Title

Quantifying aspects of lignin-cellulase interactions

Author

Pfeiffer, Katherine Anne

Publication Date

2014

Peer reviewed|Thesis/dissertation

Quantifying aspects of lignin-cellulase interactions

by

Katherine Anne Pfeiffer

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemical Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Douglas S. Clark, Chair
Professor Harvey W. Blanch
Associate Professor Michelle C. Chang

Fall 2014

Quantifying aspects of lignin-cellulase interactions

Abstract

Quantifying aspects of lignin-cellulase interactions

by

Katherine Anne Pfeiffer

Doctor of Philosophy in Chemical Engineering

University of California, Berkeley

Professor Douglas S. Clark, Chair

The transition from the production of fuels and chemicals from fossil sources to cellulosic feedstocks is an important step towards a more carbon-neutral economy. Cellulose is the most abundant polymer on the planet. Its constituent $\beta-1,4$ linked glucose units, when deconstructed to glucose monomers, can be fermented into ethanol, butanol, and other next-generation renewable products. Biomass costs less on an energy basis than oil, and some cellulosic crops require far fewer water and nitrogen inputs than corn-based bioethanol production. However, technical challenges remain, and cellulosic sugar is not yet a major platform for fuel and chemical production. The high cost of cellulolytic enzyme production, the large quantities of enzyme needed to complete the transformation of cellulose to glucose, and the presence of lignin in the heterogeneous biomass are a few major barriers to the large-scale production of cellulose-derived sugar.

Lignin inhibition via nonproductive binding of cellulase enzymes to the lignin surface decreases the overall efficiency of cellulose hydrolysis, and precludes recycling of lignin-bound enzymes. Although nonproductive binding of cellulases to lignin has been observed for many years, efforts to characterize the kinetics and structural drivers of this interaction have been lacking. In this work, lignin was isolated and characterized from acid-pretreated *Miscanthus*, in an attempt to isolate a lignin that is chemically similar to that present in an industrially relevant feedstock. Lignin films were spin-cast and characterized, in order to produce a flat, homogeneous substrate for surface binding assays. Cellulase enzymes were purified from native or heterologous expression hosts, and quartz crystal microgravimetry with dissipation monitoring (QCM-D) was used to measure the adsorption, desorption, and irreversible adsorption rates to the lignin surface. Using these tools, the several different reaction mechanisms of lignin-cellulase binding were evaluated. Although most existing lignin-binding literature uses Langmuir isotherms to quantify lignin-cellulase interactions, the data show that the mechanism that best fits the measured data is a transition model with multiple binding sites. Methods for comparing kinetic mechanisms and calculating kinetic parameters are presented.

The cellulose-degrading system of *Hypocrea jecorina* is well known as an industrial standard for enzymatic biomass hydrolysis. In this work, the lignin-binding kinetics of the four most important *H. jecorina* cellulases have been characterized. Existing literature shows that the carbohydrate binding domain is important in lignin binding. In order to quantify this effect, full-length enzymes and their isolated catalytic domains have been tested for their lignin-binding kinetics. While CBMs are responsible for fast binding of cellulases to lignin, our work implicates the catalytic domain in the irreversible aspect of lignin-binding kinetics. Furthermore, several-fold differences in adsorption rates between homologous CBMs are noted, as well as nearly 10-fold differences in adsorption rates between two homologous catalytic domains. Using this technique, we have uncovered new targets for evolution of cellulases. Finally, future directions in the study of lignin-cellulase interactions are discussed.

The final chapter in this work is an attempt to integrate lignin oxidation via oxidative enzymes into an ionic liquid pretreatment process. Certain ionic liquids are efficient solvents for whole biomass, and furthermore are able to cleave non phenolic lignin model dimer compounds, provided the dimers have been oxidized at the α-hydroxyl position. An effort to extend these model compound results to lignin oxidation and breakdown is reported.

Lignin is both an energy-dense biomass component and potential source of aromatic compounds, and a physical and competitive inhibitor of cellulase enzymes. This thesis presents efforts to decrease the cost of sugar production, both through the study of nonproductive cellulase-interactions, and through attempts to modify lignin structure for the extraction of value-added products.

"Without willing it, I had gone from being ignorant of being ignorant to being aware of being aware. And the worst part of my awareness was that I didn't know what I was aware of. I knew I knew very little, but I was certain that the things I had yet to learn wouldn't be taught to me at George Washington High School."

— Maya Angelou, *I know why the caged bird sings*

Contents

Contents **ii**

List of Figures **iv**

List of Tables **vi**

1 Lignin: a barrier to economical biomass hydrolysis **1**
1.1 Introduction 1
1.2 Lignin and the biorefinery: effect of lignin on biomass deconstruction 4
1.3 Cellulase structure, function, and diversity: effect on lignin interactions . . . 11
1.4 Direct measurement of cellulase-lignin interaction kinetics 14

Bibliography **15**

2 Development of a method to measure lignin-cellulase interaction kinetics **21**
2.1 Abstract 21
2.2 Introduction 21
2.3 Materials and methods 27
2.4 Results 32
2.5 Discussion 45
2.6 Conclusions 47

Bibliography **48**

3 Lignin-binding kinetics of four major *H. jecorina* cellulases and their catalytic domains **53**
3.1 Abstract 53
3.2 Introduction 54
3.3 Materials and methods 58
3.4 Results 60
3.5 Discussion 71
3.6 Future work: new directions in the mitigation nonproductive binding of cellulases to lignin 74

Bibliography **77**

4 Lignin decomposition: laccase-mediated α-oxidation and base-catalyzed cleavage **82**
4.1 Abstract 82
4.2 Introduction 83
4.3 Methods 86
4.4 Results 88
4.5 Discussion 96
4.6 Future work 97

Bibliography **99**

A Primers and sequences for heterologously expressed *H. jecorina* cellulases and truncated cellulases **102**

List of Figures

1.1 Lignin monomer structure and nomenclature 2
1.2 Model lignin and lignin-carbohydrate complex structure 3
1.3 Examples of type A, B, and C carbohydrate binding modules, adapted from [40]. 12

2.1 Three models of protein-surface interaction mechanism 25
2.2 2D HSQC NMR spectra of pretreated *Miscanthus* and isolated MWL 34
2.3 Milled wood lignin molecular weight profile measurement by gel permeation chromatography 35
2.4 Thickness (nm) of spin-cast lignin films 36
2.5 AFM images: height and phase of bare gold QCM-D sensor crystals 36
2.6 AFM images: height and phase of lignin coated sensors, dioxane solvent 37
2.7 AFM images: height and phase of lignin coated sensors, dioxane solvent, PDADMAC anchor 37
2.8 AFM images: height and phase of lignin coated sensors, THF solvent 37
2.9 AFM images: height and phase of lignin coated sensors, THF solvent, PDADMAC anchor layer 38
2.10 Empirical evaluation of convection limitation in measurement of adsorption rate 39
2.11 Empirical evaluation of diffusion limitation in measurement of adsorption rate . 40
2.12 Adsorption to quaternary amine SAM to directly measure bound Cel7B 41
2.13 Calculation of kinetic parameters for transition model of Cel7B on lignin 43
2.14 Evaluation of the two-step, changing footprint model 44
2.15 Evaluation of two-site transition model 45

3.1 Structural alignment of the four major *H. jecorina* binding domains 54
3.2 Amino acid alignment of the four major *H. jecorina* cellulase binding modules . 55
3.3 Amino acid alignment of the four major *H. jecorina* cellulase linkers 55
3.4 3D structures of the four main *Hypocrea jecorina* catalytic domains 57
3.5 $k_A\Gamma_{max}$, full-length cellulases 61
3.6 $k_A\Gamma_{max}$, cellulase catalytic domains 61
3.7 Maximum surface concentration of the four main *H. jecorina* cellulases. 62
3.8 Comparison of high and low temperature adsorption of 10 uM Cel7A to lignin surface 63

3.9 Lumped parameter $k_i + k_D$ plotted against experimental variables: Cel7A 65
3.10 Lumped parameter $k_i + k_D$ plotted against experimental variables: Cel6A 66
3.11 Lumped parameter $k_i + k_D$ plotted against experimental variables: Cel7B 66
3.12 Lumped parameter $k_i + k_D$ plotted against experimental variables: Cel5A 66
3.13 Summary of reversibility data and model predictions for the four full-length cellulases examined in this work 67
3.14 Washoff histories of cellulase catalytic domains, compared with the model results for the corresponding full-length enzymes 71
3.15 Graphical comparison of k_{A1} among the four full-length enzymes and their binding domains 72
3.16 Alignment of the catalytic domain crystal structure of *H. jecorina* Cel7A and Cel7B. 73

4.1 A proposed process for the oxidation and extraction of lignin as part of an ionic liquid pretreatment process. 84
4.2 Schematic for reactions between lignin, laccase, and a radical mediator. 85
4.3 Chemical structures for model lignin substrates. 86
4.4 Chemical structures for radical mediators. 87
4.5 Stability of lignin model dimers in various solvents 89
4.6 Reaction products of the deposition of non phenolic model dimer 89
4.7 Extraction of lignin-like model aromatic monomers 90
4.8 Laccase-TEMPO oxidation of four non phenolic lignin model substrates. 91
4.9 Laccase-mediator oxidation of veratryl alcohol in buffer and 20% IL 92
4.10 The effect of increasing mediator concentrations on the oxidation of non phenolic lignin dimer model. 92
4.11 Comparison of GPC molecular weight profiles of extracted lignin treated with laccase-mediator system, mediator alone, and lignin alone. 93
4.12 Oxidized mediator artifacts in GPC traces of LMS oxidation of organosolv lignin 94
4.13 GC/MS of the extract of the lignin/LMS reactions and their controls 95

List of Tables

1.1 Lignin monomer naming conventions 2
1.2 Summary of common pretreatments 6
1.3 Pretreatment advantages and disadvantages 7
1.4 Summary of lignin isolation methods 10
1.5 Organism and family names of CBMs displayed in 1.3 12

2.1 Milled wood lignin compositional analysis 32
2.2 Empirical evaluation of convection limitation in measurement of adsorption rate 39
2.3 QCM and activity-based mass comparison of bound Cel7B 41
2.4 Kinetics summary for 2-site transition model 46

3.1 Summary of Γ_{max} for each cellulase. 62
3.2 Cel7A kinetics summary 68
3.3 Cel6A kinetics summary 69
3.4 Cel7B kinetics summary 69
3.5 Cel5A kinetics summary 70
3.6 Kinetic parameters of *H. jecorina* catalytic domains 71
3.7 Optimum pH and relative activities at pH 6 and 7 of the four *H. jecorina* cellulases and *H. jecorina* β-glucosidase 75

A.1 Primers used for construction of pCDNA plasmid for heterologous expression in *H. jecorina* 103
A.2 Cel5A and Cel7B gene sequences 104

Acknowledgments

I'd like to acknowledge the efforts of the many who contributed to my graduate education. First, my advisors, Prof. Doug Clark and Prof. Harvey Blanch, have provided freedom, support, feedback, humor, and advice over the years, and this experience would have been immeasurably different without their presence.

I thank Prof. Clay Radke for offering generous access to his lab space and QCM-D equipment, and Dr. Elizabeth Schneider, who has given me much technical advice and encouragement for QCM-D experiments. Thank you to Prof. Michelle Chang for serving on both my qualifying exam and dissertation committees.

Kathryn Strobel is one of the most hard working and determined graduate students I've met, furthermore, her generosity and trust got me involved in this project in the first place. I'm not sure what this thesis would look like without Kathryn, and I'm happy to have had the chance to collaborate with her. Dr. Christy Roche is responsible for the *T. reesei* protein expression described in this work, and I thank her for her material contributions. I've also received material help from lab mates Meera Atreya and Dr. Erin Imsand, who have provided purified protein for these experiments. Dr. Hagit Sorek has contributed NMR experiments and analysis, and I thank her for her help. Thanks to Clark Lab members, especially Dr. Mike Liszka, Dr. Sarah Huffer Liszka, Dr. Kierston Shill, Dr. Christy Roche, and Dr. Jerome Fox, who have been dear friends and who have provided advice, encouragement, and criticism. Thank you to Claire Thomas, who laughs at all my jokes and is always on my side, and to the many incredible friends who have taught me so much.

I am grateful to the support staff at the College of Chemistry, including Christine Balolong, Rocio Sanchez, Fred Deakin, and Carlet Altamirano, for helping me navigate the intricacies of the UC system. The staff at the Energy Biosciences Institute, including Dr. Mara Bryan, Dr. Stefan Bauer, and Crystal Chan, keep the EBI running smoothly. The EBI has been a great place to work, not least because of the presence of these EBI staff members.

I thank the National Science Foundation, the Energy Biosciences Institute, and the state of California for funding this work.

I've had the great fortune to have a brilliant older sister whose example I have followed all my life. Thanks for everything.

Finally, I thank my parents, for a lifetime of unconditional love and support.

Chapter 1

Lignin: a barrier to economical biomass hydrolysis

1.1 Introduction

The plant cell wall polymer lignin is the second most abundant naturally occurring polymer on Earth, second only to cellulose [1]. These two polymers, lignin and cellulose, along with hemicellulose, make up the structural support system for terrestrial plants. Cellulose and hemicellulose, composed of repeating units of glucose, xylose, arabinose, and other fermentable sugars, can be deconstructed to provide a platform for the industrial fermentation of sugars to fuels, chemicals, drugs, and other products. The hydrolysis of cellulose to glucose can be carried out efficiently and specifically by a group of hydrolytic enzymes called cellulases. However, the action of these enzymes is inhibited by the presence of lignin in biomass.

Industrial-scale deconstruction of biomass is carried out by a group of cellulases produced by the filamentous fungus *Hypocrea jecorina* (also known as *Trichoderma reesei*). Lignin inhibition of these enzymes is thought to occur via both the physical obstruction of cellulose by lignin deposited on the crystalline cellulose surface, and by nonproductive adsorption, by which the enzyme becomes stuck to the solid lignin and is therefore unavailable for hydrolysis of cellulose.

Lignin and the plant cell wall

Lignin is a polymer compose of three monomer units, which are each variations on a substituted phenol ring with an aliphatic alcohol in the para position. The numbering nomenclature for the ring and aliphatic chain are given in Figure 1.1, while the naming nomenclature and structure for the three monomer types can be found in Table 1.1.

Table 1.1: Lignin monomer naming conventions and sources, adapted from [1, p.2]

Substituents	Name	Biomass source
R=R'=H	p-coumaryl alcohol (p-hydroxyphenyl)	grasses, compression wood
R=H, R'=OCH_3	coniferyl alcohol (guaiacyl)	hardwoods, softwoods, grasses
R=R'=OCH_3	sinapyl alcohol (syringyl)	hardwoods

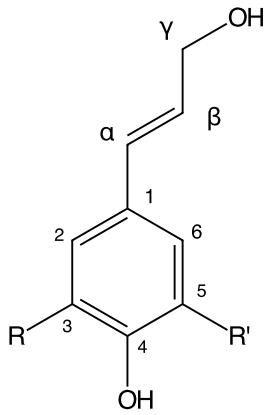

Figure 1.1: Structure and nomenclature of the three lignin monomers.

In the plant cell wall, lignin is formed by a free-radical polymerization that starts with an oxidation of the phenolic hydroxyl group (in lignin nomenclature, the C−4 position). This oxidation is carried out by a peroxidase or laccase, using hydrogen peroxide and oxygen respectively as terminal electron acceptors. The unpaired electron can delocalize at the O−4, C-1, C-3, C-5, and C-β positions, and can subsequently couple with another unpaired radical. The most common lignin linkage, the β-O-4 linkage, occurs when a phenolic hydroxy radical at the O-4 position pairs with a radical at the β position on the aliphatic tail. Similarly, common monomer linkages form at β-O-4, as well as carbon-carbon linkages between delocalized radicals. These are referred to as condensed structures (β-5, 5-5, 1-1, etc). The result is a crosslinked lignin network lacking a regular repeating structure and with a molecular weight ranging from several hundred g/mol to more than a million g/mole [1, p.7].

Lignin's role in the plant cell wall is to provide a barrier to mass transport to prevent evaporation, as well as to inhibit attack by pests and microbes [4]. It comprises $20 - 30\%$ of the dry mass of grasses and trees, depending on the type of biomass (softwood has a higher lignin content than hardwood). Lignin is found primarily in the secondary plant cell wall, where it is linked via ferrulic acid to hemicellulose, which is in turn associated with the cellulose microfibrils that provide the structural rigidity of the plant cell.

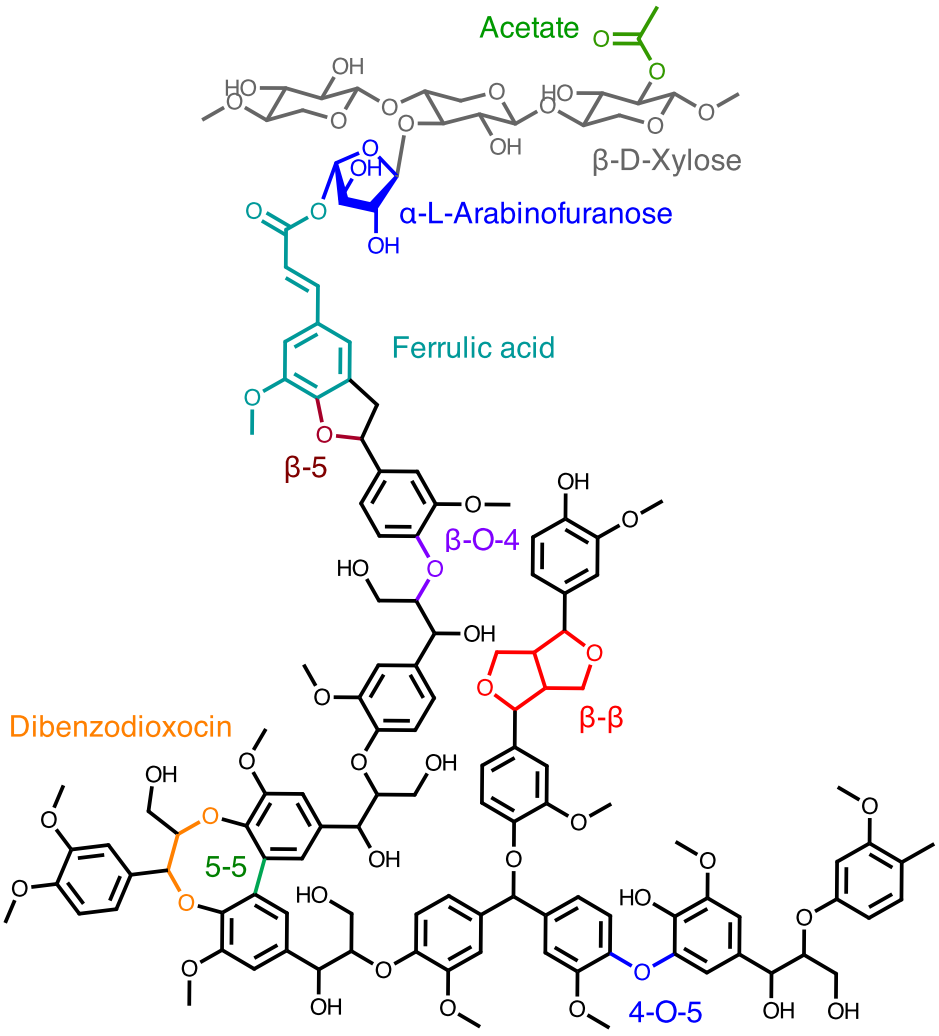

Figure 1.2: Model lignin and lignin-carbohydrate complex structure, identifying common structural features and bonds, and a lignin-carbohydrate complex through which the lignin is bound to xylan via a ferrulic acid linkage. Adapted from [2, 3]

1.2 Lignin and the biorefinery: effect of lignin on biomass deconstruction

The lignocellulosic biorefinery uses lignocellulosic biomass, such as perennial grasses, corn stover, sugar cane bagasse, woody biomass, or cellulosic post-consumer waste as an input. This material is pretreated by one or more of mechanical milling, heat, pressure changes, extreme pH, organic solvents, and oxidation with the goals of opening up the cell wall structure, decreasing the degree of crystallinity of cellulose, removing lignin and/or hemicellulose, and making the material generally more available to deconstruction by hydrolytic enzymes [5]. The solid cellulose is then hydrolyzed by a suite of enzymes including a endo- and exocellulases, as well as hemicellulases, β-glucosidases, polysaccharide monooxygenases, and other auxiliary enzymes. The product of this hydrolysis is a stream of fermentable sugars, which can be converted by microorganisms into fuel ethanol, butanol, or any of many chemical products that can be formed by the metabolism of microorganisms [6].

