An additional advantage of the coupling of HPAEC with MS is the possibility of on-line tandem mass spectrometry (MS/MS) of individual peaks. This can be used as the first step in identifying unknown carbohydrates. Although NMR is the ultimate technique for full identification of unknown compounds [15], HPAEC-MS/MS can give information on the identity of peaks quite easily. NMR will be required to further identify the detailed structure of some novel unknown compounds. HPAEC-MS provides information on the building blocks of the oligosaccharide, i.e. number of pentoses and hexoses. Figure 2A and 2B show MS/MS spectra obtained for two peaks observed with HPAEC-MS in the quality control sample with m/z 335. It can be clearly seen that one of the fragments, m/z 203, corresponds to the loss of 132 Da, which can be explained by the loss of a pentose unit. In addition, the fragment with m/z 203 corresponds to a hexose unit. Therefore, it can be concluded that the peaks with m/z 335 correspond to pentose-hexose (C5C6) disaccharides. Moreover, fragments corresponding to the loss of 60, 90 and 120 Da are visible for both C5C6 disaccharides. These losses are characteristic cross-ring cleavages, and it has been reported for hexose disaccharides that different ratios of these fragments correspond to specific linkages [13]. Although almost no reference compounds are available for C5C6 disaccharides, the MS/MS spectra in figure 2A and 2B show that for the two C5C6 disaccharides, different ratios can be observed for these fragments, indicating differences in linkages between these compounds. Figure 2C and 2D show similar MS/MS spectra of peaks with m/z 467 and 497. Based on the MS/MS fragments, these peaks can easily be identified as a C5C5C6 trimer (m/z 467) and a C5C6C6 trimer (m/z 497). Also here MS/MS spectra provide (partial) linkage information. Further research and especially data from relevant reference compounds is necessary to make more use of the MS/MS data for identification. The analysis of C5C6 oligosaccharides in biomass hydrolysates has not been reported before with HPAEC-MS, although their presence has been shown using other techniques [25,26]. These studies clearly showed the complexity of biomass hydrolysates and thus the added value of using MS as a detection method in combination with HPAEC.

HPAEC-MS method validation

It was shown above that HPAEC-MS is capable of separating and detecting complex mixtures of oligosaccharides. To apply the method in practice, it is very important to evaluate the performance of the method. As HPAEC-MS will probably be used in screening experiments where oligosaccharide patterns are compared in a quantitative manner, the method was validated on three aspects: linearity, reproducibility, and relative response

factors of various carbohydrates. The pooled hydrolysate sample was used for testing reproducibility, while purchased reference compounds were used as standards for studying linearity and response factors.

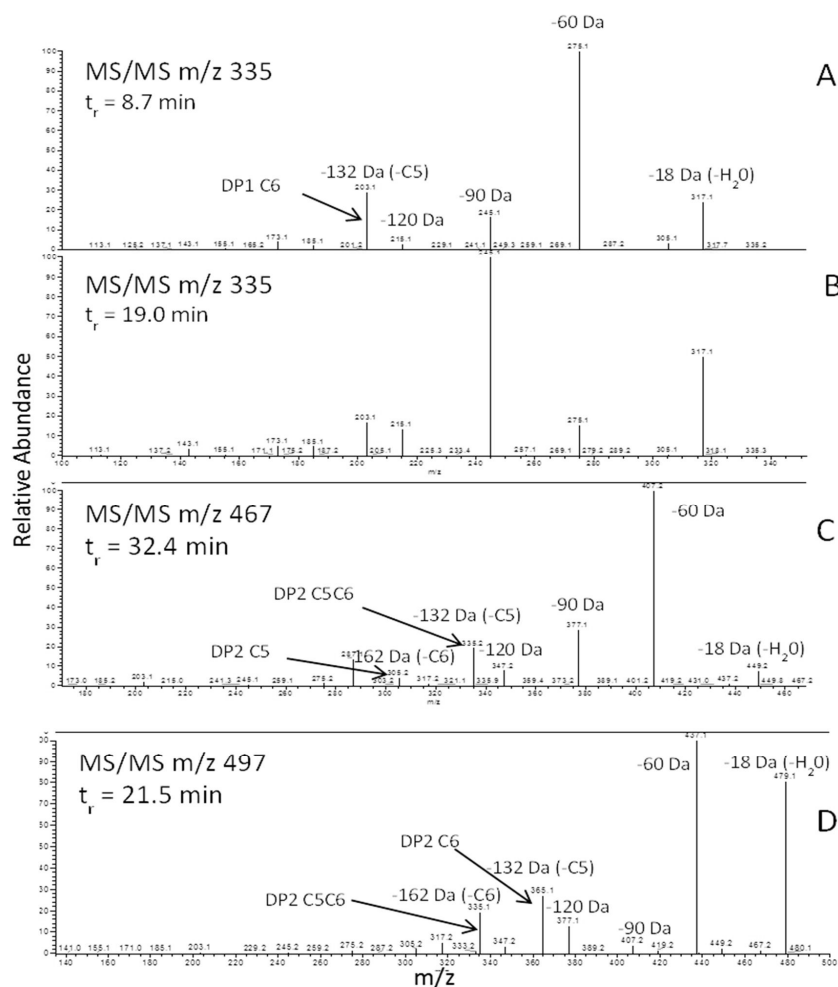


Figure 2 MS/MS spectra of peaks observed in the pooled hydrolysate sample with HPAEC-MS; (A) m/z 335, $t_r = 8.7$ min, (B) m/z 335, $t_r = 19.0$ min, (C) m/z 467, $t_r = 32.4$ min, (D) m/z 497, $t_r = 21.5$ min.

Linearity

The linearity of HPAEC-MS was studied with purchased reference compounds, as shown in Table 1. The testing standards were individual carbohydrates with eight different concentrations, ranging from 1 to 100 µg/ml. Based on the measurement results, calibration curves were generated, with which coefficient of determination (R^2) was calculated for each compound (Table 1). With all the compounds tested, the response of HPAEC-MS was linear in the range of 1-100 µg/ml, and can thus be used for quantification, although the influence of sample matrix needs to be considered with complex samples, such as biomass hydrolysates.

Reproducibility

To test the reproducibility of HPAEC-MS on biomass hydrolysates, the pooled sample was analyzed in triplicates on three different days, which resulted in nine individual measurement results. Ten carbohydrates of DP1-4 and five acidic sugars were selected for calculating their relative peak areas using the nine measurement results. These relative peak areas were in turn used for coefficient of variation (CV) calculation (Table 2). As far as the selected carbohydrates are concerned, CV values increased with increasing of DP, but were all below 10%. Compared to carbohydrates, the CV values of acidic sugars were somewhat higher, up to 16% (Table 2). These results indicate that HPAEC-MS can measure carbohydrates with reasonably high reproducibility.

Relative response factor

Typically, biomass hydrolysates contain many different carbohydrates of which many are unknown. Identification of each individual unknown carbohydrate is time-consuming and practically impossible. However, for optimizing experimental conditions, it is helpful to have estimates of concentrations of these carbohydrates. One possibility is to use so-called relative response factor, i.e. relative peak area at a certain concentration with respect to a reference compound. Therefore, the relative response factors of several carbohydrates and acidic sugars were calculated from their calibration curve, and compared with glucose-d7. As shown in Table 3, the peak areas of DP1's were almost equal to that of glucose-d7, while the relative response factors of DP2 and DP3 were ~1.7 and 2.0. Acidic sugars had response factors of ~0.5. The high response factors observed for DP2 and DP3 may due to the fact that lactose, a DP2 hexose, was used for tuning MS. These relative response factors can be used for (semi)-quantification of identified peaks, for which there are no reference compounds, or unidentified peaks when the building blocks and degree of polymerization are known.

HPAEC-MS method application

Biomass hydrolysate samples

Generally, lignocellulosic biomass hydrolysates are prepared from a diverse range of biomass types and with various methods. Different biomass are used in different areas of the world and the method chosen for biomass hydrolysate preparation differs [27]. Despite the high variety in biomass hydrolysate preparation, a main goal is to achieve high hydrolysis efficiency. This is mainly done by improving pretreatment processes and enzyme cocktail quality [8]. Analysis of biomass hydrolysate samples with HPAEC-MS provides important information on their carbohydrate composition, resulting in new targets for improving hydrolysis efficiency. To study the influence of biomass type and hydrolysate preparation method on the carbohydrate composition of biomass hydrolysates, 21 different biomass hydrolysates were generated in our lab from six types of biomass: wheat straw (WS), barley straw (BS), corn stover (CS), bagasse (Bag), oak saw dust (Oak) and willow wood chips (Willow) (Table 4). The methods used for generating these hydrolysates were (1) dilute acid (DA), (2) mild alkaline (MA), (3) alkaline/peracetic acid (PAA), and (4) concentrated acid (CA) [23]. The hydrolysates prepared with DA, MA and PAA method were enzymatically hydrolyzed, while the hydrolysates prepared with CA method were hydrolyzed with H₂SO₄.

These 21 hydrolysates were analyzed with HPAEC-MS, and the results show that their overall carbohydrate compositions are comparable. The base peaks and extracted ion chromatograms (DP1-3) of wheat straw-dilute acid (WS-DA) and wheat straw-concentrated acid (WS-CA) hydrolysates are shown here as examples (Figure 3). As expected, the highest peak was C6 DP1, glucose/galactose, while the second highest peak was C5 DP1, xylose and arabinose. Next to these peaks, several disaccharides were detected, among which were not only C6-C6 and C5-C5, but also hybrid disaccharides C6-C5 (Figure 3). DP3's and DP4's were also observed, but with significantly lower amounts. In most hydrolysates, DP4's only count for less than 1% of the detected carbohydrates (Table 4). Remarkably, the peak areas of C6-C5 disaccharides were larger in most hydrolysates compared to C6-C6 or C5-C5, suggesting the degradation of C6-C5 disaccharides is the most difficult among the three DP2 forms. Besides, of all the C6-C6 disaccharides, cellobiose only counted for 1-3% (Table 4). This result indicates that the majority of the remaining C6-C6 disaccharides detected in these hydrolysates were not the degradation products of cellulose. It can be concluded that the most abundant C6-C6 disaccharides were isomaltose, gentiobiose, and mannobiose, which are not the degradation products of hemicellulose either (Figure 1A). Therefore, it is suspected that some of the remaining DP2's are formed through transglycosylation. It is known that enzymes possess hydrolytic activity, such as β -glycosidases and β -xylosidases,

also catalyze transglycosylation reactions [28], but these reactions were not widely reported in hydrolysis processes of pretreated lignocellulosic biomass. During enzymatic hydrolysis processes, when high concentration of monomer hexoses and pentoses are present together with hydrolytic enzymes, it is highly possible that these enzymes also catalyze the transglycosylation reactions.

Table 4 Biomass hydrolysates used for HPAEC-MS analysis, and the percentage of monosaccharides and oligosaccharides to detected carbohydrates, calculated with equation:

$$\text{DPX detected carbohydrates (\%)} = \frac{\text{sum of DPX peak area}}{\text{total peak area}} \times \frac{1}{\text{DPX response factor}}$$
; response factors used: 1.0 (X=1), 1.65 (X=2), 2.0 (X=3 & 4). Hydrolysates prepared with different methods have specific distribution of disaccharide forms: C5C5, C5C6, C6C6. Grey marked area indicates the highest percentage of a disaccharide form among hydrolysates prepared with the four methods: CA, DA, MA, PAA. Bag: bagasse, BS: barley straw, CS: corn stover, Oak: oak saw dust, Willow: willow wood chips, WS: wheat straw; CA: concentrated acid, DA: dilute acid, MA: mild alkaline, PAA: peracetic acid; DC: detected carbohydrates.

	DP1/ DC	Glucose/ DP1	DP2/ DC	Cellobiose/ DP2	C5C5/ DP2	C5C6/ DP2	C6C6/ DP2	DP3/ DC	DP4/ DC
Bag-CA	53%	57%	38%	3%	13%	45%	42%	3%	0%
BS-CA	47%	62%	40%	3%	13%	45%	42%	3%	0%
CS-CA	56%	51%	35%	3%	15%	45%	40%	2%	0%
Oak-CA	53%	69%	40%	3%	11%	50%	38%	3%	0%
Willow-CA	54%	62%	34%	3%	9%	39%	52%	3%	0%
WS-CA	55%	57%	35%	3%	14%	46%	40%	2%	0%
Bag-DA	70%	57%	23%	2%	20%	45%	35%	2%	0%
BS-DA	45%	62%	35%	2%	45%	37%	18%	13%	1%
CS-DA	69%	44%	23%	1%	45%	35%	19%	3%	0%
Willow-DA	63%	51%	24%	1%	34%	35%	31%	3%	0%
WS-DA	61%	52%	27%	2%	30%	41%	29%	4%	0%
BS-MA	51%	60%	40%	0%	24%	63%	14%	7%	1%
CS-MA	54%	50%	37%	0%	25%	63%	12%	7%	0%
Oak-MA	52%	66%	43%	0%	15%	58%	27%	4%	0%
Willow-MA	55%	66%	39%	0%	23%	67%	10%	6%	0%
Bag-PAA	53%	63%	37%	1%	15%	71%	15%	6%	2%
BS-PAA	55%	70%	35%	3%	14%	65%	21%	5%	4%
CS-PAA	62%	60%	31%	1%	14%	64%	22%	3%	2%
Oak-PAA	54%	72%	41%	2%	8%	55%	37%	4%	0%
Willow-PAA	54%	59%	39%	1%	15%	61%	24%	4%	0%
WS-PAA	48%	62%	37%	1%	16%	61%	22%	6%	6%

Like in WS-CA and WS-DA hydrolysates (Figure 3), disaccharides were the most common form of residual oligosaccharides in all 21 hydrolysates (Table 4). This observation indicates that disaccharides are the bottleneck of biomass hydrolysis, caused mainly by incomplete hydrolysis, independent from biomass type and pretreatment method. However, the number and areas of the disaccharides in different hydrolysates varied, this is demonstrated with four hydrolysates, namely WS-DA, WS-CA, bagasse-dilute acid (Bag-DA), and bagasse-concentrated acid (Bag-CA) (Figure 4). It can be seen that the disaccharide pattern of WS-DA is similar to Bag-DA, while rather different from WS-CA. Also, the disaccharide pattern of WS-CA is more comparable to Bag-CA than to WS-DA. This is most probably due to WS-CA and Bag-CA were hydrolyzed with acid, while WS-DA and Bag-DA were hydrolyzed enzymatically.

Based on the observation that disaccharides were the most common form of residual oligosaccharides in all 21 hydrolysates, we used the following equation to indicate hydrolysis efficiency:

$$\frac{DP2}{\text{detected carbohydrates}} = \frac{\text{sum of DP2 peak area} / DP2 \text{ response factor}}{\text{total peak area}}$$

A larger *DP2/detected carbohydrates* value corresponds to a lower hydrolysis efficiency. It should be mentioned that not all carbohydrates present in a hydrolysate sample can be detected with HPAEC-MS method, so the *detected carbohydrates* in the equation is not the same as the total sugar amount in the hydrolysate. Therefore, the hydrolysis efficiency assessed with this equation is only a relative value, which cannot be used to compare with the hydrolysis efficiency determined with other analytical methods.

The *DP2/detected carbohydrates* values of the 21 analyzed hydrolysates were calculated (Table 4), the used *DP2 response factor* value was 1.65, which is the average value of all the disaccharide standards tested (Table 3). As shown in table 4, the hydrolysates prepared with dilute acid (DA) method had the highest C5C5 percentage, while the mild alkaline (MA) and peracetic acid (PAA) prepared hydrolysates had relatively higher percentage of C5C6. Different from enzymatically hydrolyzed hydrolysates, the most abundant disaccharide form in concentrated acid (CA) hydrolysates was C6C6. This indicates that the residual disaccharides composition in a hydrolysate was mainly influenced by hydrolysate preparation method, while biomass type played only a minor role.

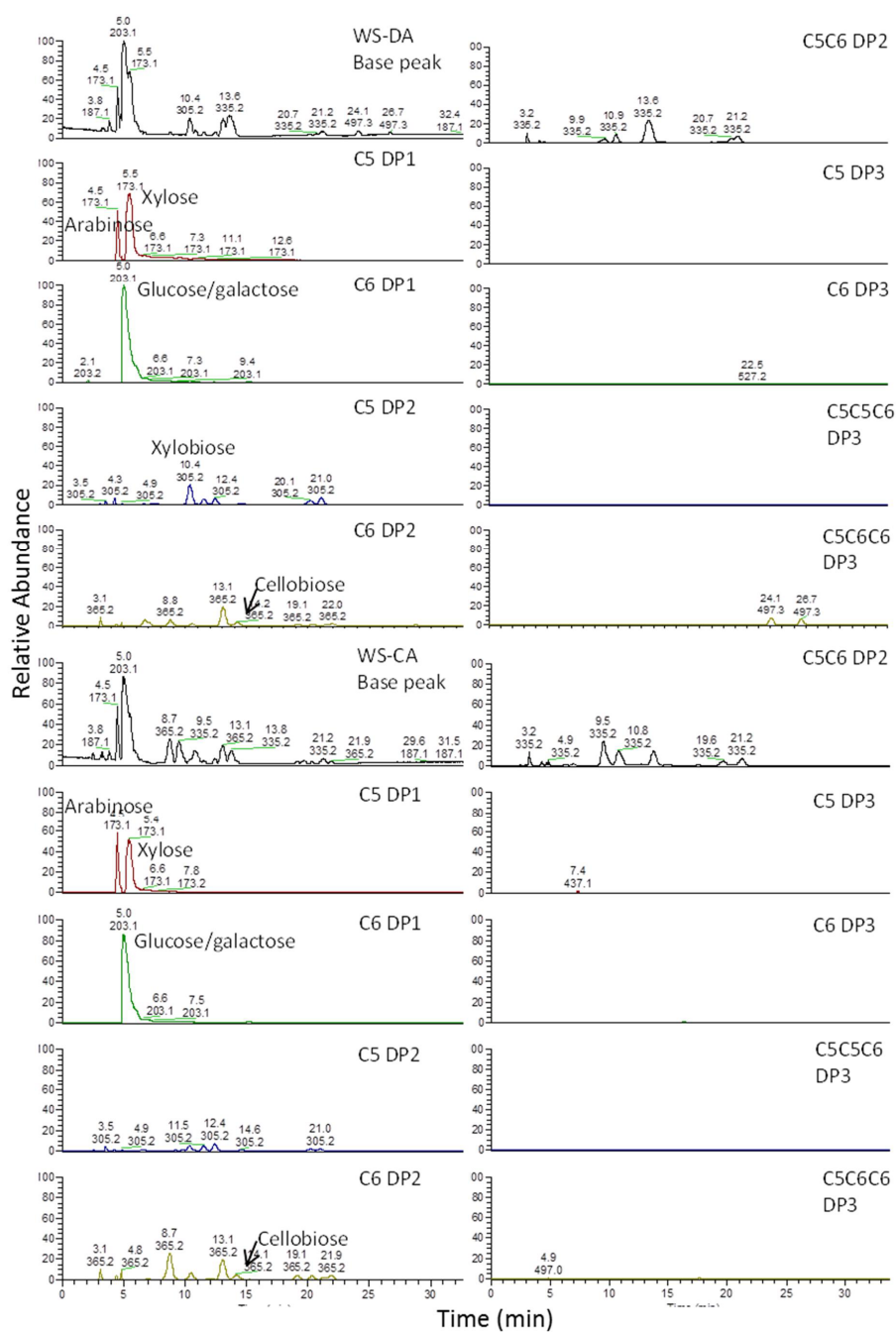
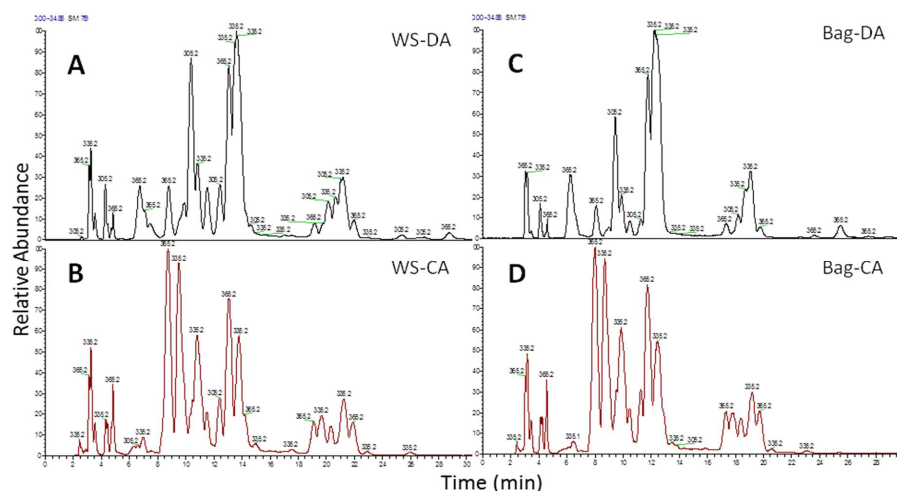


Figure 3 HPAEC–MS base peak and fixed scale extracted ion chromatograms of wheat straw–diluted acid (WS–DA) (top) and wheat straw–concentrated acid (WS–CA) (bottom) hydrolysates.



S.cerevisiae CEN.PK 113-7D, while Bag-CA was fermented with *S.cerevisiae* CEN.PK 113-7D and *P.anomala* TNO11-29. For each fermentation, three time samples were analyzed with HPAEC-MS method, which represented the beginning, the mid-point of growth phase and the stationary phase of the fermentation.

Among all the carbohydrates detected, 17 of them were chosen to analyze their trends during the fermentation process, by comparing the peak areas of each selected carbohydrate at the three fermentation time points. The 17 chosen carbohydrates are the largest peaks of DP1-3 and a sugar acid, glucuronic acid (GlcA). Table 5 shows the retention time, the putative identification, the degree of polymerization, and the peak areas after correcting with internal standards of these 17 peaks. As expected, glucose was consumed in all three fermentations, as it is the most preferred carbon source for the two yeast strains used [29,30]. Disaccharides and trisaccharides kept constant during all three fermentation processes, except maltose, cellobiose and gentiobiose. Maltose decreased in all three fermentations, while cellobiose and gentiobiose reduced only during the Bag-DA fermentation process. Maltose is thought to be consumed by both strains during the fermentation processes. It is known from literature that *S. cerevisiae* produces saccharolytic enzymes with limited activity, but due to lack of transporters, the only disaccharide the strain used as carbon source here is maltose [31,32]. The consumption of maltose was also reported in *P. anomala* strains [33]. Different from maltose, cellobiose and gentiobiose were suspected to be degraded by the residual enzymes in Bag-DA hydrolysate, since the reduction of these two disaccharides was not observed in Bag-CA hydrolysate, into which no enzymes was added.

The observation that cellobiose and gentiobiose were the only oligosaccharides degraded in the Bag-DA fermentation can be attributed to the property of the enzyme cocktails used. Accellerase 1500 mainly contains glucanase, hemi-cellulase and beta-glucosidase, suggesting that the main hydrolysis substrates are cellulose and hexose oligosaccharides. Enzymes that are responsible for hydrolyzing other oligosaccharides, such as isomaltose, mannobiose, and hybrid disaccharides, were hardly present in the enzyme cocktails used. As discussed in the previous session, hybrid disaccharides were the most common form of residual carbohydrates in hydrolysates, so it could be beneficial to include enzymes like α -D-xylosidase, β -1,4-D-galactosidase, α -L-arabinofuranosidase, and β -1,4-D-mannosidase in the enzyme cocktails [7,10,34]. These enzymes will help degrade the hybrid and α -linkage disaccharides during the hydrolysis process and/or in the fermentor later on.

It can also be seen from Table 5 that xylose was not consumed. It is reasonable since *S.cerevisiae* is not a pentose fermenting yeast [35] and the *P.anomala* strain used only

Analysis of oligosaccharides in biomass hydrolysates with HPAEC-MS

consumes xylose under aerobic conditions [36]. In this study, the peak areas of xylose even somewhat increased during all 3 fermentations, suggesting the release of xylose residues from more complex carbohydrates, apparently not detected in our HPAEC-MS analysis.

Through analyzing fermentation samples by using HPAEC-MS, the dynamics of major carbohydrates in biomass hydrolysates were monitored. This showed that HPAEC-MS is capable of providing information on the carbon sources that a specific fermenting microbe utilizes in a complex medium. A similar approach can also be applied to help indirectly analyze and identify the production of specific fungal or bacterial enzymes, with either newly isolated or genetically modified strains.

Table 5 Internal standard corrected peak areas of the selected 17 carbohydrates in fermentation samples (the decreased peak areas are marked in bold). Bag: bagasse, DA: dilute acid, CA: concentrated acid; S.c: *S. cerevisiae*, P.a: *P. anomala*; GlcA: glucuronic acid.

	RT	putative identification	Bag-DA S.c			Bag-CA S.c			Bag-CA P.a		
			T1	T2	T3	T1	T2	T3	T1	T2	T3
C5 DP1	5.2	xylose	493.2	547.7	618.4	386.6	403.2	486.4	287.1	314.8	400.0
C5 DP2 a	9.4	xylobiose	57.8	57.3	60.6	18.7	19.5	20.7	14.9	15.1	17.7
C5 DP2 b	10.4		6.2	6.7	6.1	20.2	21.3	21.8	15.6	15.7	18.8
C5 DP2 c	11.3		9.2	10.3	9.9	33.1	34.6	35.8	24.7	25.3	30.6
C5 DP3	17.1	xylotriose	1.5	1.8	2.2	---	---	---	---	---	---
C5 DP3	7.0		---	---	---	4.0	4.3	4.3	3.1	3.2	3.2
C6 DP1	4.8	glucose	775.7	215.9	24.1	527.2	351.1	69.5	407.9	180.5	15.4
C6 DP2 a	6.2	mannobiose	46.4	46.1	46.2	3.1	3.5	3.3	2.3	2.2	2.8
C6 DP2 b	8.0	isomaltose	14.8	14.1	10.0	117.5	125.4	126.4	95.1	96.0	120.7
C6 DP2 c	11.8	gentobiose	79.5	33.8	0.0	92.7	99.4	85.4	76.2	76.0	88.6
C6 DP2 d	12.9	cellobiose	9.5	3.8	3.2	22.1	26.0	23.3	19.6	17.0	21.4
C6 DP2 e	19.7	maltose	5.0	1.0	0.0	31.8	17.3	2.1	25.3	16.3	0.0
C6 DP3	20.1		5.6	5.8	5.7	---	---	---	---	---	---
C6 DP3	14.7		---	---	---	7.0	6.8	7.3	5.3	5.4	6.4
C5C6 a	12.3		211.6	222.7	218.5	80.6	83.0	82.3	60.5	63.3	81.3
C5C6 b	19.0		65.4	66.2	58.7	46.6	40.8	35.4	36.7	40.5	30.7
C6 acid	41.0	GlcA	10.1	9.7	10.1	9.3	9.2	8.6	6.0	6.2	7.7

Conclusion

To better understand complex carbohydrate samples, like lignocellulosic biomass hydrolysates, HPAEC coupled with MS was used. The HPAEC-MS method successfully separated many oligosaccharides in biomass hydrolysates and classified them based on their DP's in one experimental run. The method requires little sample work-up and provides chemical information, such as degree of polymerization and building blocks, of the detected carbohydrates. The application of HPAEC-MS was demonstrated by analyzing 21 different biomass hydrolysates. Results show that disaccharides, including C6C6, C5C5, C5C6, are the main remaining soluble sugar forms in these hydrolysates, which compositions are mainly influenced by the hydrolysate pretreatment method.

Acknowledgement

This project was (co)financed by the Netherlands Metabolomics Centre (NMC) which is part of the Netherlands Genomics Initiative / Netherlands Organization for Scientific Research.

References

1. Hahn-Hägerdal B, Galbe M, Gorwa-Grauslund MF, Lidén G, Zacchi G: Bio-ethanol--the fuel of tomorrow from the residues of today. *Trends Biotechnol.* **2006**, 24, 549-556.
2. Zhang YH: Reviving the carbohydrate economy via multi-product lignocellulose biorefineries. *J. Ind. Microbiol. Biotechnol.* **2008**, 35, 367-375.
3. Sarkar P, Bosneaga E, Auer M: Plant cell walls throughout evolution: Towards a molecular understanding of their design principles. *J. Exp. Bot.* **2009**, 60, 3615-3635.
4. Alvira P, Tomás-Pejó E, Ballesteros M, Negro MJ: Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review. *Bioresour. Technol.* **2010**, 101, 4851-4861.
5. Kuhad RC, Gupta R, Khasa YP, Singh A, Zhang YHP: Bioethanol production from pentose sugars: Current status and future prospects. *Renew. Sust. Energ. Rev.* **2011**, 15, 4950-4962.
6. Yang B, Dai Z, Ding S, Wyman CE: Enzymatic hydrolysis of cellulosic biomass. *Biofuels* **2011**, 2, 421-450.
7. van Dyk JS, Pletschke BI: A review of lignocellulose bioconversion using enzymatic hydrolysis and synergistic cooperation between enzymes-factors affecting enzymes, conversion and synergy. *Biotechnol. Adv.* **2012**, 30, 1458-1480.
8. Mathew GM, Sukumaran RK, Singhania RR, Pandey A: Progress in research on fungal cellulases for lignocellulose degradation. *J. Sci. Ind. Res.* **2008**, 67, 898-907.
9. Gibson DM, King BC, Hayes ML, Bergstrom GC: Plant pathogens as a source of diverse enzymes for lignocellulose digestion. *Curr. Opin. Microbiol.* **2011**, 14, 264-270.
10. Saha B: Hemicellulose bioconversion. *J. Ind. Microbiol. Biotechnol.* **2003**, 30, 279-291.

11. Morales V, Corzo N, Sanz ML: HPAEC-PAD oligosaccharide analysis to detect adulterations of honey with sugar syrups. *Food Chem.* **2008**, 107, 922-928.
12. Korosec M, Bertoncelj J, Gonzales AP, Kropf U, Golob U, Golob T: Monosaccharides and oligosaccharides in four types of Slovenian honey. *Acta Aliment. Hung.* **2009**, 38, 459-469.
13. Zaia J: Mass spectrometry of oligosaccharides. *Mass Spectrom. Rev.* **2004**, 23, 161-227.
14. Stefan B: Mass spectrometry for characterizing plant cell wall polysaccharides. *Front. Plant Sci.* **2012**, 3, doi: 103389/fpls201200045.
15. Coulier L, Timmermans J, Richard B, van Den Dool R, Haaksman I, Klarenbeek B, Slaghek T, van Dongen W: In-depth characterization of prebiotic galactooligosaccharides by a combination of analytical techniques. *J. Agric. Food Chem.* **2009**, 57, 8488-8495.
16. Bruggink C, Maurer R, Herrmann H, Cavalli S, Hoefler F: Analysis of carbohydrates by anion exchange chromatography and mass spectrometry. *J. Chromatogr. A* **2005**, 1085, 104-109.
17. Niessen WMA, van der Hoeven RAM, van der Greef J, Schols HA, Lucas-Lokhorst G, Voragen AGJ, Bruggink C: High-performance anion-exchange chromatography/thermospray mass spectrometry in the analysis of oligosaccharides. *Rapid Commun. Mass Spectrom.* **1992**, 6, 474-478.
18. van der Hoeven RAM, Niessen WMA, Schols HA, Bruggink C, Voragen AGJ, van der Greef J: Characterization of sugar oligomers by on-line high-performance anion-exchange chromatography-thermospray mass spectrometry. *J. Chromatogr. A* **1992**, 627, 63-73.
19. Rumbold K, Okatch H, Torto N, Siika-Aho M, Gübitz G, Robra K, Prior B: Monitoring on-line desalted lignocellulosic hydrolysates by microdialysis sampling micro-high performance anion exchange chromatography with integrated pulsed electrochemical detection/mass spectrometry. *Biotechnol. Bioeng.* **2002**, 78, 822-828.
20. Wunschel DS, Fox KF, Fox A, Nagpal ML, Kim K, Stewart GC, Shahgholi M: Quantitative analysis of neutral and acidic sugars in whole bacterial cell hydrolysates using high-performance anion-exchange liquid chromatography-electrospray ionization tandem mass spectrometry. *J. Chromatogr. A* **1997**, 776, 205-219.
21. Okatch H, Torto N, Armatefio J: Characterisation of legumes by enzymatic hydrolysis, microdialysis sampling, and micro-high-performance anion-exchange chromatography with electrospray ionisation mass spectrometry. *J. Chromatogr. A* **2003**, 992, 67-74.
22. Guignard C, Jouve L, Bogéat-Triboulot MB, Dreyer E, Hausman JF, Hoffmann L: Analysis of carbohydrates in plants by high-performance anion-exchange chromatography coupled with electrospray mass spectrometry. *J. Chromatogr. A* **2005**, 1085, 137-142.
23. Zha Y, Slomp R, Groenestijn J, Punt PJ: Preparation and Evaluation of Lignocellulosic Biomass Hydrolysates for Growth by Ethanogenic Yeasts. In *Microbial Metabolic Engineering: Methods and Protocols*, Qiong C Ed, Humana Press: New York, USA, 2012, pp. 245-259.
24. Zha Y, Muilwijk B, Coulier L, Punt PJ: Inhibitory compounds in lignocellulosic biomass hydrolysates during hydrolysate fermentation processes. *J. Bioprocess Biotechnol.* **2012**, 2, doi: 104172/2155-98211000112.
25. Bowman MJ, Dien BS, Hector RE, Sarath G, Cotta MA: Liquid chromatography-mass spectrometry investigation of enzyme-resistant xylooligosaccharide structures of switchgrass associated with ammonia pretreatment, enzymatic saccharification, and fermentation. *Bioresour. Technol.* **2012**, 110, 437-447.
26. Appeldoorn MM, Kabel MA, van Eylen D, Gruppen H, Schols HA: Characterization of oligomeric xylan structures from corn fiber resistant to pretreatment and simultaneous saccharification and fermentation. *J. Agric. Food Chem.* **2010**, 58, 11294-11301.
27. Mosier N, Wyman C, Dale B, Elander R, Lee YY, Holtzapple M, Ladisch M: Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresour. Technol.* **2005**, 96, 673-686.

28. Eneyskaya EV, Brumer III H, Backinowsky LV, Ivanen DR, Kulinskaya AA, Shabalin KA, Neustroev KN: Enzymatic synthesis of beta-xylanase substrates: transglycosylation reactions of the beta-xylosidase from *Aspergillus* sp. *Carbohydr. Res.* **2003**, 338, 313-325.
29. Rumbold K, van Buijsen HJJ, Overkamp KM, van Groenestijn JW, Punt PJ, van der Werf MJ: Microbial production host selection for converting second-generation feedstocks into bioproducts. *Microb. Cell Fact.* **2009**, 8, doi:10.1186/1475-2859-8-64.
30. Fredlund E, Druvefors U, Boysen ME, Lingsten KJ, Schnürer J: Physiological characteristics of the biocontrol yeast *Pichia anomala* J121. *FEMS Yeast Res.* **2002**, 2, 395-402.
31. Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS: Microbial cellulose utilization: Fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* **2002**, 66, 506-577.
32. Galazka JM, Tian C, Beeson WT, Martinez B, Glass NL, Cate JHD: Cellodextrin transport in yeast for improved biofuel production. *Science* **2010**, 330, 84-86.
33. Walker GM: *Pichia anomala*: Cell physiology and biotechnology relative to other yeasts. *Anton. Leeuw. Int. J. G.* **2011**, 99, 25-34.
34. de Vries R, Visser J: *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. *Microbiol. Mol. Biol. Rev.* **2001**, 65, 497.
35. Kuyper M, Winkler AA, van Dijken JP, Pronk JT: Minimal metabolic engineering of *Saccharomyces cerevisiae* for efficient anaerobic xylose fermentation: a proof of principle. *FEMS Yeast Res.* **2004**, 4, 655-664.
36. Zha Y, Hossain AH, Tobola F, Sedee N, Havekes M, Punt PJ: *Pichia anomala* 29X: a resistant strain for lignocellulosic biomass hydrolysate fermentation. *FEMS Yeast Res.* **2013**, doi: 10.1111/1567-1364.12062.

CHAPTER 4

Inhibitory Compounds in Lignocellulosic Biomass Hydrolysates during Hydrolysate Fermentation Processes

Ying Zha, Bas Mulwijk, Leon Coulier, and Peter J. Punt

Journal of Bioprocessing & Biotechniques **2012**, 2:112

Abstract

To compare the composition and performance of various lignocellulosic biomass hydrolysates as fermentation media, 8 hydrolysates were generated from a grass-like and a wood biomass. The hydrolysate preparation methods used were (1) dilute acid, (2) mild alkaline, (3) alkaline/peracetic acid, and (4) concentrated acid. These hydrolysates were fermented at 30°C, pH 5.0, using *Saccharomyces cerevisiae* CEN.PK 113-7D as model strain. The growth in different hydrolysates varied in the aspects of lag-phase, growth rate, glucose consumption rate and ethanol production rate. Subsequently, 11 potential hydrolysate inhibitors were selected, their concentrations in the time-samples of the 8 fermentations were determined using a novel analytical method, ethyl-chloroformate derivatization-GC-MS. Some of these compounds, e.g. furfural, decreased during the fermentation process, while others, such as formic and benzoic acid, remained almost constant. The 11 compounds were tested individually for their inhibitory effects on the model yeast, the results showed that most of the compounds exhibited little effect at their concentrations detected in biomass hydrolysates. Only furfural and benzoic acid clearly affected the growth of the model yeast: furfural elongated the lag-phase, while benzoic acid reduced the growth rate and biomass yield.

Introduction

Lignocellulosic biomass is the feedstock for the production of 2nd generation biofuel. The biomass, such as bagasse, corn stover, wheat straw and willow wood, is structurally composed of cellulose, hemicellulose and lignin [1,2]. To transform the biomass to liquid fermentation medium, a pretreatment and a hydrolysis step are required to break down the biomass structure and to form monomer sugars, such as glucose and xylose [3,4]. The composition of this liquid medium, named hydrolysate, is determined by the biomass type and the pretreatment-hydrolysis method used.

During the pretreatment process, various degradation products of both sugar and lignin are formed, among which are some inhibitory compounds. These compounds negatively influence the hydrolysis as well as fermentation process [5,6]. Acetic acid, furfural and 5-hydroxymethyl furfural are the most studied inhibitory compounds in hydrolysates. These compounds were also used for toxicity studies in different microorganisms [1,7,8]. The other compounds that were reported inhibitory are mainly weak acids, phenolic and aromatic species. These compounds, for example, vanillin and syringic acid, are less well studied concerning their concentrations in hydrolysate and their effects [1,5].

The effects of inhibitory compounds in a fermentation process were shown as longer lag-phase, slower growth, lower cell density and decreased ethanol productivity [9,10]. To be able to use hydrolysate for biofuel production on an industrial level, these effects need to be reduced. Several detoxification methods have been developed and applied to different hydrolysates. Activated carbon, organic solvent absorbing and extracting inhibitory compounds were proven to be effective physical detoxification methods [11,12]. The chemical detoxification methods include over-liming, reacting with reducing agent and peroxide treatment [13,14].

Since hydrolysates were made from natural materials and the preparation methods are various, the composition and performance of different hydrolysates differ. These differences are of importance for both inhibitory compounds studies and detoxification method development. Therefore, studying the similarity and difference of various hydrolysates on their composition and fermentation performance is of considerable interest. The results of these studies will provide information to analyze the relationship between hydrolysate composition and its fermentation performance as medium. For a proper study design, the selected hydrolysates should be different in their fermentation performance. This can be achieved by using different biomass types and diverse pretreatment-hydrolysis methods to

prepare the hydrolysates. In this study, we generated 8 different hydrolysates from bagasse and oak sawdust to compare their performance as fermentation media. Hydrolysates and their fermentation time-series samples were taken to study their composition and dynamics during the fermentation process. These samples were analyzed with EC-GC-MS method. This analytical method was developed to remove the sugar content in the hydrolysates and detect sugar and lignin degradation products. Among all the compounds detected, 11 were selected to be quantified in hydrolysates and their fermentation time-series samples. These selected compounds are formic acid, acetic acid, levulinic acid, furfural, furfuryl alcohol, 2-furanmethanol acetate, HMF, vanillin, syringic acid, benzoic acid and 4-hydroxybenzaldehyde. They were chosen because they were either reported as inhibitory compounds [15,16] or belong to the categories of potential inhibitors [17,18]. The concentrations of these compounds detected in the hydrolysates were used to analyze their dynamics during the fermentation process, and test their inhibitory effects individually using a screening method.

Materials and Methods

Biomass

Sugar cane bagasse (Bag) was a kind gift from ZILOR, Brazil, and oak sawdust (Oak) was obtained from ESCO, the Netherlands, a wood-flooring supplier. Both types of biomass were pre-dried at 80°C for 5 hours when received, and stored at room temperature. Sugar cane bagasse was ground to pieces with average length of 3 mm. Prior to pretreatment, the biomass was dried again at 80°C for minimum 16 hours.

Hydrolysate preparation method

Four pretreatment methods were used to prepare bagasse and oak sawdust for hydrolysis, namely dilute acid (DA) (2% H₂SO₄), mild alkaline (MA) (3% Ca(OH)₂), alkaline/peracetic acid (PAA) [19], and concentrated acid (CA) (72% H₂SO₄). The biomass pretreated with the first three methods was hydrolyzed enzymatically while the concentrated acid pretreated biomass was hydrolyzed in acid. The detailed steps of these methods are described in Zha *et al.* [20].

Strain and preculture

Saccharomyces cerevisiae CEN.PK 113-7D (CBS8340) was used as model strain in this hydrolysate study. The strain was obtained from CBS Utrecht, the Netherlands. The preculture for both fermentation and Bioscreen test was prepared in a 500 ml Erlenmeyer

shake flask with 100 ml mineral medium and 20 g/l glucose. The mineral media was prepared according to van Hoek *et al.* [21]. The preculture was inoculated with 1 ml *S. cerevisiae* CEN.PK 113-7D glycerol stock, and incubated at 30°C, 200 rpm for 20 hours.

Fermentation setup

The batch fermentation was carried out in a 2 l New Brunswick fermentor with working volume of 1 l. The fermentor, filled with 1 l demineralized water, was sterilized at 121°C. After sterilization, the fermentor was connected to the console, emptied and filled with 950 ml filter-sterilized hydrolysate. For each hydrolysate, 1 fermentation run was conducted. The fermentation temperature was set at 30°C, pH at 5.0 by adding 2 M KOH or 1 M H₂SO₄, dissolved oxygen at 0 by flushing 0.5 l/min N₂ continuously. The fermentation began at the point of inoculation. The inoculum was prepared by harvesting the cells from 50 ml preculture and re-suspending the cells in 50 ml hydrolysate. Together with inoculum, 2 ml Tween 80-Ergosterol stock were added into the fermentor. The Tween 80-Ergosterol stock contained 5.0 g/l Ergosterol and 210.0 g/l Tween 80, which were dissolved in 95% ethanol. The whole fermentation process was monitored by continuously measuring the CO₂ percentage in the off-gas. The fermentation was considered finished when the CO₂ percentage value is 0 for 10 hours. During the fermentation process, samples were taken every 60 min or 99 min. The auto-samples were directly cooled to 4°C and later stored at 0°C.

Fermentation sample analysis

The monomer sugar concentrations in the hydrolysates were determined with DIONEX ICS 3000, equipped with CarboPac PA20 carbohydrate column and plused amperometric detector. The column was operated at 30°C, with 7.5 mM NaOH as eluent, and the flow rate was 0.5 ml/min.

The optical density (OD), glucose and ethanol concentrations of the fermentation auto-samples were determined using ROCHE Cobas Mira Plus. Vortex was performed to each individual sample to reach a homogeneous cell distribution before measuring optical density at wavelength 600 nm. After optical density measurement, the samples were centrifuged at 4000 rpm for 15 min and the suspension was used for glucose and ethanol measurements. Glucose concentration was determined enzymatically, by adding reagent Glucose HK CP, purchased from ABX Pentra, and measuring formed NADH amount at wavelength 340 nm. The ethanol assay was performed by using NAD and aldehyde dehydrogenase in 0.4 M KH₂PO₄ buffer as the first reagent and alcohol dehydrogenase as the second reagent, and measuring NADH concentration at wavelength 340 nm (adapted from BIOCHEMICA © protocols).

For each fermentation, 5 auto-samples were selected, representing the following time points: directly after inoculation, end of lag-phase, growth phase, end of growth phase and stationary phase. The concentrations of formic acid and acetic acid of these samples were measured with DIONEX ICS 3000, equipped with IonPac ICE-AS6 ion-exclusion column and suppressed conductivity detector. The column was operated at 30°C, with 1.6 mM perfluorobutyric acid as eluent, and the flow rate was 1.0 ml/min.

The concentrations of furfural, furfuryl alcohol, 2-furanmethanol acetate, levulinic acid, benzoic acid, syringic acid, HMF, 4-hydroxybenzaldehyde and vanillin were analyzed with EC-GC-MS method. The method was conducted as follows. NaOH solution was added to 0.5 ml hydrolysate to bring the mixture pH above 10. Into the mixture, the labeled internal standard containing leucine-D3, succinic acid-D4 and cinnamic acid-D5 in pyridine was added. 300 µl ethanol and the injection standard containing difluorobiphenyl and dicyclohexylphthalate in pyridine were also added. The formation of the ethylesters was done by two rounds of adding 40 µl ethyl chloroformate then shaking vigorously by hand for 15 seconds. The reaction was stopped by adding 750 µl dichloromethane and 500 µl of 1 M bicarbonate buffer. The formed derivatives were extracted to the dichloromethane phase by shaking the mixture for 20 seconds. The dichloromethane phase was then transferred to another vial and dried with sodium sulfate. The dried dichloromethane phase was transferred to an auto-sample vial. The measurement was carried out by 1 µl splitless injection in the PTV injector of the AGILENT 7890A GC with AGILENT 5975C mass spectrometer as detector. A DB-1 30 m x 0.32 mm x 1 µm analytical column was used for the separation of the analytes.

Inhibitory effects test

The inhibitory effects of the selected compounds were examined by using growth tests in BIOSCREEN C Analyzer, LABSYSTEMS OY, Helsinki, Finland, as described in Zha *et al.* [20]. The compounds were added into mineral medium with 20 g/l glucose and 2 different hydrolysates, Oak-PAA and Bag-CA. The concentrations added were based on the highest levels detected in all hydrolysates, which are marked in bold in Table 5. The media pH was adjusted to 5.0±0.5 with either 3 M H₂SO₄ or 6 M KOH before inoculation. The tests were carried out in triplicates.

Results and Discussion

Biomass hydrolysates composition

Sugar cane bagasse and oak saw dust were chosen as the biomass for this study because they represent two distinct categories of biomass type, namely grass like and wood. More importantly, in a previous study, where the growth of the model yeast was screened in 24 different hydrolysates, bagasse and oak hydrolysates showed the largest diversity [20].

Both bagasse and oak were treated with the 4 different hydrolysate preparation methods. The resulting 8 hydrolysates were analyzed on their monosaccharide compositions, as shown in Table 1. Glucose and xylose were the major monomer sugars in all 8 hydrolysates, and glucose had an approximately two fold higher concentration compare to xylose. Small amounts of galactose and arabinose were detected in both bagasse and oak hydrolysates, while in oak hydrolysates, also low levels of mannose were found.

Table 1 Monomer sugar concentrations of the 8 hydrolysates (g/l)

Hydrolysates	glucose	xylose	galactose	arabinose	mannose
Bag-MA	57.9	33.6	0.1	3.3	0.0
Bag-DA	66.5	29.7	0.7	2.0	0.0
Bag-PAA	67.8	31.2	0.1	0.7	0.0
Bag-CA	107.3	62.9	1.5	4.2	0.0
Oak-MA	47.4	24.6	0.2	0.1	0.0
Oak-DA	42.1	25.0	1.0	0.6	1.3
Oak-PAA	61.0	28.1	0.2	0.1	0.0
Oak-CA	90.9	42.3	2.5	1.8	4.2

Hydrolysate fermentation

For each of the 8 hydrolysates in Table 1, a batch fermentation was carried out with the model yeast *S. cerevisiae* CEN.PK 113-7D. The fermentation performance was determined by measuring the optical density, glucose concentration and ethanol concentration of the samples taken during the whole fermentation process. The growth of the yeast varied in these 8 hydrolysates, as shown in Figure 1, as well as the glucose consumption (Figure 2) and the ethanol production (Figure 3).

Growth characteristics

The growth of the yeast cells in a fermentation process was monitored by measuring the optical density of the time samples at wavelength 600 nm. The optical density of a time sample was calculated by deducting the measured optical density value by the time-0 optical density value: $OD_{t-s} = OD_{t-m} - OD_{t-0}$.

The growth curves of the model yeast in 8 different hydrolysates and in mineral medium are shown in Figure 1. The growth of the model yeast in mineral medium in this study was highly comparable to the growth reported by Kuyper *et al.* [22]. By comparing the growth in hydrolysates and in mineral medium, it can be seen that the growth in all hydrolysates were negatively affected. This was mainly shown as slower growth, longer lag-phase and lower OD yield. It can be seen that the growth of the model yeast was similar in the hydrolysates prepared with the same method, indicating that the hydrolysate performance was mainly dependent on the pretreatment-hydrolysis method. The hydrolysates prepared by mild alkaline method resulted in the shortest lag-phase and relatively high growth rate, while the concentrated acid method prepared hydrolysates had the longest lag-phase and slower growth. The performance of the hydrolysates made by dilute acid and peracetic acid methods was in between the other two (Figure 1).

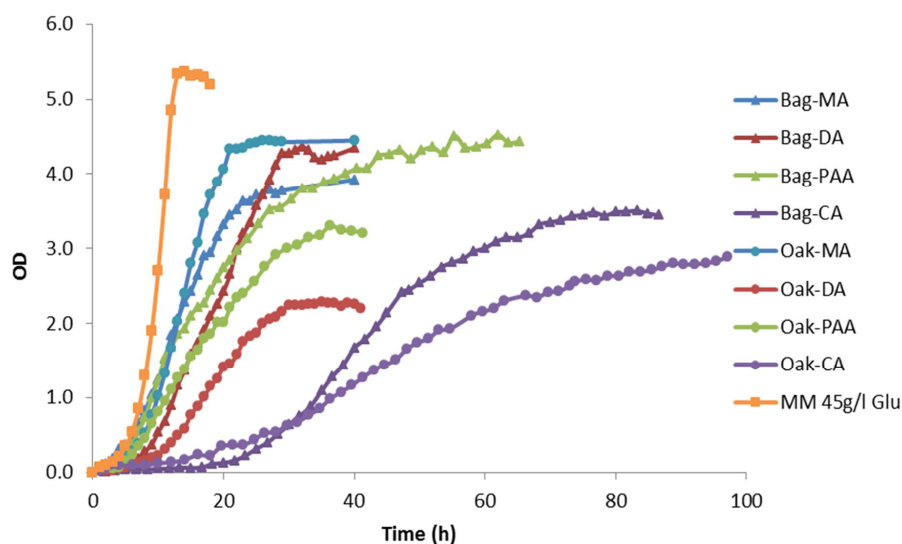


Figure 1 Growth curves of the model yeast, *Saccharomyces cerevisiae* CEN.PK113-7D, in 8 different hydrolysates. Bag: bagasse, Oak: oak sawdust; MA: mild alkaline, DA: dilute acid, PAA: alkaline/peracetic acid, CA: concentrated acid.

To quantitatively compare these 8 fermentations, lag-phase, growth rate and OD yield were defined as parameters to describe the characteristics of the fermentation performance (Table 2). By comparing the growth rates in bagasse and oak hydrolysate prepared with the same method, it is noticed that only when prepared with mild alkaline method, oak hydrolysate was with a higher growth rate than bagasse. This indicates that hydrolysates prepared from bagasse are, in general, less inhibitory than those prepared from oak. Probably, if the pretreatment method was mild and the biomass structure was relatively more difficult to break down, such as oak, there would be little inhibitory compounds released or formed. In this case, the generated hydrolysate would be less toxic. As shown, different from the growth in mineral medium, the growth in some hydrolysates slowed down several hours after the growth started (Figure 1). This phenomenon is most illustrative in PAA hydrolysates, the model yeast started with fast growth, but the growth rate dropped at a specific point, 14 h for Bag-PAA and 20 h for Oak-PAA. The 2 different growth rates shown in Table 2 are before and after the rate drop, respectively. A possible explanation for the phenomenon is that the amount of essential nutrients in these hydrolysates was limited, which could only support the growth in the first several hours. To continue growth, the yeast had to use different nutrients that were less efficient, which caused the growth rate to slow down. This explanation was consistent with the fact that the growth slowed down particularly in PAA hydrolysates. As during PAA pretreatment, 2 washing steps were involved, which removed dissolved nutrients at that moment. This possibly caused nutrient limitation in PAA hydrolysates, which lead to the reduction of the growth rate. In agreement with this, the analysis results of the hydrolysate fermentation time samples revealed that most of the amino acids present in the hydrolysates were consumed during the fermentation process (see Chapter 5).

As far as the OD yield is considered, it seems that it was related to lag-phase and growth rate, namely, long lag-phase and/or slow growth corresponded to low OD yield. For instance, the lowest OD yield was of Bag-CA and Oak-CA hydrolysate fermentations, which had the longest lag-phase (17 h) and lowest growth rate (0.035), respectively (Table 2). The differences in OD yield indicate that the yeast cells spent a higher percentage of the total energy on maintenance in hydrolysates, which maybe the result of overcoming inhibitory effect and/or using less efficient nutrients.

Inhibitors in biomass hydrolysates during fermentation processes

Table 2 Growth characteristics of the model yeast in the 8 hydrolysates

Hydrolysates	lag-phase ¹ (h)	growth rate ²	OD yield ³ (OD/ g glucose)
Bag-MA	1	0.169	0.067
Bag-DA	6	0.187	0.069
Bag-PAA	4	0.243 / 0.052	0.064
Bag-CA	17	0.085	0.033
Oak-MA	2	0.200	0.100
Oak-DA	5	0.125	0.060
Oak-PAA	4	0.190 / 0.058	0.057
Oak-CA	10	0.035	0.032
MM 45 g/l Glucose	1	0.306	0.132

¹: lag-phase is defined as the time needed to reach 2% of the maximum OD;

²: growth rate is calculated as the slope of the linear part of the *logOD* vs. *time* plot;

³: OD yield is calculated by dividing maximum OD by the amount of glucose consumed in the whole fermentation process.

Table 3 Glucose consumption and ethanol production results of the 8 fermentations

Hydrolysates	initial glucose concentration ¹ (g/l)	maximum ethanol concentration ² (g/l)	ethanol yield ³ (g/g)	maximum glucose consumption rate ⁴ (g/l/h)	maximum ethanol production rate ⁵ (g/l/h)
Bag-MA	58.8	22.7	0.39	3.3	1.3
Bag-DA	63.3	24.2	0.38	3.1	1.2
Bag-PAA	69.8	24.0	0.34	1.6	0.6
Bag-CA	104.4	34.9	0.33	2.3	0.9
Oak-MA	44.4	19.5	0.44	3.1	1.4
Oak-DA	38.2	15.3	0.40	2.0	0.8
Oak-PAA	58.0	24.2	0.42	2.4	1.0
Oak-CA	88.9	22.0	0.35	1.1	0.4
MM 45 g/l Glucose	42.5	16.0	0.38	5.5	2.0

¹: the glucose concentration of the time-0 fermentation sample;

²: the highest ethanol concentration among all fermentation samples;

³: maximum ethanol concentration divided by the total amount of glucose consumed;

⁴: the slope of the linear part of the *glucose concentration* vs. *time* plot;

⁵: the slope of the linear part of the *ethanol concentration* vs. *time* plot.

Glucose consumption and ethanol production profile

The 8 different hydrolysates differ in their initial glucose concentrations due to the diverse biomass types and hydrolysate preparation methods used, see Table 3. To analyze the effect of initial glucose concentration on growth, mineral medium with glucose concentration 20 g/l, 40 g/l, 60 g/l and 80 g/l were used to test the model strain in Bioscreen. The results showed that with glucose concentration at this range, the model yeast did not show any difference in their growth, in terms of the 3 parameters listed in Table 2 (data not shown). So it was assumed that the performance differences of the model strain in these hydrolysates were not caused by the variation of initial glucose concentration.

To present the glucose consumption and ethanol production of the 8 hydrolysate fermentations in a comparable manner, both glucose and ethanol concentrations were expressed as a percentage, with maximum value set as 100% and 0 g/l set as 0%, as shown in Figure 2. It can be seen that the hydrolysates prepared with the same method had similar pattern in both glucose consumption and ethanol production curves. This is consistent with the observation of growth curves, confirming that the hydrolysate performance was mainly determined by pretreatment-hydrolysis method rather than biomass type.

The maximum ethanol concentration and ethanol yield of the 8 different hydrolysates are listed in Table 3. The highest ethanol concentration in all fermentations was 34.9 g/l of Bag-CA hydrolysate, while the highest ethanol yield was of Oak-MA hydrolysate, 0.44 g ethanol per g glucose. This yield was 86% of the theoretical ethanol yield on glucose [23]. Furthermore, also the maximum glucose consumption rate and the maximum ethanol production rate of the 8 fermentations are compared in Table 3. It can be seen that these two rates were closely related, in general, the faster the glucose was consumed the quicker the ethanol was produced, in other words, the ethanol yields of these 8 fermentations were quite similar. Additionally, these ethanol yields were not only similar to each other, but also comparable to the one of mineral media fermentation. This suggests that ethanol yield was only slightly influenced by the inhibitory compounds in the hydrolysates, which agrees with the effect of furans and phenols on yeasts performance [17]. Since the effects of inhibitory compounds in hydrolysates were mainly on growth rate, OD yield and glucose consumption rate, it is practical to use these parameters as indicators for studying the hydrolysate inhibitory effect.