First generation sugar-based chemical processes, such as processes that convert starch or sucrose into fuels or chemicals, benefit from a homogeneous and easily deconstructed inputs, for example, corn starch and sugar cane liquor. Efforts to convert lignocellulosic materials benefit from the vast abundance and relative efficiency of producing perennial grasses, but are challenged by a more recalcitrant crystalline glucose polymer (cellulose) locked within a heterogeneous biomass.

Lignin's presence in biomass hydrolysis reactions has long been known to be inhibitory to the action of cellulase enzymes[7, 8, 9, 10, 11]. This phenomena is most apparent at low concentrations of cellulases [12], which supports the hypothesis that the major contributor to lignin inhibition is via nonproductive adsorption.

Numerous studies have attempted to quantify the extent of enzyme inhibition due to nonproductive binding using isolated lignin to supplement isolated cellulose [13, 10, 14]. The authors have shown that increasing lignin concentration in the hydrolysis reaction slows the initial hydrolysis rate [10], and decreases the final conversion of cellulose to glucose achieved by up to 80% [14]. While some researchers have shown dramatic reduction in cellulose hydrolysis upon the addition of purified lignins, others have observed a much more modest effect, if any at all [13]. Other studies have selectively removed lignin from biomass samples, for example by oxidation (perchloric acid) [15], and shown that lignin-free samples are much more readily and completely hydrolyzed than their lignin-containing counterparts.

The effect of lignin binding on cellulose hydrolysis is confounded by the inherent heterogeneity of the polymer and by the many methods used both to pretreat the biomass and isolate the lignin. The following sections will outline some sources of heterogeneity and provide justification for studying this system using lignin isolated from dilute-acid pretreated *Miscanthus*, using a milled wood lignin purification process.

Lignin chemistry: effect on lignin inhibition

Diverse lignin chemistry from diverse biomass sources

Cellulose is produced by many types of terrestrial plants, thus there is a wide variety of potential input streams in the biorefinery. The chemical makeup of lignin from different sources may be very different, beginning with the ratios of the three lignin monomers (see Table 1.1). Chemical linkages are also varied among these different lignin sources, and this has been shown to have varying effects on the cellulase inhibition.

Guo et al. (2014) extracted lignin from six different biomass sources and quantified the S/G ratio. They showed that among their samples the syringyl/guiacyl (S/G) monomer ratio was found to have a strong negative correlation with cellulase adsorption, while the hydroxyphenyl (H) content did not correlate with enzyme adsorption [16]. Other studies, however, have compared mutant *Arabidopsis* with varying S/G ratios, and saw no difference in biomass digestibility [17]. Decrease in ferulic cross-linking of arabinoxylans, on the other hand, was shown to increase the digestibility of the resulting biomass [18]. Others found that the blocking phenolic hydroxyl groups on lignin by hydroxypropylation reduced the inhibition of the model lignins vs. the unblocked lignin [12].

The degree of oxidation of lignin has also been hypothesized to be important in lignin's affinity and capacity to bind cellulases. Nakagame et al. (2011) showed that by creating dehydrogenative polymers (DHP) of monolignols with carboxylic acid functionality (ferrulic acid) and without (coniferyl alcohol).They saw that DHP made from ferrulic acid resulted in lignin with more carboxylic acid content (measured by FT-IR), and that the resulting lignin was not inhibitory when supplemented to an avicel-cellulase hydrolysis reaction [19]. They further found that when carboxylic acid content was compared across three biomass sources (corn stover, lodgepole pine, and poplar), high carboxylic acid content was correlated with high biomass digestibility [19].

Pretreatment and isolation effect on lignin

As varied as the lignin chemistry across plant species is, further variation in lignin chemistry is introduced when the biomass is pretreated. Pretreatment can break β-aryl ether moieties, result in condensed C-C bond structures, ethoxylate the α-hydroxyl group (ethanosolv process) [20], among other characterized and uncharacterized chemical changes.

Pretreatment is an essential part of enzymatic deconstruction of cellulosic biomass, and another sources of lignin modification. Pretreatment conditions exist spanning pH from 1 to 12, temperature from room temperature to around 250°C [5]. Delignification is a major goal of pretreatment, along with disruption and opening of the cell wall structure, decrystallization and depolymerization of cellulose and hemicellulose, and cleaving of lignin-carbohydrate complexes [6]. The many available pretreatment options have varying effects on lignin chemistry, content, and the resulting inhibitory effect of the lignin, summarized in Table 1.2.

Table 1.2: Summary of common biomass pretreatment processes, along with available information on the effect of the pretreatment process on lignin structure, and lignin-cellulase adsorptive capacity

Process	pH [6]	Temp. [6]	Res. lignin (%) [6]	Effect on lignin structure [21]	Cellulase-lignin adsorptive capacity [15, 22]
Steam explosion	3-5	180-290	50-60	Condensation of lignin structures, decrease in β-O-4 structures	N/A
Liquid hot water	4-5	160-230	N/A	N/A	64 mg/g
Dilute acid	1	160-220	80-90	Increase in condensed structures, decrease in β-O-4 structures	53 mg/g [23]
AFEX	12	40-180	100	Ammonolysis of ferrulic esther linkages with hemicellulose. Possible base-catalyzed depolymerization	38.7 mg/g
Liquid ammonia	9-12	160-180	15-60	Ammonolysis of ferrulic esther linkages with hemicellulose. Possible base-catalyzed dopolymerization. Extensive delignification.	42 mg/g
Oxidative Lime	7-9	25-180	40-50		65 mg/g
Ionic liquid	varies	100-180	66	Unknown	N/A
Biological oxidation	4-8	30-70		Oxidized	N/A

Table 1.3: Advantages and disadvantages of common biomass pretreatment processes. Adapted from [21]

Process	Advantages	Disadvantages
Steam explosion	Causes hemicellulose degradation and lignin transformation; cost-effective	Destruction of a portion of the xylan fraction; incomplete disruption of the lignin-carbohydrate matrix; generation of compounds inhibitory to microorganisms
Liquid hot water	Lower temperatures than steam explosion, minimal degradation products, no need for washing or neutralization, low solvent cost	Large water volumes, low product concentrations
Dilute acid	Limited chemical inputs, minimal dilution, low energy costs without recycling/environmental cost	Condensation and precipitation of lignin, destruction of hemicellulose, production of fermentation inhibitors, need to wash pretreated biomass, less effective for softwoods
AFEX	Increases accessible surface area, removes lignin and hemicellulose to an extent; does not produce inhibitors for downstream processes	Not efficient for biomass with high lignin content.
Liquid ammonia	Extensive delignification	Not effective for softwoods, cost of ammonia recycle, pH adjustment
Ionic liquid	Delignification, extensive cellulose decrystallization, fast hydrolysis rates [24]	Expensive solvent, recycle
Biological oxidation	Degrades lignin and hemicelluloses, low energy requirements	Rate of hydrolysis is very low, long residence time

Dilute acid pretreatment: Dilute acid pretreatment is currently among the most extensively studied, most common, and cost-effective pretreatment processes available on a per-gallon of ethanol basis [25]. The biomass is treated at high pressure and temperature with dilute ($0.5 - 2\%$ acid); hemicellulose is solublized and hydrolyzed while lignin has been shown to become soluble and redeposit as spherical globules on cellulose surfaces during pretreatment processes [26]. When using whole biomass to study the effect of acid pretreatment conditions, it can be difficult to decouple the effect of the pretreatment on the cellulose portion of biomass with its effect on lignin. To overcome this, the researchers pretreated isolated cellulose (cotton linters, filter paper) with and without isolated lignin, which was solubleized and visible in electron micrographs. Under certain conditions (high temperature, acidic conditions) these globules can inhibit the conversion of cellulose by 5-20% [26]. Selig et al.'s work was unable to differentiate between enzyme inhibition due to nonproductive adsorption to lignin and physical blockage of cellulose substrate.

Steam explosion and liquid hot water pretreatment: Both steam explosion and liquid hot water pretreatment use water without other chemical catalysts. A 2013 study compared the adsorption of monocomponent cellulases using quartz crystal microgravimetry and found that lignin isolated from steam explosion-pretreated substrates had a higher capacity for cellulase binding. The same study found that when Avicel was supplemented with purified lignin, a higher degree of inhibition was seen with lignin from the steam pretreated source [27]. Liquid hot water pretreatment has been studying at varying severities: researchers found that increasing the severity of the pretreatment made higher final conversions achievable, but increased the inhibitory effect of the lignin isolated from the pretreated biomass [28]. This was also shown to be driven by the adsorption of enzymes to the lignin surface, especially β-glucosidase [29].

Ammonia pretreatment: AFEX and liquid ammonia: Ammonia fiber/freeze explosion (AFEX) and liquid ammonia pretreatments are related processes, in that they both use ammonia to solublize lignin. In AFEX pretreatment, biomass is brought into contact with liquid anhydrous ammonia, held under pressures above 3 kPa, and rapidly depressurized. Liquid ammonia uses aqueous ammonia and results in significant fractionation of lignin and carbohydrate components, as lignin is dissolved and removed with the liquid ammonia stream. Both of these pretreatments are under basic conditions, and base-catalyzed cleavages as breakage of ferrulic ester links and phenolic β-O-4 cleavage occur during AFEX and liquid ammonia pretreatment [30].

Oxidative pretreatment: Oxidative pretreatments include ozonation, H_2O_2 treatment, high-pressure oxygen, oxidative lime pretreatment, and biological pretreatment utilizing lignin-degrading fungi and/or bacteria that produce oxidative enzymes.

Oxidative pretreatments may reduce the nonspecific binding of cellulases to lignin, as studies have shown that oxidation of lignin surfaces result in a lower affinity of cellulases to the surface [19]. However, like many correlations between pretreatment effect and lignin inhibition, it's very hard to deconvolute the effect of pretreatment on lignin with the effect of pretreatment on the biomass structure as a whole.

Ionic liquid pretreatment Ionic liquid pretreatment uses molten salts capable of dis-

solving cellulose, hemicellulose, and lignin in order to completely deconstruct the plant cell wall structure. It results in dramatically increased hydrolysis rates, and some separation of lignin and cellulose into separate streams. [24]. Ionic liquids such as 1-ethyl 3-methyl imidazolium acetate have a pH around 10, and have been shown to degrade β-O-4 linkages, decrease lignin molecular weight, and selective degradation of G-type lignin[31]. The effect of IL pretreatment on cellulase-lignin interactions is unknown.

Effect of lignin isolation method on lignin inhibition

When lignin is separated from its biomass source in order to better understand its inhibitory effect, there are many different ways to isolate lignin, and these methods result in varying chemical functionality and purity (common impurities include carbohydrates, protein, and surfactants). Different isolation methods, like different lignin sources, result in different cellulase binding behavior. Pareek et al. (2013) showed that when a cellulolytic enzyme mixture was exposed to six different lignins from varying sources and isolated by varying methods, the percentage of bound cellulase varied from $15 - 80\%$ [32]. Kraft lignin, a commercially available byproduct of the paper pulping industry, is a common example of a poor choice of lignin for inhibition measurements. Kraft lignin is created by a two-step process, in which lignin is first cleaved and made soluble by hot sodium hydroxide (1M), followed by nucleophilic attack using sodium sulfide. The resulting lignin in "black liquor" is precipitated by neutralization, and the lignin burned to produce energy and recover sulfur, which is chemically intercalated via sulfidolytic cleavage of β-aryl ether bonds [33].

Other common lignin isolation methods include klason lignin, isolated by treatment of biomass with high concentrations of sulfuric acid to hydrolyze any sugars present, cellulolytic enzyme lignin, in which biomass is milled and carbohydrates removed by treatment with excess amounts of cellulases, organosolv lignin, in which lignin is refluxed with organic solvents at high temperatures, and milled wood lignin, in which wood is milled and lignin extracted and purified with neutral organic solvents at room temperature. A summary of these isolation methods is presented in Table 1.2.

We have chosen a milled wood lignin isolation method, as it results in lignin free of carbohydrate or protein contamination, without chemical modification beyond that which occurs during pretreatment, and is insoluble in buffer. One potential drawback is that milled wood lignin is dissolved and re-precipitated, unlike klason lignin and cellulolytic enzyme lignin (CEL). If there is higher-order organization present in the plant cell wall lignin that affects cellulase-lignin binding, this might not be captured by MWL. There is one report in the literature of diminished nonproductive binding to CEL lignin after it has been dissolved and re-precipitated [27]. However, we believe that the other advantages of MWL, including its lack of carbohydrate and protein impurities, and lack of modification during isolation, make it the best choice.

Table 1.4: Lignin can be isolated by a variety of methods, summarized below.

Isolation method	Process description	Advantages	Drawbacks
Bjrkman milled wood lignin (MWL)	Ball-milled biomass is treated with aqueous dioxane to extract lignin	Generally considered representative of native lignin structure, low carbohydrate content	20-30% yields = incomplete lignin recovery [34, 35]
Cellulolytic enzyme lignin (CEL)	Ball-milled biomass is treated with cellulase enzymes to remove carbohydrates and proteases to remove cellulases.	High (90%) yield, minimal chemical modification	$10 - 12\%$ residual carbohydrate content, residual enzymes content varies, high MW (insoluble in DMSO, dioxane, DMF) [35]
Kraft lignin	Biomass treated with NaOH and Na_2S to solubilize lignin	Commercially available	Highly modified: high phenolic content, low methoxy content, and low MW [33]
Klason lignin	Sulfuric acid hydrolysis of sugars; lignin recovered as insoluble residue	High lignin recovery (90%), used as analytical method for lignin content	Highly condensed and altered chemical makeup [36]

1.3 Cellulase structure, function, and diversity: effect on lignin interactions

The four major cellulose-degrading enzymes of the *Hypocrea jecorina* secretome are comprised of two domains, a roughly 36-amino acid carbohydrate binding module (CBM), linked by a flexible, glycosylated linker to a catalytic domain (CD) [37].

CBM families and diversity in carbohydrate binding

Carbohydrate binding domains are a large and diverse group of non-catalytically active, polysaccharide-recognizing protein domains, with binding affinity for targets ranging from crystalline and amorphous cellulose, xylan, starch, chitin, mannan, glycogen, and cell-surface glycans [38]. Their substrate specificity is varied as well, from very specific glycan recognition to CBMs that bind a wide variety of carbohydrates. Currently, nearly 40,000 known and putative CBMs divided into 69 families (see http://www.cazy.org/Carbohydrate-Binding-Modules.html for updated count). In addition to the family organization of CBMs, which is based on fold and evolutionary relationships, CBMs are further classified by type. Type A CBMs have binding specificity for crystalline polysaccharides, including cellulose, while type B CBMs bind single polysaccharide chains and/or amorphous regions where crystalline structure is disrupted [38]. Type C CBMs exhibit lectin-like binding properties and recognize monomeric and short oligomeric saccharides, including free ends of cellulose chains [39].

Members of CBM family 1 are type A domains, include the crystalline cellulose-binding domains of the four canonical cellulases from *Hypocrea jecorina*; they are the major CBM family found in cellulose-degrading fungi [40]. These CBMs are small (about 40 amino acids), wedge-shaped, and interact with solid cellulose via three aromatic residues that form a flat face on the binding surface. Mutation of any these aromatic residues greatly diminishes cellulose binding [41]. While fungal family 1 CBMs are the most prominate examples of type A CBMs, crystalline-specific type A CBMs are also in bacterial CBMs. Examples from families 2a, 3a, 5 and 10 also display affinity for crystalline substrates. Although their folds are very different (see 1.3 for example CBM structures), they share the flat, aromatic binding face that promotes interaction with crystalline polysaccharides.

Type B CBMs span at least 12 CBM families, however they share a common β-sandwich fold structure [38]. They also show far more diversity in substrate binding specificity than fungal type 1 CBMs. Type B CBMs exist to target individual cellulose, xylose, and mannose chains, and as one might expect, are attached to catalytic domain with activity toward the targeted substrate [38] .

Type C lectin-like CBMs bind primarily to soluble sugars. Once thought not to interact with solid substrates, they were included in CBM family nomenclature because they have been discovered as domains on glycoside hydolases. It's now known that some Type C CBMs bind to the chain ends of polysaccharides, however their interaction with crystalline cell walls is minimal [40]. Type A, B, and C domains spanning several CBM families are illustrated

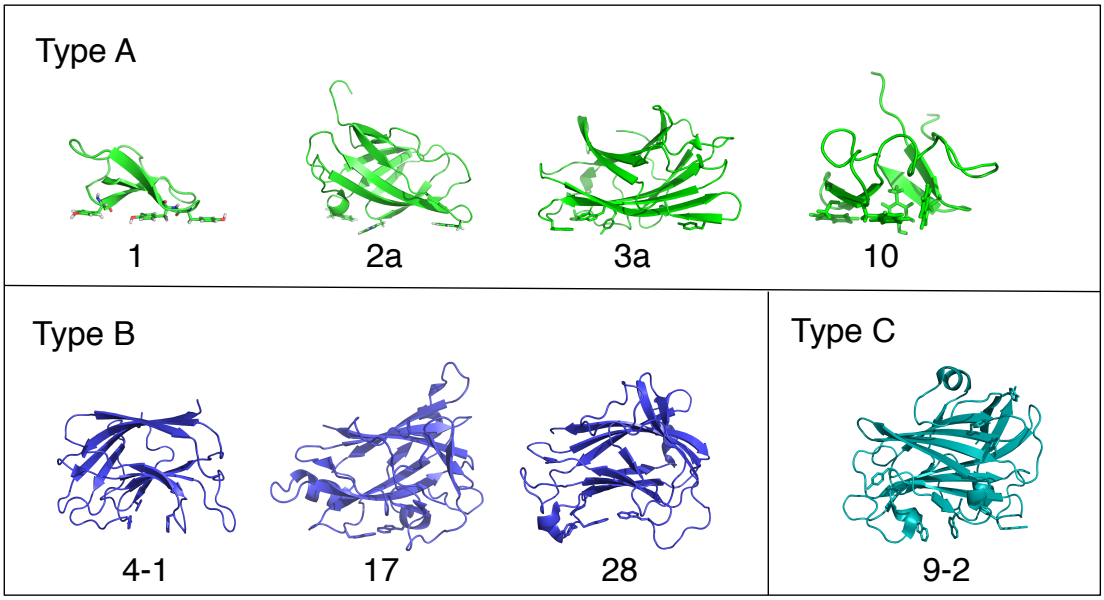

Figure 1.3: Examples of type A, B, and C carbohydrate binding modules, adapted from [40].