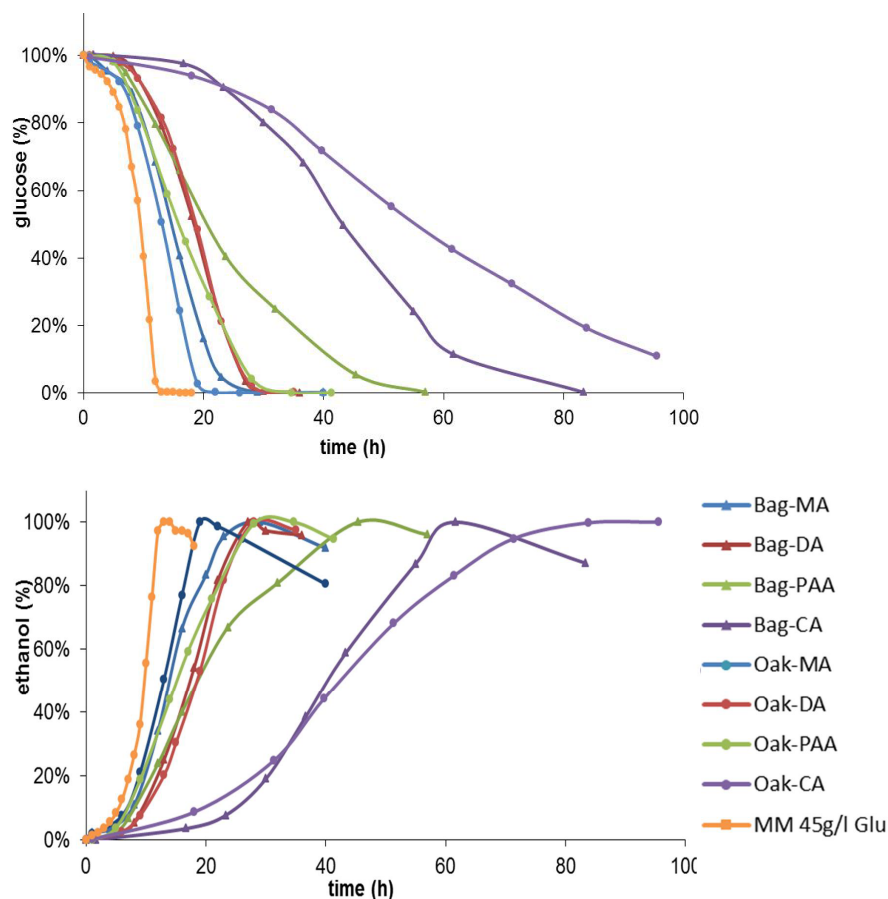


Figure 2 Glucose consumption (top) and ethanol production (bottom) curves of the 8 fermentations in percentage. Bag: bagasse, Oak: oak sawdust; MA: mild alkaline, DA: dilute acid, PAA: alkaline/peracetic acid, CA: concentrated acid.

Selection and quantification of inhibitory compounds in hydrolysate samples

To identify the role of specific hydrolysate inhibitors on fermentation performance, exo-metabolomics analysis was carried out. As a first step in interpreting this type of analysis, a group of 11 compounds were selected and quantified in their hydrolysate fermentation samples. Based on quantification results, the dynamics of these compounds during a fermentation process could be determined. This analysis will also allow tests of inhibitory effects of these compounds at concentrations present in the hydrolysates. The selected compounds were formic acid, acetic acid, levulinic acid, furfural, furfuryl alcohol, 2-

furanmethanol acetate, HMF, vanillin, syringic acid, 4-hydroxybenzaldehyde and benzoic acid. The selection was made based on data reported in literature and observations made in our preliminary studies (Table 4).

Furfural and HMF are both furan compounds, and were identified as potential inhibitors in biomass hydrolysates [17,24,25]. Furfural was pointed to be the key inhibitor in hydrolysates by Heer *et al.* in 2008 [15]. It was known that furfural was converted to furfuryl alcohol by yeast as a detoxification mechanism [26]. The inhibitory effect of furfural was reported as increasing lag-phase [15] and reducing specific growth rate [6].

Formic, acetic and levulinic acid are the weak acids formed in most of the biomass hydrolysis preparation process [1,5] and their inhibitory effects and mechanism on yeasts have been studied in the past several years [6,27]. It was suggested that these weak acids reduce yeast growth and ethanol yield by causing intracellular anion accumulation, which is pH dependent [1,16]. Recently, Sanda *et al.* reported that both formic acid and acetic acid affect the utilization of xylose in recombinant xylose-fermenting strain [27].

Vanillin, syringic acid and 4-hydroxybenzoic acid were characterized phenolic compounds in hydrolysates [5,17]. The inhibitory effects of 4-hydroxybenzoic acid have been studied with several different yeast strains [18], the study concluded that the compound showed little effects on the yeasts used. In this study, the two closely related compounds, benzoic acid and 4-hydroxybenzaldehyde were chosen to be quantified in hydrolysates and tested on their effects on the growth of the model yeast.

As summarized in Table 4, formic acid, acetic acid, levulinic acid, furfural, HMF, vanillin, syringic acid, 4-hydroxybenzaldehyde and benzoic acid are the characterized degradation products in biomass hydrolysates; while furfuryl alcohol and 2-furanmethanol acetate are the possible conversion products of furfural and/or HMF.

These selected compounds were analyzed and quantified in both hydrolysates and their fermentation samples. For each fermentation, 5 samples were chosen according to the following criteria: (1) directly after inoculation, (2) end of lag-phase, (3) growth phase, (4) end of growth phase, (5) stationary phase. The concentrations of these selected compounds in the fermentation samples are listed in Table 5.

Formic acid and acetic acid were detected in all hydrolysates and their fermentation samples. In general, acetic acid concentrations were 10-15 times higher than that of formic acid. The

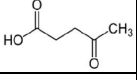
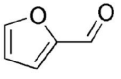
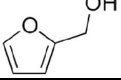
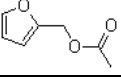
highest concentrations of these two acids were found in CA hydrolysates, 0.57 g/l of formic acid and 8.0 g/l of acetic acid. These concentrations are comparable with the ones detected in acid pretreated spruce and bagasse hydrolysates [14]. During fermentation processes, no obvious consumption of either acid was observed, though both fluctuated slightly. Unlike formic acid and acetic acid, levulinic acid was only present in CA hydrolysates with a concentration of 1.2 g/l, without a decrease during fermentation.

Vanillin, syringic acid and 4-hydroxybenzaldehyde were present mainly in MA and DA hydrolysates although in rather low amounts, 30-50 mg/l. These concentrations are similar with those detected previously [1,5]. In contrast to syringic acid, both vanillin and 4-hydroxybenzaldehyde decreased during fermentation, suggesting their conversion or consumption. Due to its presence as a preservative in the enzyme cocktails of about 2.0 g/l [28], benzoic acid was detected in all enzymatic hydrolyzed hydrolysates, namely MA, DA and PAA hydrolysates, with a similar concentration of 150 mg/l. During the whole fermentation process, the level of benzoic acid did not change (Table 5). Surprisingly, in PAA treated bagasse hydrolysate, benzoic acid was apparently converted into its corresponding ethanol before the starting of the fermentation. It is unclear why this conversion took place specifically in Bag-PAA hydrolysate.

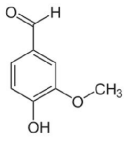
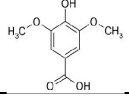
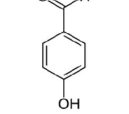
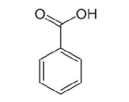
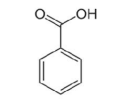
It can be seen in Table 5 that furfural was found at considerable levels in CA hydrolysates, and at low amounts in DA hydrolysates. In both DA and CA hydrolysate fermentations, the furfural concentration rapidly decreased at the onset of the fermentation until levels of about 30 mg/l, with exception of Bag-DA hydrolysate. Correspondingly, the concentration of furfuryl alcohol increased in the same time frame. This suggests that furfural was converted to furfuryl alcohol in the lag-phase of the fermentation, which agrees with the report of Palmqvist *et al.* [26]. Different from the observation in *A. niger* [29], furfuryl alcohol was not further converted into furoic acid. Furfural was also found in Bag-PAA hydrolysate at 30 mg/l, but it was not converted during the whole fermentation process. Similar to furfural, HMF was also found in DA and CA hydrolysates, but with a much lower amount. The HMF concentration reduced gradually in both lag-phase and growth-phase of these hydrolysate fermentations.

Interestingly, 2-furanmethanol acetate showed similar pattern as furfuryl alcohol, the compound increased with the decrease of furfural and HMF in the fermentation lag-phase (Table 5). Based on the structure of 2-furanmethanol acetate, it is suspected that the compound was the reaction product of furfuryl alcohol and acetic acid. From this result, we suggest that furfuryl alcohol was possibly partially converted to 2-furanmethanol acetate by reacting with acetic acid.

Table 4 A summary of the selected compounds: their concentrations detected in various hydrolysates, and the concentrations at which inhibitory effects were shown on *S.cerevisiae*.

compound	structure	hydrolysates		concentrations in hydrolysates		concentrations showed effects on <i>S.cerevisiae</i>	
		biomass	preparation method	mg/l	ref	mg/l	ref
formic acid		Corn stover	Acid/temperature	130-310	[12]	4000	[16]
		Spruce/bagasse	Acid/temperature	600-800	[14]	2700	[30]
		Hardwood chips	Autohydrolysis/temperature	4000-4600	[31]		
		Corn stover	Steam explosion	6800	[16]		
acetic acid		Bagasse	Acid hydrolysis	2400	[32]	5000	[32]
		Bagasse	Enzyme hydrolysis	2100	[32]		
		Wheat straw	Acid hydrolysis	1300	[32]	6000	[16]
		Wheat straw	Enzyme hydrolysis	900	[32]	7500	[30]
		Corn stover	Acid hydrolysis	2300	[32]	>10000	[33]
		Willow wood	Acid hydrolysis	2200	[32]		
		Yellow polar wood	Organosolv	900-4900	[34]		
		Spruce/bagasse	Acid/temperature	3100-5200	[14]		
		Corn stover	Acid/temperature	2270-3740	[12]		
		Hardwood chips	Autohydrolysis/temperature	4500-5800	[31]		
levulinic acid		Spruce/bagasse	Acid/temperature	200-300	[14]		
		Corn stover	Acid/temperature	130-410	[12]		
furfural		Yellow polar wood	Organosolv	0.2-35.2	[34]	1000	[32]
		Wheat straw	Alkaline/oxidation	0-146*	[5]	>800	[33]
		Bagasse	Acid hydrolysis	410	[32]		
		Wheat straw	Acid hydrolysis	270	[32]	>4000	[16]
		Corn stover	Acid hydrolysis	510	[32]		
		Willow wood	Acid hydrolysis	500	[32]		
		Hardwood chips	Autohydrolysis/temperature	510-780	[31]		
		Corn stover	Acid/temperature	570	[12]		
		Corn stover	Steam explosion	710	[16]		
		Wheat straw	Acid steam explosion	480-680	[15]		
		Spruce	Acid steam	1100	[15]		
		Barley straw	Acid steam	2880	[15]		
		Spruce/bagasse	Acid/temperature	600-1200	[14]		
furfuryl alcohol		N.A.					
2-furan methanol acetate		N.A.					

Inhibitors in biomass hydrolysates during fermentation processes

HMF		Wheat straw	Alkaline/oxidation	0-16*	[5]	1000	[32]
		Yellow polar wood	Organosolv	0-56.5	[34]	>3000	[33]
		Hardwood chips	Autohydrolysis/temperature	80-130	[31]	>4000	[16]
		Bagasse	Acid hydrolysis	70	[32]		
		Wheat straw	Acid hydrolysis	60	[32]		
		Corn stover	Acid hydrolysis	100	[32]		
		Willow wood	Acid hydrolysis	140	[32]		
		Corn stover	Acid/temperature	50-140	[12]		
		Wheat straw	Acid steam explosion	177-277	[15]		
		Spruce	Acid steam	2140	[15]		
		Barley straw	Acid steam	996	[15]		
		Corn stover	Steam explosion	560	[16]		
vanillin		Wheat straw	Alkaline/oxidation	8-96*	[5]	4000	[16]
		Wheat straw	Acid steam explosion	91-122	[15]		
		Spruce	Acid steam	152	[15]		
		Barley straw	Acid steam	106	[15]		
		Willow	Dilute acid	430	[1]		
syringic acid		Corn stover	Steam explosion	4000	[16]		
4-hydroxy benzaldehyde		Wheat straw	Wet oxidation	22	[1]		
		Wheat straw	Alkaline/oxidation	6-52*	[5]		
benzoic acid		Wheat straw	Wet oxidation	21	[1]		
		Willow	Dilute acid	10	[1]		
		Wheat straw	Alkaline/oxidation	12-59*	[5]		
benzoic acid		Corn stover	Steam explosion	900	[16]	2000	[16]

*: these values are expressed as g/100g straw; N.A.: Not Available.

Inhibitory effect of the selected compounds tested in mineral medium

The quantification results of the selected compounds in hydrolysates provided reference concentrations to test their inhibitory effects. For each compound, the highest concentration detected among all samples, marked as bold in Table 5, was used as the initial testing value. Based on initial test results, the concentrations were increased or decreased up to 5-10 folds for the actual test. The medium used here was mineral medium with 20 g/l glucose.

Table 5 The concentrations of the selected 11 compounds in the samples of the 8 hydrolysate fermentations (mg/l)

fermentation samples	formic acid	acetic acid	levulinic acid	furfural	furfuryl alcohol	2-furanmethanol acetate	HMF	vanillin	syringic acid	4-hydroxybenzaldehyde	benzoic acid
Bag-MA 1	93	1342	<25	<10	<10	<1	<8	34	7	42	139
Bag-MA 2	83	1190	<25	<10	<10	<1	<8	26	6	40	143
Bag-MA 3	77	1164	<25	<10	<10	<1	<8	3	6	5	143
Bag-MA 4	63	1026	<25	<10	<10	<1	<8	3	6	5	140
Bag-MA 5	56	978	<25	<10	<10	<1	<8	3	6	5	143
Bag-DA 1	184	1816	<25	30	21	<1	11	19	8	23	141
Bag-DA 2	173	1750	<25	<10	42	1	10	5	9	13	141
Bag-DA 3	165	1661	<25	<10	44	1	<8	4	9	6	145
Bag-DA 4	147	1535	<25	<10	46	1	<8	5	9	10	149
Bag-DA 5	153	1520	<25	<10	48	2	<8	5	9	10	153
Bag-PAA 1	16	241	<25	27	<10	<1	<8	3	<5	4	<10*
Bag-PAA 2	16	273	<25	27	<10	<1	<8	3	<5	4	<10*
Bag-PAA 3	0	128	<25	28	<10	<1	<8	3	<5	6	<10*
Bag-PAA 4	0	28	<25	29	<10	<1	<8	3	<5	9	<10*
Bag-PAA 5	0	48	<25	30	<10	<1	<8	3	<5	10	<10*
Bag-CA 1	568	7234	1148	579	97	12	57	<1	<5	<1	<10
Bag-CA 2	528	7049	1159	32	750	98	29	<1	<5	<1	<10
Bag-CA 3	552	6922	1206	28	730	99	<8	<1	<5	<1	<10
Bag-CA 4	533	6460	1297	30	739	97	<8	<1	<5	<1	<10
Bag-CA 5	534	6469	1314	29	747	99	<8	<1	<5	<1	<10
Oak-MA 1	133	1198	<25	<10	<10	<1	<8	13	7	<1	128
Oak-MA 2	135	1310	<25	<10	<10	<1	<8	3	7	<1	130
Oak-MA 3	173	1679	<25	<10	<10	<1	<8	2	6	<1	129
Oak-MA 4	151	1547	<25	<10	<10	<1	<8	2	7	3	132
Oak-MA 5	154	1562	<25	<10	<10	<1	<8	2	7	3	130
Oak-DA 1	330	3490	<25	60	30	1	17	14	46	<1	157
Oak-DA 2	318	3420	<25	28	94	6	14	4	44	<1	159
Oak-DA 3	302	3228	<25	26	106	7	8	3	46	<1	163
Oak-DA 4	282	3003	<25	26	107	7	<8	3	47	<1	162
Oak-DA 5	278	3051	<25	27	112	7	<8	3	45	<1	164
Oak-PAA 1	36	560	<25	<10	<10	<1	<8	4	<5	<1	144
Oak-PAA 2	39	592	<25	<10	<10	<1	<8	3	<5	<1	139
Oak-PAA 3	33	603	<25	<10	<10	<1	<8	3	<5	<1	141
Oak-PAA 4	34	474	<25	<10	<10	<1	<8	3	<5	<1	144
Oak-PAA 5	37	500	<25	<10	<10	<1	<8	3	<5	<1	141
Oak-CA 1	492	7994	1082	431	95	12	55	<1	7	<1	<10
Oak-CA 2	454	7877	1119	50	603	84	37	<1	6	<1	<10
Oak-CA 3	499	7869	1198	34	640	93	10	<1	7	<1	<10
Oak-CA 4	479	7591	1324	33	684	97	<8	<1	7	<1	<10
Oak-CA 5	509	7901	1360	34	698	100	<8	<1	7	1	<10

*: Instead of benzoic acid, benzylalcohol peak was found in Bag-PAA hydrolysate samples. Since benzylalcohol and several unknown peaks that may relate to benzoic acid were unique to Bag-PAA samples, it is possible that the benzoic acid presented in Bag-PAA hydrolysate was converted to several related compounds.

Furfural and benzoic acid

Furfural and benzoic acid clearly affected the growth of the model yeast at concentrations as were present in the hydrolysates, see Figure 3. The inhibitory effect of furfural displayed mainly as longer lag-phase. The lag-phase started to elongate already at a very low furfural concentration, 0.06 g/l, and increased from about 5 hours to 15 hours at a concentration of 0.6 g/l, which was about the concentration in CA hydrolysates (Table 5). It was observed that furfural concentration reduced mainly in the lag-phase during the fermentation process. This suggests that the presence of furfural obstructed the growth of the model yeast, and only when its concentration in the medium dropped below a threshold, the growth could start. It is suspected that this threshold was 0.03 g/l, as the growth commenced in most hydrolysates at this furfural concentration.

Unlike furfural, the inhibitory effect of benzoic acid was lowering the growth rate and final optical density level of the model yeast, as shown in Figure 3. At the concentration of 0.16 g/l, which was also the highest benzoic acid concentration detected in the hydrolysates, the growth rate decreased more than 60% compared to the reference medium, and the final optical density level dropped from 1.28 to 0.65. It seems that the inhibitory effect of benzoic acid was closely related to its concentration present in the medium.

The combination effect of furfural and benzoic acid on the model yeast seems to be addable, as shown in Figure 4. That is to say, the lag-phase and the growth rate in the medium with both furfural and benzoic acid were very similar to which in the medium with furfural and with benzoic acid, respectively. Apparently, the inhibition by furfural and benzoic acid takes place at different stages of the growth process, namely, furfural before growth started and benzoic acid after. This indicates the inhibitory mechanisms of furfural and benzoic acid were different.

5-hydroxymethyl furfural (HMF)

HMF was frequently mentioned as a inhibitor next to furfural in hydrolysates [24,35], but seemed to have a milder inhibitory effect [7,23]. The highest HMF concentration present in the 8 hydrolysates in this study was 0.06 g/l, which did not give any effect on growth when added into mineral medium (data not shown). The inhibitory effect of HMF only became visible when its concentration reached 0.6 g/l and enhanced strongly when it was increased to 1.2 g/l, see Figure 5. In contrast to furfural, the inhibitory effect of HMF was mainly shown as slower growth and lower final optical density, next to elongated lag-phase.

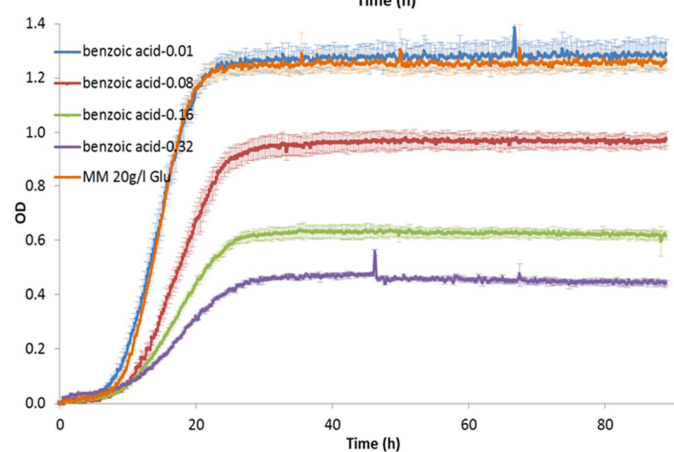
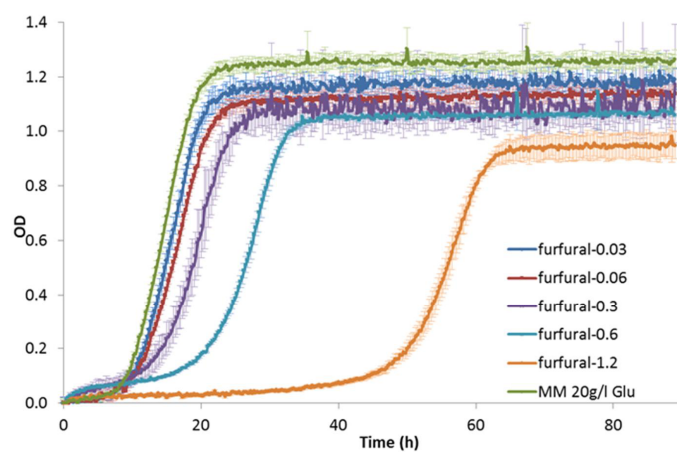


Figure 3
Inhibitory effect of furfural (top) and benzoic acid (bottom) on the model yeast in MM with 20 g/l glucose. The values in the label are the compound concentrations in media (g/l).

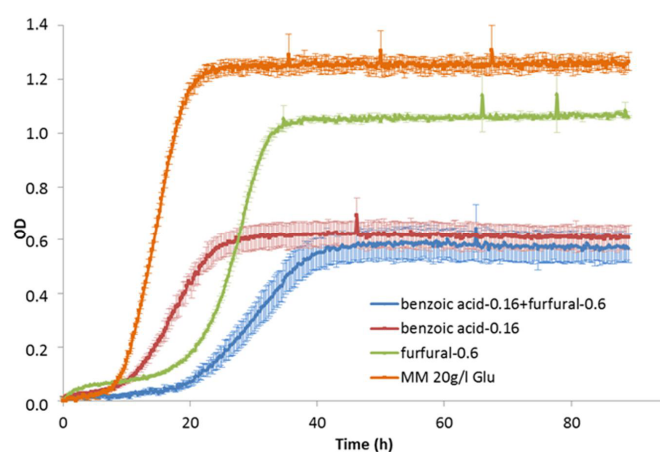


Figure 4
Combined inhibitory effects of furfural and benzoic acid on the model yeast in MM with 20 g/l glucose. The values in the label are the compound concentrations in media (g/l).

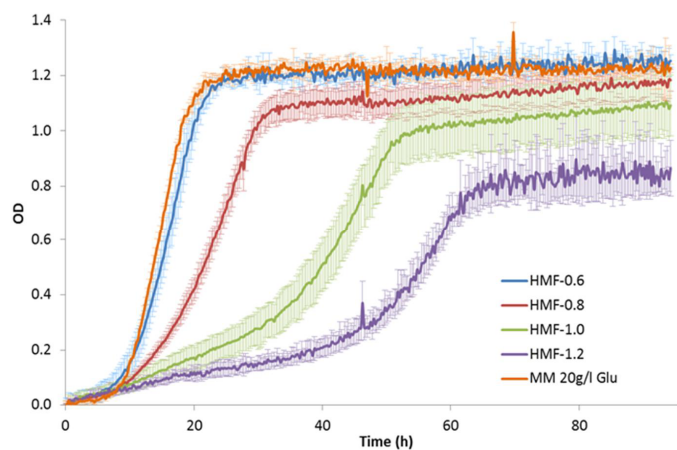


Figure 5
Inhibitory effects of HMF on the model yeast in MM with 20 g/l glucose. The values in the label are the compound concentrations in media (g/l).

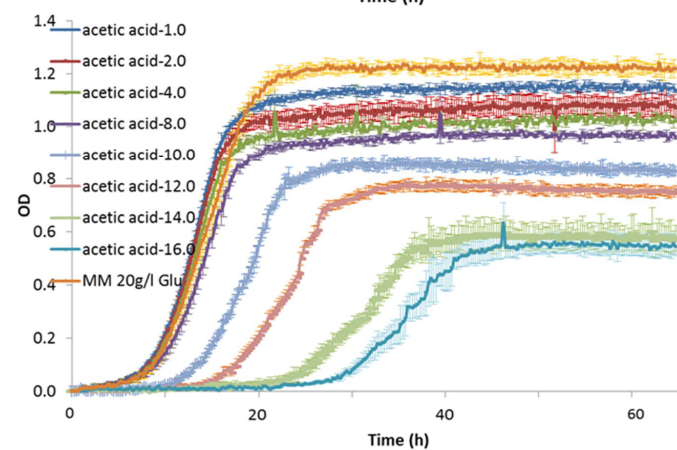
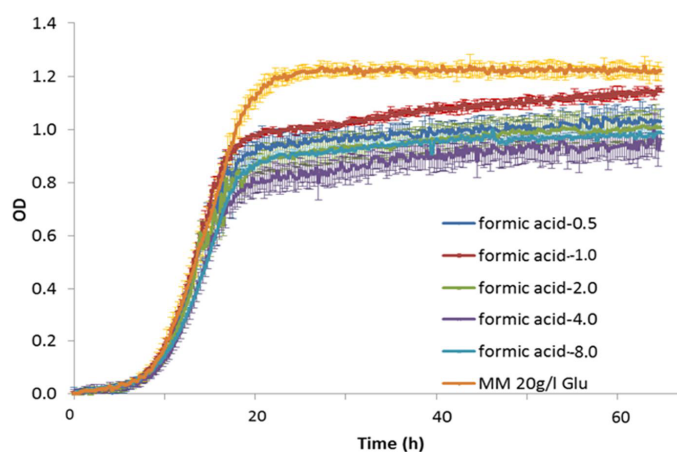


Figure 6
Inhibitory effects of formic acid (top) and acetic acid (bottom) on the model yeast in MM with 20 g/l glucose. The values in the label are the compound concentrations in media (g/l).

Formic acid and acetic acid

The presence of formic acid at 0.5 g/l had little effect on the growth rate of the model yeast, but reduced final optical density slightly, see Figure 6. Increasing the formic acid concentration in mineral medium from 0.5 g/l to 8.0 g/l hardly enhanced this effect. The influence of acetic acid on the growth of the model yeast was similar to formic acid up to 8.0 g/l. Only when acetic acid concentration exceeded 8.0 g/l, both growth rate and final optical density were reduced significantly, and the lag-phase was clearly elongated, similar as described previously [16,36]. The highest concentrations of formic acid and acetic acid found in hydrolysates were 0.6 g/l and 8.0 g/l, respectively (Table 5). At these concentrations, the inhibitory effects of both acids were only marginal. To reach severe inhibitory effect, the level of formic acid needs to be enhanced by more than 13 folds, while the acetic acid levels are close to the inhibiting concentration. From this point of view, acetic acid is more likely to be an inhibitor in hydrolysates than formic acid.

Levulinic acid, syringic acid, vanillin and 4-hydroxybenzaldehyde

Though reported as inhibitory compounds in the hydrolysates [17], the inhibitory effects of levulinic acid, syringic acid, vanillin and 4-hydroxybenzaldehyde were only marginal. The four compounds were tested by adding them individually into mineral medium according to their highest concentrations detected in the hydrolysates. Only at 10 fold increased levels, effects became visible for these compounds, although still mild, see Figure 7.

Among the 4 compounds, vanillin with concentration 0.5 g/l gave the most inhibitory effect, which was mainly on lag-phase. Levulinic acid showed similar effect on growth as formic and acetic acid, but at a much higher concentration, 15.0 g/l. Since these 4 compounds only started to affect the growth of the model strain at a 10-fold concentration compare to their highest concentrations in hydrolysates, they are thought to be none-inhibitory in the hydrolysates.

Inhibitory effect test in hydrolysates

The inhibitory effect tests of the selected compounds in mineral medium suggested that furfural and benzoic acid were the most important inhibitory compounds. They affected the growth of the model strain considerably at their concentrations presented in the hydrolysates. As hydrolysates have a total different matrix compare to mineral medium, it is interesting to test if these two compounds display similar inhibitory effect in hydrolysates.

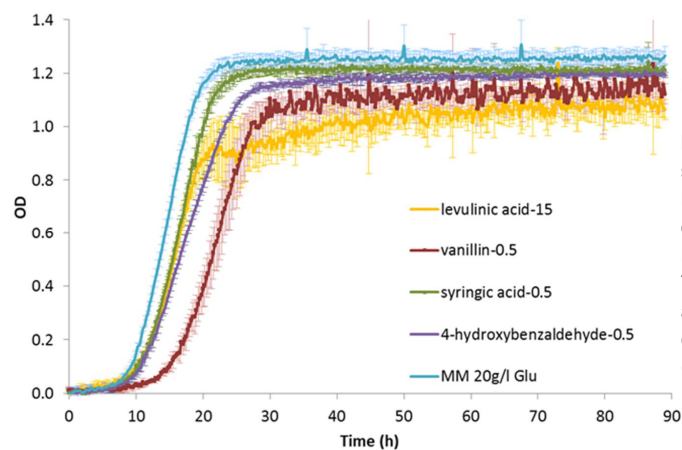


Figure 7
Inhibitory effects of levulinic acid, vanillin, syringic acid and 4-hydroxybenzaldehyde on the model yeast in MM with 20 g/l glucose. The values in the label are the compound concentrations in media (g/l).