Table 1.5: Organism and family names of CBMs displayed in 1.3

CBM Type	Fold family	Organism	PDB ID
A	1	*Hypocrea jecorina*	1CBM
A	2a	*Cellulomonas fimi*	1EXG
A	3a	*Clostridium cellulolyticum*	1G43
A	10	*Pseudomonas fluorescens*	1QLD
B	4-1	*Cellulomonas fimi*	1GU3
B	17	*Clostridium cellulovorans*	1J83
B	28	*Bacillus akibai*	1UWW
C	9-2	*Thermotoga maritima*	1I8U

in Figure 1.3. Solvent-exposed aromatic residues that have been shown or hypothesized to interact with the pyranose rings on solid substrates are shown as side chains.

The carbohydrate binding domain was shown to drive the interaction of *Hypocrea jecorina* cellulases with cellulose in 1988, when selective proteolysis of the linker domain was found to diminish cellulase binding [37]. The interaction of CBM1 from Cel7A was shown to be mediated by the three planar aromatic residues of the binding domain several years later, when mutations of those residues resulted in diminish binding of the CBM to crystalline cellulose [42, 41]. Since then, other studies have distinguished between crystalline and amorphous-specific CBMs [40, 43], and affinity for different crystal faces of cellulose [44].

Lignin-CBM interaction: contributing factors & mitigation

Like cellulose-cellulase binding, lignin-cellulase interaction has also been hypothesized to occur via the carbohydrate binding domain. Palonen et al. (2004) showed that on several model lignin substrates (Kraft lignin and cellulolytic enzyme lignin from steam pretreated spruce), *Hj*Cel7A bound to lignin only when the CBM was present; isolated catalytic domains produced no detectable binding. A similar disparity was noted with *Hj*Cel5A, which bound to lignin both as a full enzyme, and to a lesser extent as an isolated CD [45].

Other researchers tested the hypothesis that the same aromatic residues that mediate cellulose binding are also responsible for lignin binding. *Hj*Cel7A CBM has three tyrosine residues that make up a flat binding face (see Figure 1.3A). When one of these residues was mutated to an alanine, lignin affinity measured by a Lanmuir-type isotherm decreased. When another of the tyrosines was mutated to a tryptophan, the lignin affinity increased [27].

The three aromatic residues on the CBM surface form a hydrophobic patch on the flat binding face, and researchers have hypothesized that a measure of the prevalence of hydrophobic patches, rather than total surface hydrophobicity, may be a better way to predict enzyme interaction with lignin [46]. These researchers found that neither the total exposed hydrophobic area, nor the isoelectric point, was correlated with adsorption to a lignin film. However, when they used a hydrophobic patch score (as described in [47]), the adsorption rate and total surface capacity of 8 different enzymes was well correlated [46].

Although the CBM has been implicated as a major contributor to lignin-binding behavior, some cellulases are natively expressed without CBMs. Two natively CBM-free endoglucanases were tested for their adsorption to lignin, and the authors showed that bound to lignin to varying degrees and hydrolysis was inhibited by lignin adsorption [48].

Thermal stability has also been suggested as a potential factor for susceptibility to nonproductive binding on lignin. More thermostable enzymes may be less likely to unfold on the lignin surface, and thus may bind more reversibly. This hypothesis is supported by work comparing chimeric enzymes of *Hj*Cel7A CBM and several catalytic domains of varying thermostability. The most thermostable enzyme was least affected by the presence of lignin in a biomass hydrolysis reaction [49].

pH may also play a role in the nonproductive binding of lignin and cellulases. While the *Hypocrea jecorina* cellulases are most active on purified cellulose and small molecule soluble substrates around pH 4.85, a higher pH may produce better results when whole woody biomass is used [50]. Lan et al. (2013) showed that significantly higher biomass yields could be achieved by increasing the pH of the reaction to between 5.2 and 6.2, and suggested that lignin binding might play a role in this higher pH optimum. When either an endoglucanase catalytic domain or the same CD fused to the *Hj*Cel7A CBM was exposed to isolated lignin at varying pH (between 4 and 7), the cellulase affinity for lignin dropped by up to 15-fold [51]. Other researchers showed the opposite effect with xylanase: as the pH was increased, the fraction of xylanase adsorbing to kraft lignin increased [52].

Finally, the role of the linker domain, an unstructured, heavily glycosylated region that

joins the catalytic domain and the CBM, might also contribute to the affinity of cellulases for lignin. Mutations to the linker have been published in a patent [53], claiming that decreasing the pI of the linker and increasing the serine:thronine ratio in the linker both result in decreased lignin binding.

Methods of minimizing lignin's effect on cellulases

As discussed above, research is ongoing to attempt to reduce cellulase affinity for lignin through mutagenesis, evolution, and understanding structural contributions to lignin binding. Several other methods have also proven effective in mitigating nonproductive binding to lignin. Addition of surfactants, which may reduce lignin binding, increase enzyme stability, or decrease interfacial deactivation, have been studied extensively.

The addition of non-ionic surfactants decrease the overall binding affinity of Cel7A to steam pretreated spruce by about 25%, and increased the final substrate by over 50% [54]. Because the surfactants did not affect thermal stability, the authors concluded that reducing nonproductive binding was the mechanism by which the surfactant increased the efficiency of the hydrolysis reaction. Contact angle measurements on pretreated corn stover, compared with biomass treated with surfactants, show a decreased contact angle, along with increased glucose yield [55]. Other competitive binding blockers, such as the proteins ovalbumin and BSA, have also been observed to bind to lignin and to increase glucose yields [46, 56].

1.4 Direct measurement of cellulase-lignin interaction kinetics

Direct measurement of cellulase-lignin binding interactions can make it possible to understand the effect of the many factors that can contribute to this phenomenon. Previous researchers have constructed Langmuir isotherms [45] and measured maximum surface capacity of cellulase on lignin [15]. These equilibrium measurements, while informative, can't give us access to binding kinetics or reaction mechanisms that include an irreversible component.

There are a number of experimental methods for tracking surface changes, including quartz crystal microgravimetry, which is discussed extensively in Chapter 2. In Chapter 3, we extend the methods developed in Chapter 2 to the four most important enzymes in the *H. jecorina* cellulose degradation system and their isolated catalytic domains. Using this information, we can identify new targets for enzyme evolution, provide kinetic parameters for modeling, and develop a greater understanding of the structural underpinnings of lignin-cellulose nonproductive binding.

Bibliography

[1] Cyril Heitner, Don Dimmel, and John A. Schmidt. *Lignin and lignans : advances in chemistry*. Taylor & Francis, Boca Raton, 2010.

[2] Dylan Dodd and Isaac KO Cann. Enzymatic deconstruction of xylan for biofuel production. *GCB Bioenergy*, 1(1):2–17, 2009.

[3] W. Graham Forsythe, Mark D. Garrett, Christopher Hardacre, Mark Nieuwenhuyzen, and Gary N. Sheldrake. An efficient and flexible synthesis of model lignin oligomers. *Green Chemistry*, 15(11):3031–3038, 2013.

[4] Ruben Vanholme, Brecht Demedts, Kris Morreel, John Ralph, and Wout Boerjan. Lignin biosynthesis and structure. *Plant Physiology*, 153(3):895–905, 2010.

[5] A. T. W. M. Hendriks and G. Zeeman. Pretreatments to enhance the digestibility of lignocellulosic biomass. *Bioresour Technol*, 100(1):10–8, 1 2009.

[6] Shishir PS Chundawat, Gregg T. Beckham, Michael E. Himmel, and Bruce E. Dale. Deconstruction of lignocellulosic biomass to fuels and chemicals. *Annual review of chemical and biomolecular engineering*, 2:121–145, 2011.

[7] A. O. Converse, H. Ooshima, and D. S. Burns. Kinetics of enzymatic hydrolysis of lignocellulosic materials based on surface area of cellulose accessible to enzyme and enzyme adsorption on lignin and cellulose. *Applied Biochemistry and biotechnology*, 24(1):67–73, 1990.

[8] H. Ooshima, D. S. Burns, and A. O. Converse. Adsorption of cellulase from Trichoderma reesei on cellulose and lignacious residue in wood pretreated by dilute sulfuric acid with explosive decompression. *Biotechnol Bioeng*, 36(5):446–52, 8 1990.

[9] Yanpin Lu, Bin Yang, David Gregg, John N. Saddler, and Shawn D. Mansfield. Cellulase adsorption and an evaluation of enzyme recycle during hydrolysis of steam-exploded softwood residues. *Appl Biochem Biotechnol*, 98-100:641–54, 2002.

[10] Alex Berlin, Mikhail Balakshin, Neil Gilkes, John Kadla, Vera Maximenko, Satoshi Kubo, and Jack Saddler. Inhibition of cellulase, xylanase and beta-glucosidase activities by softwood lignin preparations. *J Biotechnol*, 125(2):198–209, 9 2006.

[11] Seiji Nakagame, Richard P. Chandra, and Jack N. Saddler. The effect of isolated lignins, obtained from a range of pretreated lignocellulosic substrates, on enzymatic hydrolysis. *Biotechnol Bioeng*, 105(5):871–9, 4 2010.

[12] Xuejun Pan. Role of functional groups in lignin inhibition of enzymatic hydrolysis of cellulose to glucose. *Journal of biobased materials and bioenergy*, 2(1):25–32, 2008.

[13] L. Meunier-Goddik and M. H. Penner. Enzyme-catalyzed saccharification of model celluloses in the presence of lignacious residues. *J Agric Food Chem*, 47(1):346–51, 1 1999.

[14] Jenni Rahikainen, Saara Mikander, Kaisa Marjamaa, Tarja Tamminen, Angelos Lappas, Liisa Viikari, and Kristiina Kruus. Inhibition of enzymatic hydrolysis by residual lignins from softwood–study of enzyme binding and inactivation on lignin-rich surface. *Biotechnol Bioeng*, 108(12):2823–34, 12 2011.

[15] Rajeev Kumar and Charles E. Wyman. Cellulase adsorption and relationship to features of corn stover solids produced by leading pretreatments. *Biotechnol Bioeng*, 103(2):252–67, 6 2009.

[16] Fenfen Guo, Wenjing Shi, Wan Sun, Xuezhi Li, Feifei Wang, Jian Zhao, and Yinbo Qu. Differences in the adsorption of enzymes onto lignins from diverse types of lignocellulosic biomass and the underlying mechanism. *Biotechnol Biofuels*, 7(1):38, 2014.

[17] Hans-Joachim G. . J. G. Jung, Weiting Ni, Clint C. S. Chapple, and Knut Meyer. Impact of lignin composition on cell-wall degradability in an arabidopsis mutant. *Journal of the Science of Food and Agriculture*, 79(6):922–928, 1999.

[18] John H. Grabber, John Ralph, and Ronald D. Hatfield. Ferulate cross-links limit the enzymatic degradation of synthetically lignified primary walls of maize. *Journal of Agricultural and Food Chemistry*, 46(7):2609–2614, 1998.

[19] Seiji Nakagame, Richard P. Chandra, John F. Kadla, and Jack N. Saddler. Enhancing the enzymatic hydrolysis of lignocellulosic biomass by increasing the carboxylic acid content of the associated lignin. *Biotechnol Bioeng*, 108(3):538–48, 3 2011.

[20] Stefan Bauer, Hagit Sorek, Valerie D. Mitchell, Ana B. Ibáñez, and David E. Wemmer. Characterization of Miscanthus giganteus lignin isolated by ethanol organosolv process under reflux condition. *J Agric Food Chem*, 60(33):8203–12, 8 2012.

[21] Parveen Kumar, Diane M. Barrett, Michael J. Delwiche, and Pieter Stroeve. Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production. *Industrial & Engineering Chemistry Research*, 48(8):3713–3729, 2009.

[22] Dahai Gao, Shishir P. S. Chundawat, Nirmal Uppugundla, Venkatesh Balan, and Bruce E. Dale. Binding characteristics of Trichoderma reesei cellulases on untreated, ammonia fiber expansion (AFEX), and dilute-acid pretreated lignocellulosic biomass. *Biotechnol Bioeng*, 108(8):1788–800, 8 2011.

[23] Poulomi Sannigrahi, Arthur J. Ragauskas, and Stephen J. Miller. Effects of two-stage dilute acid pretreatment on the structure and composition of lignin and cellulose in loblolly pine. *BioEnergy Research*, 1(3-4):205–214, 2008.

[24] Kierston Shill, Sasisanker Padmanabhan, Qin Xin, John M. Prausnitz, Douglas S. Clark, and Harvey W. Blanch. Ionic liquid pretreatment of cellulosic biomass: enzymatic hydrolysis and ionic liquid recycle. *Biotechnol Bioeng*, 108(3):511–20, 3 2011.

[25] Valery B. Agbor, Nazim Cicek, Richard Sparling, Alex Berlin, and David B. Levin. Biomass pretreatment: fundamentals toward application. *Biotechnol Adv*, 29(6):675–85, 2011.

[26] Michael J. Selig, Sridhar Viamajala, Stephen R. Decker, Melvin P. Tucker, Michael E. Himmel, and Todd B. Vinzant. Deposition of lignin droplets produced during dilute acid pretreatment of maize stems retards enzymatic hydrolysis of cellulose. *Biotechnol Prog*, 23(6):1333–9, 2007.

[27] Jenni L. Rahikainen, Raquel Martin-Sampedro, Harri Heikkinen, Stella Rovio, Kaisa Marjamaa, Tarja Tamminen, Orlando J. Rojas, and Kristiina Kruus. Inhibitory effect of lignin during cellulose bioconversion: The effect of lignin chemistry on non-productive enzyme adsorption. *Bioresource technology*, 2013.

[28] Ja Kyong Ko, Youngmi Kim, Eduardo Ximenes, and Michael R. Ladisch. Effect of liquid hot water pretreatment severity on properties of hardwood lignin and enzymatic hydrolysis of cellulose. *Biotechnol Bioeng*, 7 2014.

[29] Ja Kyong Ko, Eduardo Ximenes, Youngmi Kim, and Michael R. Ladisch. Adsorption of enzyme onto lignins of liquid hot water pretreated hardwoods. *Biotechnol Bioeng*, 8 2014.

[30] Venkatesh Balan, Leonardo da Costa a. . C. Sousa, Shishir P. S. Chundawat, Derek Marshall, Lekh N. Sharma, C. Kevin Chambliss, and Bruce E. Dale. Enzymatic digestibility and pretreatment degradation products of AFEX-treated hardwoods (Populus nigra). *Biotechnol Prog*, 25(2):365–75, 2009.

[31] Jia-Long . L. Wen, Tong-Qi . Q. Yuan, Shao-Long . L. Sun, Feng Xu, and Run-Cang . C. Sun. Understanding the chemical transformations of lignin during ionic liquid pretreatment. *Green Chemistry*, 16(1):181–190, 2014.

[32] Nidhi Pareek, Thomas Gillgren, and Leif J. Jönsson. Adsorption of proteins involved in hydrolysis of lignocellulose on lignins and hemicelluloses. *Bioresour Technol*, 148:70–7, 11 2013.

[33] Josef Gierer. Chemical aspects of kraft pulping. *Wood Science and Technology*, 14(4):241–266, 1980.

[34] K. Lundquist and R. Simonson. Lignin preparations with very low carbohydrate content. *Svensk Papperstidning*, 78(11):360, 1975.

[35] Kevin M. Holtman, Hou-Min M. Chang, and John F. Kadla. Solution-state nuclear magnetic resonance study of the similarities between milled wood lignin and cellulolytic enzyme lignin. *J Agric Food Chem*, 52(4):720–6, 2 2004.

[36] Thomas Q. Hu. *Chemical modification, properties, and usage of lignin*. Kluwer Academic/Plenum Publishers, New York, 2002.

[37] P. Tomme, H. Van Tilbeurgh, G. Pettersson, J. Van Damme, J. Vandekerckhove, J. Knowles, T. Teeri, and M. Claeyssens. Studies of the cellulolytic system of Trichoderma reesei QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis. *Eur J Biochem*, 170(3):575–81, 1 1988.

[38] Alisdair B. Boraston, David N. Bolam, Harry J. Gilbert, and Gideon J. Davies. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J*, 382(Pt 3):769–81, 9 2004.

[39] V. Notenboom, A. B. Boraston, D. G. Kilburn, and D. R. Rose. Crystal structures of the family 9 carbohydrate-binding module from Thermotoga maritima xylanase 10A in native and ligand-bound forms. *Biochemistry*, 40(21):6248–56, 5 2001.

[40] Anthony W. Blake, Lesley McCartney, James E. Flint, David N. Bolam, Alisdair B. Boraston, Harry J. Gilbert, and J. Paul Knox. Understanding the biological rationale for the diversity of cellulose-directed carbohydrate-binding modules in prokaryotic enzymes. *J Biol Chem*, 281(39):29321–9, 9 2006.

[41] Markus Linder, Maija-Liisa Mattinen, Maarit Kontteli, Gunnar Lindeberg, Jerry Ståhlberg, Torbjörn Drakenberg, Tapani Reinikainen, GöRan Pettersson, and Arto Annila. Identification of functionally important amino acids in the cellulose-binding domain of Trichoderma reesei cellobiohydrolase I. *Protein Sci*, 4(6):1056–1064, 6 1995.

[42] T. Reinikainen, L. Ruohonen, T. Nevanen, L. Laaksonen, P. Kraulis, T. A. Jones, J. K. Knowles, and T. T. Teeri. Investigation of the function of mutated cellulose-binding domains of Trichoderma reesei cellobiohydrolase I. *Proteins*, 14(4):475–82, 12 1992.

[43] G. Carrard, A. Koivula, H. Söderlund, and P. Béguin. Cellulose-binding domains promote hydrolysis of different sites on crystalline cellulose. *Proc Natl Acad Sci U S A*, 97(19):10342–7, 9 2000.

[44] Janne Lehtiö, Junji Sugiyama, Malin Gustavsson, Linda Fransson, Markus Linder, and Tuula T. Teeri. The binding specificity and affinity determinants of family 1 and family 3 cellulose binding modules. *Proc Natl Acad Sci U S A*, 100(2):484–9, 1 2003.

[45] Hetti Palonen, Folke Tjerneld, Guido Zacchi, and Maija Tenkanen. Adsorption of Trichoderma reesei CBH I and EG II and their catalytic domains on steam pretreated softwood and isolated lignin. *Journal of Biotechnology*, 107(1):65–72, 1 2004.

[46] Deanne W. Sammond, John M. Yarbrough, Elisabeth Mansfield, Yannick J. Bomble, Sarah E. Hobdey, Stephen R. Decker, Larry E. Taylor, Michael G. Resch, Joseph J. Bozell, Michael E. Himmel, Todd B. Vinzant, and Michael F. Crowley. Predicting enzyme adsorption to lignin films by calculating enzyme surface hydrophobicity. *J Biol Chem*, 289(30):20960–9, 7 2014.

[47] Ron Jacak, Andrew Leaver-Fay, and Brian Kuhlman. Computational protein design with explicit consideration of surface hydrophobic patches. *Proteins*, 80(3):825–38, 3 2012.

[48] Alex Berlin, Neil Gilkes, Arwa Kurabi, Renata Bura, Maobing Tu, Douglas Kilburn, and John Saddler. Weak lignin-binding enzymes: a novel approach to improve activity of cellulases for hydrolysis of lignocellulosics. *Appl Biochem Biotechnol*, 121-124:163–70, 2005.

[49] Jenni Liisa Rahikainen, Ulla Moilanen, Susanna Nurmi-Rantala, Angelos Lappas, Anu Koivula, Liisa Viikari, and Kristiina Kruus. Effect of temperature on lignin-derived inhibition studied with three structurally different cellobiohydrolases. *Bioresource Technology*, 146(0):118 – 125, 2013.