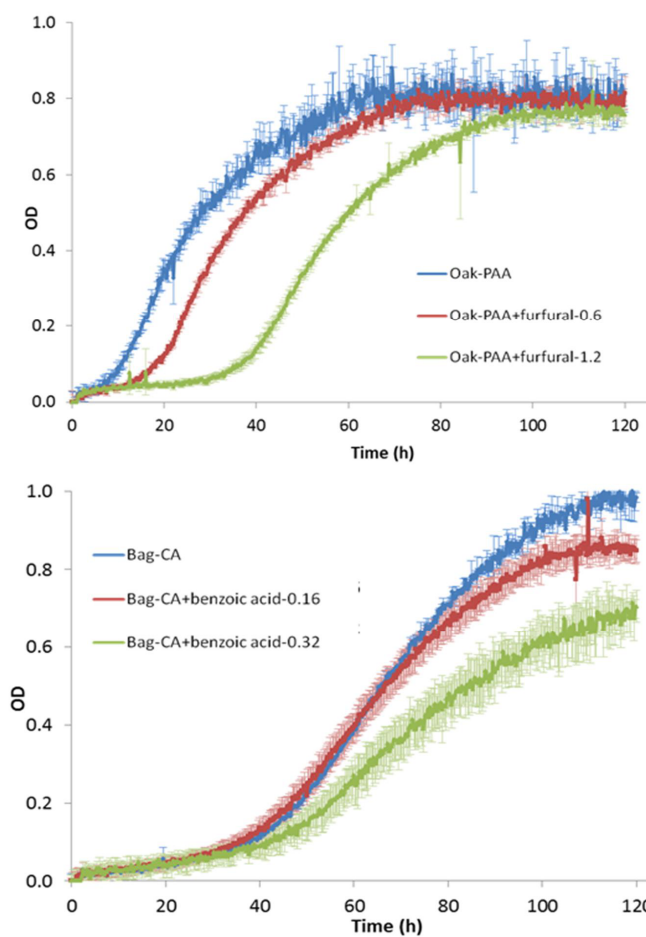


Figure 8
Inhibitory effects of furfural in Oak-PAA hydrolysate (top) and benzoic acid in Bagasse-CA hydrolysate (bottom), on the model yeast. The values in the label are the compound concentrations in media (g/l).

Furfural and benzoic acid were tested in Oak-PAA and Bagasse-CA hydrolysate, respectively. The reason for using these two hydrolysates was that Oak-PAA was a furfural free hydrolysate and Bagasse-CA was benzoic acid free. The testing concentrations of both compounds were one and twofold of their highest levels in the hydrolysates, 0.6 g/l and 1.2 g/l of furfural, and 0.16 g/l and 0.32 g/l of benzoic acid (Figure 8).

As shown in Figure 8, the inhibition effects of both compounds were also observed in hydrolysates and similar to those seen in mineral medium. The presence of furfural lengthened the lag-phase of the growth in Oak-PAA hydrolysate, and benzoic acid affected the growth rate and the final optical density level in Bag-CA hydrolysate. However, these effects were milder than in mineral medium, as by adding 1.2 g/l furfural in mineral medium, the lag-phase increased to 40 h, while the lag-phase enhanced to only 30 h when added into hydrolysate. For benzoic acid, in mineral medium the optical density dropped to half of that in reference medium when 0.16 g/l was present, while in hydrolysate the optical density decreased less than 10% (Figure 2, 8). These results indicate that the hydrolysate matrix buffers inhibitory effects. It can also be seen in Figure 8 that the pattern of the growth curve of both hydrolysates changed little by adding either furfural or benzoic acid. This suggests that the growth curve pattern of a hydrolysate is determined by the combined structure of most or all the compounds present in it.

Conclusion

This study showed that the fermentation performance of different hydrolysates varied in lag-phase, growth rate and biomass yield, as well as their composition as far as the selected compounds are considered. These differences among hydrolysates seem to be caused mainly by hydrolysate preparation method, and secondly by biomass type. The detection of the 11 selected compounds in fermentation samples revealed that the levels of most compounds changed during fermentation process. Remarkably, furfural was converted to furfuryl alcohol and possibly also 2-furan methanol acetate in the fermentation lag-phase. The toxicity test of the 11 selected compounds showed that furfural and benzoic acid exhibited clear inhibitory effects on model yeast *S.cerevisiae* CEN.PK 113-7D at their concentrations detected in hydrolysates, while the effects of acetic acid and HMF were minor, but enhanced dramatically at the increase of concentration.

Acknowledgement

This project was (co)financed by the Netherlands Metabolomics Centre (NMC) which is part of the Netherlands Genomics Initiative / Netherlands Organisation for Scientific Research.

References

1. Almeida JRM, Modig T, Petersson A, Hahn-Hägerdal B, Lidén G, Gorwa-Grauslund MF: Increased tolerance and conversion of inhibitors in lignocellulosic hydrolysates by *Saccharomyces cerevisiae*. *J. Chem. Technol. Biot.* **2007**, 82, 340-349.
2. da Silva AS, Inoue H, Endo T, Yano S, Bon EPS: Milling pretreatment of sugarcane bagasse and straw for enzymatic hydrolysis and ethanol fermentation. *Bioresource Technol.* **2011**, 101, 7402-7409.
3. Sun Y, Cheng J: Hydrolysis of lignocellulosic materials for ethanol production, A review. *Bioresource Technol.* **2002**, 83, 1-11.
4. Alvira P, Tomás-Pejó E, Ballesteros M, Negro MJ: Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis, A review. *Bioresource Technol.* **2010**, 101, 4851-4861.
5. Klinker HB, Ahring BK, Schmidt AS, Thomsen AB: Characterization of degradation products from alkaline wet oxidation of wheat straw. *Bioresource Technol.* **2002**, 82, 15-26.
6. Palmqvist E, Hahn-Hägerdal B: Fermentation of lignocellulosic hydrolysates. II, Inhibitors and mechanisms of inhibition. *Bioresource Technol.* **2000**, 74, 25-33.
7. Martin C, Jönsson LJ: Comparison of the resistance of industrial and laboratory strains of *Saccharomyces* and *Zygosaccharomyces* to lignocellulose-derived fermentation inhibitors. *Enzyme Microb. Tech.* **2003**, 32, 386-395.
8. Franden MA, Pienkos PT, Zhang M: Development of a high-throughput method to evaluate the impact of inhibitory compounds from lignocellulosic hydrolysates on the growth of *Zymomonas mobilis*. *J. Biotechnol.* **2009**, 144, 259-267.
9. Albers E, Larsson C: A comparison of stress tolerance in YPD and industrial lignocellulose-based medium among industrial and laboratory yeast strains. *J. Ind. Microb. Biot.* **2009**, 36, 1085-1091.
10. Modig T, Almeida JR, Gorwa-Grauslund MF, Lidén G: Variability of the response of *Saccharomyces cerevisiae* strains to lignocellulose hydrolysate. *Biotechnol. Bioeng.* **2008**, 100, 423-429.
11. Mussatto SI, Roberto IC: Alternatives for detoxification of diluted-acid lignocellulosic hydrolyzates for use in fermentative processes, A review. *Bioresource Technol.* **2004**, 93, 1-10.
12. Zhu J, Yong Q, Xu Y, Yu S: Detoxification of corn stover prehydrolyzate by trialkylamine extraction to improve the ethanol production with *Pichia stipitis* CBS 5776. *Bioresource Technol.* **2011**, 102, 1663-1668.
13. Jönsson LJ, Palmqvist E, Nilvebrant NO, Hahn-Hägerdal B: Detoxification of wood hydrolysates with laccase and peroxidase from the white-rot fungus *Trametes versicolor*. *Appl. Microb. Biot.* **1998**, 49, 691-697.
14. Alriksson B, Cavka A, Jönsson LJ: Improving the fermentability of enzymatic hydrolysates of lignocellulose through chemical in-situ detoxification with reducing agents. *Bioresource Technol.* **2011**, 102, 1254-1263.

15. Heer D, Sauer U: Identification of furfural as a key toxin in lignocellulosic hydrolysates and evolution of a tolerant yeast strain. *Microb. Biotechnol.* **2008**, 1, 497-506.
16. Huang H, Guo X, Li D, Liu M, Wu J, Ren H: Identification of crucial yeast inhibitors in bio-ethanol and improvement of fermentation at high pH and high total solids. *Bioresource Technol.* **2011**, 102, 7486-7493.
17. Klinke HB, Thomsen AB, Ahring BK: Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass. *Appl. Microb. Biot.* **2004**, 66, 10-26.
18. Palmqvist E, Grage H, Meinander NQ, Hahn-Hägerdal B: Main and interaction effects of acetic acid, furfural, and *p*-hydroxybenzoic acid on growth and ethanol productivity of yeasts. *Biotechnol. Bioeng.* **1999**, 63, 46-55.
19. Zhao X, Peng F, Cheng K, Liu D: Enhancement of the enzymatic digestibility of sugarcane bagasse by alkali-peracetic acid pretreatment. *Enzyme Microb. Tech.* **2009**, 44, 17-23.
20. Zha Y, Slomp R, Groenestijn J, Punt PJ: Preparation and Evaluation of Lignocellulosic Biomass Hydrolysates for Growth by Ethanogenic Yeasts. In *Microbial Metabolic Engineering, Methods and Protocols*, Qiong C Ed., Humana Press: New York, USA, 2012, pp. 245-259.
21. van Hoek P, van Dijken JP, Pronk JT: Effect of specific growth rate on fermentative capacity of baker's yeast. *Appl. Environ. Microb.* **1998**, 64, 4226-4233.
22. Kuyper M, Winkler AA, van Dijken JP, Pronk JT: Minimal metabolic engineering of *Saccharomyces cerevisiae* for efficient anaerobic xylose fermentation, a proof of principle. *FEMS Yeast Res.* **2004**, 4, 655-664.
23. Kwon YJ, Ma AZ, Li Q, Wang F, Zhuang GQ, Liu CZ: Effect of lignocellulosic inhibitory compounds on growth and ethanol fermentation of newly-isolated thermotolerant *Issatchenkia orientalis*. *Bioresource Technol.* **2011**, 102, 8099-8104.
24. Wierckx N, Koopman F, Bandounas L, de Winde JH, Ruijsenaars HJ: Isolation and characterization of *Cupriavidus basilensis* HMF14 for biological removal of inhibitors from lignocellulosic hydrolysate. *Microb. Biotechnol.* **2009**, 3, 336-343.
25. Palmqvist E, Hahn-Hägerdal B: Fermentation of lignocellulosic hydrolysates. I, inhibition and detoxification. *Bioresource Technol.* **2000**, 74, 17-24.
26. Palmqvist E, Almeida JS, Hahn-Hägerdal B: Influence of furfural on anaerobic glycolytic kinetics of *Saccharomyces cerevisiae* in batch culture. *Biotechnol. Bioeng.* **1999**, 62, 447-454.
27. Sanda T, Hasunuma T, Matsuda F, Kondo A: Repeated-batch fermentation of lignocellulosic hydrolysate to ethanol using a hybrid *Saccharomyces cerevisiae* strain metabolically engineered for tolerance to acetic and formic acids. *Bioresource Technol.* **2011**, 102, 7917-7924.
28. Hazan R, Levine A, Abeliovich H: Benzoic acid, a weak organic acid food preservative, exerts specific effects on intracellular membrane trafficking pathways in *Saccharomyces cerevisiae*. *Appl. Environ. Microb.* **2004**, 70, 4449-4457.
29. Rumbold K, van Buijsen HJJ, Gray VM, van Groenestijn JW, Overkamp KM, Slomp RS, van der Werf MJ, Peter PJ: Microbial renewable feedstock utilization, A substrate-oriented approach. *Bioeng. Bugs* **2010**, 1, 359-366.
30. Maiorella B, Blanch HW, Wilke CR: By-product inhibition effects on ethanolic fermentation by *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **1983**, 25, 103-121.
31. Sannigrahi P, Pu Y, Ragauskas A: Cellulosic biorefineries--unleashing lignin opportunities. *Curr. Opin. Environ. Sustain.* **2010**, 2, 383-393.
32. Rumbold K, van Buijsen HJJ, Overkamp KM, van Groenestijn JW, Punt PJ, van der Werf MJ: Microbial production host selection for converting second-generation feedstocks into bioproducts. *Microb. Cell. Fact.* **2009**, 8, 64.
33. Keating JD, Panganiban C, Mansfield SD: Tolerance and adaptation of ethanogenic yeasts to lignocellulosic inhibitory compounds. *Biotechnol. Bioeng.* **2006**, 93, 1196-1206.

34. Koo BW, Park N, Jeong HS, Choi JW, Yeo H, Choi I-G: Characterization of by-products from organosolv pretreatments of yellow poplar wood (*Liriodendron tulipifera*) in the presence of acid and alkali catalysts. *J. Ind. Eng. Chem.* **2011**, 17, 18-24.
35. da Cunha-Pereira F, Hickert LR, Sehnem NT, de Souza-Cruz PB, Rosa CA, Ayub MA: Conversion of sugars present in rice hull hydrolysates into ethanol by *Spathaspora arborariae*, *Saccharomyces cerevisiae*, and their co-fermentations. *Bioresource Technol.* **2011**, 102, 4218-4225.
36. Casey E, Sedlak M, Ho NW, Mosier NS: Effect of acetic acid and pH on the cofermentation of glucose and xylose to ethanol by a genetically engineered strain of *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **2010**, 10, 385-393.

CHAPTER 5

Identifying Inhibitory Compounds in Lignocellulosic Biomass Hydrolysates using An Exometabolomics Approach

Ying Zha, Johan A. Westerhuis, Bas Mulwijk, Karin M. Overkamp,
Bernadien M. Nijmeijer, Leon Coulier, Age K. Smilde, and Peter J. Punt

This chapter has been submitted to '*BMC Biotechnology*'

Abstract

Background

During the pretreatment process of lignocellulosic biomass, inhibitors are formed that reduce the fermentation performance of the fermenting yeast. To systematically identify inhibitors in lignocellulosic biomass hydrolysates, an exometabolomics approach was applied.

Results

We studied the composition and fermentability of 24 different biomass hydrolysates. To create diversity, the 24 hydrolysates were prepared from six different biomass types, and with four different pretreatment methods. Their composition and that of fermentation samples generated with these hydrolysates were analyzed with two GC-MS methods. To preclude sugars, which obscure the detection of less abundant compounds, either ethyl acetate extraction or ethyl chloroformate derivatization was used to treat samples before conducting GC-MS. Through establishing relationship between fermentability and composition of the hydrolysates, using multivariate PLS-2CV and nPLS-2CV data analysis models, potential inhibitors were identified. These identified compounds were tested for their effects on the growth of the model yeast, confirming that the majority of the identified compounds were indeed inhibitors.

Conclusion

Using a non-targeted systematic approach, metabolomics, inhibitory compounds in lignocellulosic biomass hydrolysates were successfully identified. The identified inhibitors include both known ones, such as furfural, HMF and vanillin, and novel inhibitors, namely sorbic acid and phenylacetaldehyde.

Background

Lignocellulosic biomass, like bagasse, wheat straw, and corn stover, is the 2nd generation feedstock for biofuel production. Compared to fossil fuel, it is abundant, renewable and environmental friendly, while compared to 1st generation feedstock, like corn, it does not compete with world food supply [1,2]. Lignocellulosic biomass is composed of cellulose, hemicellulose and lignin, of which cellulose is the homopolymer of glucose, while hemicellulose is a heteropolymer mainly composed of glucose and xylose [3,4]. To produce biofuel from lignocellulosic biomass, a pretreatment step is required to break down its structure and expose cellulose for enzymatic hydrolysis [5,6]. The hydrolysis product, so-called biomass hydrolysate, is used as substrate for biofuel production through fermentation processes [7]. During most biomass pretreatment processes, harsh conditions, like high temperature and high pressure, were adopted. This causes sugars and lignin in biomass hydrolysates to degrade, forming products that possess inhibitory effects towards fermenting hosts, thus resulting in reduced growth and productivity [8-11].

Research has been conducted to identify the compounds in the biomass hydrolysates that cause inhibitory effects [12-14]. In these studies, it was found that the inhibitors are mainly of the following three categories, weak acids, furans and phenolic compounds, and the most frequently studied representatives are acetic acid, furfural, HMF and vanillin, respectively [15-17]. A variety of experimental and analytical methods were used in these studies for identifying inhibitory compounds, and a common feature of these studies was that the approach was targeted [18]. In another words, a group of compounds were selected prior to hydrolysate compositional analysis, based on knowledge of lignocellulosic biomass structure and previous research. The selected compounds were analyzed for their presence in the biomass hydrolysates and their toxicity towards the fermenting microorganisms [11,19,20].

Besides the identified inhibitors, evidence was obtained showing that other compounds present in biomass hydrolysates also display inhibitory effects [21,22]. They were observed as unknown peaks in hydrolysate compositional analysis results, which reduced in size after detoxification [23]. To identify novel inhibitory compounds in biomass hydrolysates, in this study a non-targeted exometabolomics approach was applied. Generally, metabolomics is one of the 'omics' tools that studies the performance of research objects by analyzing their overall compositions [24,25]. In this study, research objects are lignocellulosic biomass hydrolysates, which are used as fermentation media for bioethanol production. The performance of biomass hydrolysates as fermentation media vary due to the difference in their compositions, i.e. inhibitory compounds and their concentrations. Through establishing

the relation between the composition and the performance of different biomass hydrolysates statistically, compounds that possess inhibitory effect could be indicated in an unbiased way (Figure 1).

In metabolomics, the search for important metabolites responsible for a certain response, e.g. fermentability, is often performed with multivariate data analysis methods [26,27]. These multivariate methods are able to search for the interactions between metabolites that are responsible for the response that is modeled. Partial least squares (PLS) is a multivariate data analysis method that is commonly used in metabolomics to search for the important metabolites [28]. As an extension of the PLS method, also n-way PLS may be used when the data-set consists of a time series such as in the case of our metabolomics experiments. As multivariate data analysis methods may lead to false positive correlations, rigorous validation of these models is necessary [29,30]. Therefore, we decided to use double cross validation (2CV) to obtain unbiased prediction errors for the (n)PLS models [31,32].

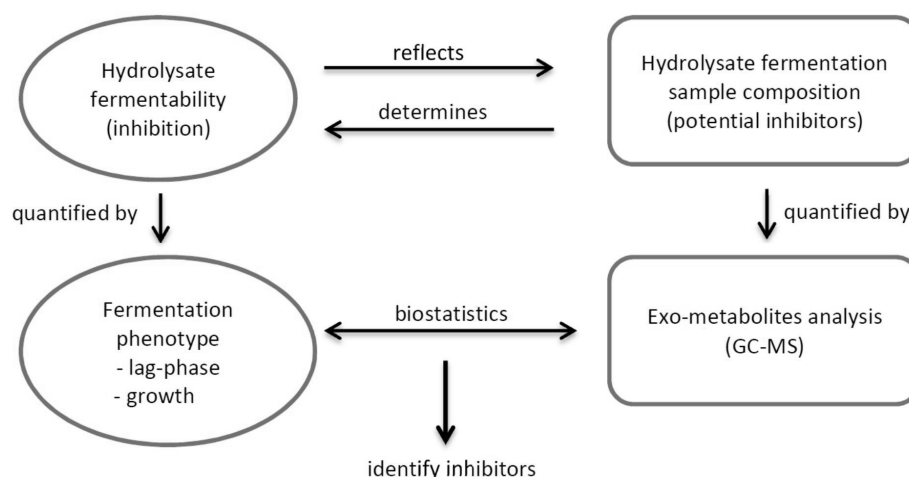


Figure 1 Graphic illustration of the concept of the exometabolomics approach.

We report here the detailed procedure and the results of using the exometabolomics approach for identifying inhibitory compounds in lignocellulosic biomass hydrolysates. This includes the batch fermentability of 24 different biomass hydrolysates using baker's yeast, *S. cerevisiae* CEN.PK113-7D, and the analysis results of the fermentation samples by two GC-MS methods; the statistical model building procedure for identifying potential inhibitory

compounds, and the toxicity testing results of the suggested potential inhibitors. The results of this study show that of the potential inhibitory compounds indicated by the statistical models, a large fraction indeed exhibited inhibitory effects on the growth of fermenting yeast. These compounds consist of both known inhibitors, such as furfural and HMF, and novel inhibitors.

Materials and Methods

Biomass hydrolysate preparation and fermentation

24 different hydrolysates were prepared from six types of biomass, by using four different hydrolysate preparation methods. The six types of biomass were sugar cane bagasse (Zillor, Brazil), corn stover (University of Cape Town, South Africa), wheat straw (Oostwaardshoeve, The Netherlands), barley straw (Oostwaardshoeve, The Netherlands), willow wood chips (Oostwaardshoeve, The Netherlands) and oak sawdust (wood-flooring supplier ESCO, The Netherlands). Prior to pretreatment, biomass (except oak sawdust) was ground to pieces of average length 3 mm and dried at 80°C for at least 16 hours. To prepare 1 l hydrolysate, 300 g dried biomass was used. The four hydrolysate pretreatment methods were dilute acid (2% H₂SO₄), mild alkaline (3% Ca(OH)₂), alkaline/peracetic acid and concentrated acid (72% H₂SO₄). The biomass pretreated with the first three methods was hydrolyzed enzymatically, using Accellerase 1500 (Genencor®), while acid hydrolysis was used for biomass pretreated with concentrated acid (40% and 15% H₂SO₄). The detailed pretreatment and hydrolysis procedure was described in Zha *et al.* [33]. After hydrolysis, solid content was separated from the hydrolysate by filtration, and the filtrated hydrolysate was sterilized using filter sterilization and stored at 4°C before use.

Batch fermentations were carried out in 2 l New Brunswick fermentors, using 1 l of sterilized hydrolysate as substrate. The fermenting yeast was *Saccharomyces. cerevisiae* CEN.PK 113-7D (CBS 8340), and the inoculum was prepared in a 500 ml Erlenmeyer flask. The cells were harvested by centrifugation after incubating overnight in mineral medium (MM) [34] with 20 g/l glucose, and inoculated into fermentors with density of 0.1 g cell dry weight per 1 l hydrolysate. All fermentations were carried out at 30°C, under anaerobic conditions by sparging 0.5 l/min N₂ continuously, and pH was set at 5 by adding 1 M H₂SO₄ or 2 M KOH.

For each of the 24 hydrolysates, one batch fermentation was conducted after checking its reproducibility [18]. During the whole fermentation process, CO₂ concentration in the off-gas

was monitored automatically and samples were taken at fixed time intervals. These samples were kept at 4°C and used to measure optical density (OD), glucose and ethanol concentration with either Cobas® Mira Plus (Roche) or Arena® 20 Analyzer (Thermo Scientific).

Hydrolysate fermentation sample analysis

For each of the selected hydrolysate fermentations, cell free time samples were chosen for analyzing their overall compositions. Two GC-MS methods, namely ethyl acetate extraction (EA)-GC-MS and ethyl chloroformate derivatization (EC)-GC-MS, were used to analyze the fermentation samples.

For EA-GC-MS, the extraction was done by adding 550 µl ethyl acetate into 0.5 ml sample and vortex for 2 min. The mixture was centrifuged to separate ethyl acetate fraction, of which 400 µl was transferred to a vial and dried by blowing N₂. The following internal standards in ethyl acetate were added to the same vial: phenylethanol-D5, cinnamic acid-D5 and hydroxybenzaldehyd-D4. The extraction and centrifugation process was repeated, and from the ethyl acetate fraction, another 400 µl was transferred to the same vial, after drying with N₂, the following internal standards in pyridine were added: alanine-D4 and citric acid-D4. The extract was then oxidized by adding 30 µl 56 mg/ml ethoxyamine-HCl, and incubating at 40°C for 90 min. Followed by adding dicyclohexylphthalate (DCHP) and difluorobiphenyl (DFB) in pyridine as injection standard, the oxidized extract was silylated by adding 100 µl *N*-methyl-*N*-trimethylsilyl-trifluoroacetamide (MSTFA), and incubating at 40°C for 50 min. Measurement was carried out by 1 µl splitless injection in the PTV injector of an Agilent® 7890A GC with an Agilent® 5975C MS as detector. The analytical column used was HP-5MS 30 m x 0.25 mm x 0.25 µm.

For EC-GC-MS, the sample pH was brought above 10 by adding NaOH solution, followed by the addition of following internal standards in pyridine: leucine-D3, succinic acid-D4 and cinnamic acid-D5. The injection standards, DCHP and DFB in pyridine, and 300 µl ethanol were also added to the sample. Then the ethylesters formation was done by two rounds of adding 40 µl ethyl chloroformate into the sample and shaking it vigorously by hand for 15 sec. The reaction was stopped by adding 750 µl dichloromethane and 500 µl 1 M bicarbonate buffer. The formed derivates were extracted with dichloromethane, and the extraction was dried with Na₂SO₄. The measurement was carried out the same way as in EA-GC-MS method. The analytical column used was DB-1 30m x 0.32mm x 1 µm.

The analysis results of EA-GC-MS and EC-GC-MS were reported separately in data-sets, with detected peaks as row and fermentation sample as column. The reported values were areas of the detected peaks after correction with internal standards.

Statistical model building

The two statistical models used were partial least square with double cross validation (PLS-2CV) [31] and n-way PLS with double cross validation (nPLS-2CV) [35]. The 2CV version of the nPLS model was developed in house. The models were written as m-files in MATLAB environment (R2012a) with PLS toolbox 2.0 (Eigenvector).

PLS-2CV models

PLS-2CV is a linear regression model, which predicts the fermentation phenotypes with the GC-MS analysis results of the fermentation samples (data-sets). The PLS-2CV models were assessed by calculating the so-called Q^2 values, which indicate the prediction ability of the data-sets for a specific phenotype [31]. The maximum value of Q^2 is 1, representing that the model could perfectly predict the phenotypes. Generally, models with $Q^2 \geq 0.5$ were selected for analyzing the selectivity ratios (SR) assigned to each peaks in the data-sets. Similar to regression coefficient ('reg'), SR is a measure for variable importance in discrimination models. Contrary to 'reg', SR is corrected for the influence of interfering compounds that are not related to the modeled response [36,37]. Peaks with the highest SR values were considered having the primary contribution to the model building. Among these peaks, the identified ones were selected as potential inhibitory compounds, and tested in Bioscreen C Analyzer for their effects on the fermenting yeast.

To model lag-phase, the data-sets containing the first two fermentation samples (t1 and t2) were used. As listed in Table 1, the difference as well as the combination of t1 and t2 data-sets were used to build PLS-2CV model. EA-GC-MS and EC-GC-MS data-sets were modeled both separately and combined. Thus, for lag-phase, in total six PLS-2CV models were built (Table 1). These data-sets were preprocessed by conducting a 'square-root' transformation to reduce the nonsymmetrical distributions of the peak areas for all compounds, and this also homogenizes the heteroscedastic measurement error. Afterwards, an 'auto-scaling' was carried out to reduce the effect that compounds with large peak areas would dominate the regression models [38,39]. The phenotype values were 'mean-centered' before data analysis.

To model glucose consumption rate (Glu CR) and ethanol production rate (EtOH PR) (see Eq2 to 4), the data-sets of all five fermentation samples were used individually (t1 to t5, see Table 3). The data preprocessing was conducted in the same way as by lag-phase data-sets (Table 1).

Table 1 Data-sets used for building PLS-2CV models.

The two data preprocessing methods were symbolized by ' $\sqrt{}$ ' (square-root), and '*auto*' (autoscaling); ':' indicates that the corresponding data-sets were combined; EA_{tx}: EA-GC-MS data-set of time sample tx; EC_{tx}: EC-GC-MS data-set of time sample tx (x represents one of the five fermentation samples).

Lag-phase model 1	$auto(\sqrt{EA_{t2}} - \sqrt{EA_{t1}})$
Lag-phase model 2	$auto(\sqrt{EC_{t2}} - \sqrt{EC_{t1}})$
Lag-phase model 3	$auto((\sqrt{EA_{t2}} - \sqrt{EA_{t1}}):(\sqrt{EC_{t2}} - \sqrt{EC_{t1}}))$
Lag-phase model 4	$auto(\sqrt{EA_{t1}:EA_{t2}})$
Lag-phase model 5	$auto(\sqrt{EC_{t1}:EC_{t2}})$
Lag-phase model 6	$auto(\sqrt{EA_{t1}:EA_{t2}:EC_{t1}:EC_{t2}})$
Glu CR and EtOH PR	$auto(\sqrt{EA_{tx}})$ or $auto(\sqrt{EC_{tx}})$

nPLS-2CV models

N-way PLS (nPLS) handles multiway data-sets, and was used to model glucose consumption rate (Glu CR) and ethanol production rate (EtOH PR). In this study, the data-sets were three-way, the three ways were (1) fermentation batch, (2) time samples of each batch, and (3) analysis results of each sample. The analysis results of EA-GC-MS and EC-GC-MS methods were used both separately and combined. Similar to PLS-2CV model, the data-sets were arranged in two way and preprocessed by conducting 'square-root' and 'auto-scaling' before transforming to the three-way structure. The phenotype values were 'mean-centered' before model building. The nPLS-2CV models were assessed by calculating the Q^2 values. In most cases, models with $Q^2 \geq 0.5$ were selected for analyzing the regression coefficient ('reg') of each peak in the data-sets, as SR for nPLS has not yet been developed. Peaks with highest absolute 'reg' values were considered having the most contribution for predicting the phenotypes. Among these peaks, the identified ones were selected as potential inhibitory compounds, and tested in Bioscreen C Analyzer for their effects on the fermenting yeast.

Potential inhibitory compound test

Solutions of potential inhibitory compounds were prepared in both MM with 20 g/l glucose and YPD (Yeast extract Peptone Dextrose) medium with concentrations of 0.2, 0.5 and 1.0 g/l. If a compound was saturated at 1.0 g/l, the solutions were made with 20%, 50% and 100% of the saturated concentration. These compounds are shown in bold type in Table 5.

The prepared solutions were used as media in the growth test of the fermenting yeast, *S. cerevisiae* CEN.PK 113-7D. The growth test was conducted in triplets in honeycomb plates, using Bioscreen C Analyzer (Labsystems OY). The detailed procedure of Bioscreen test is described in Zha *et al.* [33].

Results

Biomass hydrolysates preparation

To successfully identify inhibitory compounds in biomass hydrolysates with statistical models, acquiring hydrolysates with diverse performance is of importance [18]. To achieve this, 24 different hydrolysates were prepared from six different biomass and by using four hydrolysate preparation methods. Among the six biomass, wheat straw, barley straw and corn stover are agricultural wastes, bagasse is sugar industry byproduct, and willow and oak are wood products. Each of the six biomass was pretreated with four different methods, which used 2% sulfuric acid, 72% sulfuric acid, lime, and peracetic acid, respectively. The resulting 24 hydrolysates were tested for their performance as fermentation media on a small scale (ml), showing that there was a significant diversity among these 24 hydrolysates [33]. For the exometabolomics study, these hydrolysates were prepared in larger quantity (l). For each hydrolysate, a batch fermentation of 1 l working volume was carried out based on previously developed procedures [11].

Defining phenotypes

Identical batch fermentations were carried out for each of the 24 different hydrolysates generated. The fermentability was monitored by measuring OD, glucose and ethanol concentrations of the samples taken with a fixed time interval. To quantify the fermentability of the hydrolysates, four phenotypes were defined, which are lag-phase, glucose consumption rate (Glu CR), ethanol production rate (EtOH PR) and ethanol yield (EtOH Y).

The definition of these four phenotypes are given in equation 1 to 4 (Eq1 to Eq4), and the measurement results of the fermentation samples were used to calculate these phenotypes.

Eq1: lag-phase = time to reach 2% (OD_{max}-OD_{min}) (h)

Eq2: Glu CR = the slope of the linear part of the glucose consumption curve (g/l/h)

Eq3: EtOH PR = the slope of the linear part of the ethanol production curve (g/l/h)

Eq4: EtOH Y= EtOH_{max}/initial glucose concentration (g/g)

As shown in the phenotype definitions, lag-phase has time as unit (Eq1), which represents the duration before growth began. Since during lag-phase, the fermenting yeast adapt to the media composition for growth [40], a longer lag-phase indicates the presence of compounds that restrain the starting of growth. Glucose consumption rate (Glu CR) is an indicator of the growth rate of the fermenting yeast, while ethanol production rate (EtOH PR) and ethanol yield (EtOH Y) describe the productivity of the fermenting yeast in a specific hydrolysate. For each of the 24 fermentations, these four phenotypes were calculated (Table 2). It should be mentioned that growth rate is one of the most commonly used phenotypes describing the performance of fermenting hosts. In this study, instead of using growth rate, we chose Glu CR to describe growth. This is because OD measurement results were not easily comparable due to sample characteristics, such as color differences among hydrolysates, and flocculation. To confirm that Glu CR is a good indicator of growth performance, OD% was used to calculate tentative growth rate (Table 2). It can be seen that the tentative growth rates have very similar trend compared to Glu CR (Figure 2). Since glucose measurements are more accurate than OD, we have decided to use Glu CR as an indicator of growth rate.

As shown in Table 2, all 24 hydrolysates had different glucose concentrations, indicating that biomass type as well as pretreatment method influenced the biomass hydrolysis efficiency. In general, mild alkaline (MA) pretreated biomass resulted in relatively low glucose concentration, while concentrated acid (CA) lead to higher hydrolysis efficiency [33]. Based on our previous results, the variation in glucose concentration in the range observed in Table 3 was of no influence on fermentation performance (results not shown).

The performance of the 24 hydrolysate varied significantly as fermentation media, which was consistent with the screening experiments on milliliter scale [33]. As far as lag-phase is considered, hydrolysates like Oak-CA and Willow-CA supported growth almost immediately after inoculation, while the fermenting yeast needed an adaptation period of as long as 10 hours in CS-CA and WS-CA hydrolysates. The Glu CR of the 24 hydrolysates ranged from 0.80 (Oak-CA) to 4.63 (WS-CA), which trend was comparable to that of EtOH PR (Figure 3).

This resulted in very similar ethanol yield among the hydrolysates, around 0.4 g/g (Table 2), which was also the ethanol yield of *S.cerevisiae* CEN.PK113-7D in mineral medium with 20 g/l glucose [41]. This observation suggested that under anaerobic conditions, the effect of inhibitory compounds in hydrolysates had little effect on the ethanol yield of the fermenting yeast. Therefore, this phenotype was not used in building statistical models for the purpose of identifying hydrolysate inhibitors.

Table 2 Fermentability of the 24 biomass hydrolysates, expressed as the calculation results of the defined phenotypes.

Hydrolysate	Glucose ¹ (g/l)	Ethanol ² (g/l)	Lag-phase ³ (h)	Glu CR ⁴ (g/l/h)	EtOH PR ⁵ (g/l/h)	EtOH Y ⁶ (g/g)	Tentative growth rate ⁷
Bag-CA	67.39	20.61	7.5	1.42	0.44	0.306	2.42
Bag-DA	63.33	24.20	6.0	3.64	1.52	0.382	4.61
Bag-MA	58.82	22.71	2.0	3.84	1.58	0.386	5.73
Bag-PAA	52.48	19.87	3.0	2.52	0.71	0.379	2.76
BS-CA	67.45	30.92	7.5	4.57	1.73	0.458	7.39
BS-DA	49.87	20.95	6.5	3.63	1.42	0.420	5.66
BS-MA	42.56	18.40	6.0	3.05	1.41	0.432	*
BS-PAA	53.50	22.22	3.0	2.96	1.03	0.415	5.12
CS-CA	65.63	26.62	10.5	3.21	1.32	0.406	4.73
CS-DA	42.80	18.74	5.5	3.43	1.49	0.438	6.98
CS-MA	32.83	15.85	6.5	3.35	1.32	0.483	7.92
CS-PAA	50.29	20.84	3.5	2.38	1.03	0.414	4.53
Oak-CA	66.72	12.06	1.5	0.80	0.29	0.181	2.1
Oak-DA	38.22	15.27	5.0	2.41	0.98	0.400	5.37
Oak-MA	44.35	19.49	2.5	3.43	1.55	0.439	7.52
Oak-PAA	60.80	25.97	3.0	2.73	1.12	0.427	3.75
Willow-CA	31.58	13.60	1.0	4.26	1.10	0.431	14.04
Willow-DA	45.15	17.68	7.5	2.74	1.14	0.392	5.72
Willow-MA	23.50	10.76	4.5	2.68	1.29	0.458	*
Willow-PAA	51.30	22.81	5.5	2.45	1.05	0.445	5.03
WS-CA	60.54	24.71	9.0	4.63	1.87	0.408	7.6
WS-DA	58.29	24.83	4.5	3.47	1.64	0.426	6.05
WS-MA	32.12	13.95	6.5	4.01	1.92	0.434	11.37
WS-PAA	51.94	21.61	3.5	3.03	1.27	0.416	5.48

¹: glucose concentration of the 24 hydrolysates; ²: final ethanol concentration; ³: (Eq1);

⁴: glucose consumption rate (Eq2); ⁵: ethanol production rate (Eq3); ⁶: ethanol yield (Eq4);

⁷: the slope of the linear part of the OD% curve. *: OD measurement was not possible due to flocculation; **bold**: fermentations that are selected for sample compositional analysis.

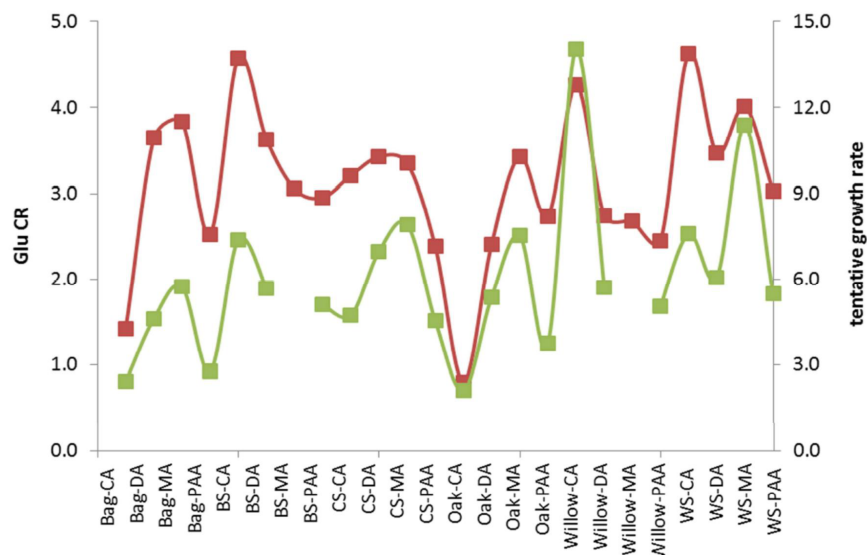


Figure 2 Trends comparison between tentative growth rate and glucose consumption rate. Tentative growth rate: the slope of the linear part of the OD% curve (green); Glucose consumption rate (Glu CR): the slope of the linear part of the glucose consumption curve (Eq2) (red).

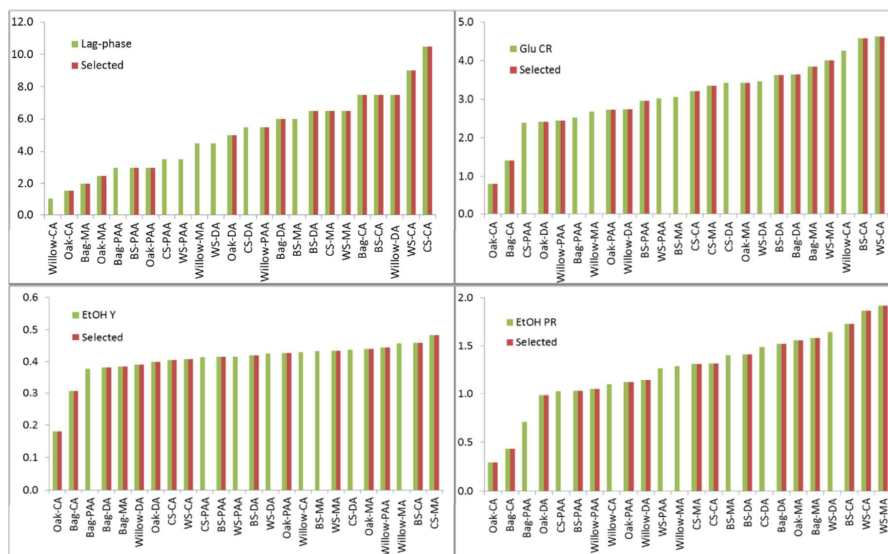


Figure 3 The calculation results of the four phenotypes, sorted from the smallest to the largest. The 'red' bars are the selected 16 fermentations for exometabolomics analysis.

Among the 24 hydrolysate fermentations, some had similar performance in terms of the phenotypes calculated (Figure 3). Since the statistical models to be used for analyzing the relationship between fermentability and sample composition were based on linear regression, it is important to reduce overrepresentation of certain phenotype classes. In addition, it is also beneficial to minimize the amount of samples for exometabolomics analysis. Therefore, from the 24 fermentations, 16 were selected based on the variations in their phenotypes, biomass type and pretreatment method. The selected 16 hydrolysates contain all six biomass types and all four biomass pretreatment methods (Table 2), and the fermentability of these selected hydrolysates show a more or less even spread of the fermentation phenotypes (Figure 3).

Hydrolysate fermentation sample analysis

After quantifying the performance of the hydrolysate fermentations with the four phenotypes, cell free time samples of the 16 selected fermentations were analyzed for their overall compositions. These samples were chosen based on the criteria that they should uniquely represent the whole fermentation process. As each fermentation can be divided into three phases, namely lag-phase, growth-phase and stationary-phase, five samples were selected, as shown in Table 3. The division of the three fermentation phases was consistent with the definition of the phenotypes, i.e. the end of lag-phase is when OD reaches 2% of the maximum OD, the end of growth-phase is when glucose consumption is completed, and the duration of stationary-phase is set at 10 hours. In this way, a total of 80 samples from 16 hydrolysate fermentations were selected for compositional analysis.

Table 3 The five fermentation time samples for compositional analysis with the two GC-MS methods.

t1	beginning of fermentation	immediately after inoculation
t2	end of lag-phase	time needed to reach 2% (OD _{max} -OD _{min})
t3	growth mid-point	time needed to consume half of the initial glucose
t4	growth end point	time needed to consume all glucose
t5	stationary phase	10 hours after growth end point

The focus of the compositional analysis was the potential inhibitory compounds in hydrolysate samples, which are believed to be mainly non-sugar compounds, such as weak acids, furans and phenols [8,9,15]. GC-MS was chosen as the analytical tool, as the method is capable of detecting a wide range of these compounds, including many unknowns [20,42].

A crucial point in analyzing hydrolysate samples with GC-MS was to remove sugars, which are present in large quantities in the samples and severely interfere with the detection of non-sugar compounds [13,43]. For this purpose, two sample work-up methods were developed, namely, ethyl acetate extraction and ethyl chloroformate derivatization.

Ethyl acetate extraction GC-MS (EA-GC-MS) was previously described by Heer *et al.* [13]. The method uses ethyl acetate (EA) as solvent, in which compounds that are apolar, e.g. with aromatic rings, are dissolved, while polar compounds, like sugars, remain in the water phase. In current study, the hydrolysate samples were extracted twice with EA to allow adequate recovery of the extracts. After removing EA phase from the water phase, it was dried through evaporation, thus concentrated and ready for analysis with GC-MS.

Due to the nature of this sample work-up method, only compounds dissolvable in EA were analyzed, and since EA was removed through evaporation, volatile compounds were partially lost. This makes recovery an important issue in EA-GC-MS method, which was assessed with a group of furans and aromatic compounds before analyzing hydrolysates. It was found that sample pH influences the extraction, when pH was raised to above 6.0, significant decrease of recovery was observed with multiple aromatic standards. Therefore, all hydrolysate samples were extracted with EA at pH 5.0. In doing so, the recovery of aromatic compounds was satisfying, above 90%, while the recovery of furans was rather low and inconsistent due to evaporation. So the analysis results of aromatic compounds were considered more reliable than furans.

To complement EA-GC-MS method, ethyl chloroformate derivatization GC-MS (EC-GC-MS) was developed in our lab. Ethyl chloroformate (EC) was used to convert acids to their ethyl ester form, thus compounds like carboxylic acids, amino acids, aromatic compounds and furans could be detected by MS. EC-GC-MS therefore has a larger coverage of compounds compared to EA-GC-MS, and is easier to operate. But due to the diverse reactivity of compounds with EC, it is possible that compounds present with high concentration could only be detected with low signal. The involvement of a derivatization step could also cause a single compound to have more than one derivatization product, which complicates the characterization of the compound. EC-GC-MS method not only complemented EA-GC-MS by detecting small carboxylic acids and furans, but also overlapped with EA-GC-MS by detecting aromatic compounds. As far as aromatic compounds are concerned, it seems that the results of EA-GC-MS were more reliable due to the reactivity diversity issue occurring in EC-GC-MS.

After analyzing all 80 samples with both methods, a 'compound list' was generated for each method by listing all peaks clearly visible in the chromatograms. By comparing the mass spectra of these peaks with the existing GC-MS compound library in our lab, identities were assigned to some of the peaks. With EA-GC-MS method, in total 129 peaks were detected, among which 44 were identified; while in EC-GC-MS results, there were 114 detected peaks, of which 56 were identified. From the identified compounds, the majority detected by EC-GC-MS method were acids, including carboxylic acids, such as levulinic acid and succinic acid, phenolic acids, like phenylacetic acid and syringic acid, and 18 amino acids (Appendix 1). EA-GC-MS mainly detected phenolic compounds, containing phenolic aldehydes, alcohols and acids (Appendix 1).

To all detected peaks from both identified compounds and unknowns, pseudo-quantities were assigned by integrating their peak areas. To correct sample matrix effect, internal standards were measured in both blank and hydrolysate sample. The peak area difference between blank and hydrolysate sample of the internal standards was calculated as a correction-factor, and was used to correct all the integrated peak areas of the same hydrolysate type. Thus, compound lists based on corrected peak areas were formed for both analytical methods.

Statistical model building

To identify inhibitory compounds in biomass hydrolysates, relationship between hydrolysate fermentability and fermentation sample composition was studied by building statistical models. The models used in this study were partial least square (PLS) and n-way PLS (nPLS), validated by conducting double cross validation (2CV), which was done by leave-one-out in the inner and outer loop [31,35]. The purpose of both models was to point to compounds that are most responsible for a certain fermentation phenotype. This was done by predicting the phenotypes using the data-sets formed through analyzing fermentation samples with the two GC-MS methods.

Lag-phase

Lag-phase is the period before growth takes place in a fermentation process (Eq1), it is mainly influenced by the initial media composition. During lag-phase, the fermenting yeast adapts to the hydrolysate by adjusting its composition, thus some compounds will be degraded or converted [44,45]. Therefore, it is reasonable to describe lag-phase by comparing the composition difference between sample t1 and t2 (model 1, 2 and 3 in Table 1). In addition, the composition of sample t1 and t2 represents the beginning and the end

point of the lag-phase (Table 3), which was also used to build models for predicting lag-phase (model 4, 5 and 6 in Table 1).

In total, six models were constructed for lag-phase (Table 1), of which the data-sets of EA-GC-MS and EC-GC-MS methods were used both separately and combined. This is because the effect of fusing these two data-sets was unknown. The prediction results of the six models are shown in Table 4a. It can be seen that among the six models, only 'model 2' and 'model 5' had a Q^2 value above 0.5, indicating that these two models could be used to predict lag-phase. As shown in Table 1 that the inputs of both 'model 2' and 'model 5' were from the EC-GC-MS data-set, suggesting that the compounds detected by EC-GC-MS method had more influence on lag-phase compared to those measured with EA-GC-MS.

Table 4a Lag-phase prediction results and Q^2 values of the PLS-2CV models shown in Table 1.
Bold: models selected for analyzing the SR of the peaks in the data-sets.

PLS-2CV	lag-phase	prediction					
		model 1	model 2	model 3	model 4	model 5	model 6
Bag-CA	7.5	6.13	9.28	7.43	5.79	6.18	4.96
Bag-DA	6.0	3.36	5.24	4.61	6.79	4.95	5.53
Bag-MA	2.0	5.02	3.55	4.27	4.21	3.08	2.90
BS-CA	7.5	7.14	7.67	7.45	7.42	9.28	8.44
BS-DA	6.5	8.99	5.29	8.23	9.49	6.57	7.71
BS-PAA	3.0	3.61	4.12	3.82	3.86	4.67	3.86
CS-CA	10.5	11.71	9.44	10.91	10.32	10.61	11.33
CS-MA	6.5	-6.10	6.28	-1.36	4.40	6.66	6.25
Oak-CA	1.5	4.25	4.48	4.43	6.27	6.02	6.19
Oak-DA	5.0	4.42	5.70	4.74	6.16	5.94	5.81
Oak-MA	2.5	3.81	4.00	3.99	-3.28	2.00	-0.54
Oak-PAA	3.0	5.39	4.38	5.05	3.47	2.99	2.55
Willow-DA	7.5	6.83	8.95	7.61	5.19	6.70	6.68
Willow-PAA	5.5	3.61	6.08	4.36	7.12	2.19	4.51
WS-CA	9.0	7.93	8.17	8.00	6.95	6.86	6.18
WS-MA	6.5	6.58	2.01	3.97	4.38	7.95	7.59
Q^2		-1.01	0.54	0.06	0.05	0.51	0.47

EA-GC-MS data-set failed to predict lag-phase properly ('model 1' and 'model 4' in Table 4a), but when combined with EC-GC-MS data-set, the prediction improved, resulting in a model with Q^2 value of 0.47 ('model 6' in Table 4a). As the Q^2 value of 'model 6' is very close to 0.5, this model was still selected, together with model 2 and 5, to calculate the selectivity ratios (SR) assigned to each peaks in these data-sets.

For each detected peaks in EC-GC-MS data-set, the SR values of the three models in bold in Table 4a were summed, and ranked based on their SR-sum values; while for each detected peaks in EA-GC-MS data-set, the SR value of 'model 6' were ranked. The top 40 peaks with the highest SR-sum values, 20 from EC-GC-MS data-set and 20 from EA-GC-MS data-set, were considered as the main contributors in predicting lag-phase. Among these 40 peaks, the ones with identity were tested for their effects on the fermenting yeast (Section 'potential inhibitory compound testing', Table 5a). The detailed ranking procedure of lag-phase model SR is shown in Appendix 2-1.

Glu CR and EtOH PR

Different from lag-phase, Glu CR and EtOH PR could be influenced by all five fermentation time points according to their definitions (Eq2 and Eq3). These phenotypes were modeled by the data-sets of the five fermentation samples both individually, using the PLS-2CV model, and collectively, using the nPLS-2CV model.

PLS-2CV modeled Glu CR and EtOH PR with the data-sets of individual fermentation samples, which reveals the influence of these single time points on these two phenotypes. The modeling results show that EC-GC-MS data-sets failed to predict Glu CR and EtOH PR, as the resulting Q^2 values were all negative (data not shown). On the contrary, the EA-GC-MS data-sets of sample t3, t4 and t5 successfully modeled the two phenotypes, as shown in Table 4b, the resulting Q^2 values were above 0.5. This suggests that, different from lag-phase, Glu CR and EtOH PR were relating to the compounds detected with EA-GC-MS method. Moreover, the prediction became meaningful only after time point t2 (Table 4b, $Q^2 > 0$), indicating that Glu CR and EtOH PR were not affected by the initial hydrolysate composition, but influenced by the composition after lag-phase and during growth. This confirms that the data-sets of time point t1 and t2 possess a different structure compared to the other three time points. This structure contains information that could properly describe lag-phase (Table 4a), which ends after time point t2, but failed to predict Glu CR and EtOH PR, which describe a different phase of the fermentation process.

To include the effect of hydrolysate composition change during the fermentation process on Glu CR and EtOH PR, the five time-point samples were also analyzed collectively, with the nPLS-2CV model. Consistent with the PLS-2CV models, the prediction was only valid with EA-GC-MS data-set (Table 4c). Since it was known from PLS-2CV models that data-set of sample t1 gave negative Q^2 values (Table 4b), nPLS-2CV models were also built with the data-set of sample t2 to t5. As shown in Table 4c, the predictions of Glu CR and EtOH PR were improved when sample t1 was excluded from the data-set, indicating that the input of sample t1 data-set was negative.

Table 4b Glu CR and EtOH PR prediction results and Q^2 values of the PLS-2CV models.
Bold: models selected for analyzing the SR of the peaks in the EA-GC-MS data-sets.

PLS-2CV	Glu CR	prediction					EtOH PR	prediction				
		t1	t2	t3	t4	t5		t1	t2	t3	t4	t5
Bag-CA	1.42	2.86	2.30	2.08	1.77	1.08	0.44	1.04	0.92	0.81	0.81	0.47
Bag-DA	3.64	2.62	3.07	3.00	3.01	3.56	1.52	0.93	1.15	1.18	1.18	1.39
Bag-MA	3.84	3.12	3.28	3.64	3.45	3.22	1.58	1.38	1.13	1.51	1.51	1.40
BS-CA	4.57	2.38	3.94	4.15	3.71	3.88	1.73	0.91	1.61	1.70	1.70	1.61
BS-DA	3.63	4.60	3.99	3.73	3.93	3.45	1.42	1.99	1.72	1.57	1.57	1.47
BS-PAA	2.96	2.39	2.80	2.73	2.81	2.65	1.03	1.04	1.20	1.13	1.13	1.13
CS-CA	3.21	2.98	3.51	3.71	3.24	3.51	1.32	1.12	1.31	1.44	1.44	1.35
CS-MA	3.35	4.21	4.52	4.00	4.23	4.31	1.32	1.85	2.03	1.81	1.81	1.89
Oak-CA	0.80	2.88	2.76	1.77	1.94	2.00	0.29	1.03	1.09	0.57	0.57	0.66
Oak-DA	2.41	2.54	3.04	2.63	1.83	2.35	0.98	1.23	1.46	1.22	1.22	1.22
Oak-MA	3.43	2.67	2.70	3.03	2.98	3.12	1.55	1.01	1.04	1.20	1.20	1.25
Oak-PAA	2.73	2.95	3.01	2.63	2.59	2.66	1.12	1.13	1.18	0.99	0.99	1.00
Willow-DA	2.74	2.78	2.61	2.89	3.31	2.83	1.14	1.15	0.99	1.16	1.16	1.13
Willow-PAA	2.45	3.70	3.06	3.21	3.48	3.39	1.05	1.46	1.23	1.32	1.32	1.35
WS-CA	4.63	3.43	3.74	3.49	3.67	3.49	1.87	0.93	1.42	1.33	1.33	1.32
WS-MA	4.01	4.10	3.29	3.47	3.43	3.37	1.92	1.63	1.29	1.41	1.41	1.37
Q^2		-0.16	0.37	0.68	0.57	0.61		-0.32	0.04	0.54	0.50	0.56

Thus, for Glu CR and EtOH PR, three PLS-2CV models and a nPLS-2CV model were selected respectively for analyzing the contributions of the peaks in EA-GC-MS data-set to model predictions (models with 'Bold' in Table 4b and Table 4c). With PLS-2CV models, similar to lag-phase, SR of the peaks were summed and ranked. The top 40 peaks with the highest SR values were considered as the main contributors of PLS-2CV models of either Glu CR or EtOH PR. While with nPLS-2CV models, the regression coefficient ('reg') values were used for ranking. The top 40 peaks with the highest absolute 'reg', 20 with positive values and 20 with negative values, were considered as the main contributor of nPLS-2CV model of either Glu CR or EtOH PR. Among the selected peaks, the ones with identity were tested for their effects on the fermenting yeast (Section 'potential inhibitory compound testing', Table 5a). The detailed ranking and selection procedure of the testing compounds are shown in Appendix 2-2. Interestingly, more than 80% of the compounds suggested by Glu CR and EtOH PR models are identical. This indicates, from a statistical point of view, the correlation between Glu CR and EtOH PR.

Table 4c Glu CR and EtOH PR prediction results and Q² values of the nPLS-2CV models.

Bold: models selected for analyzing the SR of the peaks in the EA-GC-MS data-sets.

nPLS-2CV	Glu CR	prediction		EtOH PR	prediction	
		t1 - t5	t2 - t5		t1 - t5	t2 - t5
Bag-CA	1.42	1.55	1.41	0.44	1.17	0.59
Bag-DA	3.64	2.87	3.07	1.52	1.05	1.16
Bag-MA	3.84	3.16	3.25	1.58	1.27	1.32
BS-CA	4.57	3.76	3.67	1.73	1.53	1.51
BS-DA	3.63	4.02	3.91	1.42	1.69	1.64
BS-PAA	2.96	2.79	2.88	1.03	1.20	1.23
CS-CA	3.21	3.49	3.54	1.32	1.35	1.36
CS-MA	3.35	4.40	4.37	1.32	2.00	2.00
Oak-CA	0.80	1.96	1.86	0.29	0.65	0.62
Oak-DA	2.41	2.28	2.35	0.98	1.17	1.08
Oak-MA	3.43	2.71	2.86	1.55	1.05	1.11
Oak-PAA	2.73	2.81	2.78	1.12	1.08	1.07
Willow-DA	2.74	2.83	2.90	1.14	1.16	1.18
Willow-PAA	2.45	3.68	3.55	1.05	1.49	1.42
WS-CA	4.63	3.70	3.64	1.87	1.40	1.39
WS-MA	4.01	3.58	3.46	1.92	1.41	1.37
Q ²		0.526	0.580		0.182	0.419

Potential inhibitory compound testing

Through constructing statistical models and analyzing the compounds that contribute the most to the models with valid phenotype predictions, two groups of potential inhibitory compounds were identified. To study the effect of these compounds on the fermenting yeast, *S.cerevisiae* CEN.PK113-7D, growth tests were conducted in mineral medium (MM) with 20 g/l glucose. The potential inhibitory compounds were added individually with the following three concentrations, 0.2, 0.5 and 1.0 g/l, respectively.

It should be noted that these testing concentrations could be much higher compared to that in actual biomass hydrolysates, i.e. less than 0.1 g/l [11,14,46,47]. Due to the synergistic effects present in biomass hydrolysates, the toxicity threshold of a specific compound can be much lower compared to that was tested in synthetic medium. Therefore, though the testing concentrations were higher compared to that in biomass hydrolysates, the testing results are still valuable.

The first group of compounds shown in Table 5a were identified by all three phenotype models, among which, furfural resulted in longer lag-phase at all three concentrations tested, while sorbic acid and syringaldehyde reduced growth rate. Suberic acid exhibited positive effect towards the fermenting yeast, mainly through shortening lag-phase. Since this phenomenon was only observed in MM, but not in YPD, which is a much richer medium compared to MM, we reason that the acid was probably used as a nutrient by the yeast. HMF, though known as an important inhibitor in biomass hydrolysates [48-50], only exhibited inhibitory effect at 1.0 g/l on the growth rate of the fermenting yeast (Table 5a). However, HMF seems to elongate lag-phase when tested together with other compounds suggested by the lag-phase models. It can be seen that HMF triggered synergistic effect with levulinic acid, 2-furoic acid and pantoic acid, respectively, at 0.5 g/l (Table 5b). This may be the reason why HMF was identified, though little effect was observed when tested individually.

Furfural was identified as a key toxin in biomass hydrolysates [13,51], and consistent with the current study, its main inhibitory effect was elongating lag-phase [52-54]. It was reported earlier that furfural as well as HMF are converted to their alcohol form (furfuryl alcohol and HMF alcohol) and eventually acid form (furoic acid and HMF acid) by the fermenting yeast due to detoxification [44,45]. This was also observed in this study. During lag-phase, the concentration of furfural and HMF reduced, while their alcohols and acids were formed. Since the concentration of furfuryl alcohol and 2-furoic acid is showing an opposite pattern

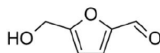
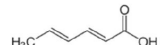
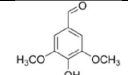
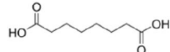
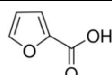
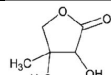
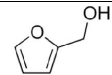
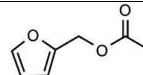
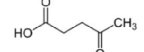
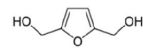
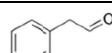
compared to furfural, and HMF alcohol to HMF, as could be expected, these compounds were also identified by analyzing the lag-phase models (Table 5a).