[50] T. Q. Lan, Hongming Lou, and J. Y. Zhu. Enzymatic saccharification of lignocelluloses should be conducted at elevated pH 5.2–6.2. *BioEnergy Research*, 6(2):476–485, 2013.

[51] Jenni Liisa Rahikainen, James David Evans, Saara Mikander, Anna Kalliola, Terhi Puranen, Tarja Tamminen, Kaisa Marjamaa, and Kristiina Kruus. Cellulase-lignin interactions–the role carbohydrate-binding module and pH in non-productive binding. *Enzyme and Microbial Technology*, 2013.

[52] Keungarp Ryu and Youngon Kim. Adsorption of a xylanase purified from pulpzyme hc onto alkali-lignin and crystalline cellulose. *Biotechnology letters*, 20(10):987–990, 1998.

[53] Brian R. Scott, Patrick St-Pierre, James Lavigne, Nabil Masri, Theresa C. White, and John J. Tomashek. Novel lignin-resistant cellulase enzymes, 2 2010.

[54] Torny Eriksson, Johan Börjesson, and Folke Tjerneld. Mechanism of surfactant effect in enzymatic hydrolysis of lignocellulose. *Enzyme and Microbial Technology*, 31(3):353–364, 2002.

[55] Qing Qing, Bin Yang, and Charles E. Wyman. Impact of surfactants on pretreatment of corn stover. *Bioresour Technol*, 101(15):5941–51, 8 2010.

[56] Bin Yang and Charles E. Wyman. Bsa treatment to enhance enzymatic hydrolysis of cellulose in lignin containing substrates. *Biotechnol Bioeng*, 94(4):611–7, 7 2006.

Chapter 2

Development of a method to measure lignin-cellulase interaction kinetics

2.1 Abstract

The kinetics and mechanism of protein interaction with solid surfaces is a broadly relevant phenomena, important to fields as diverse as industrial biocatalysis, biomedical applications, food science, and cell biology. The nonproductive interaction of cellulase enzymes with lignin, a plant cell wall polymer, inhibits the large-scale conversion of biomass to soluble sugars. Cellulase has been shown to interact with lignin, but heterogeneity of lignin surfaces, challenges in measuring irreversible components of these interactions, and fast reaction kinetics have made quantifying the kinetics of these reactions difficult.

This work uses quartz crystal microgravimetry with dissipation monitoring (QCM-D) to take real-time measurements of adsorbed mass on a flat lignin surface. We have developed a method for casting homogeneous lignin films that are chemically similar to lignin found in pretreated biomass, and used QCM-D to compare three models of reversible-irreversible binding behavior: a transition model, a a two-state model with changing adsorbate footprint, and a multiple-site transition model. We find that of the three models tested, the multiple-site transition model is the only plausible kinetic mechanism for the behavior of Cel7B binding to lignin.

While the implications of lignin-cellulase interactions may be limited to fields of renewable energy and green chemistry, the analytical and experimental methods demonstrated are relevant to any system in which the kinetics and reaction mechanism of reversible and irreversible adsorption at a solid-liquid interface are important.

2.2 Introduction

Proteins are known to adsorb at solid-liquid interfaces, both productively and nonproductively, and an understanding of the underlying kinetics and mechanisms of these reactions

is central to understanding phenomena such as surface fouling, competitive adsorption, heterogeneous catalysis, bioseparations, and cell signaling. Equilibrium measurements of bound and unbound species, as in the creation of a Langmuir isotherm, can model equilibrium behavior of equilibrium reactions. When there is the potential for irreversibly bound species on the surface, or when competitive adsorption makes binding kinetics important, quantifying surface interactions becomes considerably more difficult. This work demonstrates the evaluation of three models that account for surface adsorption, desorption, and irreversible adsorption of cellulase enzyme to lignin, a plant cell wall polymer.

The conversion of lignocellulosic biomass to soluble sugars using hydrolytic enzymes is platform for renewable fuels and chemicals. The action of these enzymes is impeded by the presence of lignin [1]. Lignin is a major component of the plant cell wall, second only to cellulose, and has a well-documented and deleterious effect on the enzymatic conversion of cellulose to sugars through physical blockage of cellulose [2, 3], and via nonproductive adsorption of hydrolytic enzymes to lignin [4, 5], reviewed in 2011 [6].

Past efforts to measure cellulase-lignin binding have been limited to either indirect measurements (by tracking changes in cellulose hydrolysis of biomass with varying lignin content, for example) [7, 8, 1, 9, 5, 10, 11], and to equilibrium binding studies [11, 12, 4]. Irreversible binding of cellulases to lignin has been detected previously[12, 13]. Equilibrium measurements fail to account for irreversible surface interaction, and are insufficient to capture the rate by which enzymes adsorb to the surface. Industrial interest in lignin-cellulase interactions, together with the limitations of existing methods for quantifying this phenomenon, prompted us to test several kinetic models for these interactions.

We propose to directly measure binding kinetics by tracking cellulase mass on a smooth lignin surface using quartz crystal microgravimetry with dissipation monitoring (QCM-D). Cellulase is adsorbed to and desorbed from a lignin-coated sensor while varying bulk enzyme concentration, flow rate, and contact time. The QCM-D measures surface mass by recording the frequency change of a piezoelectric quartz crystal sensor as mass binds to the surface; this frequency shift is related to the surface mass. From these measurements we can evaluate kinetic models for reversible and irreversible enzyme adsorption, and calculate rate constants for each reaction in the model. To correctly estimate kinetic rate constants, we must 1) develop homogeneous, reproducible, representative lignin surfaces from a representative isolated lignin, 2) evaluate and account for potential artifacts in the QCM-D measurement, including mass transfer limitations of the QCM-D and the extent of hydration of the protein film, and 3) specify and test likely kinetic models.

Preparation of lignin surfaces

Lignin is a cross linked, free-radical polymerized polymer network of phenylpropane units. It provides the plant cell wall with structural rigidity, hydrophobicity, and protection from pests. Not only does lignin structure vary among plant types, but its structure can be further altered by isolation and pretreatment processes [11, 14, 15, 16]. These changes in chemical functionality can affect cellulase adsorption: for example, lignin isolated from a

variety of feedstocks and with varying pretreatment and isolation conditions were shown to inhibit cellulose hydrolysis to varying degrees [4, 10]. Contamination by carbohydrates or protein can result from some enzymatic lignin isolation processes and should be avoided. We seek to develop a tool to evaluate lignin's effect on cellulases in the biorefinery, therefore have isolated lignin from pretreated biomass. We chose the milled wood lignin extraction process to isolate lignin, as it results in an isolated lignin free of carbohydrates, protein contamination, and is chemically similar to lignin found in the plant cell wall [17, 18, 19].

Examples lignin films suitable for QCM-D applications are present in the literature [20, 21]. However, the diversity in chemical composition and functionality that arises from the wide variety of biomass sources, pretreatment conditions, and isolation methods can result in large differences in surface coating, even when the same method for dissolving lignin and spin coating are used. Considerable method development must be done with individual lignin preparations to ensure thin and homogeneous surfaces are produced.

Understanding potential artifacts in QCM-D measurement

Measurement of reaction kinetics using QCM-D, or with any flow-cell based detection system such as surface plasmon resonance or other microfluidics-based biosensors, is often complicated by the potential for concentration gradients within the bulk solution. Fast kinetics, low flow rates, and laminar flow make diffusion limitations more likely within the sensor cell [22]. Damköhler numbers on the order of 1 have been predicted for enzyme binding reactions in QCM-D and SPR systems [23, 24, 25]. This means that that the system is neither fully mass-transfer limited nor fully reaction limited, but that concentration gradients are likely to exist, dependent upon the flow rate, analyte concentration, diffusion rate, and intrinsic reaction kinetics.

When attempting to measure the intrinsic reaction rate constants, diffusion limitations must be either accounted for within the kinetic model, or the device operated in flow and concentration regimes in which the reaction rate is not limited by mass transfer.

Another complication in measuring protein-surface binding kinetics with QCM-D is the fact that adsorbed protein is hydrated. The frequency change that the instrument measured is a function of the total mass on the surface: both the protein molecules adsorbed, as well as the water that hydrates the protein film [26]. In order to get an accurate measurement of the molar concentration of protein on the surface, a secondary measurement of protein concentration is needed. A 2004 study using simultaneous QCM and surface plasmon resonance showed that up to 90% of the apparent adsorbed mass can be attributed to adsorbed water, and that the amount of adsorbed water was dependent on the surface concentration of the analyte [27]. In this work, we have estimated the molar concentration of adsorbed protein films by binding and eluting the protein from a surface, and measuring enzyme activity in the eluate.

Kinetic models

When protein adsorption to solid surfaces was reviewed, it was noted that the vast majority of protein-surface adsorption interactions that had been studied were at least partially irreversible [28]. As such, the Langmuir model, which considers adsorption and desorption of indiscriminate particles to distinct binding sites, is insufficient to describe the adsorption behavior of proteins [29]. Complications to the Langmuir model such as irreversible binding, changes in protein conformation on the surface, surface clustering, and cooperativity in binding behavior are reviewed in [30]. A number protein adsorption studies have used two-state models to explain protein adsorption and irreversible binding behavior. A 2005 study of the adsorption of carbonic anhydrase to functionalized surfaces showed that the stability of mutant enzymes affected the degree of irreversibility of adsorption [31]. The authors were not, however, able to measure the kinetics of either the irreversible reaction or of the desorption reaction, instead, their results were based on reversibility measurements obtained at a single time point. Other studies have investigated reversible and irreversible adsorption and modeled reaction kinetics [32, 33], and adsorbed and desorbed protein structure [34, 35]. We have built upon the methods described in these earlier works, to design a set of experiments to distinguish between different models of protein-surface binding.

In this paper, we consider three models of reversible-irreversible binding behavior: a transition model, which states that upon adsorption, conformational changes occur such that the affinity or reversibility of the adsorption changes over time, a a two-state model with changing adsorbate footprint size over time, and a multiple-site transition model, in which multiple adsorption sites exist on the surface, with correspondingly different adsorption rates. Upon adsorption, a protein can either desorb or adsorb irreversibly. See Figure 2.1 for a summary.

Transition model

We begin with the transition model, discussd in detail in [33] and [32], illustrated in Figure 2.1A. Using this model and associated reaction rate constants, a transient mass balance equations can be written for each species adsorbed to the surface.

$$\frac{d\Gamma_E}{dt} = k_a[E_{bulk}]\Gamma_0 - k_d\Gamma_E - k_i\Gamma_E \tag{2.2.1}$$

$$\Gamma_{max} = \Gamma_0 + \Gamma_E + \Gamma_i \tag{2.2.2}$$

$$\frac{d\Gamma_i}{dt} = k_i\Gamma_E \tag{2.2.3}$$

where Γ_E and Γ_i are the mass-based surface concentrations of reversibly and irreversibly bound enzyme respectively, Γ_0 is the concentration of free sites on the surface, Γ_{max} is the maximum surface concentration, k_a, k_d, and k_i are the adsorption, desorption, and irreversible adsorption rate constants respectively, and $[E]_{bulk}$ is the bulk enzyme concentration.

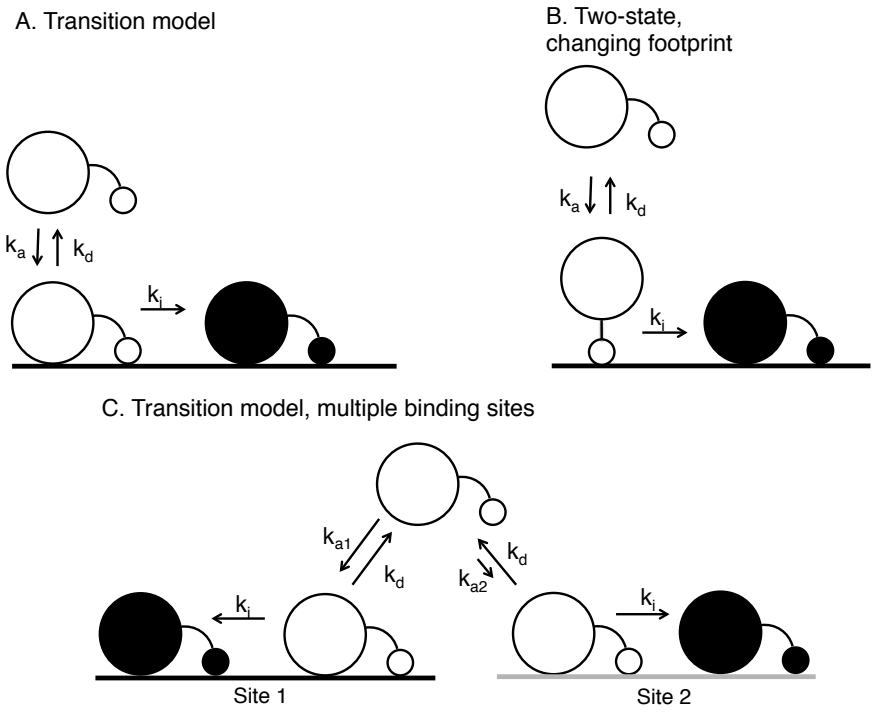

Figure 2.1: Models of protein adsorption, desorption, and irreversible adsorption considered. A. Transition model. Cel7B adsorbs reversibly to the surface, and then undergoes an irreversible reaction on the surface, rendering it adsorbed permanently. B. Two-state, changing footprint model. Cel7B adsorbs to the surface with a small footprint (perhaps via the CBM) and undergoes an irreversible reaction that increases its footprint on the surface and renders it irreversibly adsorbed. C. Two-site transition model. Cel7B adsorbs reversibly with two different adsorption rates to two distinct surface sites. Once bound to the surface, the protein can desorb or undergo irreversible adsorption, as in the transition model. This simplest two-site model does not include multiple desorption or irreversible adsorption rates.

To measure the adsorption rate constant k_a, Equation (2.2.1) is evaluated when $t \approx 0$, and the surface concentrations Γ_E and Γ_i are very small. Equation (2.2.1) simplifies to:

$$\left.\frac{d\Gamma_E}{dt}\right|_{t\approx 0} = k_a[E]_{bulk}\Gamma_{max} \tag{2.2.4}$$

By varying $[E]_{bulk}$ and measuring the initial rate of the resulting surface concentration data, a linear relationship between initial binding rate and enzyme concentration is produced; the slope of this line intersects the y-axis at zero and is equal $k_a\Gamma_{max}$.

In order to measure the kinetic rate constants for desorption and irreversible adsorption, enzyme is flowed across the surface for a fixed amount of time, referred to here as contact time (t_c). The enzyme solution is then replaced with a buffer solution, and washoff commences. We define a new time parameter, t_w, and set it equal to zero as washoff begins. As described

in detail in [32], the mass balance equations for reversibly and irreversibly adsorbed enzyme are written for the washoff conditions, i.e. ($[E_{bulk}] = 0$, and each is integrated with respect to t_w. The sum of these two integrals is equal to the total surface concentration as a function of t_w. We evaluate this integral between $t_w = 0$ and $t_w = \infty$ to get a total surface concentration as a function of washoff time and the lumped kinetic parameter ($k_d + k_i$), as show below in Equation (2.2.5).

$$\frac{\Gamma_T(t_w) - \Gamma_{T,\infty}}{\Gamma_{T,0} - \Gamma_{T,\infty}} = e^{-(k_d+k_i)t_w} \tag{2.2.5}$$

The two lumped rate constants ($k_d + k_i$) cannot be individually solved analytically; instead, we numerically solve the mass balance equations, together with the washoff data that specifies the surface concentration at the end of washoff for each contact time (t_c). By minimizing the sum of square error between the model's prediction and the data collected, solve for k_d and k_i (see Figure 2.13B).

Two-state, changing footprint model

In the two-state, changing footprint model (Figure 2.1B), enzyme adsorbs reversibly to the surface and undergoes a reaction that changes the size of the footprint of the enzyme on the surface and renders it irreversibly bound [36, 37]. In this model, Γ_{max} is again equal to the maximum surface capacity of the substrate, on a mass basis. If we use the same naming conventions as in the transition model, reversibly and irreversibly bound enzyme have a footprint areas A_E and A_i respectively. The transient mass balance equations are identical to Equations (2.2.1) and (2.2.3), but the site balance is represented by Equation (2.2.6).

$$\Gamma_0 = \Gamma_{max} - a\Theta_E\Gamma_{max} - \Theta_i\Gamma_{max} \tag{2.2.6}$$

where Θ_E and Θ_i are the mass fraction of enzyme bound in each state (Γ_E/Γ_{max} and Γ_i/Γ_{max} respectively, and a is equal to the relative footprint area of the reversibly and irreversibly bound species ($a = A_E/A_i$).

Equation (2.2.6) shows that the available surface capacity is dependent on the proportion of enzyme bound in each state. As the binding reaction progresses, more and more enzyme is bound in an irreversible, large-footprint state and the overall capacity decreases. When the bulk enzyme concentration is low, surface-bound enzyme will transition to the large-footprint, irreversibly bound state and decrease the overall surface capacity. When the bulk enzyme concentration is high, a higher surface capacity will be achieved, as more enzyme can bind in the small-footprint configuration.

Two-site transition model

The two-site transition model shown in Figure 2.1C introduces a second adsorption rate and a second distinct binding site, but the desorption and irreversible rates are assumed to be

equal between the two sites. The mass balance equation on reversibly bound enzyme then becomes:

$$\frac{d\Gamma_E}{dt} = k_{a1}[E_{bulk}]\Gamma_{0,1} + k_{a2}[E_{bulk}]\Gamma_{0,2} - k_d\Gamma_E - k_i\Gamma_E \tag{2.2.7}$$

where k_{a1}, k_{a2}, $\Gamma_{0,1}$, and $\Gamma_{0,2}$ are adsorption rate constants and available binding sites of each respective affinity.

$$\Gamma_{max} = \Gamma_{0,1} + \Gamma_{0,2} + \Gamma_E + \Gamma_i \tag{2.2.8}$$

$$\Gamma_{max} = \Gamma_{max,1} + \Gamma_{max,2} \tag{2.2.9}$$

The evaluation of k_d and k_i remains same as the simple transition model, as no adsorption takes place during washoff.

2.3 Materials and methods

Materials

All reagents were analytical grade (Sigma). Acid-pretreated *Miscanthus* was produced by two-step dilute acid pretreatment on the pilot scale by Andritz, Glens Falls, NY. The first step used 0.5% sulfuric acid, 158°C for 20 minutes; the second step used 1% sulfuric acid at 180°C for 4 minutes. Solids were washed with water to neutral pH and lyophilized before lignin extraction.

Lignin isolation

For our purposes, an ideal lignin would reflect the functionality of the lignin present in pretreated *Miscanthus*. Lignin was isolated from a stock of *Miscanthus* pretreated in a 2-step dilute acid process at Andritz, Glen Falls, NY. The pretreatment process was as follows. The first step (2 runs) used 0.5% sulfuric acid, 158°C for 20 minutes. In the second step, the combined material from the first run was further treated with 1% sulfuric acid, 180°C for 4 minutes. From this pretreated material, lignin was isolated using a modification of the milled wood lignin technique, which is a dioxane-based extraction under mild conditions.

Milled wood lignin is an isolated lignin free of carbohydrates, protein contamination, and insoluble material, and is chemically similar to lignin found in situ [17, 18, 19]. The lignin isolation procedure is adapted from [18, 19]. The pretreated biomass was rinsed and lyophilized, followed by ball milling for 5 minutes (Kleco ball mill). Lignin was then extracted from biomass with 96% dioxane for 24 hours. The extract was filtered (Whatman #1 filter paper) and the solvent removed by rotary evaporation at 50°C.

The residue was dissolved in a 9:1:4 volume ratio of pyridine/glacial acetic acid/water solution at room temperature (2 mL pyridine solution per gram biomass), followed by the

addition of chloroform (5 mL per gram biomass). The phase-separated mixture settled overnight, and the lignin-rich bottom phase was filtered through Whatman #1 filter paper and evaporated at 40° C. The volume was reduced to approximately 0.5 mL per gram biomass, and was precipitated into a vigorously stirred solution of acidified water (5 mM H_2SO_4). The precipitate was collected, rinsed with deionized water to neutral pH, and lyophilized.

Analysis of chemical composition and functionality of isolated lignin

Compositional analysis of isolated lignin

Compositional analysis of pretreated biomass and extracted lignin was performed as described in [38]. Polysaccharide content was determined by acid hydrolysis of cellulose and hemicellulose; glucose is formed by acid hydrolysis of cellulose and hemicellulose hydrolysis results in soluble xylose and arabinose. To hydrolyze sample, 50-200 mg solid biomass or lignin were weighed into a text tube, to which 0.5 mL 72% sulfuric acid was added. After 1 hour, 14.0 mL ultrapure water were added, vials capped, and the samples were autoclaved for 1 hour. Clear supernatant was filtered (0.45μm). Sugar concentrations were determined by Dionex HPLC. A CarboPac PA20, using a 150 mm by 3 mm column equipped with a guard column and amperometric detector. NaOH solution (18 mM) was used as the mobile phase at 0.4 mL/min at 308°C. The injection volume was 20 mL with a run time of 25 min. Mixed sugar standards were used for quantification of glucose, xylose, galactose, arabinose, and mannose in the samples. Galactose and mannose, which make up 1-2% of *Miscanthus* hemicellulose sugars, coelute with xylose and arabinose. Because their response factors are very similar, the error in calculating hemicellulose content is insignificant.

Functional analysis of lignin structure using 2D nuclear magnetic resonance

2D-NMR analysis for the pretreated biomass was performed using the method previously reported [39]. In brief, ball milled acid pretreated *Miscanthus* (15 mg) was dissolved in 0.6 mL DMSO-d6 containing 10 μL of [Emim][OAc]-d13. In the case of the MWL, 15 mg was dissolved in 0.6 mL DMSO-d6. 2D NMR spectra were acquired on a Bruker AVANCE 600 MHz spectrometer equipped with a gradient 5-mm TXI 1H/13C/15N cryoprobe. The chemical shifts were referenced to the central DMSO solvent peak (δC 39.5 ppm, δH 2.49 ppm). 13C-1H correlation spectra (HSQC: heteronuclear single quantum coherence) were measured with a Bruker standard pulse sequence. The experiments had the following parameters: acquired from 10-0 ppm in F2 (1H) by using 2,048 data points for an acquisition time (AQ) of 170 ms, 210-0 ppm in F1 (13C) by using 256 increments (F1 acquisition time 4.0 ms) of 200 scans with a 1s interscan delay (D1). HSQC cross-peaks were assigned by combining the results and comparing them with the literature. Volume integration of contours in HSQC plots were calculated using Bruker?s TopSpin 3.1 software. For quantification of S/G distri-

butions, only the C2-H2 correlations from G units and the C2-H2/C6-H6 correlations from S units were used, and the G integrals were logically doubled. For rough estimation of the various inter-unit linkage types, the following well-resolved contours were integrated: Aα, Bα, Cα, and their percentages are expressed as a percentage of the total common inter-units (A-C). 2D NMR analysis was done in collaboration with Dr. Hagit Sorek, who carried out the experiments.

Lignin molecular weight determination by size exclusion chromatography

Gel permeation chromatography was used to approximate the molecular weight of isolated lignin. PL-GPC 50 from A Varian, Inc. with two MesoPore Columns (each 300mm length, 7.5mm diameter) and a UV-vis detector at 280nm were used. The flow rate was 1.0 mL/min; THF was the sample-containing mobile phase. The column temperature was 30°C. Polystyrene standards were used for calibration.

Preparation of lignin films

Prior to spin coating, gold Qsense sensors were plasma-cleaned (Harrick PDC-32G; Pleasantville, NY, USA) to remove surface organic residue. Next, the sensors were soaked in a 5:1:1 solution of water, 30% v/v ammonium hydroxide, and 30% w/v H_2O_2, rinsed with water, and finally plasma cleaned again. Where indicated, the sensors were soaked in a 0.1% v/v solution of polydiallyldimethylammonium chloride (PDADMAC), a cationic polymer that adsorbs to the gold surface and provides a charged anchoring layer.

Milled wood lignin from acid-pretreated *Miscanthus* was dissolved in a variety of solvents (0.75M ammonium hydroxide, tetrahydrofuran, and 1,4 dioxane (96%). Lignin concentration varied from $1 - 3\%$ w/v, and the solutions were spin cast (Laurel Technologies WS-400B-6NPP/LITE; North Wales, PA, USA) at 3000 rpm for 1 minute.

Characterization of lignin films

Surface roughness characterization by atomic force microscopy

Surface height and phase information was taken with an MFP-3D atomic force microscope (Asylum Research, Santa Barbara, CA) in tapping mode.

Film thickness

Film thickness was measured by taking frequency and dissipation measurements of the dry QCM sensors before and after spin coating. Frequency and dissipation measurements were taken at overtones 1-11; the 7th overtone was used for analysis. Minimal differences in frequency and dissipation were observed between each of the overtones, and the total dissipation was <5% of the frequency change. The Sauerbrey equation was used to calculate the thickness of the lignin layer. (See below for additional discussion of QCM-D measurements).

Adsorbed protein measurement by quartz crystal microgravimetry with dissipation monitoring (QCM-D)

A lignin-coated sensor crystal was placed in the flow cell of a quartz crystal microbalance with dissipation monitoring (QCM-D; Q-Sense E4, Vaštra Frölunda, Sweden). The sensor was equilibrated by flowing buffer (50 mM acetate buffer, pH 4.85) through the cell until the frequency stabilized (10-60 minutes). Following equilibration, purified cellulase enzyme in the same buffer solution was introduced and the frequency change tracked. Following enzyme adsorption, buffer was again introduced to wash off reversibly adsorbed enzymes. For a sufficiently rigid adsorbed layer, frequency shift is proportional to adsorbed mass, as described by the Sauerbrey equation:

$$\Delta f = \frac{2f_0^2}{A\sqrt{\rho_q \mu_q}} \Delta m \tag{2.3.1}$$

where Δf is the frequency shift measured by the QCM; Δm is the change in resonating mass associated with the sensor surface; f_0 is the resonance frequency of the quartz crystal (5 MHz); A is the sensor surface area (1.539 cm^2), ρ_q is the density of quartz (2.648 g/cm^3), and μ_q is the shear modulus of quartz for AT cut crystal (29.47 GPa).

QCM frequency data was collected for the fundamental frequency of the crystal vibration, as well as the 3rd, 5th, 7th, 9th, and 11th overtones. The Sauerbrey equation is valid when the dissipation is low compared to the frequency shift (<5%) and the frequency changes of each overtone are the same. These requirements are met for the measurements taken in this work, thus the Sauerbrey equation is used to calculate surface mass. When adlayers become very thick or are highly viscoelastic, dissipation forces become more significant and higher order modeling of the layer is required; this was not necessary for this work.

Evaluation of mass transfer limitations in quartz crystal microgravimetry adsorption rate measurements

In order to ensure that the measured rates are not limited by either convective or diffusive mass transfer limitations, several QCM-D measurements were taken as described in the previous section. First, the concentration of analyte was decreased until the linear portion of the adsorption curve was an order of magnitude longer than the residence time in the cell ($t_r = V_c/v$), where V_c is the volume of the flow cell and v is the volumetric flow rate. This ensures that the concentration is stable in the flow cell during the initial rate measurement.

To evaluate the effect of the diffusion limitation, the flow rate is varied while the analyte concentration is held constant. As the flow rate is increased, the thickness of the diffusion layer decreases, and the adsorption rate dominates.

Molar concentration of adsorbed cellulase film

In order to measure the amount of enzyme present in adsorbed cellulase films, the enzyme was bound and eluted from a quaternary amine-functionalized thiol self-assembled monolayer (SAM). The eluted protein was then quantified by activity measurement on a fluorescent soluble substrate, and compared to the mass detected by QCM-D.

The quaternary amine-functionalized thiol $HS(CH_2)_{11}NMe_3^+$ was synthesized as described in [40]. Identity and purity of resulting substance was confirmed by NMR. The self assembled monolayer was formed from a 1 mM solution of $HS(CH_2)_{11}NMe_3^+$ (Q-thiol) in 100% ethanol. QCM-D sensor crystals (bare gold) were soaked in the Q-thiol solution for 18 hours, rinsed with ethanol followed by water, and loaded into the QCM for equilibration.

Data taken for quantification of adsorbed water were taken at 15°C with a flow rate of 200 uL/min. The sensors were equilibrated in 50 mM acetate buffer pH 4.85 until a stable baseline was recorded (10-30 minutes). *Hj*Cel7B was flowed across the sensors, followed by washing with equilibration buffer to remove residual Cel7A from the flow cells. After several minutes of washing, the adsorbed Cel7A was eluted with 50 mM acetic acid, pH 4.0, with 100 mM NaCl. Fractions were collected at the outlet of each chamber. When elution was complete, equilibration buffer was once again flowed across the sensor surface. Because buffer pH, ionic strength, and composition affects the frequency and dissipation of the sensor vibration, the re-equilibration allows a final adsorbed mass after elution to be measured.

Active Cel7B present in the eluate fractions was quantified by testing the activity on the soluble cellulase substrate 4−methylumbelliferyl cellobioside. Eluate fractions were tested by adding 10 uL eluate to 90 uL of 1 mM substrate in 50 mM acetate buffer, pH 4.85. The plate was sealed and incubated in a thermal cycler at 50°C. After incubation, 100 uL of 1.0 N NaOH was added, and the fluorescence was measured at 365/445 nm on a Spectramax M2 plate reader (Molecular Devices, Sunnyvale, CA). Fluorescence was correlated with cellulase concentration by performing a calibration on each plate read.

Protein expression

Cel7B was recombinately expressed in its native host, *Trichoderma reesei*, as described previously [41]. Cel7B was purified from the culture supernatant by ammonium sulfate precipitation (65% saturation, on ice). The precipitated Cel7B was centrifuged and resuspended into 20 mM tris buffer, pH 7.0, desalted by 10 kDa spin filtration, and loaded onto a Q sepharose FF column and eluted with a gradient elution of 100 mM NaCl in tris buffer, pH 7.0. Cel7B-containing fractions were pooled and buffer exchanged to acetate buffer, pH 4.85.

2.4 Results

Analysis of chemical composition and functionality of isolated lignin

Compositional analysis

Compositional analysis was performed on dried *Miscanthus*, acid-pretreated *Miscanthus*, and milled wood lignin extracted from the acid-pretreated material.

Table 2.1: Compositional analysis of *Miscanthus* dry biomass, acid-pretreated *Miscanthus*, and lignin extracted from the pretreated biomass with the milled wood lignin process described in this work. One standard deviation is presented where available.

Material	Glucan (%)	Xylan (%)	Arabinan (%)	Acetyl (%)	Klason lignin (%)	Ash (%)
Dry *Miscanthus*	40.0	24.0	NA	NA	25.9	3.5*
Acid-pretreated biomass	55.6 ± 0.179	0.54 ± 0.015	0.11 ± 0.008	0.11 ± 0.002	33.6 ± 0.133	9.73 ± 0.041
Milled wood lignin	0.08 ± 0.025	0.034 ± 0.002	0.10 ± 0.006	0.16 ± 0.006	91.0 ± 0.58	ND

*Additional unaccounted-for components are extractables removed before analysis.

The lack of carbohydrate impurities in the purified lignin stream is notable. In each case, the small amount of detected carbohydrate is roughly equal to the limit of quantification of the HPLC detection method. The lack of carbohydrate in the samples means that carbohydrate-binding behavior will not be confused with lignin-binding behavior when quantifying the lignin-binding kinetics of cellulase enzymes.

Structural comparison of isolated milled wood lignin with acid-pretreated *Miscanthus* by 2D-NMR

To obtain the structural features of the isolated milled wood lignin we used 2D HSQC NMR. This method allowed us to obtain a comprehensive structural characterization of the lignin, used in our experiments, and also to compare the isolated lignin structural features to the source lignin in the acid pretreated *Miscanthus*. The lignin structural features are apparent in two regions: the aliphatic-oxygenated (δC/?H 50-90/2.5-6.0 ppm) and aromatic 13C-1H correlations (δC/δH 100-150/6-8 ppm) regions. NMR spectra of the acid pretreated *Miscanthus* and its isolated MWL are shown in Figure 2.2. The structural type of inter-unit linkage patterns of lignin can be found in the aliphatic side chain region. In the case of the acid pretreated *Miscanthus*, polysaccharide signals dominated the spectrum, and partially overlapped with some lignin signals. However, at least one of the correlations for each of

the structures in lignin is well isolated. On the other hand, the spectrum of the MWL presented mostly lignin signals that, in general terms, matched those observed in the HSQC spectrum of the pretreated *Miscanthus*, but with improved detection of the more minor lignin structures.

The aliphatic-oxygenated region of the spectra gave information about the different inter-unit linkages present in the lignin. In this region, correlation peaks were observed from methoxyl groups, β-ether (β-O-4) units (A), phenylcoumaran (β-5) units (B) and resinol (β-β) units (C). In general, substructures were more clearly visible in the HSQC spectrum of the MWL since the sample is free from carbohydrate impurities. The relative abundances of the main lignin inter-unit linkages, estimated from volume integration of contours in the HSQC spectra, show similarity between the isolated lignin and the pretreated biomass. The data indicated that the structures of both lignins consist of $\sim 60\%$ β-O-4 linkages, $\sim 30 - 35\%$ phenylcoumaran and a lower amount of resinol. The main correlation peaks in the aromatic region of the HSQC spectra (Figure 2.2) corresponded to the aromatic rings and unsaturated side chains of the different lignin units and hydroxycinnamates. This region emphasizes the differences in the guaiacyl:syringyl (G:S) distribution in the lignin polymer. Signals from guaiacyl (G), and syringyl (S) units were observed almost equivalently in the spectra of the pretreated Miscanthus and in the isolated MWL, due to this region being purely derived from such structures (and completely free of the polysaccharide correlations). The signals revealed some heterogeneity among the G and S units that affected the correlation, possibly due to different substituents (phenolic or etherified in different substructures), though when comparing the aromatic region with the whole cell wall NMR of miscanthus1, it is apparent that those structures were formed during the pretreatment and presumably indicate on condensed lignin. The S/G ratio estimated from the HSQC (S/G 0.5) is similar for both MWL and pretreated *Miscanthus*, although signals of H-lignin units were not detected in the HSQC spectrum of the isolated MWL. Accordingly, it could be concluded that the common inter-unit linkages were still preserved in the isolated lignin and that the isolation process simply removed the carbohydrates contamination and did not alter the lignin structure.

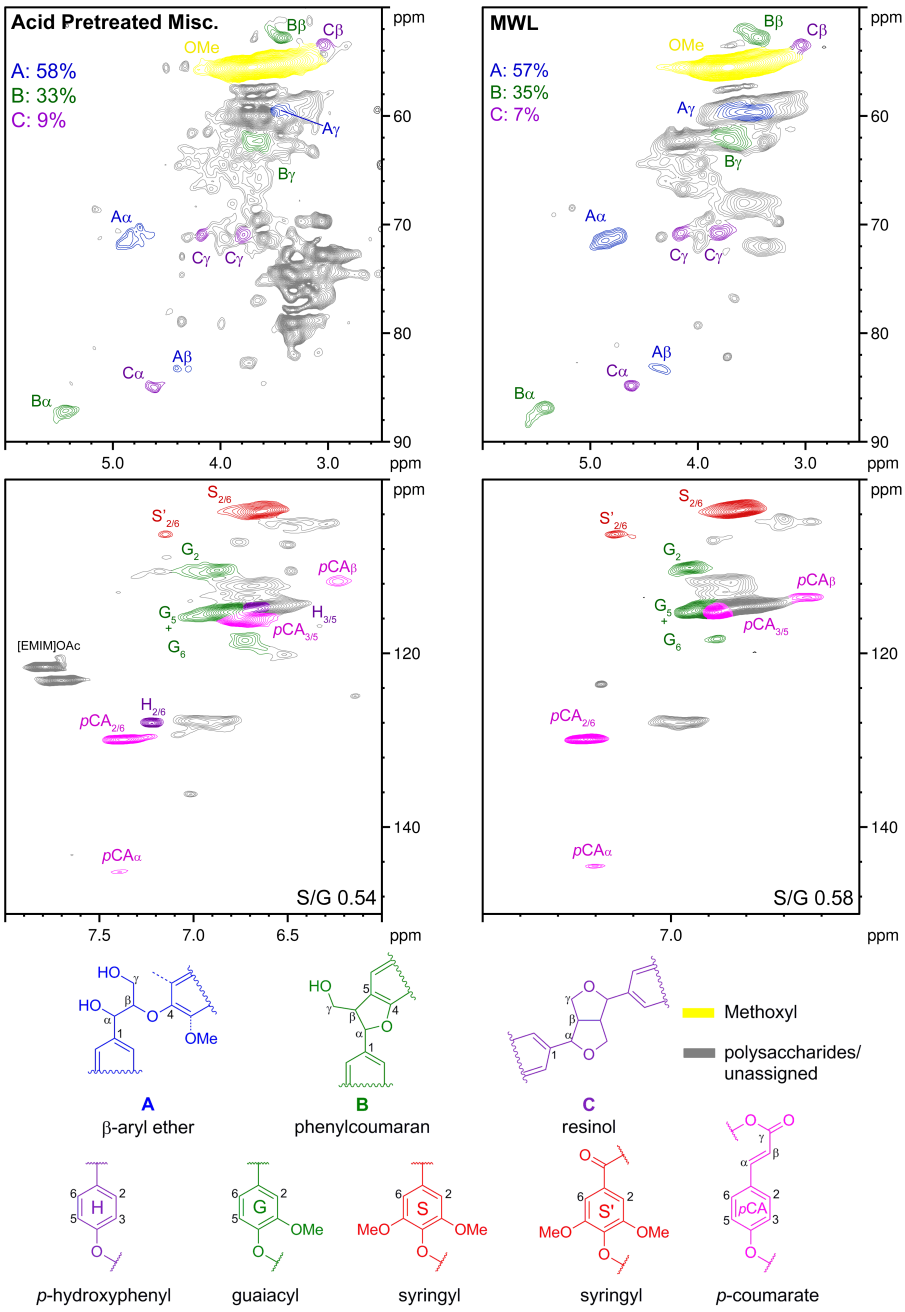

Figure 2.2: Side chain (δC/δH 50-90/2.5-6.0) and aromatic (δC/δH 90-150/5.5-8.0) regions in the 2D HSQC NMR spectra of pretreated *Miscanthus* (left) and of the isolated MWL (right). Color-coded structures correspond to the major resonances in the spectra.

Lignin molecular weight profile by GPC

The molecular weight of isolated milled wood lignin from acid-pretreated *Miscanthus* was determined by gel permeation chromatography with polystyrene molecular weight standards. The weight average molecular weight (Mw) was 2880 g/mole; this is smaller than some reports for milled wood lignin [42], likely due to fragmentation that occurs during pretreatment.