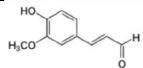
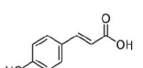
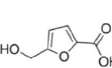
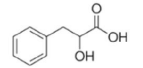
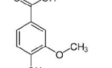
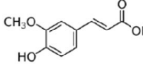
The potential inhibitors identified by Glu CR and EtOH PR models were mainly phenolic compounds (Table 5a). It is known from previous research that the toxic form of a phenolic compound is often the aldehyde, which is converted to its alcohol during the fermentation process due to detoxification [20,22,55]. Therefore, possible conversion of the phenolic alcohols suggested by the models was checked. For those phenolic alcohol compounds with increased concentrations during the fermentation process, the aldehyde forms were used in the growth tests, assuming that the alcohols were the conversion products. These phenolic aldehydes are marked in *italic* in Table 5a. In agreement with former studies, the compounds exhibited inhibitory effects were mostly aldehydes and acids (Table 5a). The major inhibitory effects were reduced growth rate and lower final OD. Phenylacetaldehyde, vanillin and conifer aldehyde caused growth deficiency at 1.0 g/l (0.5 g/l for phenylacetaldehyde, Table 5a).

Besides the compounds listed in Table 5a, another group suggested by the models were the amino acids, of which concentrations decreased during the fermentation process. This provides the possibility that the depletion of amino acids in hydrolysates worsened the fermentation performance of the fermenting yeast. However, as growth of the fermenting yeast in hydrolysates was not improved when amino acids were added (data not shown), this was apparently not the case. Another explanation would be that the presence of amino acids and possibly other nutrients compensates the inhibitory effects of the inhibitors. This assumption was verified by comparing the inhibitory effects of the compounds listed in Table 5a in MM and YPD medium, which contains abundant peptides and nutrients compared to MM. The inhibitory effects of all the tested compounds alleviated in YPD medium, particularly, the effects underlined in Table 5a were absent in YPD. This observation indicates that the toxicity of inhibitors was culture medium dependent, suggesting that the fermentability of biomass hydrolysates could be improved by adding extra nutrients like yeast extract [56].

Identifying inhibitors in biomass hydrolysates through exometabolomics

Table 5a Inhibitory effects of the compounds suggested by lag-phase, Glu CR and EtOH PR models.

reference medium (MM with 20 g/l glucose)		LP GR OD 7 h 0.105 1.2										
compounds	structure	0.2 g/l			0.5 g/l			1.0 g/l				
		LP	GR	OD	LP	GR	OD	LP	GR	OD		
compounds identified by all 3 phenotype models												
furfural		10 h	---	---	15 h	---	---	30 h	no growth			
HMF		---	---	---	---	---	---	---	< 20%	---		
sorbic acid		---	< 60%	< 80%	---	< 60%	< 80%	no growth				
syringaldehyde		---	---	---	---	---	---	---	< 60%	< 80%		
suberic acid								<u>3 h</u>	<u>110 %</u>	<u>110 %</u>		
no effect												
2-furoic acid					pantolactone							
compounds identified by lag-phase model												
benzoic acid		---	< 60%	< 80%	---	< 40%	< 80%	---	< 40%	< 60%		
no effect												
furfuryl alcohol					2-furanmethanol acetate							
levulinic acid					HMF alcohol							
compounds identified by Glu CR and EtOH PR models												
Phenylacet aldehyde		20 h	< 80%	---	no growth			no growth				
vanillin		---	---	---	11 h	< 80%	---	30 h	no growth			
4-hydroxybenzaldehyde		---	---	---	<u>9.5 h</u>	< 80%	---	<u>11 h</u>	< 60%	< 80%		
dihydroxy benzene		---	---	---	---	---	---	---	< 80%	---		

conifer aldehyde		---	---	---	23 h	---	---	no growth
no effect								
p-coumaric acid		homovanillic acid		HMF acid				
3-phenyllactic acid		4-hydroxybenzoic acid		vanillic acid				
compounds identified by Glu CR model								
ferulic acid		---	---	---	---	---	---	---
							$\leq$ 40%	$\leq$ 80%
benzaldehyde		---	---	---	9.5 h	< 80%	---	13.5 h < 80% < 80%
compounds identified by EtOH PR model								
salicylic acid		---	$\leq$ 80%	$\leq$ 80%	---	$\leq$ 60%	$\leq$ 80%	---
							$\leq$ 60%	$\leq$ 60%
not tested								
4-hydroxyphenylaldehyde	phloretic acid			5-HMF methyl keton				

LP: lag-phase: time needed to reach 2% (OD_{max}-OD_{min}) (h), GR: growth rate: the slope of the linear part of the OD curve (OD/h), OD: final OD. Values with % are relative growth rate and final OD compared to that in reference medium. '---': no effect compared to reference medium.

The compounds indicated in 'italic' were originally identified by their (converted and less toxic) alcohol forms; the compounds indicated in 'bold' were saturated when 1 g/l solutions were prepared at the fermentation temperature, in these cases, besides the saturated solution, a 2- and 5- fold dilution was used, represented in the 0.5 and 0.2 g/l columns, respectively; the 'underlined effects' were NOT observed when tested in YPD medium.

Furfural and HMF are the two most studied inhibitors in biomass hydrolysates, including their inhibitory effects as well as their conversion pathways [13,45,50]. However, the synergistic effects of these two compounds with other potential inhibitors in hydrolysates were seldom tackled. In this study, the joint inhibitory effects of furfural or HMF with one other potential inhibitory compound were tested at 0.5 g/l in MM with 20 g/l glucose, and the compounds gave synergistic effect with either furfural or HMF are listed in Table 5b. It can be seen that HMF caused a notable synergistic effect with levulinic acid, 2-furoic acid, pantoyllacton and syringaldehyde, respectively. These compounds showed no inhibitory effect individually at 0.5 g/l, but when added together with HMF, they elongated lag-phase as well as reduced

Identifying inhibitors in biomass hydrolysates through exometabolomics

growth rate (Table 5b). Compared to HMF, the synergistic effect caused by furfural was minor, since no significant lag-phase increase or growth rate reduction was observed when an extra compound was added (Table 5b).

Table 5b Compounds that caused synergistic effect with furfural or HMF at 0.5 g/l.

reference medium (MM with 20 g/l glucose)		LP 7 h	GR 0.105	OD 1.2					
furfural 0.5 g/l 15 h 90% 100%	+	compounds			=	mixture			
		LP	GR	OD		LP	GR	OD	
		HMF	---	---		---	<u>19 h</u>	<u>< 80%</u>	---
		HMF acid	---	---		---	16.5 h	---	---
		salicylic acid	---	< 60%		< 80%	15 h	<u>< 40%</u>	< 60%
		vanillin	11 h	< 80%		---	17 h	< 80%	---
		syringaldehyde	---	---		---	16.5 h	<u>< 80%</u>	---
HMF 0.5 g/l 7 h 100% 100%	+	levulinic acid	---	---	---	=	9 h	<u>< 80%</u>	---
		2-furoic acid	---	---	---		8.5 h	<u>< 80%</u>	---
		pantoyllacton	---	---	---		9 h	<u>< 80%</u>	---
		salicylic acid	---	< 60%	< 80%		8.5 h	<u>< 40%</u>	< 60%
		vanillin	11 h	< 80%	---		11 h	<u>< 60%</u>	---
		syringaldehyde	---	---	---		<u>11 h</u>	---	---

LP: lag-phase: time needed to reach 2% (ODmax-ODmin) (h), GR: growth rate: the slope of the linear part of the OD curve (OD/h), OD: final OD. Values with % are relative growth rate and final OD compared to that in reference medium. '---': no effect compared to reference medium; the 'underlined effects' were relatively significant.

Discussion

Lignocellulosic biomass is a natural resource that has the potential to become the major feedstock for biofuel production [57,58]. To identify inhibitory compounds in biomass hydrolysates, a metabolomics approach was adopted in this study. Compared to targeted methods, no compound pre-selection was made with the metabolomics approach, so that the inhibitor identification was not influenced by prior knowledge [18,26]. The study results show that the metabolomics approach successfully identified compounds that influence the growth of the fermenting yeast, *Saccharomyces cerevisiae* CEN.PK 113-7D. Some compounds elongated lag-phase, like furfural and vanillin, while others reduced growth, such as HMF and benzaldehyde. Interestingly, without pre-selection, compounds that were previously known as inhibitors in biomass hydrolysates were identified in this study. This confirms that metabolomics is a relevant approach in studying the composition and identifying inhibitors of lignocellulosic biomass hydrolysates.

As the analysis targets were potential inhibitory compounds in biomass hydrolysates, which are weak acids, furans and phenolic compounds [15-17], GC-MS was chosen as the analytical tool [18]. To remove sugars in hydrolysate samples, ethyl acetate (EA) extraction or ethyl chloroformate (EC) derivatization was conducted prior to sample analysis. Due to the property difference of these two sample preparation methods, their target compound groups were also different. The EA method had reliable measurement for aromatic compounds, while the EC methods mainly detected carboxylic acids and furans. Remarkably, this difference in analytical method in relation to metabolomics results was also seen during statistical model building, as EA-GC-MS data-sets could predict Glu CR and EtOH PR properly, but failed to model lag-phase on its own, which was validly predicted by EC-GC-MS data. Accordingly, furans were mainly identified to elongate lag-phase, and aromatic compounds were mostly responsible for reduced growth. These results suggest that in a metabolomics study, it is important to have a wide coverage of detectable compounds, so that the chance of overlooking potential target compounds can be reduced [27,59]. And one way of achieving this is to use multiple analytical tools for measuring the same sample.

Furfural and HMF were reported as the two most important inhibitors in biomass hydrolysates, which delay as well as reduce growth [13,45,53,60]. In the growth test of this study, it was found that furfural indeed elongated lag-phase at a concentration of 0.2 g/l, but HMF did not display any inhibitory effect until its concentration reached above 0.5 g/l (Table 5a). However, when tested jointly, HMF enhanced the negative effect of furfural on lag-phase, and reduced growth rate. Besides, when HMF was tested together with other compounds,

which showed no effect individually, like levulinic acid, 2-furoic acid and pantoyllacton, inhibition took place (Table 5b). These observations suggest that HMF probably functions as a co-inhibitor in biomass hydrolysate, for which inhibition is mainly the result of synergistic effects. Furthermore, synergistic effect reduces the threshold concentration for inhibition. For instance, both HMF and syringaldehyde showed toxicity only at 1.0 g/l towards the fermenting yeast, but when tested jointly, the inhibitory effect was present at 0.5 g/l (Table 5b). So it is possible that when multiple inhibitors are present, the toxicity threshold of HMF and syringaldehyde reduce to below 0.1 g/l, which is close to their reported concentration in biomass hydrolysates [11,14,46,47].

Among the compounds identified with Glu CR and EtOH PR models, a group showed no effect in the growth test. Noticeably, this group of compounds is all aromatic acids (Table 5a). Earlier studies demonstrated that aldehyde was the most toxic form of aromatic compounds, the corresponding acids were less, while the alcohol form was the least toxic [20,22,55]. This was confirmed in this study, and was clearly illustrated with vanillin and vanillin acid, of which the acid form had no effect, while the aldehyde form almost abolished growth at 1.0 g/l (Table 5a). Besides the identification of previously reported inhibitors in biomass hydrolysates [9,15,17,55,60,61], two new compounds were found to be toxic, which are sorbic acid and phenylacetaldehyde. As shown in Table 5a, both compounds already showed significant inhibitory effect on growth at 0.2 g/l. The high toxicity towards the fermenting yeast indicates that these two compounds are important inhibitors in biomass hydrolysates. Though not recorded as hydrolysate inhibitors, sorbic acid was described as a preservative weak acid, which disturbs yeast growth through uncoupling mechanism [8,60,62,63], while phenylacetaldehyde was known of having antibiotic activity in maggot therapy [64]. It should be mentioned that the enzyme cocktail used in this study also contains sorbic acid, so the sorbic acid detected in biomass hydrolysates was partially from addition of the hydrolyzing enzyme in most feedstock hydrolysates.

Of the potential inhibitory compounds suggested by the statistical models, about half are unknowns. Some of these compounds are on the very top of the ranking lists, see Appendix 2-1, 2-2. Since most of the known compounds suggested by the models showed inhibitory effect towards the fermenting yeast in growth tests, it is expected that there are also important / novel inhibitors among the unknown compounds. To verify this, identification needs to be conducted for these unknown compounds, which will be the next step in identifying lignocellulosic biomass hydrolysate inhibitors.

Through applying metabolomics approach, the inhibition property of these compounds was connected to their presence in lignocellulosic biomass hydrolysates. To our knowledge, this is the first systematic study on identifying inhibitory compounds in lignocellulosic biomass hydrolysates using a non-targeted approach.

Conclusion

Inhibitory compounds in lignocellulosic biomass hydrolysates were successfully identified through applying an exometabolomics approach. The identification was conducted by relating the fermentability of biomass hydrolysates with their composition using statistical models, (n)PLS-2CV. The non-sugar composition of biomass hydrolysates were analyzed with two GC-MS methods, using ethyl acetate extraction and ethyl chloroformate derivatization to remove sample sugar contents, respectively. Among the identified compounds, besides the known inhibitors, sorbic acid and phenylacetaldehyde were for the first time identified as inhibitors in lignocellulosic biomass hydrolysates.

Acknowledgments

This project was (co) financed by the Netherlands Metabolomics Centre (NMC) which is part of the Netherlands Genomics Initiative / Netherlands Organization for Scientific Research.

References

1. Hahn-Hägerdal B, Galbe M, Gorwa-Grauslund M,F, Lidén G, Zacchi G: Bio-ethanol -- the fuel of tomorrow from the residues of today. *Trends Biotechnol.* **2006**, 24, 549-556.
2. González-García S, Moreira MT, Feijoo G: Comparative environmental performance of lignocellulosic ethanol from different feedstocks. *Renew. Sust. Energ. Rev.* **2010**, 14, 2077-2085.
3. Jørgensen H, Kristensen JB, Felby C: Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities. *Biofuel Bioprod. Bior.* **2007**, 1, 119-134.
4. van Dyk JS, Pletschke BI: A review of lignocellulose bioconversion using enzymatic hydrolysis and synergistic cooperation between enzymes-factors affecting enzymes, conversion and synergy. *Biotechnol. Adv.* **2012**, 30, 1458-1480.
5. Alvira P, Tomás-Pejó E, Ballesteros M, Negro MJ: Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review. *Bioresource Technol.* **2010**, 101, 4851-4861.
6. Gupta R, Khasa YP, Kuhad RC: Evaluation of pretreatment methods in improving the enzymatic saccharification of cellulosic materials. *Carbohydr. Polym.* **2011**, 84, 1103-1109.

7. Sun Y, Cheng J: Hydrolysis of lignocellulosic materials for ethanol production: A review. *Bioresource Technol.* **2002**, 83, 1-11.
8. Palmqvist E, Hahn-Hägerdal B: Fermentation of lignocellulosic hydrolysates. II: Inhibitors and mechanisms of inhibition. *Bioresource Technol.* **2000**, 74, 25-33.
9. Klinke HB, Thomsen AB, Ahring BK: Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass. *Appl. Microbiol. Biot.* **2004**, 66, 10-26.
10. Panagiotopoulos IA, Bakker RR, de Vrije T, Koukios EG: Effect of pretreatment severity on the conversion of barley straw to fermentable substrates and the release of inhibitory compounds. *Bioresource Technol.* **2011**, 102, 11204-11211.
11. Zha Y, Muilwijk B, Coulier L, Punt PJ: Inhibitory compounds in lignocellulosic biomass hydrolysates during hydrolysate fermentation processes. *J. Bioprocess Biotechnol.* **2012**, 2, doi: 10.4172/2155-9821.1000112.
12. Huang H, Guo X, Li D, Liu M, Wu J, Ren H: Identification of crucial yeast inhibitors in bio-ethanol and improvement of fermentation at high pH and high total solids. *Bioresource Technol.* **2011**, 102, 7486-7493.
13. Heer D, Sauer U: Identification of furfural as a key toxin in lignocellulosic hydrolysates and evolution of a tolerant yeast strain. *Microb. Biotechnol.* **2008**, 1, 497-506.
14. Thomsen MH, Thygesen A, Thomsen AB: Identification and characterization of fermentation inhibitors formed during hydrothermal treatment and following SSF of wheat straw. *Appl. Microbiol. Biotechnol.* **2009**, 83, 447-455.
15. Taherzadeh MJ, Karimi K: Fermentation Inhibitors in Ethanol Processes and Different Strategies to Reduce Their Effects. In *Biofuels: Alternative Feedstocks and Conversion Processes*. Pandey A, Larroche C, Ricke SC, Dussap C, Gnansounou E, Eds., Academic Press: USA, 2011, pp. 287-311.
16. Almeida JRM, Modig T, Petersson A, Hahn-Hägerdal B, Lidén G, Gorwa-Grauslund MF: Increased tolerance and conversion of inhibitors in lignocellulosic hydrolysates by *Saccharomyces cerevisiae*. *J. Chem. Technol. Biot.* **2007**, 82, 340-349.
17. Jönsson L, Alriksson B, Nilvebrant N: Bioconversion of lignocellulose: inhibitors and detoxification. *Biotechnol. Biofuels* **2013**, 6, 16, doi:10.1186/1754-6834-6-16.
18. Zha Y, Punt PJ: Exometabolomics approaches in studying the application of lignocellulosic biomass as fermentation feedstock. *Metabolites* **2013**, 3, 119-143.
19. Chen SF, Mowery RA, Castleberry VA, van Walsum G, Chambliss CK: High-performance liquid chromatography method for simultaneous determination of aliphatic acid, aromatic acid and neutral degradation products in biomass pretreatment hydrolysates. *J. Chromatogr. A* **2006**, 1104, 54-61.
20. Klinke HB, Olsson L, Thomsen AB, Ahring BK: Potential inhibitors from wet oxidation of wheat straw and their effect on ethanol production of *Saccharomyces cerevisiae*: Wet oxidation and fermentation by yeast. *Biotechnol. Bioeng.* **2003**, 81, 738-747.
21. Larsson S, Palmqvist E, Hahn-Hägerdal B, Tengborg C, Stenberg K, Zacchi G, Nilvebrant NO: The generation of fermentation inhibitors during dilute acid hydrolysis of softwood. *Enzyme Microb. Technol.* **1999**, 24, 151-159.
22. Jönsson LJ, Palmqvist E, Nilvebrant N-O, Hahn-Hägerdal B: Detoxification of wood hydrolysates with laccase and peroxidase from the white-rot fungus *Trametes versicolor*. *Appl. Microbiol. Biotechnol.* **1998**, 49, 691-697.
23. Martinez A, Rodriguez ME, York SW, Preston JF, Ingram LO: Effects of Ca(OH)₂ treatments ('overliming') on the composition and toxicity of bagasse hemicellulose hydrolysates. *Biotechnol. Bioeng.* **2000**, 69, 526-536.
24. Roessner U, Bowne J: What is metabolomics all about? *Biotechniques*. **2009**, 46, 363-365.
25. Fiehn O: Metabolomics -- the link between genotypes and phenotypes. *Plant. Mol. Biol.* **2002**, 48, 155-171.

26. van der Werf MJ, Jellema RH, Hankemeier T: Microbial metabolomics: replacing trial-and-error by the unbiased selection and ranking of targets. *J. Ind. Microbiol. Biot.* **2005**, 32, 234-252.
27. Braaksma M, Bijlsma S, Coulter L, Punt PJ, van der Werf MJ: Metabolomics as a tool for target identification in strain improvement: the influence of phenotype definition. *Microbiology* **2011**, 157, 147-159.
28. Krishnan A, Williams LJ, McIntosh AR, Abdi H: Partial Least Squares (PLS) methods for neuroimaging: A tutorial and review. *Neuroimage* **2011**, 56, 455-475.
29. Bro R: Multiway calibration. Multilinear PLS. *J. Chemometrics* **1996**, 10, 47-61.
30. Smilde AK: Comments on multilinear PLS. *J. Chemometrics* **1997**, 11, 367-377.
31. Westerhuis J, Hoefsloot H, Smit S, Vis D, Smilde A, van Velzen E, van Duijnhoven J, van Dorsten F: Assessment of PLSDA cross validation. *Metabolomics* **2008**, 4, 81-89.
32. Smit S, van Breemen MJ, Hoefsloot HCJ, Smilde AK, Aerts JMFG, de Koster CG: Assessing the statistical validity of proteomics based biomarkers. *Anal. Chim. Acta.* **2007**, 592, 210-217.
33. Zha Y, Slomp R, Groenestijn J, Punt PJ: Preparation and Evaluation of Lignocellulosic Biomass Hydrolysates for Growth by Ethanologenic Yeasts. In *Microbial Metabolic Engineering: Methods and Protocols*. Qiong C Ed., Humana Press: New York, USA, 2012, pp. 245-259.
34. van Hoek P, van Dijken JP, Pronk JT: Effect of specific growth rate on fermentative capacity of baker's yeast. *Appl. Environ. Microbiol.* **1998**, 64, 4226-4233.
35. Andersson CA, Bro R: The n-way toolbox for MATLAB. *Chemometrics Intellig. Lab. Syst.* **2000**, 52, 1-4.
36. Rajalahti T, Arneberg R, Berven FS, Myhr K, Ulvik RJ, Kvalheim OM: Biomarker discovery in mass spectral profiles by means of selectivity ratio plot. *Chemometrics Intellig. Lab. Syst.* **2009**, 95, 35-48.
37. Kvalheim OM: Interpretation of partial least squares regression models by means of target projection and selectivity ratio plots. *J. Chemometrics* **2010**, 24, 496-504.
38. Hendriks MMWB, van Eeuwijk FA, Jellema RH, Westerhuis JA, Reijmers TH, Hoefsloot HCJ, Smilde AK: Data-processing strategies for metabolomics studies. *TrAC Trend Anal. Chem.* **2011**, 30, 1685-1698.
39. van den Berg R, Hoefsloot HCJ, Westerhuis J, Smilde A, van der Werf MJ: Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* **2006**, 7, 142.
40. Brejning J, Jespersen L: Protein expression during lag phase and growth initiation in *Saccharomyces cerevisiae*. *Int. J. Food Microbiol.* **2002**, 75, 27-38.
41. Kuyper M, Winkler AA, van Dijken JP, Pronk JT: Minimal metabolic engineering of *Saccharomyces cerevisiae* for efficient anaerobic xylose fermentation: a proof of principle. *FEMS Yeast Res.* **2004**, 4, 655-664.
42. Humpala JF, Chundawat SPS, Vismeh R, Jones AD, Balan V, Dale BE: Rapid quantification of major reaction products formed during thermochemical pretreatment of lignocellulosic biomass using GC-MS. *J. Chromatogr. B* **2011**, 879, 1018-1022.
43. Ranatunga TD, Jervis J, Helm RF, McMillan JD, Hatzis C: Identification of inhibitory components toxic toward *Zymomonas mobilis* CP4(pZB5) xylose fermentation. *Appl. Biochem. Biotech.* **1997**, 67, 185-198.
44. Taherzadeh MJ, Gustafsson L, Niklasson C, Lidén G: Conversion of furfural in aerobic and anaerobic batch fermentation of glucose by *Saccharomyces cerevisiae*. *J. Biosci. Bioeng.* **1999**, 87, 169-174.
45. Liu ZL, Slininger PJ, Dien BS, Berhow MA, Kurtzman CP, Gorsich SW: Adaptive response of yeasts to furfural and 5-hydroxymethylfurfural and new chemical evidence for HMF conversion to 2,5-bis-hydroxymethylfuran. *J. Ind. Microbiol. Biot.* **2004**, 31, 345-352.

46. Klinké HB, Ahring BK, Schmidt AS, Thomsen AB: Characterization of degradation products from alkaline wet oxidation of wheat straw. *Bioresource Technol.* **2002**, 82, 15-26.
47. Millati R, Niklasson C, Taherzadeh MJ: Effect of pH, time and temperature of overliming on detoxification of dilute-acid hydrolysates for fermentation by *Saccharomyces cerevisiae*. *Process Biochem.* **2002**, 38, 515-522.
48. Almeida JRM, Modig T, Röder A, Lidén G, Gorwa-Grauslund M: Pichia stipitis xylose reductase helps detoxifying lignocellulosic hydrolysate by reducing 5-hydroxymethyl-furfural (HMF). *Biotechnol. Biofuels* **2008**, 1, 12.
49. Wierckx N, Koopman F, Bandounas L, de Winde JH, Ruijsenaars HJ: Isolation and characterization of *Cupriavidus basilensis* HMF14 for biological removal of inhibitors from lignocellulosic hydrolysate. *Microb. Biotechnol.* **2010**, 3, 336-343.
50. Koopman F, Wierckx N, de Winde JH, Ruijsenaars HJ: Identification and characterization of the furfural and 5-(hydroxymethyl)furfural degradation pathways of *Cupriavidus basilensis* HMF14. *P. Natl. A. Sci.* **2010**, 107, 4919-4924.
51. Tofighi A, Azin M, Assadi MM, Assadi-rad MHA, Nejadshattari T, Fallahian MR: Inhibitory effect of high concentrations of furfural on industrial strain of *Saccharomyces cerevisiae*. *Int. J. Environ. Res.* **2010**, 4, 137-142.
52. Li Q, Matthew Lam LK, Xun L: Biochemical characterization of ethanol-dependent reduction of furfural by alcohol dehydrogenases. *Biodegradation* **2011**, 22, 1227-1237.
53. Allen S, Clark W, McCaffery JM, Cai Z, Lanctot A, Slininger P, Liu ZL, Gorsich S: Furfural induces reactive oxygen species accumulation and cellular damage in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* **2010**, 3, 2.
54. Ask M, Bettiga M, Mapelli V, Olsson L: The influence of HMF and furfural on redox-balance and energy-state of xylose-utilizing *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* **2013**, 6, 22.
55. Larsson S, Quintana-Sáinz A, Reimann A, Nilvebrant NO, Jönsson LJ: Influence of lignocellulose-derived aromatic compounds on oxygen-limited growth and ethanolic fermentation by *Saccharomyces cerevisiae*. *Appl. Biochem. Biotech.* **2000**, 84-86, 617-632.
56. Jørgensen H: Effect of nutrients on fermentation of pretreated wheat straw at very high dry matter content by *Saccharomyces cerevisiae*. *Appl. Biochem. Biotechnol.* **2009**, 153, 44-57.
57. Mabee WE, McFarlane PN, Saddler JN: Biomass availability for lignocellulosic ethanol production. *Biomass Bioenergy* **2011**, 35, 4519-4529.
58. Zhang YHP: What is vital (and not vital) to advance economically-competitive biofuels production. *Process Biochem.* **2011**, 46, 2091-2110.
59. Dunn WB, Ellis DI: Metabolomics: Current analytical platforms and methodologies. *TrAC Trend Anal. Chem.* **2005**, 24, 285-294.
60. Keating JD, Panganiban C, Mansfield SD: Tolerance and adaptation of ethanologenic yeasts to lignocellulosic inhibitory compounds. *Biotechnol. Bioeng* **2006**, 93, 1196-1206.
61. Palmqvist E, Hahn-Hägerdal B: Fermentation of lignocellulosic hydrolysates. I: inhibition and detoxification. *Bioresource Technol.* **2000**, 74, 17-24.
62. Mira NP, Lourenço AB, Fernandes AR, Becker JD, Sá-Correia I: The RIM101 pathway has a role in *Saccharomyces cerevisiae* adaptive response and resistance to propionic acid and other weak acids. *FEMS Yeast Res.* **2009**, 9, 202-216.
63. Stratford M, Plumridge A, Archer DB: Decarboxylation of sorbic acid by spoilage yeasts is associated with the PAD1 gene. *Appl. Environ. Microb.* **2007**, 73, 6534-6542.
64. Pavillard ER, Wright EA: An antibiotic from maggots. *Nature* **1957**, 180, 916-917.

CHAPTER 6

Pichia anomala 29X: A Resistant Strain for Lignocellulosic Biomass Hydrolysate Fermentation

Ying Zha, Abeer H. Hossain, Felix Tobola, Norbert Sedee,
Mieke Havekes, and Peter J. Punt

FEMS Yeast Research **2013**, DOI: 10.1111/1567-1364.12062

Abstract

To efficiently use lignocellulosic biomass hydrolysates as fermentation media for bioethanol production, besides being capable of producing significant amount of ethanol, the fermenting host should also meet the following two requirements: (1) resistant to the inhibitory compounds formed during biomass pretreatment process, (2) capable of utilizing C5 sugars, such as xylose, as carbon source. In our lab, a screening was conducted on microorganisms collected from environmental sources for their tolerance to hydrolysate inhibitors. A unique resistant strain was selected and identified as *Pichia anomala* (*Wickerhamomyces anomalus*), deposited as CBS 132101. The strain is able to produce ethanol in various biomass hydrolysates, both with and without oxygen. Besides, the strain could assimilate xylose and use nitrate as N-source. These physiological characteristics make *Pichia anomala* an interesting strain for bioethanol production from lignocellulosic biomass hydrolysates.