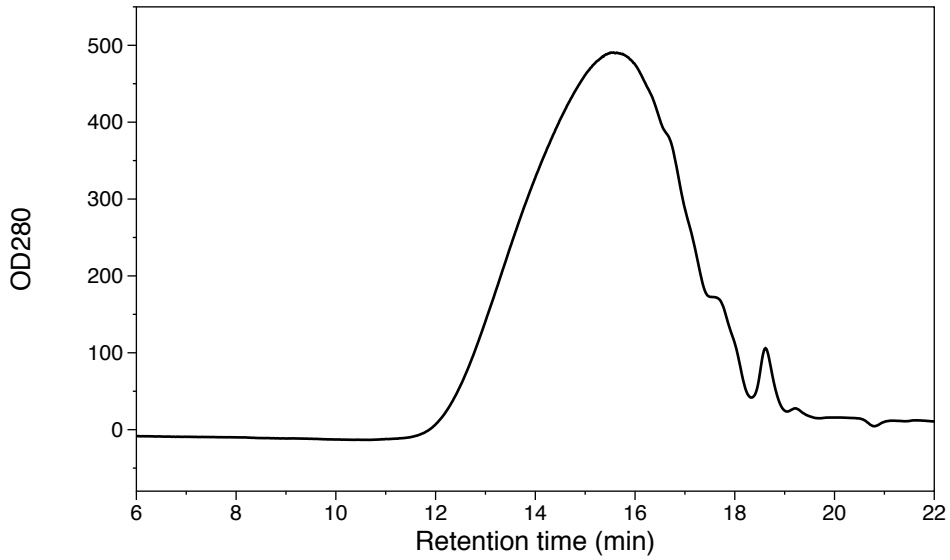

Figure 2.3: Milled wood lignin molecular weight profile measurement by gel permeation chromatography

Lignin film thickness and surface morphology

Lignin film thickness was measured using QCM-D as described above. Frequency and dissipation measurements were taken before and after spin casting, and the change in frequency was compared. Frequency shifts across the third through 11th overtone were within 5% of each other before and after spin coating, and the dissipation shifter upon the addition of the lignin layer was also small compared to the frequency shift. Because these criteria for use of the Sauerbrey equation were met, the frequency shift (7th overtone) was converted to surface mass. Using an assumed lignin density of 1.3 g/L [43], the film thickness was calculated.

At 1.0% w/v, the lignin film thickness ranges from 70-80 nm. This is thick enough to ensure uniform and complete coverage of the substrate, and thin enough to allow for sensitive mass measurements on the surface of the film as proteins adsorb. All subsequent measurements were taken using films cast from 1% lignin dissolved in THF.

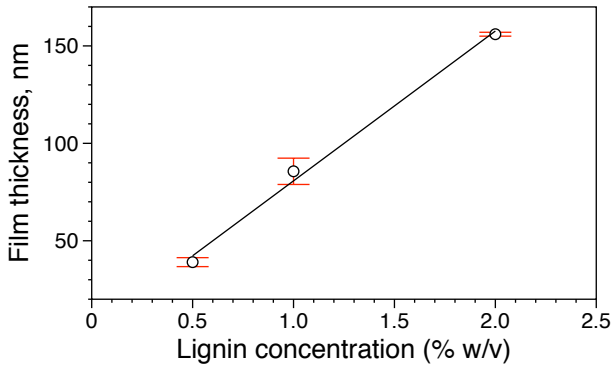

Figure 2.4: Film thickness (nm) of spin-cast lignin films. The solvent is THF.

The lignin films produced when spin coated from a variety of solvents were analyzed by atomic force microscopy. Previous reports suggested that dilute ammonium hydroxide (0.5 − 3.2M NH_4OH) is a suitable solvent for producing lignin films [20], however, dynamic light scattering data indicated that milled wood lignin from acid pretreated biomass was not fully soluble in this solvent. No aggregation was detected in either THF or 96% dioxane. Films case from lignin dissolved in either THF or 96% dioxane, with and without an anchoring layer of adsorbed polydiallyldimethylammonium chloride are presented in Figures 2.5 - 2.9.

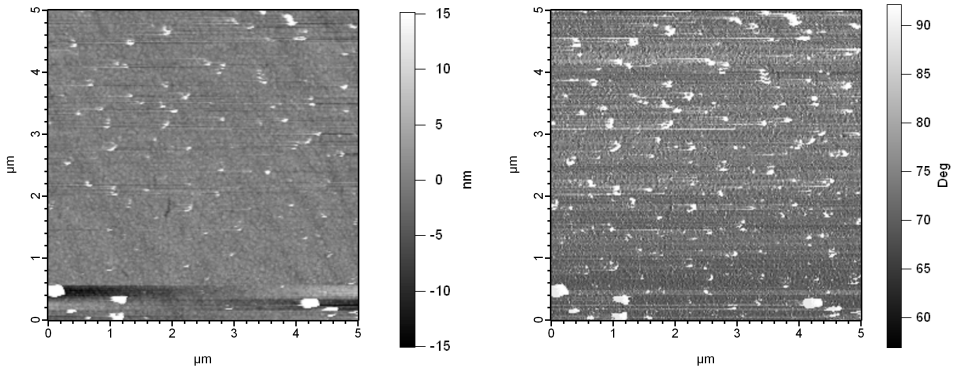

Figure 2.5: Atomic force microscopy height and phase images of bare gold QCM-D sensor crystals.

While 96% dioxane is an excellent solvent for milled wood lignin (it is, after all, the solvent by which the lignin was extracted from biomass), films cast from this solvent exhibit micron-scale heterogeneity in both height and phase data, with features protruding up from the surface. Films cast from THF onto a bare gold substrate look much different: the heterogeneities present are in the form of holes in the surface. The imaging technique does not allow us to be sure whether the bare gold is exposed at the bottom of these holes.

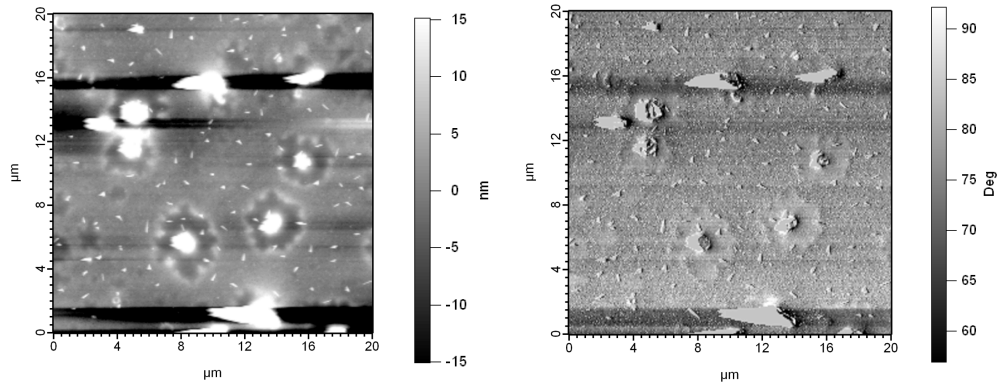

Figure 2.6: Atomic force microscopy height and phase images of QCM-D sensor crystals spin coated with lignin, dissolved in 1,4 dioxane.

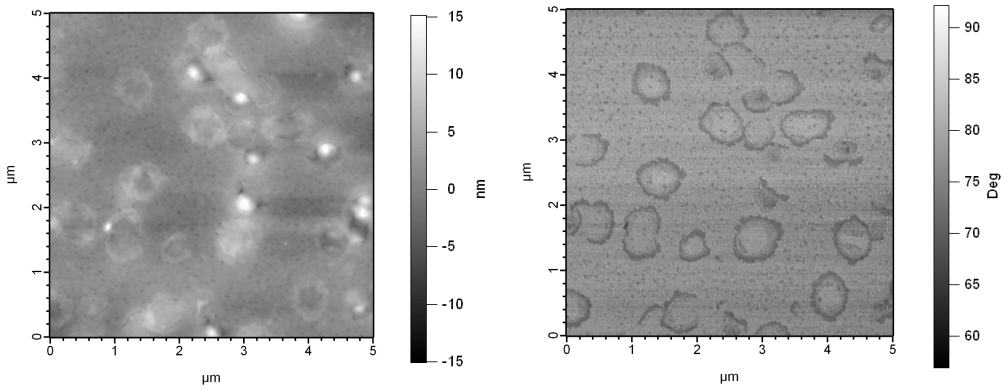

Figure 2.7: Atomic force microscopy height and phase images of QCM-D sensor crystals, spin coated with lignin, dissolved in 1,4 dioxane, PDADMAC anchor layer.

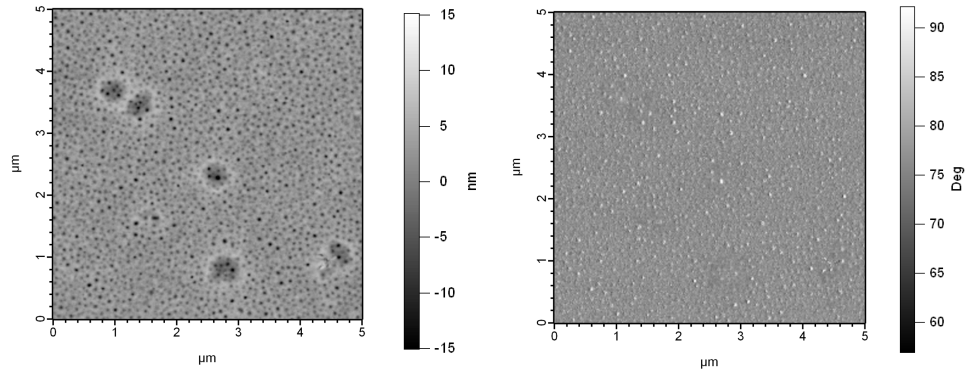

Figure 2.8: Atomic force microscopy height and phase images of QCM-D sensor crystals, spin coated with lignin, dissolved in tetrahydrofuran.

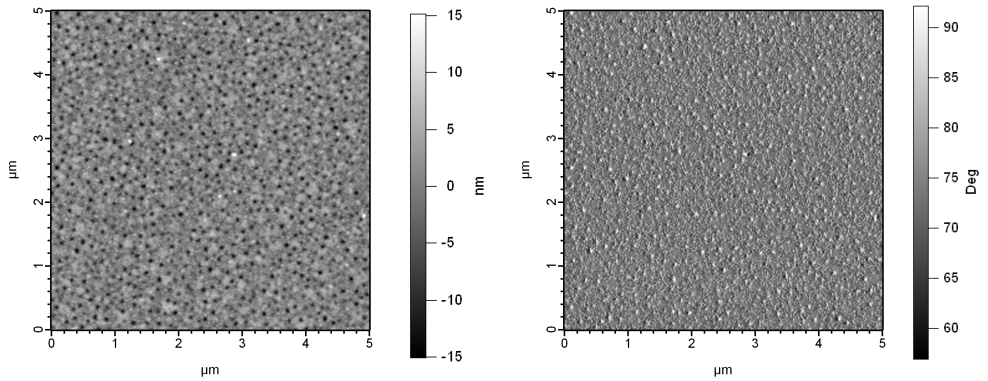

Figure 2.9: Atomic force microscopy height and phase images of QCM-D sensor crystals, spin coated with lignin, dissolved in tetrahydrofuran, with PDADMAC anchor layer.

Addition of the cationic anchoring layer polyDADMAC was done in the hopes of improving the adhesion of the lignin and the gold substrate during spin coating. Films cast from dioxane, while exhibiting less extreme roughness, continued to show significant heterogeneity in the phase measurement, indicating some micron-scale ordering. When THF was used to cast the films, the surface was much smoother and without either holes or roughness seen in the non-anchored films. While the surface does exhibit some small variation in height and phase, it is the smoothest and most homogeneous surface cast.

Mass transfer limitations of adsorption rate measurement of *Hj*Cel7B

Convection and diffusion limitations were tested by adjusting the flow rate and analyte concentration such that the initial rate measured by the QCM-D apparatus was insensitive to the flow rate of analyze through the flow cell.

Convection limitations

The QCM-D flow cell has a volume of 40μL above the sensor cell. When the flowing solution changes from buffer to a protein solution, the liquid residence time t_r is the amount of time it takes to fill the flow cell with the new analyte solution. If the protein concentration is high, the sensor surface is saturated before the flow cell concentration has stabilized. Figure 2.10 and Table 2.2 demonstrate the presence of a convection limitation under high analyte concentrations. When the concentration of quickly-adsorbing enzyme is high, the initial rate portion of the mass profile occurs only during the amount of time it takes the enzyme solution to fill the flow cell.

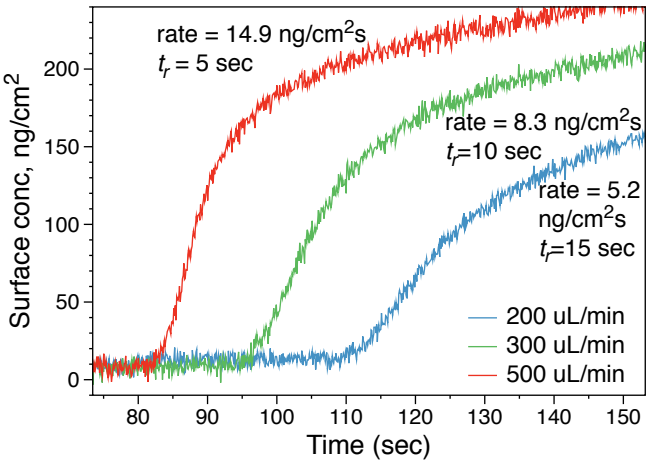

Figure 2.10: High concentration, low flow rate leads to initial rate measurements that are limited by the residence time in the flow cell (noted on figure).

Table 2.2: At high concentrations, the initial rate measurements that are limited by the residence time in the flow cell. Each of the three adsorption curves were produced at the same analyte concentration, and in each case, the initial rate duration is approximately equal to t_r. The concentration in the flow cell is not stable during initial rate measurement, and is convection limited.

Concentration (μM)	Flow rate (μL/min)	Residence time t_r (s)	Observed rate (ng/cm^{2}s)	Linear rate duration (s)
10	200	15	5.2	12
10	300	10	8.3	8
10	500	5	14.9	5

Diffusion limitiation

When analyte concentration is reduced several orders of magnitude, the observed initial rate lasts for minimally several times the residence time (t_r) of liquid in the cell. However, when the concentration is decreased, the potential for a diffusion limitation grows. The Damköhler number is a dimensionless number that can be used to compare the relative rates of a reaction (in this case, the surface reaction wherein enzyme binds to the surface) and diffusion (through the liquid boundary layer at the surface from which analyte is rapidly depleted). For the QCM-D system measuring typical enzyme binding rate constants at recommended flow rates, Da on the order of 1 have been estimated [24, 25], meaning that the system is neither fully mass-transfer limited, nor reaction limited, but significant concentration gradients are likely to exist, dependent upon the flow rate, analyte concentration, diffusion rate, and intrinsic reaction kinetics. As is shown in Chapter 3, the adsorption rate constant for Cel7A is greater

than that for Cel7B. This is reflected in the diffusion limitation as well, as the transition from diffusion limited to adsorption rate limited occurs at a faster flow rate for Cel7A.

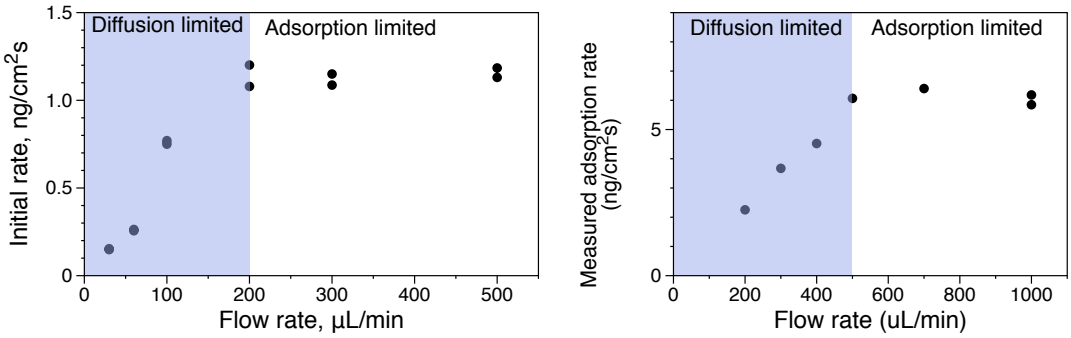

Figure 2.11: Empirical evaluation of diffusion limitation in measurement of adsorption rate, for Cel7B (left) and Cel7A (right). Diffusion limitations are more dramatic as the adsorption rate constant increases ($k_{A,Cel7A} > k_{A,Cel7B}$).

Hydration measurement of adsorbed Cel7B film

The QCM-D technique is based upon the change in mechanical oscillation of a piezoelectric crystal, caused by surface adsorption or structural changes in adsorbed layers. When QCM-D is applied in aqueous solutions, the change in oscillation frequency detected is a result of the total surface mass, including both biomolecules adsorbed to the surface, and water bound or kinetically coupled to the biomolecules [26, 44]. This is different from spectral techniques, such as surface plasmon resonance and ellipseometry, in which coupled water is not detected. This presents a problem for experimentalists who seek to use QCM-D to measure the kinetics of surface adsorption: the adsorbed biomolecular mass must be decoupled from the hydrated mass of the protein and coupled water.

In the past, this has been done by taking simultaneous QCM-D and SPR measurements, and correlating their measurements. In the absence of custom instrumentation as used previously [27], we seek to decouple adsorbed protein and adsorbed water using only the QCM-D instrument, by binding and eluting Cel7B to a surface, collecting the elute, and comparing the eluted mass with the mass tracked by the QCM-D. During the elution, the QCM-detected mass decreases as expected, and when the eluate is assayed, Cel7B is detected. However, the mass detected in the eluate solution is less than the mass detected by QCM-D; the ratio of the two masses is the ratio of the protein film that is not attributable to coupled water.

Cel7B was bound and eluted to a QCM-D sensor coated with a quaternary-amine functionalized monolayer. Data from seven QCM runs were collected and eluate fractions assayed. On average, 10.5% (standard deviation = 1.31%) of the QCM-reported mass is attributable to *Hj*Cel7B eluted from the surface. This is in reasonable agreement with published work using other methods to estimate the amount protein sensed by QCM-D [27].

Table 2.3: QCM and activity-based mass comparison of bound Cel7B

QCM eluate mass, ng	Active Cel7B in eluate, ng	Active mass/QCM mass
197.4	19.8	0.100
199.3	21.3	0.107
204.2	27.2	0.133
189.7	18.6	0.098
208.7	21.9	0.104
193.6	18.7	.096
202.4	19.3	.095
	Average	0.105
	Standard deviation	0.013

This method is not without shortcomings. The quaternary amine surface is clearly chemically different than a lignin film; the conformation and degree of hydration of cellulase adsorbed to this surface may also be different between the two surfaces. QCM-D offers the ability to track both the adsorbed mass via frequency (Δf) and viscoelastic properties of the adlayer via the dissipation measurement (ΔD). Previous work has demonstrated that conformational changes in adsorbed protein layers can be detected by changes in the slope of ΔD vs. Δf [45]. If we compare Cel7B adsorbed on a lignin surface with Cel7B adsorbed on the quaternary amine SAM, a plot of Δf vs. ΔD shows a similar linear trend between the two layers (slopes of 0.037 and 0.036, respectively). This is evidence that although the lignin and amine surfaces are functionally different, the conformation of the adsorbed protein is likely to be similar.

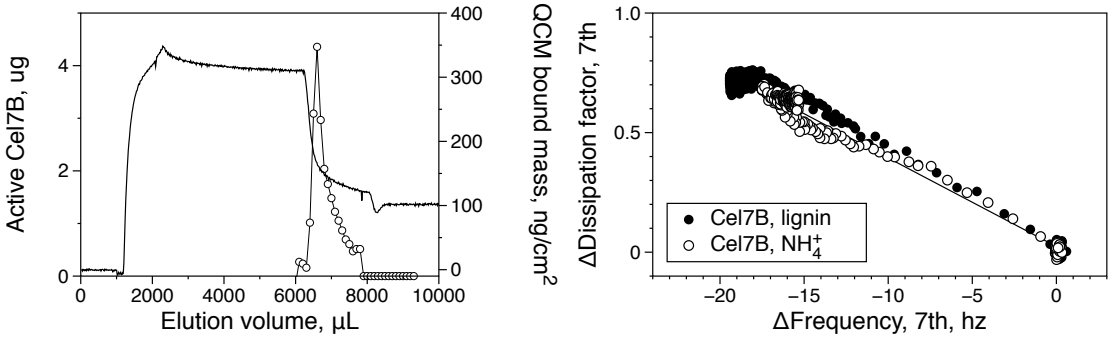

Figure 2.12: A. Binding and elution profile of Cel7B to quaternary amine based SAM. QCM-measured mass is depicted as a solid line, and active Cel7B detected in elution fractions are displayed as an open circle. B. ΔD vs. ΔF comparison: lignin surface (filled circles) and NH_4^+ SAM surface (open circles). The slopes of the resulting ΔF vs. ΔD lines are the same, providing support for the hypothesis that their bound conformation and degree of hydration is similar between the two surfaces.

Reaction kinetics

Transition model

After collecting adsorption and washoff data at varying bulk enzyme concentrations and contact times, we first tested the transition model, the simplest of the three models summarized in Figure 2.1.

Initial rate of cellulase adsorption to lignin were measured by varying Cel7B concentration between 0.05 μM to 0.30μM, and measuring initial adsorption rates. The slope of the resulting line is equal to the adsorption rate constant k_A times the maximum adsorptive capacity Γ_{max}, as seen in Figure 2.13 A.

In order to deconvolute k_D and k_i, a series of washoff histories were collected. Enzyme was flowed over the lignin-coated sensors for a set contact time, t_c, after which the sensors were rinsed with buffer to remove any reversibly-bound enzyme from the surface. A washoff summary, comprised of the fraction of bound enzyme that remained bound irreversibly as a function of contact time. From this washoff summary, combined with the mass balance equations, we numerically solved for the individual rate constants (see Figure 2.13B).

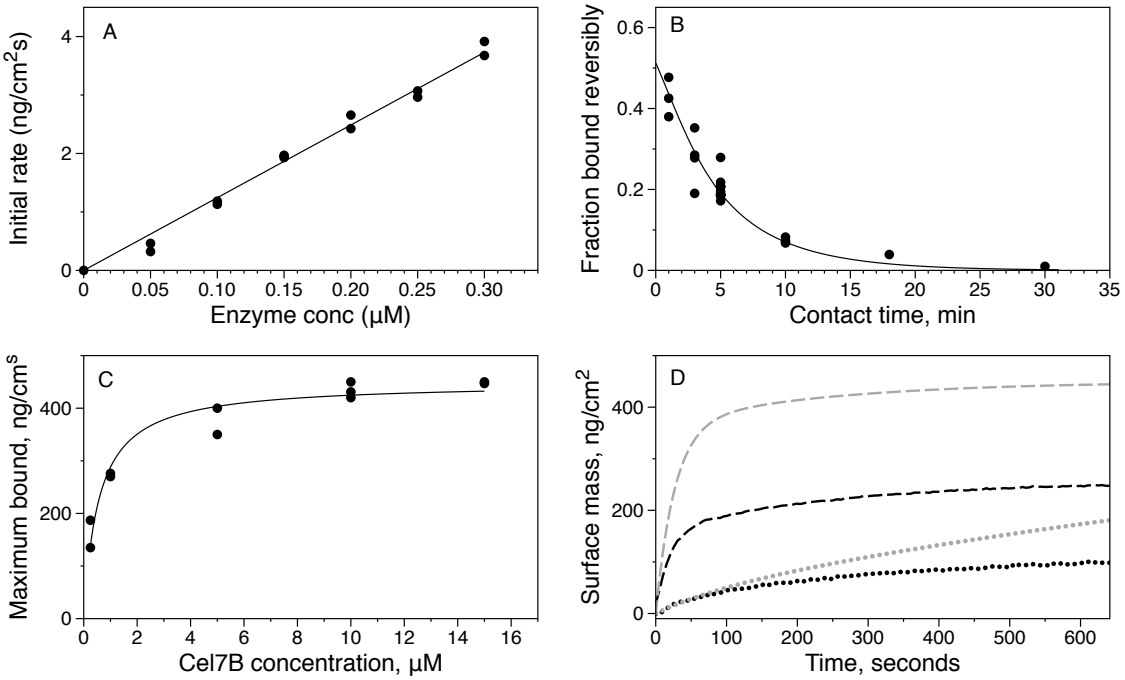

Figure 2.13: Evaluation of the transition model for Cel7B adsorption to lignin. A. Initial rate of Cel7B adsorption to milled wood lignin isolated from acid-pretreated *Miscanthus*. The slope of the fit line is equal to the kinetic parameter $k_A * \Gamma_{max} = 12.4$ ng/(cm^{2}s*μM). $R^2 = .99$. B. The kinetic parameters k_i and k_d are obtained as a lumped parameter by fitting the washoff curves to Equation (2.2.5). C. The kinetic parameter Γ_{max} is determined by adsorbing Cel7B to the surface at increasing concentrations for up to four hours. The maximum surface concentration after four hours is shown vs. Cel7B concentration, and reaches a maximum of 450 ng/cm^2. Line is drawn to guide the eye. D. When the measured Γ_{max} is included in the model (model shown in gray lines, experimental data in black), it fails to fit the binding behavior for either 0.05 uM Cel7B (dots) or 1.0 uM Cel7B (dashes). Only the initial adsorption rate is correctly reflected by the model.

The total number of available surface sites, Γ_{max}, was determined by adsorbing Cel7B to the sensor surface at increasing concentrations for up to four hours. This was done at 15°C to minimize instrument drift over long adsorption times. This analysis results in a maximum adsorptive capacity measured by QCM-D of 450 ng/cm^2, as shown in Figure 2.13C.

However, when this parameter is applied in the transition model and compared with experimental data at low and moderate concentrations (1-2 orders of magnitude lower than maximum concentrations used to obtain Γ_{max}), it is clear that the model does not fit the data (see Figure 2.13D) and results in drastically overestimated surface mass. The one-site transition model cannot explain the observed data.

Two-state, changing footprint model

In the two-state model, protein adsorbs to the surface with with a small footprint, and undergoes a surface rearrangement to a larger footprint. Proteins have been shown to rearrange on surfaces, such that they occupy more space on the surface [46, 47]. This hypothesis is intuitively plausible for cellulases, which bind to lignin via small carbohydrate binding modules [12] and may be able to undergo a second step by which the catalytic domain binds or unfolds on the surface.

We tested this hypothesis is tested by exposing the surface to a low concentration of Cel7B, followed by exposure to a higher concentration of Cel7B. When the surface concentration achieved by this two-step adsorption process is compared with a surface that was exposed only to the high Cel7B concentration, we see that the maximum capacity reached is indistinguishable (see Figure 2.4. This shows that the maximum capacity is independent of the rate of adsorption. Furthermore, a major change in surface structure would be expected to have an accompanying change in QCM dissipation factor. In fact, we see a linear relationship between Δf and ΔD, regardless of the kinetics of surface adsorption (data not shown).

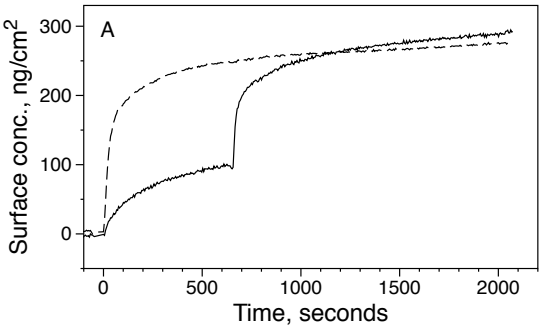

Figure 2.14: When the surface is exposed first to a low concentration (0.05 μM) of Cel7B, followed by a higher concentration (1.0 μM, solid line) the final surface concentration reached is virtually the same as when the surface is exposed to only the high concentration (dashed line). The two-step, changing footprint model fails to explain the observed kinetics.

Two-site transition model

We instead hypothesize that lignin, a heterogeneous polymer with a wide variety of chemical functionality, may display binding sites with varying affinity for Cel7B. If we consider a two-site transition model, in which the sum of two adsorption rates $k_{A1}\Gamma_{max1}$ and $k_{A2}\Gamma_{max2}$, is equal to the initial rate measured in Figure 2.13A. If we further assume that $k_{A1}\Gamma_{max1} \gg k_{A2}\Gamma_{max2}$, we can fit the parameter Γ_{max1} from low-concentration, early-time data series, because the binding due to $k_{A1}\Gamma_{max1}$ would dominate. Because $\Gamma_{max1} + \Gamma_{max2} = \Gamma_{max}$, we can solve for k_{A2} by minimizing the sum squared error when we fit the model to the

adsorption data at a concentration of 1 μM. We then iterate this solution to solve for k_d, k_i, and Γ_{max2}. The kinetic parameters used to model the system are summarized in Table 2.4.

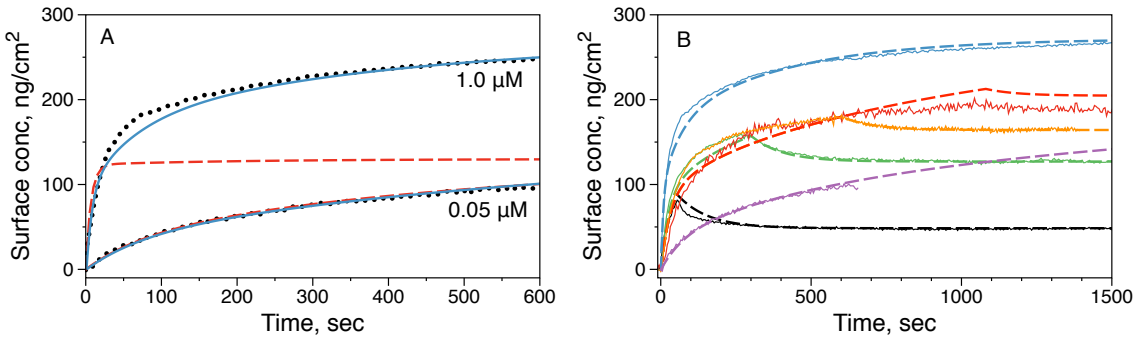

Figure 2.15: A. When the model is modified to include two simultaneous adsorption rates with two distinct surface sites, we can fit both high and low concentration data with the addition of the second rate. Black dots = QCM-D data, red dashed lines = one-site transition model, Γ_{max} = 130, blue line = two-site transition model. B. When the two-site transition model is applied to adsorption and desorption curves of varying concentrations and contact time, both adsorption and desorption can be modeled accurately. Dotted lines = model prediction, solid lines = empirical data. Purple = 0.05 μM; black = 0.25 μM, t_c = 1 min, green = 0.3 μM, t_c = 3 min, orange = 0.25 μM, t_c = 10 min, red = 0.25 μM, t_c = 18 min, blue = 1.0 μM Cel7A.

The value that we obtain for Γ_{max} when we fit to the 1 μM adsorption data is still considerably less than the maximum value determined in Figure 2.13 C. We hypothesize that rather than two discrete binding sites, there are in fact many, and as the bulk concentration increases, these lower-affinity sites are filled. It is also a possible that along with multiple adsorption rates, their are also multiple desorption and irreversible binding rates for each site; desorption and irreversible binding behavior was not explored at concentrations high enough to explore desorption of Cel7B from these lower-affinity sites.

2.5 Discussion

Our investigation of these three models for cellulase adsorption to lignin supports the hypothesis that Cel7B adsorption occurs via a multiple binding site transition model. This model provides a better fit to both high and low concentration adsorption and desorption data than either a transition model with a single binding site, or a transition model that incorporates a changing footprint size. Readers may note that this model includes two additional parameters, Γ_1 and k_2, vs. the single site transition model, and we should thus expect a better fit to experimental data. However, we also believe that the multiple-site transition model is inherently plausible for the interaction of Cel7B and lignin.

Table 2.4: Kinetics summary for 2-site transition model

Parameter	Value (± 1 standard deviation)	Units	R^2
$(k_{a1\|} + k_{a2})\Gamma_{max}$	13.0	ng/(cm^{2}s μM)	0.988 vs. initial rates
$k_{a1}\Gamma_{max1}$	11.55	ng/(cm^{2}s μM)	
$k_{a2}\Gamma_{max1}$	1.45	ng/(cm^{2}s μM)	
Γ_{max1}	98[a]	ng/cm^2	
Γ_{max2}	172[a]	ng/cm^2	
Γ_{max}, measured	440 (13.5)[a]	ng/cm^2	
$k_i + k_d$	0.0098 (0.0026)	s^{-1}	N/A
k_d	0.00515	s^{-1}	0.92 vs. washoff histories (Fig. 2.13B
ki	0.00467	s^{-1}	0.92 vs. washoff histories
QCM mass fraction Cel7B	10.5 ±1.33%	%	N/A

[a] QCM-measured mass; 10.5% of this mass is measured as active Cel7B.

Lignin is a chemically heterogeneous surface, composed of three different monomer building blocks, and is polymerized via a free-radical process to create lignin, linked to hemicellulose via ferrulic acid strucures. As such, there are a wide variety of chemical functionalities present in lignin, especially after pretreatment and isolation. Lignins rich in certain structural components have been observed, either directly or indirectly, to have a higher affinity for lignin. For example, the carboxylic acid content of isolated lignins is correlated with the digestibility of Avicel in the presence of isolated lignins [11]. Lower S/G ratios are also correlated with higher lignin adsorption [48]. Observations of cellulase adsorption to lignin at varying pH indicate that ionizable groups on the lignin surface may be responsible for a decrease in adsorption at increasing pH [49]; this provides more evidence for a heterogeneous surface that could display functionality with varying cellulase affinity.

Cellulase enzymes, and specifically the family 1 carbohydrate binding module that mediate *Hj*Cel7B interaction with cellulose, have been shown to have varying affinities for different crystal forms and faces of cellulose. A 2012 molecular dynamics simulation study suggested that a Family 1 CBM binds preferentially to the hydrophobic (110) surface of a cellulose microfibril, and that CBMs bound to either of the hydrophilic surfaces could diffuse to their preferential binding site [50]. A single-molecule fluorescence study reveals preferential binding of family 1 CBMs to crystalline cellulose surfaces (vs. amorphous surfaces) [51], though amorphous cellulose surfaces (from dissolved and precipitated cellulose, for example) have also shown to adsorb cellulases via Family 1 CBMs [32]. It's clear, therefore, that CBM1 has the capacity to bind different cellulase crystal structures with varying affinity, and it's reasonable to hypothesize that the same is true for lignin.

2.6 Conclusions

The adsorption, desorption, and irreversible binding of *Hj*Cel7B to lignin has been modeled and tested. Lignin from acid-pretreated biomass was isolated using a milled wood lignin purification protocol, and spin-cast into thin, homogeneous films suitable for surface-based assays such as quartz crystal microgravimetry with dissipation monitoring. We calibrated the QCM-D measurements to account for protein film hydration by binding, eluting, and quantifying Cel7B to the surface. Finally, we have investigated several models for Cel7B adsorption: a transition model, wherein reversibly adsorbed protein undergoes a conversion to an irreversibly bound state, a multi-state model in which the adsorbed footprint changes over time, and finally we find that the best fit for experimental data is a transition model with multiple adsorption rates. Experimental data was fit to a model with two adsorption rates, but evidence of additional lower-affinity sites exists, as higher surface concentrations are reached when protein concentrations are increased outside of the concentrations modeled in this work.

Bibliography

[1] V. S. Chang and M. T. Holtzapple. Fundamental factors affecting biomass enzymatic reactivity. *Appl Biochem Biotechnol*, 84-86:5–37, 2000.

[2] Valdeir Arantes and Jack N. Saddler. Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. *Biotechnol Biofuels*, 3:4, 2010.

[3] Joseph A. Rollin, Zhiguang Zhu, Noppadon Sathitsuksanoh, and Y-H Percival H. Zhang. Increasing cellulose accessibility is more important than removing lignin: a comparison of cellulose solvent-based lignocellulose fractionation and soaking in aqueous ammonia. *Biotechnol Bioeng*, 108(1):22–30, 1 2011.

[4] H. Ooshima, D. S. Burns, and A. O. Converse. Adsorption of cellulase from Trichoderma reesei on cellulose and lignacious residue in wood pretreated by dilute sulfuric acid with explosive decompression. *Biotechnol Bioeng*, 36(5):446–52, 8 1990.

[5] Alex Berlin, Mikhail Balakshin, Neil Gilkes, John Kadla, Vera Maximenko, Satoshi Kubo, and Jack Saddler. Inhibition of cellulase, xylanase and beta-glucosidase activities by softwood lignin preparations. *J Biotechnol*, 125(2):198–209, 9 2006.

[6] Seiji Nakagame, Richard P. Chandra, and Jack N. Saddler. *The Influence of Lignin on the Enzymatic Hydrolysis of Pretreated Biomass Substrates*, pages 145–167. 11 2011.

[7] A. O. Converse, H. Ooshima, and D. S. Burns. Kinetics of enzymatic hydrolysis of lignocellulosic materials based on surface area of cellulose accessible to enzyme and enzyme adsorption on lignin and cellulose. *Applied Biochemistry and biotechnology*, 24(1):67–73, 1990.

[8] V. J. H. Sewalt, W. G. Glasser, and K. A. Beauchemin. Lignin impact on fiber degradation. 3. Reversal of inhibition of enzymatic hydrolysis by chemical modification of lignin and by additives. *Journal of Agricultural and Food Chemistry*, 45(5):1823–1828, 1997.

[9] Alex Berlin, Neil Gilkes, Arwa Kurabi, Renata Bura, Maobing Tu, Douglas Kilburn, and John Saddler. Weak lignin-binding enzymes: a novel approach to improve activity of cellulases for hydrolysis of lignocellulosics. *Appl Biochem Biotechnol*, 121-124:163–70, 2005.

[10] Seiji Nakagame, Richard P. Chandra, and Jack N. Saddler. The effect of isolated lignins, obtained from a range of pretreated lignocellulosic substrates, on enzymatic hydrolysis. *Biotechnol Bioeng*, 105(5):871–9, 4 2010.

[11] Seiji Nakagame, Richard P. Chandra, John F. Kadla, and Jack N. Saddler. Enhancing the enzymatic hydrolysis of lignocellulosic biomass by increasing the carboxylic acid content of the associated lignin. *Biotechnol Bioeng*, 108(3):538–48, 3 2011.

[12] Hetti Palonen, Folke Tjerneld, Guido Zacchi, and Maija Tenkanen. Adsorption of Trichoderma reesei CBH I and EG II and their catalytic domains on steam pretreated softwood and isolated lignin. *Journal of Biotechnology*, 107(1):65–72, 1 2004.

[13] Jenni Liisa Rahikainen, Ulla Moilanen, Susanna Nurmi-Rantala, Angelos Lappas, Anu Koivula, Liisa Viikari, and Kristiina Kruus. Effect of temperature on lignin-derived inhibition studied with three structurally different cellobiohydrolases. *Bioresource Technology*, 146(0):118 – 125, 2013.