Introduction

The future feedstock for bioethanol production is lignocellulosic biomass, which is abundant, renewable and inexpensive [1]. The typical procedure for bioethanol production from biomass consists biomass pretreatment and hydrolysis, hydrolysate fermentation, and ethanol separation [2]. In the pretreatment process, harsh conditions are used to break down the structure of biomass and expose cellulose and hemicellulose for hydrolysis. The harsh conditions also cause the formation of compounds that inhibit the hydrolysis as well as fermentation process [3-5]. After the pretreatment, the biomass is hydrolyzed into a mixture containing high concentrations of glucose and xylose, low amounts of galactose, arabinose and mannose, and inhibitory compounds [6]. This mixture, so-called hydrolysate, is used as fermentation medium for ethanol production.

To transform sugars in hydrolysates to ethanol with high efficiency, selection of the appropriate microorganism(s) is of high importance. Based on the hydrolysate composition, it is suggested that the desired microorganism(s) is to meet the requirements of being able to utilize glucose as well as xylose as carbon source and being resistant to hydrolysate inhibitors. The current ethanol production host is baker's yeast *Saccharomyces cerevisiae*, which converts glucose to ethanol with high rate [7]. However, native *S.cerevisiae* is not able to use xylose, and although efforts have been made to engineer this host for xylose consumption [8,9], until now the efficiency remained unsatisfying in feedstock hydrolysates [10-12]. In contrast to *S.cerevisiae*, *Scheffersomyces stipitis* is able to use xylose as carbon source for ethanol production [13,14], but its growth rate is very low when hydrolysate inhibitors are present [15,16]. These limitations make it interesting to discover new microorganisms, which meet both requirements for fermenting biomass hydrolysates.

In this study, we attempted to isolate strains from diverse environmental sources, and to select the ones that show resistance to hydrolysate inhibitors. Subsequently, the identification of the selected strains was carried out and their fermentation performance was analyzed.

Materials and Methods

Strain isolation

The environmental sources for isolating microorganisms are listed in Table 1. Briefly, these sources were suspended in physiological salt solution individually and incubated at 25°C, 150 rpm for 2 hours. The incubated suspensions were filtrated, 40 ml filtrates were collected and centrifuged. After centrifugation, the pellets were re-suspended in 4 ml fresh physiological salt solution. These solutions were then plated on plates with tryptone soya agar (Oxoid CM0131), Schaedler agar (Oxoid CM0437), and Oxytetracycline glucose yeast extract agar (Oxoid CM0545). The tryptone soya agar and oxytetracycline glucose yeast extract agar plates were incubated at 30°C aerobically, while schaedler agar and oxytetracycline glucose yeast extract agar plates were used to cultivate microaerophilic microorganisms. In total 184 colonies were selected and inoculated in 96-well microtiter plates with 0.5 ml tryptone soya broth as medium. The colonies that grew under aerobic conditions were incubated aerobically, and the microaerophilic colonies were incubated under microaerobic conditions.

The resistance of the isolated microorganisms to hydrolysate inhibitors was tested using 96-well microtiter plates. Mineral medium (MM) [17] was supplied with a mixture of 10 mM furfural, 10 mM HMF, 10 mM acetic acid, and 10% wheat straw hydrolysate prepared with a concentrated acid method [18]. The pH of the culture medium was adjusted to 5.0 before inoculation. The tests were conducted in duplicate under both aerobic and microaerobic conditions dependent on the original selection condition. The plates were incubated at 30°C, 100 rpm for 5 days, and the growth was checked either visually or by measuring optical density (OD) of each well using Infinite® F500 (TECAN). The strains growing in this medium were analyzed for their identities by sequencing the D2 domain of large subunit (26S) rRNA genes (the D2-LSU procedure: BASECLEAR, Leiden, the Netherlands).

Strains, biomass hydrolysates and growth test

Pichia anomala (*Wickerhamomyces anomalus*) 29 (CBS134880), 32, 35 and *Pichia burtonii* (*Hyphopichia burtonii*) were the strains isolated in this study. *P. anomala* 29X (CBS132101), was selected as a xylose consuming derivative from *P. anomala* 29. *P. anomala* CBS 1984 and CBS 5759, and *Saccharomyces cerevisiae* CEN.PK 113-7D (CBS8340) were purchased from CBS (Utrecht, the Netherlands), and *P. anomala* J121 (CBS 100487) was kindly provided by Dr. Passoth (SLU, Sweden).

The biomass used were sugar cane bagasse (Bag) (Zilor, Brazil), corn stover (CS) (University of Cape Town, South Africa), wheat straw (WS) (Oostwaardshoeve, The Netherlands), and oak sawdust (Oak) (wood-flooring supplier ESCO, The Netherlands). Prior to pretreatment, biomass (except Oak) was ground to pieces of average length 3 mm and dried at 80 °C for at least 16 hours. To prepare 1 l hydrolysate, 300 g dried biomass was used. The pretreatment methods were dilute acid (DA) (2% H₂SO₄), and concentrated acid (CA) (72% H₂SO₄). The pretreated biomass was hydrolyzed enzymatically, using Accellerase 1500 (Genencor®) [18]. After hydrolysis, solid content was separated from the hydrolysate by filtration, and the filtrated hydrolysate was sterilized using filter sterilization and stored at 4 °C before use.

Growth tests were conducted in honeycomb plates, using Bioscreen C Analyzer (Labsystems OY). MM with 20 g/l glucose was used as reference medium, into which the selected inhibitors were added individually with the concentrations listed in Table 2. The detailed procedure of Bioscreen C test is described in Zha *et al.* [18]. All tests were carried out in triplicate.

Fermentation

Batch fermentations were carried out in 2 l New Brunswick fermentors, using 1 l of either MM with 60 g/l glucose and 30 g/l xylose or sterilized hydrolysate as substrate. The strain used for fermentation was *P. anomala* 29X (CBS 132101), and the inoculum was prepared in a 500 ml Erlenmeyer flask. The cells were harvested by centrifugation after incubating 40 hours in MM with 20 g/l xylose as carbon source, and inoculated into fermentors with density of 0.2 g cell dry weight per 1 l medium. All fermentations were carried out at 30 °C, and pH 5 by adding 1 M H₂SO₄ or 2 M KOH. The anaerobic condition was created by sparging 0.5 l/min N₂ continuously, and 20% air was mixed into the gas-in flow when oxygen was required. During the whole fermentation process, CO₂ concentration in the off-gas was monitored automatically and samples were taken at fixed time intervals. These samples were kept at 4 °C and used to measure OD, glucose and ethanol concentration with Arena® 20 Analyzer (Thermo Scientific), and xylose with YSI 2700 SELECT Biochemistry Analyzer.

Results and Discussion

Resistant strain isolation

With the purpose of discovering natural isolates that are resistant to the inhibitory compounds in biomass hydrolysates, diverse environmental sources were collected and explored to obtain bacterial, yeast and fungal colonies. In total, 184 colonies were obtained from 10 sources, as listed in Table 1.

Table 1 Colonies obtained from the 10 environmental sources collected in our lab.

source type	bacteria		yeast / fungi	
	aerobic	microaerobic	aerobic	microaerobic
fresh cow rumen fluid	7	7	0	0
fermented cow rumen fluid	6	6	0	0
garden compost	6	7	5	5
vegetable/fruit compost	6	4	0	3
mushroom compost	13	10	3 (1) ^a	1 (1)
grass silage	8	8	6 (2^b)	4 (2)
goat manure	3	4	3	4 (1)
sewage	7	6	4 (1)	4 (1)
industrial anaerobic waste water	6	4	1	2 (1)
industrial aerobic waste water	8	8	2	3 (1)
total	70	64	24	26

^a Strains in () grew with addition of 10 mM furfural, 10 mM HMF, 10 mM acetic acid and 10% wheat straw hydrolysate; ^b The four strains in **bold** showed the highest OD increase.

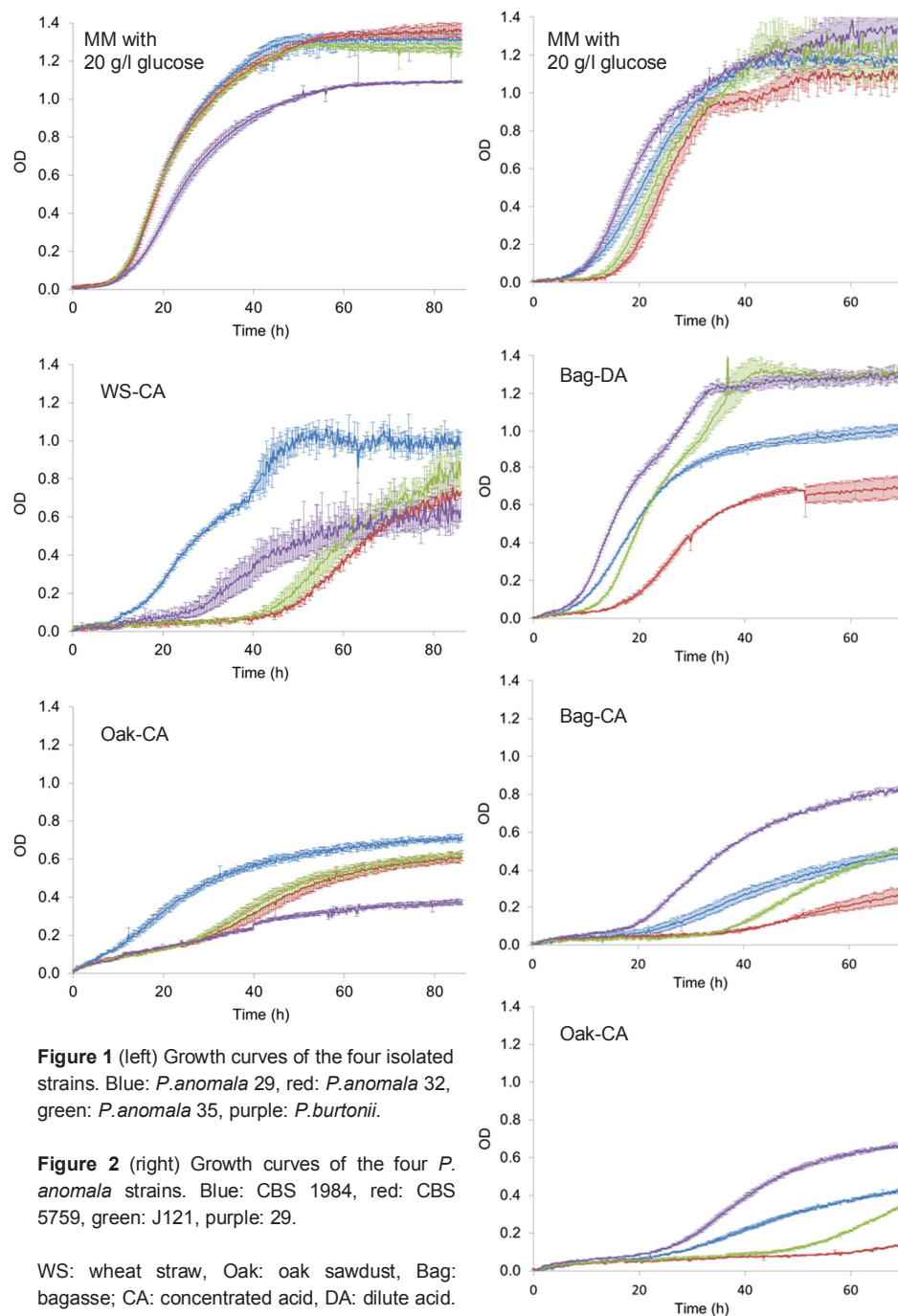
To select colonies that show resistance to biomass hydrolysate inhibitors, furfural, HMF and acetic acid were used as carbon source in MM for growth screening. These three compounds are the representatives of inhibitors in biomass hydrolysates [19-21]. In a final screen, 11 strains grew in a mixture of furfural, HMF, acetic acid and 10% wheat straw hydrolysate, among which four showed clear OD increase (Table 1). Interestingly, the corresponding colonies of these four strains were all isolated from grass silage. Of these four, based on partial 26S sequencing (2D-LSU), three isolates (#29,32,35) were identified as *Pichia anomala*, (*Wickerhamomyces anomalus*), one as *Pichia burtonii* (*Hyphopichia burtonii*; isolate #34). DNA sequence Accession numbers of the sequenced partial 26S rRNA genes of the four isolates are HF952836 (isolate #29), HF952837 (isolate #32), HF952838 (isolate #35), and HF952839 (isolate #34).

Resistant strain selection in hydrolysates

Interestingly, of the four most resistant strains isolated from grass silage, three were identified as *P. anomala* (*Wickerhamomyces anomalus*). This yeast was also found on insects and fruits [22,23], and is known for its ethanol and ethyl acetate production [24,25]. One of the most studied *P. anomala* strain was J121, its physiological characteristics and metabolite profiles were discussed in detail by Fredlund *et al.* [26,27]. Although there was a brief description on ethanol production by a *P. anomala* strain in alkaline hydrolyzed citrus peel [28], its ability to ferment diverse biomass hydrolysates and its resistance to biomass hydrolysate inhibitors was not studied.

In our study, the selected three *P. anomala* strains #29,32,35 and the *P. burtonii* strain were tested for their growth in hydrolysates prepared from WS and Oak. To demonstrate the resistance and difference of the four strains, hydrolysates pretreated with CA method were used as media, since it was shown that CA hydrolysates were most toxic to baker's yeast [18]. The growth tests were conducted in Bioscreen C Analyzer at 30°C in triplicate, and the results are shown in Figure 1. It can be seen that for all four strains, growth took place in both hydrolysates but with a lower rate than in MM with 20 g/l glucose. Strains were more sensitive in Oak-CA hydrolysate, which is more toxic compared to WS-CA ([18] and Figure 1).

Among the four *Pichia* strains, *P. anomala* 29 exhibited a unique tolerance to hydrolysate inhibitors, as its growth was considerably better. *P. burtonii* strain already showed a poorer growth performance in MM with 20 g/l glucose compared to the three *P. anomala* strains, which was even more the case in hydrolysates, indicating that *P. burtonii* is not a suitable potential fermenting yeast for biomass hydrolysates. The growth rate of *P. anomala* 32 and 35 was highly similar, but lower compared to strain 29 in both WS-CA and Oak-CA hydrolysates (Figure 1). Since *P. anomala* 29 performed better in hydrolysates compared to strain 32 and 35, it was suspected that *P. anomala* 29 has unique resistance to hydrolysate inhibitors. To verify this, three other *P. anomala* strains from culture collections, namely CBS 1984, CBS 5759 and J121, were tested for their growth in hydrolysates, together with *P. anomala* 29. CBS 5759 and its haploid derivative CBS 1984 were analyzed for their fermentation characteristics showing that their central carbon metabolism regulation was mainly influenced by oxygen availability [29]. *P. anomala* J121 was isolated from airtight stored grain and is used as a biocontrol yeast for preservation of moist grain [26].



The growth of these *P. anomala* strains was compared in MM with 20 g/l glucose and several different hydrolysates, as shown in Figure 2. In MM, their growth performance was comparable, though *P. anomala* 29 grew slightly faster. Different from in MM, in hydrolysates, clear growth variations were observed. In all tested hydrolysates, *P. anomala* 29 performed the best, had the shortest lag-phase, the fastest growth rate and reached the highest final OD. This indicates that *P. anomala* 29 is unique in its high resistance to the inhibitory compounds in biomass hydrolysates. It is also noticed that in hydrolysates with higher toxicity, i.e. ones prepared with CA method, the growth difference between *P. anomala* 29 and the other three strains was larger. In DA prepared hydrolysate, which is less toxic, the performance of J121 was similar to that of *P. anomala* 29. These results suggest that especially in hydrolysates with high toxicity, *P. anomala* 29 is a suitable fermenting yeast.

Xylose utilization

It is known from previous research that *Pichia anomala* is able to use xylose as carbon source [22]. Since xylose is the second most abundant sugar form in lignocellulosic biomass hydrolysates, it is important to study if *P. anomala* 29 is also able to utilize xylose.

For this purpose, *P. anomala* 29 was cultivated in shake flask with MM and 20 g/l xylose as sole carbon source. No growth was observed in the first 48 hours, after which a slight OD increase was shown. This indicates that *P. anomala* 29 is able to use xylose as carbon source, but with a much lower efficiency compared to glucose.

To improve the growth of *P. anomala* 29 on xylose, 2 ml of the xylose culture was transferred into another shake flask with fresh MM and 20 g/l xylose for growth. This procedure was repeated until growth was observed within 20 hours after inoculation. The *P. anomala* 29 derivative, referred to as 29X, with improved xylose consumption ability was deposited in CBS under number 132101, and used as the inoculum for all fermentations described below.

Fermentation

The small scale (400 μ l) growth tests demonstrated that *Pichia anomala* 29 is a potential resistant strain for biomass hydrolysate fermentation. Besides, the shake flask experiment showed that the strain is able to use xylose as carbon source, which is a highly favorable characteristic in fermenting biomass hydrolysates [30,31]. Another potential advantage of using *P. anomala* as fermentation host is that the contamination occurrence may be reduced due to its production of killer toxin [26,32]. Furthermore, the isolated *P. anomala* strain is able

to use NO_3 as N-source (data not shown), making it even more versatile in using substrates with low NH_4 or amino acid levels [26,33]. To initially understand the growth property of *P. anomala* 29X in a controlled environment, batch fermentations with working volume of 1 l were conducted in both MM with 60 g/l glucose and 30 g/l xylose, and hydrolysates, prepared with the CA method.

MM fermentations

Pichia anomala was described as having a Pasteur effect, which makes it oxygen sensitive instead of glucose sensitive, in terms of product formation [27,29]. Due to this, we initially conducted two parallel batch fermentations in MM, one anaerobically (100% N_2) and one, with 20% air (80% N_2). The carbon sources of these fermentations were 60 g/l glucose and 30 g/l xylose, which mimics the sugar composition of biomass hydrolysates [3]. The fermentations were monitored by measuring OD, glucose, xylose and ethanol concentrations of the samples taken with a fixed interval during the process, as shown in Figure 3 [26].

First of all, it can be seen that the strain is capable of growing both with and without oxygen supply, though the growth was much faster when oxygen was present (Figure 3a). The glucose consumption was completed within 20 hours when 20% air was mixed into the gas-in flow, while under strict anaerobic conditions, it took about 45 hours to consume all the glucose in MM (Figure 3b).

Secondly, a similar ethanol production pattern was observed in both fermentations, with and without oxygen supply. The amount of ethanol produced was also comparable. This challenges the claim that *P. anomala* as a species showing Pasteur effect [27,29], since no suppression on ethanol formation was observed when 20% air was added into the gas-in flow. However, as soon as glucose was depleted, ethanol was subsequently consumed when oxygen was present. It was suspected that besides producing biomass, ethanol was partially converted into ethyl acetate (EA), since small amount of EA was detected in the fermentor (~ 5 g/l). The production of EA by *P. anomala* was also reported by previous studies as a mechanism of forming anti-fungal compounds [34,35] or producing flavor component during wine making process [36,37]. EA was also produced as the main product through biotechnological process using *P. anomala* on a pilot scale, as EA is a widely used environmentally friendly solvent [38]. Therefore, besides ethanol, *P. anomala* can also be used potentially for EA production with biomass hydrolysate as substrate.

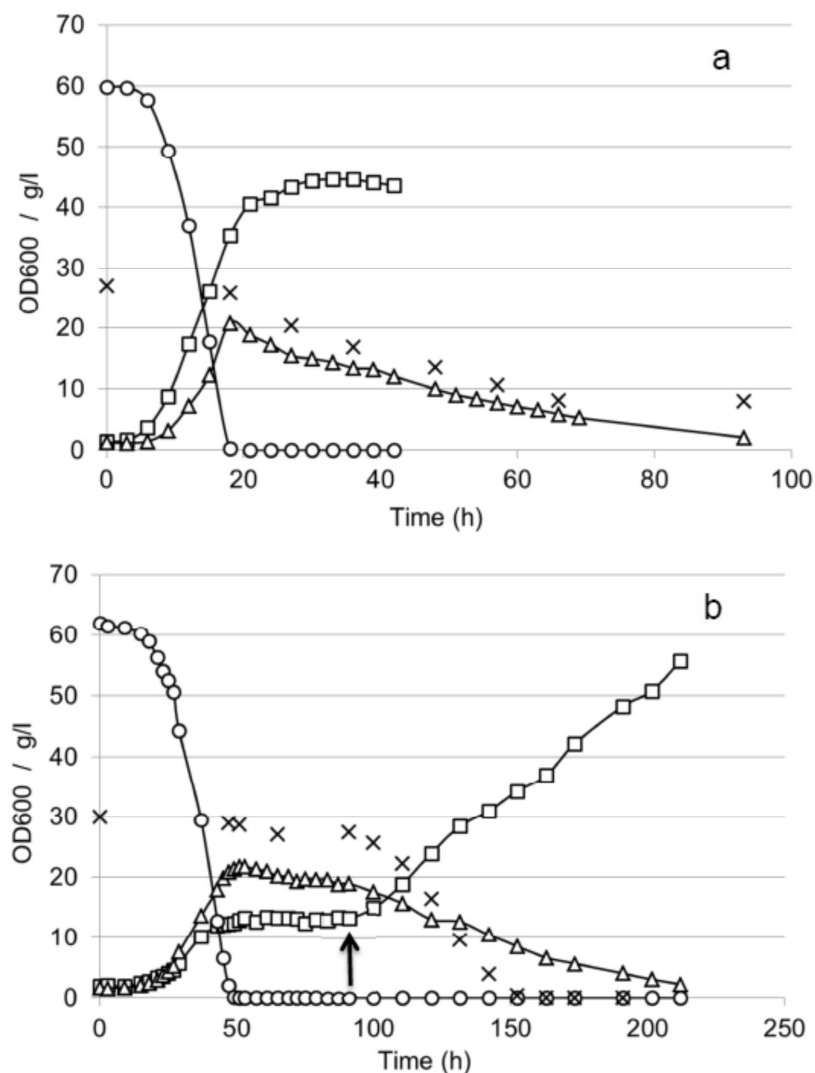


Figure 3 Fermentation results of *Pichia anomala* 29X (CBS132101) in MM (60 g/l glucose and 30 g/l xylose), square: OD, circle: glucose, cross: xylose, triangle: ethanol. (a) 20% air + 80% N₂, (b) 100% N₂ to 20% air + 80% N₂, the time-point when 20% air was added into the gas-in flow is indicated by an arrow.

Thirdly, the strain was capable of assimilating xylose after glucose was depleted (Figure 3a), which is a characteristic that was reported for some but not all *P. anomala* isolates earlier [39]. We confirmed the presence of xylose reductase, xylitol dehydrogenase and D-xylulokinase in *P. anomala* strain by detecting homologs of the corresponding genes of *P. stipitis* in its

genome. (xylose reductase: GenBank ADQ89193.1, xylitol dehydrogenase: GenBank AAD28251.1, D-xylulokinase: GenBank EAZ63302.2; *P. anomala* genome database: <http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=885923>). However, no xylose consumption was observed under anaerobic conditions, and only when air was mixed into the gas-in flow, a simultaneous xylose decrease and OD increase took place (Figure 3b). This indicates that the *P. anomala* strain is not xylose fermenting, probably due to the redox imbalance that caused by the overproduction of NADH in the xylose reduction step [40]. This feature is undesirable in the context of ethanol production from biomass hydrolysates.

Finally, as described above, the produced ethanol was consumed / converted after glucose was depleted in the fermentor, which took place simultaneously with the consumption of xylose, after air was added into the gas-in flow (Figure 3b). To prevent losing ethanol, the following strategies could be considered (1) keep the fermentation anaerobic, and engineer and/or adapt the strain to ferment xylose under anaerobic conditions; (2) stop the fermentation process as soon as glucose consumption is complete, by temperature drop for instance, then recover ethanol from the fermentation broth, and continue fermenting the remaining xylose by a pentose fermenting strain [41]; (3) *in-situ* ethanol recovery, that is to remove ethanol during the fermentation process continuously. It is highly likely that an evolutionary adaptation process is also needed to improve the ethanol productivity of the *P. anomala* strain [42].

Hydrolysate fermentations

From the fermentation results of *P. anomala* 29X in MM with 60 g/l glucose and 30 g/l xylose, it was clear that the strain produced ethanol using glucose both with and without oxygen, while only when oxygen was present, xylose and ethanol consumption took place (Figure 3). Based on these observations, a two-stage fermentation procedure was adopted to test the performance of *P. anomala* 29X in hydrolysates. The two-stage fermentation started with anaerobic conditions (100% N₂) to allow the production of ethanol with glucose present in the hydrolysates. As soon as glucose was depleted, indicated by a sudden CO₂ drop in the off-gas, 20% air was mixed into the gas-in to allow the utilization of xylose.

To test the resistance of *P. anomala* 29X to the inhibitory compounds in hydrolysates, CA pretreated hydrolysates were used as fermentation media. Since CA hydrolysates had higher inhibitory effects compared to the ones prepared with other methods, like mild alkaline and diluted acid [18], the two CA hydrolysates used here were made from WS and CS, as shown in Figure 4.

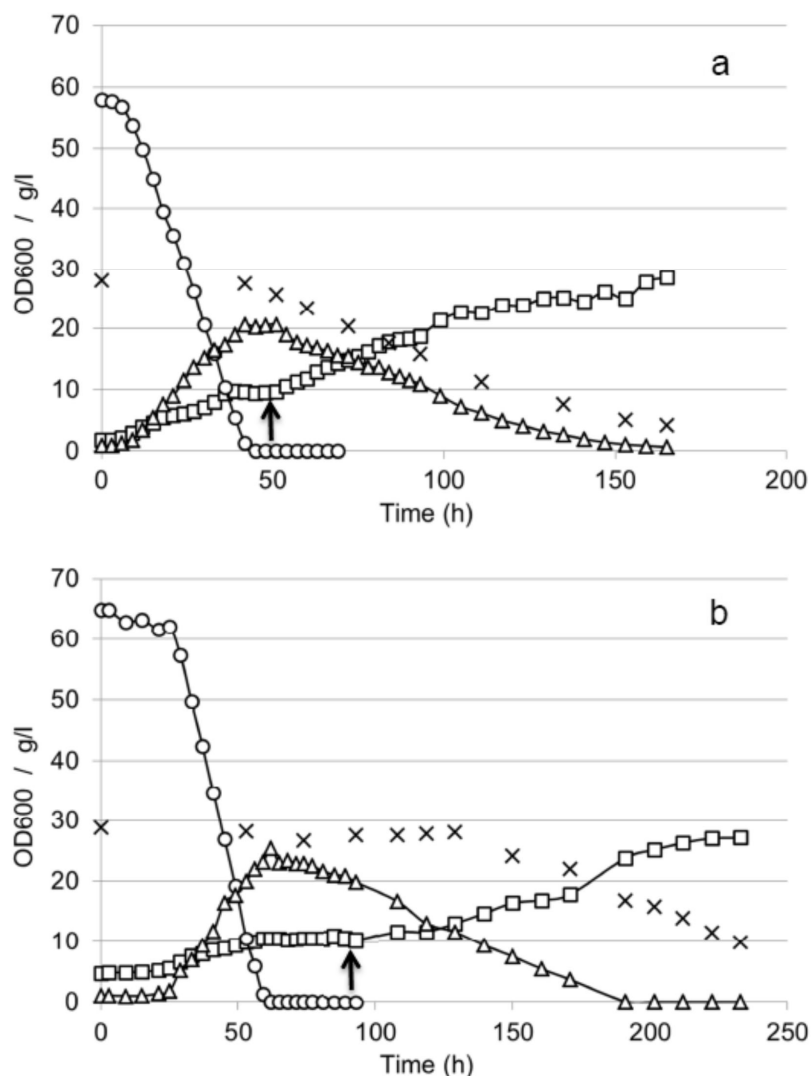


Figure 4 Fermentation results of *Pichia anomala* 29X in concentrated acid prepared hydrolysates, aeration conditions: 100% N₂ to 20% air + 80% N₂, the time-point when 20% air was added into the gas-in flow are indicated by arrows, square: OD, circle: glucose, cross: xylose, triangle: ethanol. (a) wheat straw hydrolysate (WS-CA), (b) corn stover hydrolysate (CS-CA).

It can be seen that very similar to its performance in MM (Figure 3b), *P. anomala* 29X strain rapidly consumed all the glucose in WS-CA hydrolysate under anaerobic conditions, no xylose or ethanol consumption was observed at this stage (Figure 4a). As soon as air was mixed into the gas-in flow, indicated by the black arrow in Figure 4a, growth took place again,

while ethanol as well as xylose started to decrease. The similar fermentation performance as in MM suggests that *P. anomala* 29X had considerably high resistance to the inhibitors in WS-CA hydrolysate, in terms of ethanol production from glucose.

As far as xylose utilization is concerned, the strain seems to be more sensitive to the inhibitors, since the xylose consumption was much slower compared to that in MM (Figure 4a and 3b). This was also observed with CS-CA hydrolysate, of which the xylose consumption was still incomplete after 233 hours (Figure 4b). The low inhibitor tolerance during the xylose fermentation process was reported in recombinant yeasts previously [11,43,44]. For instance, when less diluted hydrolysate was used as fermentation medium, xylose consumption rate decreased significantly, while glucose level was almost not influenced. Since this phenomenon was also observed with *P. anomala* 29X, a xylose utilizing isolate, it is suspected that the inhibitory mechanism may relate to the enzymes involved in the xylose consumption pathways.

Compared to WS-CA, CS-CA was a more toxic hydrolysate when tested with *S.cerevisiae* (see Chapter 5). In this study, the *P. anomala* 29X strain showed also less resistance to the CS-CA hydrolysate, indicated by a longer initial lag-phase and a slower glucose consumption rate (Figure 4b). Moreover, after air was mixed into the gas-in flow, no growth took place for a period of 27 hours, and only after 67 hours, xylose started to decrease (Figure 4b). This suggests that the resistance of *P. anomala* 29X is hydrolysate dependent, which was also shown in small scale growth experiments (Section Resistant strain selection in hydrolysates). Therefore, when hydrolysates with very high inhibitor concentration are used as fermentation substrates, dilution and/or detoxification still need to be applied to allow a reasonable performance of the yeast strain.

Resistance test

To further understand the resistance of *P. anomala* 29X strain to individual inhibitory compounds in biomass hydrolysates, an initial growth test was conducted using MM with 20 g/l glucose as reference medium, and *S.cerevisiae* CEN.PK 113-7D as reference strain. A series of potential inhibitory compounds were chosen, as shown in Table 2. They were either quantitatively detected in biomass hydrolysates [3], or were shown to be decreasing in hydrolysate fermentation with *S.cerevisiae* CEN.PK 113-7D (unpublished data). The concentrations used were relevant to that in actual hydrolysates [3] (Table 2).

Compared to *S.cerevisiae*, *P.anomala* 29X exhibited higher tolerance to furfural, benzoic acid, ferulic acid and salicylic acid (shaded with grey in Table 2), which are common inhibitors in biomass hydrolysates [5,45-47]; while for the other tested compounds, both strains showed similar resistance. This suggests again that *P.anomala* 29X is a potentially interesting strain for bio-product formation using lignocellulosic biomass hydrolysates as feedstock.

Table 2 Results of strain resistance test.

		<i>P. anomala</i> 29X			<i>S. cerevisiae</i> CEN.PK 113-7D		
In reference medium ^a		LP ^b 5.5 h	GR ^c 0.067	OD ^d 1.2	LP 6 h	GR 0.108	OD 1.2
Inhibitor	Concentration (g/l)	LP	GR	OD	LP	GR	OD
Acetic acid	4.0	8 h	< 60% ^e	--- ^f	9.5 h	< 60%	< 80%
Benzoic acid	0.2	10 h	< 40%	< 60%	9 h	< 20%	< 60%
Ferulic acid	0.2	---	< 80%	---	---	< 40%	---
Formic acid	0.6	15 h	< 60%	< 80%	10 h	< 60%	---
Furfural	0.6	8.5 h	---	---	10.5 h	---	---
Furoic acid	0.2	---	---	---	---	< 80%	---
Phenylacetic acid	0.2	---	< 80%	---	---	< 80%	---
Salicylic acid	0.2	---	< 60%	< 80%	---	< 40%	< 60%
Vanillin	0.2	---	< 80%	---	7.5 h	---	---
2,4-dihydroxycinnamic acid	0.2						
3,4-dihydroxycinnamic acid	0.2						
4-hydroxybenzaldehyde	0.2						
4-Hydroxycinnamic acid	0.2						
4-hydroxyphenylacetate	0.2			No Effect			
HMF	0.2						
Levulinic acid	1.5						
Vanillic acid	0.2						

^a MM with 20 g/l glucose; ^b LP: lag-phase: time needed to reach 2% of the maximum OD (h); ^c GR: growth rate: determined from the linear part of the OD curve (OD/h); ^d OD: final OD; ^e values in <% are relative growth rate and final OD compared to the corresponding host in reference medium; ^f '---': no effect compared to reference medium.

Conclusion

A strain, resistant to biomass hydrolysate inhibitors, was isolated from grass silage and identified as *Pichia anomala* (*Wickerhamomyces anomalus*). We have shown that the strain is able to produce ethanol in multiple biomass hydrolysates with different toxicity levels, is capable of utilizing xylose for growth when supplied with air, and can use nitrate as N-source. These characteristics makes *Pichia anomala* a potential ethanol production host using lignocellulosic biomass as feedstock. Further studies on the physiology of the yeast will help improve its fermentation performance in biomass hydrolysates.

Acknowledgement

This project was (co) financed by the Netherlands Metabolomics Centre (NMC) which is part of the Netherlands Genomics Initiative / Netherlands Organization for Scientific Research.

References

1. Hahn-Hägerdal B, Galbe M, Gorwa-Grauslund M,F, Lidén G, Zacchi G: Bio-ethanol -- the fuel of tomorrow from the residues of today. *Trends Biotechnol.* **2006**, *24*, 549-556.
2. Limayem A, Ricke SC: Lignocellulosic biomass for bioethanol production: Current perspectives, potential issues and future prospects. *Prog. Enegr. Combust.* **2012**, *38*, 449-467.
3. Zha Y, Muilwijk B, Coulier L, Punt PJ: Inhibitory compounds in lignocellulosic biomass hydrolysates during hydrolysate fermentation processes. *J. Bioprocess Biotechniq.* **2012**, *2*, 112 doi: 10.4172/2155-9821.1000112.
4. Bellido C, Bolado S, Coca M, Lucas S, González-Benito G, García-Cubero MT: Effect of inhibitors formed during wheat straw pretreatment on ethanol fermentation by *Pichia stipitis*. *Bioresource Technol.* **2011**, *102*, 10868-10874.
5. Taherzadeh MJ, Karimi K: Fermentation Inhibitors in Ethanol Processes and Different Strategies to Reduce Their Effects. In *Biofuels: Alternative Feedstocks and Conversion Processes*. Pandey A, Larroche C, Ricke SC, Dussap C, Gnansounou E Eds., Academic Press, USA, 2011, pp. 287-311.
6. Jørgensen H, Kristensen JB, Felby C: Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities. *Biofuel Bioprod. Bior.* **2007**, *1*, 119-134.
7. van Maris AJA, Abbott DA, Bellissimi E, van den Brink J, Kuyper M, Luttik MAH, Wisselink HW, Scheffers WA, van Dijken JP, Pronk JT: Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: current status. *Anton. Leeuw. Int. J. G.* **2006**, *90*, 391-418.
8. Hahn-Hägerdal B, Karhumaa K, Fonseca C, Spencer-Martins I, Gorwa-Grauslund MF: Towards industrial pentose-fermenting yeast strains. *Appl. Microbiol. Biot.* **2007**, *74*, 937-953.
9. Chu BCH, Lee H: Genetic improvement of *Saccharomyces cerevisiae* for xylose fermentation. *Biotechnol. Adv.* **2007**, *25*, 425-441.

10. Öhgren K, Bengtsson O, Gorwa-Grauslund MF, Galbe M, Hahn-Hägerdal B, Zacchi G: Simultaneous saccharification and co-fermentation of glucose and xylose in steam-pretreated corn stover at high fiber content with *Saccharomyces cerevisiae* TMB3400. *J. Biotechnol.* **2006**, 126, 488-498.
11. Erdei B, Frankó B, Galbe M, Zacchi G: Glucose and xylose co-fermentation of pretreated wheat straw using mutants of *S. cerevisiae* TMB3400. *J. Biotechnol.* **2013**, 164, 50-58.
12. Martín C, Marcet M, Almazán O, Jönsson LJ: Adaptation of a recombinant xylose-utilizing *Saccharomyces cerevisiae* strain to a sugarcane bagasse hydrolysate with high content of fermentation inhibitors. *Bioresource Technol.* **2007**, 98, 1767-1773.
13. Toivola A, Yarrow D, Van Den Bosch E: Alcoholic fermentation of D-xylose by yeasts. *Appl. Environ. Microbiol.* **1984**, 47, 1221-1223.
14. Hahn-Hägerdal B, Jeppsson H, Skoog K, Prior BA: Biochemistry and physiology of xylose fermentation by yeasts. *Enzyme Microb. Technol.* **1994**, 16, 933-943.
15. Cho DH, Shin SJ, Bae Y, Park C, Kim YH: Ethanol production from acid hydrolysates based on the construction and demolition wood waste using *Pichia stipitis*. *Bioresource Technol.* **2011**, 102, 4439-4443.
16. Huang CF, Lin TH, Guo GL, Hwang WS: Enhanced ethanol production by fermentation of rice straw hydrolysate without detoxification using a newly adapted strain of *Pichia stipitis*. *Bioresource Technol.* **2009**, 100, 3914-3920.
17. van Hoek P, van Dijken JP, Pronk JT: Effect of specific growth rate on fermentative capacity of baker's yeast. *Appl. Environ. Microbiol.* **1998**, 64, 4226-4233.
18. Zha Y, Slomp R, Groenestijn J, Punt PJ: Preparation and Evaluation of Lignocellulosic Biomass Hydrolysates for Growth by Ethanologenic Yeasts. In *Microbial Metabolic Engineering: Methods and Protocols*. Qiong C Ed., Humana Press, New York, USA, 2012, pp. 245-259.
19. Tofighi A, Azin M, Assadi MM, Assadi-rad MHA, Nejadshattari T, Fallahian MR: Inhibitory effect of high concentrations of furfural on industrial strain of *Saccharomyces cerevisiae*. *Int. J. Environ. Res.* **2010**, 4, 137-142.
20. Wierckx N, Koopman F, Bandounas L, de Winde JH, Ruijsenaars HJ: Isolation and characterization of *Cupriavidus basilensis* HMF14 for biological removal of inhibitors from lignocellulosic hydrolysate. *Microb. Biotechnol.* **2010**, 3, 336-343.
21. Huang H, Guo X, Li D, Liu M, Wu J, Ren H: Identification of crucial yeast inhibitors in bio-ethanol and improvement of fermentation at high pH and high total solids. *Bioresource Technol.* **2011**, 102, 7486-7493.
22. Walker GM: *Pichia anomala*: Cell physiology and biotechnology relative to other yeasts. *Anton. Leeuw. Int. J. G.* **2011**, 99, 25-34.
23. Pérez-Mendoza F, Ruiz-Terán F, Abarca BE, Navarro-Ocaña A, Aguilar-Uscanga G, Valerio-Alfaro G: Oxoester oxidoreductase activities in new isolates of *Pichia anomala* from apple, grape and cane juices. *FEMS Yeast Res.* **2005**, 5, 685-690.
24. Passoth V, Eriksson A, Sandgren M, Ståhlberg J, Piens K, Schnürer J: Airtight storage of moist wheat grain improves bioethanol yields. *Biotechnol. Biofuels.* **2009**, 2, 16 doi: 10.1186/1754-6834-2-16.
25. Haïssam JM: *Pichia anomala* in biocontrol for apples: 20 years of fundamental research and practical applications. *Anton. Leeuw. Int. J. G.* **2011**, 99, 93-105.
26. Fredlund E, Druvefors U, Boysen ME, Lingsten KJ, Schnürer J: Physiological characteristics of the biocontrol yeast *Pichia anomala* J121. *FEMS Yeast Res.* **2002**, 2, 395-402.
27. Fredlund E, Broberg A, Boysen ME, Kenne L, Schnürer J: Metabolite profiles of the biocontrol yeast *Pichia anomala* J121 grown under oxygen limitation. *Appl. Microbiol. Biot.* **2004**, 64, 403-409.
28. Tao N, Gao Y, Liu Y: Isolation and characterization of a *pichia anomala* strain: A promising candidate for bioethanol production. *Brazilian J. Microbiol.* **2011**, 42, 668-675.

29. Fredlund E, Blank LM, Schnürer J, Sauer U, Passoth V: Oxygen- and glucose-dependent regulation of central carbon metabolism in *Pichia anomala*. *Appl. Environ. Microb.* **2004**, 70, 5905-5911.
30. Sedlak M, Ho NWY: Production of ethanol from cellulosic biomass hydrolysates using genetically engineered *Saccharomyces* yeast capable of cofermenting glucose and xylose. *Appl. Biochem. Biotech.* **2004**, 114, 403-416.
31. Katahira S, Mizuike A, Fukuda H, Kondo A: Ethanol fermentation from lignocellulosic hydrolysate by a recombinant xylose- and cellobiosaccharide-assimilating yeast strain. *Appl. Microbiol. Biot.* **2006**, 72, 1136-1143.
32. Passoth V, Fredlund E, Druvefors UÅ, Schnürer J: Biotechnology, physiology and genetics of the yeast *Pichia anomala*. *FEMS Yeast Res.* **2006**, 6, 3-13.
33. de Barros Pita W, Leite FCB, de Souza Liberal AT, Simões DA, de Morais Jr. MA: The ability to use nitrate confers advantage to *Dekkera bruxellensis* over *S. cerevisiae* and can explain its adaptation to industrial fermentation processes. *Anton. Leeuw. Int. J. G.* **2011**, 100, 99-107.
34. Fredlund E, Druvefors UÅ, Olstorp MN, Passoth V, Schnürer J: Influence of ethyl acetate production and ploidy on the anti-mould activity of *Pichia anomala*. *FEMS Microbiol. Lett.* **2004**, 238, 133-137.
35. Druvefors UA, Schnürer J: Mold-inhibitory activity of different yeast species during airtight storage of wheat grain. *FEMS Yeast Res.* **2005**, 5, 373-378.
36. Yilmaztekin M, Cabaroğlu T, Erten H: The effect of temperature and aeration on ethyl acetate production by *Pichia* yeasts. *Curr. Opin. Biotechnol.* **2011**, 22, S91-S92.
37. Rojas V, Gil JV, Piñaga F, Manzanares P: Studies on acetate ester production by non-*Saccharomyces* wine yeasts. *Int. J. Food. Microbiol.* **2001**, 70, 283-289.
38. Löser C, Urit T, Stukert A, Bley T: Formation of ethyl acetate from whey by *Kluyveromyces marxianus* on a pilot scale. *J. Biotechnol.* **2013**, 163, 17-23.
39. Kurtzman CP: *Pichia E.C. Hansen emend. Kurtzman*. In *The Yeasts, a Taxonomic Study*. Kurtzman CP, Fell JW Eds., Elsevier Science BV, Amsterdam, The Netherlands, 1998, pp. 352.
40. Bruinenberg P, Bot PM, Dijken J, Scheffers WA: NADH-linked aldose reductase: the key to anaerobic alcoholic fermentation of xylose by yeasts. *Appl. Microbiol. Biotechnol.* **1984**, 19, 256-260.
41. Li Y, Park JY, Shiroma R, Tokuyasu K: Bioethanol production from rice straw by a sequential use of *Saccharomyces cerevisiae* and *Pichia stipitis* with heat inactivation of *Saccharomyces cerevisiae* cells prior to xylose fermentation. *J. Biosci. Bioeng.* **2011**, 111, 682-686.
42. Matsushika A, Oguri E, Sawayama S: Evolutionary adaptation of recombinant shochu yeast for improved xylose utilization. *J. Biosci. Bioeng.* **2010**, 110, 102-105.
43. Helle S, Cameron D, Lam J, White B, Duff S: Effect of inhibitory compounds found in biomass hydrolysates on growth and xylose fermentation by a genetically engineered strain of *S. cerevisiae*. *Enzyme Microb. Tech.* **2003**, 33, 786-792.
44. Björling T, Lindman B: Evaluation of xylose-fermenting yeasts for ethanol production from spent sulfite liquor. *Enzyme Microb. Tech.* **1989**, 11, 240-246.
45. Heer D, Sauer U: Identification of furfural as a key toxin in lignocellulosic hydrolysates and evolution of a tolerant yeast strain. *Microb. Biotechnol.* **2008**, 1, 497-506.
46. Klink HB, Thomsen AB, Ahring BK: Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass. *Appl. Microbiol. Biot.* **2004**, 66, 10-26.
47. Larsson S, Quintana-Sáinz A, Reimann A, Nilvebrant NO, Jönsson LJ: Influence of lignocellulose-derived aromatic compounds on oxygen-limited growth and ethanolic fermentation by *Saccharomyces cerevisiae*. *Appl. Biochem. Biotech.* **2000**, 84-86, 617-632.

SUMMARY

Lignocellulosic biomass is the 2nd generation feedstock for biofuel production through fermentation processes. The resources are mostly agricultural and industrial wastes, which are inexpensive, environmental friendly and not competitive with world food-supply. Lignocellulosic biomass has a rigid structure, which is mainly composed of cellulose, hemicellulose and lignin. To break down the structure and expose cellulose and hemicellulose for hydrolysis, a pretreatment procedure needs to be applied. There are various ways to pretreat biomass, which differ in their severity and efficiency. A mild pretreatment method results in incomplete releasing of cellulose, while a harsh method may lead to the releasing and forming of toxic compounds through sugar and lignin degradation. These toxic compounds inhibit the growth of the fermenting micro-organism(s), which results in reduced productivity.

To improve the fermentability of the hydrolysis products of pretreated lignocellulosic biomass, the so called biomass hydrolysates, identifying their inhibitory compounds is of great importance. **Chapter 1** of this thesis reviews the approaches and techniques that have been used to study the inhibitors in various biomass hydrolysates, and introduces a non-targeted methodology to systematically identify biomass hydrolysate inhibitors: the exometabolomics approach. The four steps involved in this approach are: (i) defining research question, (ii) experimental design, (iii) sample selection and analysis and (iv) data analysis and interpretation.

To identify hydrolysate inhibitors through an exometabolomics approach, a wide range of biomass hydrolysates need to be prepared. Four pretreatment methods were developed and applied to six different biomass types, respectively. The fermentability of the resulting biomass hydrolysates were tested with the fermenting yeast, *Saccharomyces cerevisiae*. The detailed procedures of these four pretreatment methods are described in **Chapter 2**, which also reports the difference in fermentability among the generated biomass hydrolysates.

The hydrolysis efficiency of the pretreated biomass is further discussed in **Chapter 3**. To study hydrolysis beyond the formation of monosaccharides and understand the main oligosaccharide forms in biomass hydrolysates, high-performance anion-exchange chromatography coupled with mass spectrometry (HPAEC-MS) was applied. The method is able to separate and detect many oligosaccharides in one experimental run, and by using estimated response factors, the relative quantities of the detected oligosaccharides were

assessed. The analysis results revealed that, besides the monomers, disaccharides were the main remaining sugar form in all biomass hydrolysates generated by the pretreatment methods described in Chapter 2.

Chapters 4 to 6 are focused on identifying inhibitory compounds in lignocellulosic biomass hydrolysates and studying their effects on fermenting yeasts during fermentation processes. During biomass pretreatment processes hydrolysate inhibitors are released or formed, amongst which are mainly weak acids, furans and phenolic compounds (examples are acetic acid, furfural and vanillin, respectively). The presence and concentration of the inhibitors are biomass type and pretreatment method dependent, and inhibition effects include longer lag-phase, lower growth rate, and reduced productivity.

Chapter 4 reports the examination of the presence and dynamics of a target group of inhibitors in several different biomass hydrolysates. The fermentability of the hydrolysates were tested by conducting batch fermentations using baker's yeast, during which time-series samples were taken for their non-sugar composition analysis. The actual concentrations of the pre-selected inhibitors were determined, which are valuable references for toxicity tests of relevant potential inhibitory compounds.

Chapter 5 reports the detailed experimental procedure and results, in terms of the exometabolomics approach introduced in Chapter 1. The research question, identifying inhibitors in biomass hydrolysates, was answered by statistically correlating the fermentability of 16 different biomass hydrolysates with their non-sugar compositions. The composition analysis was conducted with two GC-MS methods to reach a high compound coverage, and the data analysis was realized by building statistical models. A list of compounds was identified as inhibitors in biomass hydrolysates, including novel compounds. The results suggest that metabolomics is a relevant approach in target identification in complex systems like lignocellulosic biomass hydrolysates.

Finally, in **Chapter 6**, we describe the isolation of microorganisms, from different environmental sources, that exhibit high resistance to biomass hydrolysate inhibitors. A unique strain was isolated and identified as *Pichia anomala*. This yeast showed relative high resistance in the tested hydrolysates compared to some other *P. anomala* strains, and was capable of utilizing xylose as C-source and nitrate as N-source. Through further research and possibly genetic modifications, this strain has the potential to become a suitable yeast for fermenting lignocellulosic biomass hydrolysates.

SAMENVATTING

Lignocellulolytische biomassa is dé 2e generatie feedstock voor bio-ethanol productie met behulp van fermentatie processen. Biomassa-bronnen worden met name gevonden onder de afvalproducten van landbouw en/of industrie. Deze zijn goedkoop, milieu-vriendelijk en niet concurrerend met het gebruik als menselijk of dierlijk voedsel. De structuur van lignocellulose is zeer rigide en bestaat uit de componenten cellulose, hemicellulose en lignine. Om de starre structuur open te breken, zodat de cellulose en hemicellulose toegankelijk worden voor hydrolyse, is het nodig een voorbehandelingsmethode toe te passen. Hier bestaan diverse varianten van die verschillend zijn qua openbrekend vermogen en qua efficiëntie. Een milde voorbehandelingsmethode resulteert in het niet volledig toegankelijk worden van cellulose, terwijl een agressieve methode kan leiden tot het ontstaan van ongewenste componenten door omzetting van suikers en lignine in giftige afbraakproducten. Deze giftige componenten remmen de groei van het micro-organisme dat het fermentatieproces uitvoert, wat resulteert in een verlaagde productiviteit.

Om de fermenteerbaarheid te verbeteren van het lignocellulose hydrolyseproduct (met andere woorden het biomassa hydrolysaat), is het belangrijk om te identificeren welke componenten remmend zijn. **Hoofdstuk 1** van dit proefschrift beschrijft de in deze studie gebruikte aanpak en introduceert een breed-toepasbare methodiek waarmee systematisch remmende componenten in biomassa hydrolysaten geïdentificeerd kunnen worden: de 'exometabolomics aanpak'. Deze exometabolomics aanpak bestaat uit vier stappen, te weten: (i) definiëren van de onderzoeksvraag, (ii) ontwerpen van de experimentele opzet, (iii) selectie en analyse van monsters, en (iv) data analyse en interpretatie.

Er is een brede verscheidenheid aan biomassa hydrolysaten nodig om met behulp van de exometabolomics aanpak de remmende componenten in die hydrolysaten te kunnen identificeren. Vier verschillende voorbehandelingsmethoden werden uitgevoerd en toegepast op zes verschillende soorten biomassa. De fermenteerbaarheid van de resulterende vierentwintig verschillende biomassa hydrolysaten werd getest met de ethanol-producerende gist *Saccharomyces cerevisiae*. In **Hoofdstuk 2** worden de gedetailleerde protocollen van de vier voorbehandelingsmethoden beschreven en ook de verschillen in fermenteerbaarheid van de met deze methoden geproduceerde biomassa hydrolysaten.

Het onderwerp van discussie in **Hoofdstuk 3** is de hydrolyse efficiëntie van de voorbehandelde biomassa. Om het hydrolyseproces breder te bestuderen dan alleen maar

Samenvatting

te kijken naar het vrijmaken van monosacchariden, en ook om begrip te krijgen van welke oligosacchariden voorkomen in biomassa hydrolysaten, is een specifieke meetmethode toegepast genaamd high-performance anion-exchange chromatography + mass spectrometry (HPAEC-MS). Deze methode is in staat vele oligosacchariden binnen één meting te scheiden en individueel te detecteren. Bovendien kan door het schatten van de response factor per oligosaccharide een overzicht gemaakt worden van de relatieve hoeveelheden van de gedetecteerde oligosacchariden. Het resultaat van deze analyse was dat naast monomeren de resterende suikers in hoofdzaak als disacchariden aanwezig waren in alle biomassa hydrolysaten zoals gemaakt met behulp van de voorbehandelingsmethoden beschreven in Hoofdstuk 2.

Hoofdstuk 4 tot en met 6 zijn gericht op het identificeren van remmende componenten in lignocellulolytische biomassa hydrolysaten en het bestuderen van hun effecten op de ethanol-producerende gist gedurende het fermentatieproces. Dit soort remmende componenten worden vrijgemaakt of gevormd gedurende het biomassa voorbehandelingsproces en zijn voornamelijk onder te verdelen in de groepen zwakke zuren, furanen en fenolische stoffen (voorbeelden respectievelijk azijnzuur, furfural en vaniline). De aanwezigheid en concentratie van de remmende componenten worden bepaald door het type biomassa en door de toegepaste voorbehandelingsmethode. Het remmingseffect van deze componenten kan zich uiten in onder andere een langere lag-fase, een lagere groeisnelheid en een gereduceerde productiviteit.

Hoofdstuk 4 beschrijft onderzoek naar de aanwezigheid en dynamiek van een geselecteerde groep remmers in een aantal verschillende biomassa hydrolysaten. De fermenteerbaarheid van de hydrolysaten werd getest door batch-fermentaties met bakkersgist uit te voeren en gedurende dit hele proces tijdsmonsters te nemen waarvan vervolgens de samenstelling werd bepaald (met uitsluiting van suikerverbindingen). Van een geselecteerde groep remmers werden de daadwerkelijke concentraties bepaald. De gevonden waarden zijn daarna gebruikt voor het opzetten van een giftigheidstest van mogelijk remmende componenten.

Hoofdstuk 5 beschrijft de gedetailleerde experimentele procedure en resultaten van de exometabolomics aanpak zoals geïntroduceerd in Hoofdstuk 1. De onderzoeksvraag, in dit geval de identificatie van remmers in biomassa hydrolysaten, werd beantwoord door met behulp van biostatistische methoden correlaties te identificeren tussen de fermenteerbaarheid van 16 verschillende biomassa hydrolysaten en hun chemische samenstelling. Om een zo breed mogelijk spectrum van componenten te kunnen detecteren

werd de samenstelling bepaald met behulp van twee verschillende GC-MS methoden. Analyse van de meetresultaten werd uitgevoerd door deze in statistische modellen te beschrijven. Het eindresultaat was een lijst van componenten met remmende eigenschappen in biomassa hydrolysaten, waaronder ook niet eerder hiervoor beschreven componenten. Dit resultaat geeft aan dat metabolomics een relevante aanpak kan zijn bij het vinden van specifieke componenten in complexe systemen zoals lignocellulolytische biomassa hydrolysaten

Ter afsluiting, in **Hoofdstuk 6** wordt beschreven hoe we uit verschillende milieus micro-organismen hebben geïsoleerd die restitentie vertonen tegen de remmende effecten van biomassa hydrolysaat componenten. Een micro-organisme met unieke kwaliteiten op dit gebied werd geïdentificeerd als *Pichia anomala*. Deze specifieke gist stam vertoonde in de geteste hydrolysaten een hoge resistentie tegen groeiremming in vergelijking met andere *P. anomala* stammen, en was ook in staat xylose als koolstofbron en nitraat als stikstofbron te gebruiken. Als deze stam, door verder onderzoek en genetische modificatie, nog verder geoptimaliseerd kan worden in deze eigenschappen, dan is deze stam in potentie een goede kandidaat voor commerciële vergisting van lignocellulolytische biomassa hydrolysaten.

LIST OF PUBLICATIONS

Zha Y, Slomp R, Groenestijn J, Punt PJ: Preparation and Evaluation of Lignocellulosic Biomass Hydrolysates for Growth by Ethanologenic Yeasts. In *Microbial Metabolic Engineering: Methods and Protocols*. Qiong C Ed., Humana Press: New York, USA, 2012, pp. 245-259.

Zha Y, Muilwijk B, Coulier L, Punt PJ: Inhibitory compounds in lignocellulosic biomass hydrolysates during hydrolysate fermentation processes. *J. Bioprocess Biotechnol.* **2012**, 2:112, doi: 10.4172/2155-9821.1000112.

Zha Y, Punt PJ: Exometabolomics approaches in studying the application of lignocellulosic biomass as fermentation feedstock. *Metabolites* **2013**, 3, 119-143.

Coulier L, Zha Y, Bas R, Punt PJ: Analysis of oligosaccharides in lignocellulosic biomass hydrolysates by high-performance anion-exchange chromatography coupled with mass spectrometry (HPAEC–MS). *Bioresource Technol.* **2013**, 133, 221-231.

Gaykawad SS, Zha Y, Punt PJ, van Groenestijn JW, van der Wielen LAM, Straathof AJJ: Pervaporation of ethanol from lignocellulosic fermentation broth. *Bioresource Technol.* **2013**, 129, 469-476.

Zha Y, Hossain AH, Tobola F, Sedee N, Havekes M, Punt PJ: *Pichia anomala* 29X: A resistant strain for lignocellulosic biomass hydrolysate fermentation. *FEMS Yeast Res.* **2013**, doi: 10.1111/1567-1364.12062.

Zha Y, Westerhuis JA, Muilwijk B, Overkamp KM, Nijmeijer BM, Coulier L, Amilde AK, Punt PJ: Identifying inhibitory compounds in lignocellulosic biomass hydrolysates using an exometabolomics approach. Submitted to *BMC Biotechnol.*

CURRICULUM VITAE

Ying Zha was born on the 30th December 1983 in Yinchuan, the capital city of Ningxia province, China, where she attended Tanglai high school in the period of 1996 to 2002. In 2002, she entered Zhejiang University in Hangzhou, the capital city of Zhejiang province, and majored in Environmental Engineering. During her undergraduate education, she studied as an exchange student in The Hong Kong University of Science & Technology for a year, majored in Chemical Engineering. She obtained her bachelor degree in 2006 and moved to the Netherlands to follow the master program of Life Science & Technology in TUDelft. During her master education, she conducted an internship at DSM Food Specialties, studied the influence of media composition on morphology development of *Aspergillus niger*. Her master thesis was carried out in the Industrial Microbiology group of TUDelft, on the topic of the regulation of the *TIR1* gene under anaerobic conditions in *Saccharomyces cerevisiae*, and the involvement of the Rim101 pathway.

After obtaining her master diploma in July 2008, she started to work as a PhD candidate in TNO Quality of Life (the current department of Microbiology & Systems Biology), in Zeist, under the supervision of Professor Peter Punt. Her research project, which is within the framework of Netherlands Metabolomics Centre (NMC), is focused on identifying inhibitors in lignocellulosic biomass hydrolysates using an exometabolomics approach, as illustrated in detail in this thesis. Starting from August 2013, she is employed as assistant scientist at DSM Biotechnology Center, in Delft.