[14] Stefan Bauer, Hagit Sorek, Valerie D. Mitchell, Ana B. Ibáñez, and David E. Wemmer. Characterization of Miscanthus giganteus lignin isolated by ethanol organosolv process under reflux condition. *J Agric Food Chem*, 60(33):8203–12, 8 2012.

[15] Jenni L. Rahikainen, Raquel Martin-Sampedro, Harri Heikkinen, Stella Rovio, Kaisa Marjamaa, Tarja Tamminen, Orlando J. Rojas, and Kristiina Kruus. Inhibitory effect of lignin during cellulose bioconversion: The effect of lignin chemistry on non-productive enzyme adsorption. *Bioresource technology*, 2013.

[16] Nidhi Pareek, Thomas Gillgren, and Leif J. Jönsson. Adsorption of proteins involved in hydrolysis of lignocellulose on lignins and hemicelluloses. *Bioresour Technol*, 148:70–7, 11 2013.

[17] Kevin M. Holtman, Hou-Min M. Chang, and John F. Kadla. Solution-state nuclear magnetic resonance study of the similarities between milled wood lignin and cellulolytic enzyme lignin. *J Agric Food Chem*, 52(4):720–6, 2 2004.

[18] Anders Bjorkman. Studies on finely divided wood. Part 1: Extraction of lignin with neutral solvents. *Svensk Papperstidning*, 59(7):477–485, 1956.

[19] K. Lundquist and R. Simonson. Lignin preparations with very low carbohydrate content. *Svensk Papperstidning*, 78(11):360, 1975.

[20] Magnus Norgren, Shannon M. Notley, Andrea Majtnerova, and Göran Gellerstedt. Smooth model surfaces from lignin derivatives. i. preparation and characterization. *Langmuir*, 22(3):1209–14, 1 2006.

[21] Raquel Martín-Sampedro, Jenni L. Rahikainen, Leena-Sisko S. Johansson, Kaisa Marjamaa, Janne Laine, Kristiina Kruus, and Orlando J. Rojas. Preferential adsorption and activity of monocomponent cellulases on lignocellulose thin films with varying lignin content. *Biomacromolecules*, 14(4):1231–9, 4 2013.

[22] Thomas Gervais and Klavs F. Jensen. Mass transport and surface reactions in microfluidic systems. *Chemical Engineering Science*, 61(4):1102–1121, 2 2006.

[23] Martin L. Yarmush, Dhananjay B. Patankar, and David M. Yarmush. An analysis of transport resistances in the operation of biacoretm; implications for kinetic studies of biospecific interactions. *Molecular Immunology*, 33(15):1203–1214, 6 1996.

[24] Hyun J. Kwon, Craig K. Bradfield, Brian T. Dodge, and George S. Agoki. Study of simultaneous fluid and mass adsorption model in the qcm-d sensor for characterization of biomolecular interactions. In *Proceedings of the COMSOL Conference*, 2009.

[25] Ajit Sadana. *Engineering biosensors: kinetics and design applications*. Access Online via Elsevier, 2001.

[26] Matthew C. Dixon. Quartz crystal microbalance with dissipation monitoring: enabling real-time characterization of biological materials and their interactions. *J Biomol Tech*, 19(3):151–8, 7 2008.

[27] Erik Reimhult, Charlotte Larsson, Bengt Kasemo, and Fredrik Höök. Simultaneous surface plasmon resonance and quartz crystal microbalance with dissipation monitoring measurements of biomolecular adsorption events involving structural transformations and variations in coupled water. *Anal Chem*, 76(24):7211–20, 12 2004.

[28] K. Nakanishi, T. Sakiyama, and K. Imamura. On the adsorption of proteins on solid surfaces, a common but very complicated phenomenon. *J Biosci Bioeng*, 91(3):233–44, 2001.

[29] Irving Langmuir. Vapor pressures, evaporation, condensation, and adsorption. *Journal of the American Chemical Society*, 54(7):2798–2832, 1932.

[30] Michael Rabe, Dorinel Verdes, and Stefan Seeger. Understanding protein adsorption phenomena at solid surfaces. *Adv Colloid Interface Sci*, 162(1-2):87–106, 2 2011.

[31] Martin Karlsson, Johan Ekeroth, Hans Elwing, and Uno Carlsson. Reduction of irreversible protein adsorption on solid surfaces by protein engineering for increased stability. *J Biol Chem*, 280(27):25558–64, 7 2005.

[32] Samuel A. Maurer, Claire N. Bedbrook, and Clayton J. Radke. Competitive sorption kinetics of inhibited endo-and exoglucanases on a model cellulose substrate. *Langmuir*, 28(41):14598–14608, 2012.

[33] Luis G. Cascão Pereira, Andrea Hickel, Clayton J. Radke, and Harvey W. Blanch. A kinetic model for enzyme interfacial activity and stability: pa-hydroxynitrile lyase at the diisopropyl ether/water interface. *Biotechnol Bioeng*, 78(6):595–605, 6 2002.

[34] Flora Felsovalyi, Tushar Patel, Paolo Mangiagalli, Sanat K. Kumar, and Scott Banta. Effect of thermal stability on protein adsorption to silica using homologous aldo-keto reductases. *Protein Sci*, 21(8):1113–25, 8 2012.

[35] Keitaro Yoshimoto, Motohiko Nishio, Hiroaki Sugasawa, and Yukio Nagasaki. Direct observation of adsorption-induced inactivation of antibody fragments surrounded by mixed-peg layer on a gold surface. *J Am Chem Soc*, 132(23):7982–9, 6 2010.

[36] Joseph McGuire, Marie C. Wahlgren, and Thomas Arnebrant. Structural stability effects on the adsorption and dodecyltrimethylammonium bromide-mediated elutability of bacteriophage T4 lysozyme at silica surfaces. *Journal of Colloid and Interface science*, 170(1):182–192, 1995.

[37] Ingemar Lundström and Hans Elwing. Simple kinetic models for protein exchange reactions on solid surfaces. *Journal of colloid and interface science*, 136(1):68–84, 1990.

[38] Kierston Shill, Sasisanker Padmanabhan, Qin Xin, John M. Prausnitz, Douglas S. Clark, and Harvey W. Blanch. Ionic liquid pretreatment of cellulosic biomass: enzymatic hydrolysis and ionic liquid recycle. *Biotechnol Bioeng*, 108(3):511–20, 3 2011.

[39] Kun Cheng, Hagit Sorek, Herbert Zimmermann, David E. Wemmer, and Markus Pauly. Solution-state 2D NMR spectroscopy of plant cell walls enabled by a dimethylsulfoxide-d6/1-ethyl-3-methylimidazolium acetate solvent. *Anal Chem*, 85(6):3213–21, 3 2013.

[40] Joe Tien, Andreas Terfort, and George M. Whitesides. Microfabrication through electrostatic self-assembly. *Langmuir*, 13(20):5349–5355, 1997.

[41] Harshal Chokhawala, Christine Roche, Tae Wan Kim, Meera Atreya, Neeraja Vegesna, Craig M. Dana, Harvey W. Blanch, and Douglas S. Clark. Mutagenesis of Trichoderma reesei endoglucanase I: Impact of expression host on activity and stability at elevated temperatures. *In preparation*, 2015.

[42] Allison Tolbert, Hannah Akinosho, Ratayakorn Khunsupat, Amit K. Naskar, and Arthur J. Ragauskas. Characterization and analysis of the molecular weight of lignin for biorefining studies. *Biofuels, Bioproducts and Biorefining*, 2014.

[43] Thomas Q. Hu. *Chemical modification, properties, and usage of lignin*. Kluwer Academic/Plenum Publishers, New York, 2002.

[44] F. Höök, B. Kasemo, T. Nylander, C. Fant, K. Sott, and H. Elwing. Variations in coupled water, viscoelastic properties, and film thickness of a mefp-1 protein film during adsorption and cross-linking: a quartz crystal microbalance with dissipation monitoring,

ellipsometry, and surface plasmon resonance study. *Anal Chem*, 73(24):5796–804, 12 2001.

[45] Wendy Y. X. Peh, Erik Reimhult, Huey Fang Teh, Jane S. Thomsen, and Xiaodi Su. Understanding ligand binding effects on the conformation of estrogen receptor alpha-dna complexes: a combinational quartz crystal microbalance with dissipation and surface plasmon resonance study. *Biophys J*, 92(12):4415–23, 6 2007.

[46] GJ J. Szöllsi, I. Derényi, and J. Vörös. Reversible mesoscopic model of protein adsorption: From equilibrium to dynamics. *Physica A: Statistical Mechanics and its Applications*, 343:359–375, 11 2004.

[47] Christian F. Wertz and Maria M. Santore. Fibrinogen adsorption on hydrophilic and hydrophobic surfaces: geometrical and energetic aspects of interfacial relaxations. *Langmuir*, 18(3):706–715, 2002.

[48] Fenfen Guo, Wenjing Shi, Wan Sun, Xuezhi Li, Feifei Wang, Jian Zhao, and Yinbo Qu. Differences in the adsorption of enzymes onto lignins from diverse types of lignocellulosic biomass and the underlying mechanism. *Biotechnol Biofuels*, 7(1):38, 2014.

[49] Jenni Liisa Rahikainen, James David Evans, Saara Mikander, Anna Kalliola, Terhi Puranen, Tarja Tamminen, Kaisa Marjamaa, and Kristiina Kruus. Cellulase-lignin interactions–the role carbohydrate-binding module and pH in non-productive binding. *Enzyme and Microbial Technology*, 2013.

[50] Mark R. Nimlos, Gregg T. Beckham, James F. Matthews, Lintao Bu, Michael E. Himmel, and Michael F. Crowley. Binding preferences, surface attachment, diffusivity, and orientation of a family 1 carbohydrate-binding module on cellulose. *Journal of Biological Chemistry*, 287(24):20603–20612, 2012.

[51] Jerome M. Fox, Phillip Jess, Rakesh B. Jambusaria, Genny M. Moo, Jan Liphardt, Douglas S. Clark, and Harvey W. Blanch. A single-molecule analysis reveals morphological targets for cellulase synergy. *Nat Chem Biol*, 9(6):356–61, 6 2013.

Chapter 3

Lignin-binding kinetics of four major *H. jecorina* cellulases and their catalytic domains

3.1 Abstract

Hypocrea jecorina is a filamentous fungus that secretes a suite of hydrolytic enzymes capable of efficiently degrading cellulose. It's the industrial standard for cellulase production, and four cellulase enzymes make up the majority of its cellulose-degrading capability. In this chapter, the kinetics of nonproductive lignin adsorption for each of these cellulases, along with their catalytic domains, have been measured and modeled.

These four enzymes each have a two-domain structure with homologous Family 1 carbohydrate binding domains. Their catalytic domains fall into three different glycosyl hydolase families, with exoglucanase Cel7A and endoglucanase Cel7B being homologous catalytic domains. Like several existing studies, we find that the rate of adsorption to the lignin surface is enhanced by the presence of a CBM. However, among the 4 homologous CBMs in this work, a 3-fold difference in adsorption rate constant is observed. Similarly, a nearly 10-fold difference in adsorption rate is observed for the isolated catalytic domains of the homologous Cel7A and Cel7B.

While the adsorption rate is governed by the CBM, the irreversible component of the binding appears to be controlled by the catalytic domain. The thermostable Cel5A enzyme is the most resistant to irreversible binding. Qualitatively, the reversibility trends exhibited by the four full-length enzymes are reproduced by their isolated catalytic domains. Measurement of the lignin binding kinetics of these four enzymes and their catalytic domains reveals structural contributions to lignin binding, and reveals new targets for enzyme evolution and improvement.

3.2 Introduction

The four main cellulases from the filamentous fungus *Hypocrea jecorina* are present in concentrations such that their individual hydrolytic activities contribute synergystically to biomass deconstruction [1]. As such, differential inhibition via nonproductive binding to lignin may have an outsized effect on the overall hydrolyzing ability of the cellulase mixture.

Many experiments have been done over the years in an attempt to quantify the effect of lignin on biomass hydrolysis. Their results have varied widely. The simplest method to quantify lignin's nonproductive binding effect is to supplement a mixture of Avicel or microcrystalline cellulose with lignin. Results of this type of experiment range from no effect of nonproductive binding to overall extent of hydrolysis [2], to an 8-fold reduction in 48-hour hydrolysis [3] to a roughly 4-fold reduction in V_0 [4].

As early as 1990, researchers recognized that in order to make useful predictions about the effect of lignin adsorption, reaction dynamics, as well as adsorption equilibrium, are needed [5]. While some of these kinetic measurements have been attempted on cellulosic substrates [6, 7], and while mechanistic modeling of cellulose degradation has made much progress in recent years [8], kinetic measurements to probe the reaction pathways, kinetics, and structural contributions to lignin-cellulase interaction are lacking.

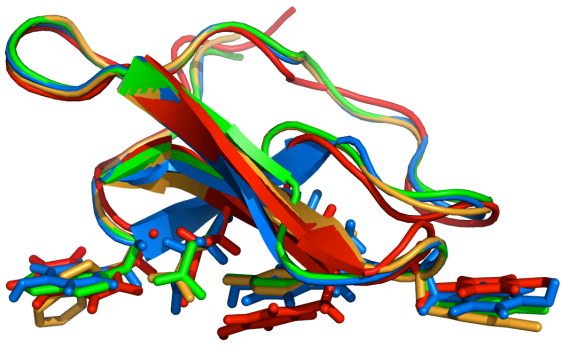

Figure 3.1: The binding domains of the four major *H. jecorina* cellulases are overlaid, with structures aligned to the Cel7A CBM (NMR structure solved in [9]) using the align command in Pymol. Blue: Cel7A, green: Cel6A, red: Cel7B, orange: Cel5A.

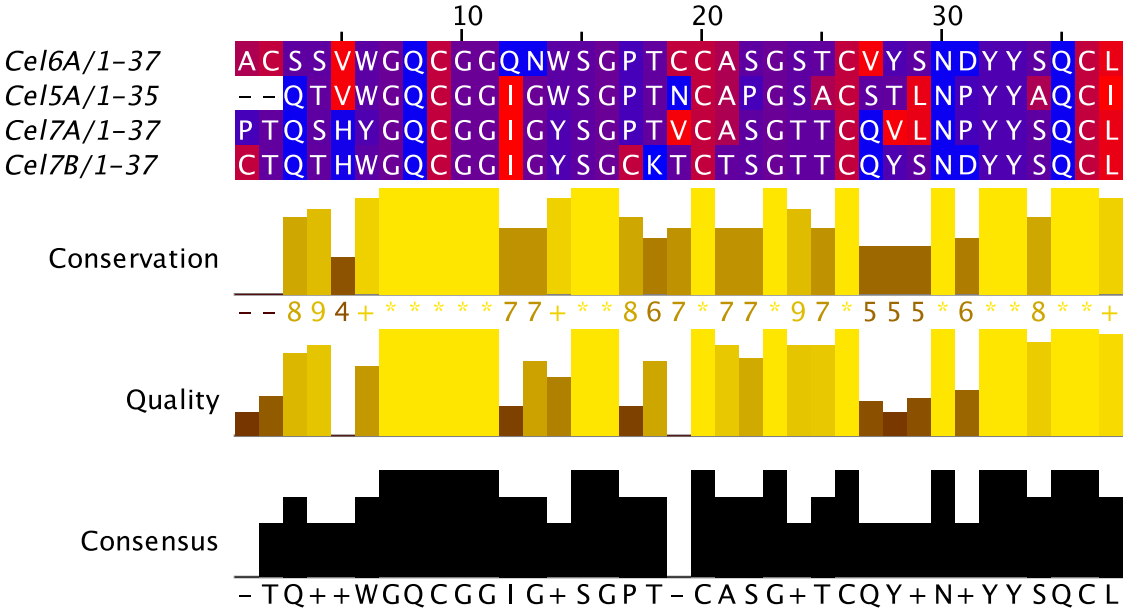

Figure 3.2: The binding domains of the four major *H. jecorina* cellulases are aligned, with conserved residues and a consensus sequence shown. Alignment of the four CBMs was done using BLASTp, with the data visualization by Jalview.

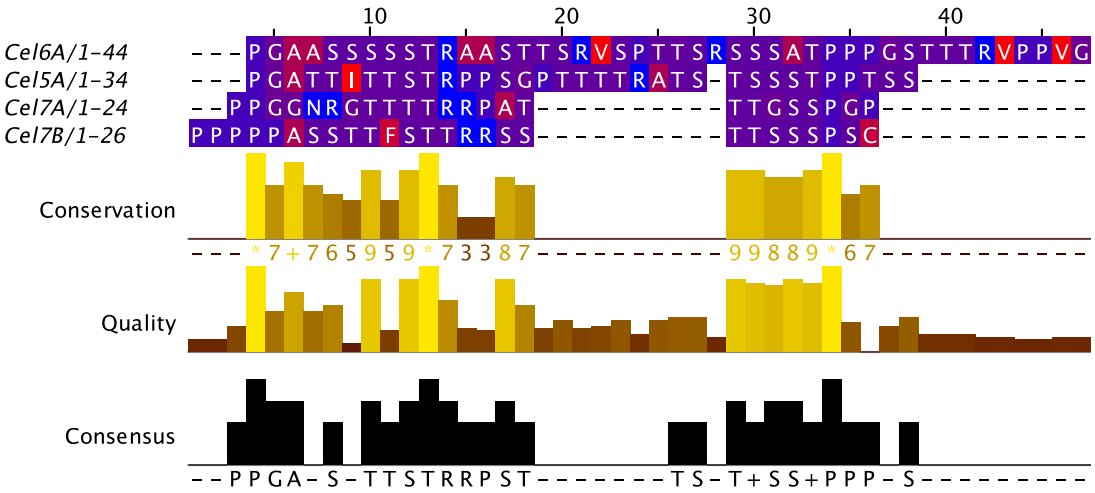

Figure 3.3: The linker regions of the four major *H. jecorina* cellulases are aligned, with conserved residues and a consensus sequence shown. Alignment of the four CBMs was done using BLASTp, with the data visualization by Jalview.

The purpose of this study is to gain a detailed understanding of the lignin binding kinetics of the four main cellulose-degrading in the *Hypocrea jecorina* secretome. While these four enzymes have remarkably similar binding domains (see Figures 3.2 and 3.2, we hope to understand whether the the existing diversity in Family 1 CBM structures, as well as differences i the linkers or catalytic domains, that have large effects on the different aspects of lignin interaction.

There is reason to believe that lignin binding differences might exist among the *H. jecorina* CBMs based on their cellulose binding behavior. When the CBMs of Cel7A and Cel7B were synthesized by solid state peptide synthesis, the Cel7B CBM had a higher affinity for cellulose than that of Cel7A [10]. When Cel7A CBM variants were made, most of the difference in affinity could be realized by a single amino acid change on the flat binding face by substituting a tyrosine for a tryptophan residue.

Carrard and Linder (1999) showed that under certain circumstances Cel7A binds cellulose reversibly, while Cel6A binds irreversibly, despite very similar structures. This irreversible binding behavior in Cel6A could be diminished by a point mutation of the tryptophan surface residue on Cel6A to tyrosine, as is present in Cel7A, or by disrupting a disulfide bond that is present in Cel6A but not Cel7A [11].

Other researchers have documented different substrate preferences among the *H. jecorina* CBMs. Tomme (1988) showed that the CBM of Cel6A confers affinity to both crystalline and amorphous cellulose, while Cel7A's CBM confers affinity to only the crystalline substrate. When the CBM was removed, a marked decrease was seen in Cel7A's affinity for crystalline cellulose, while its affinity for amorphous substrate remained unchanged. Cel6A saw a decrease in affinity for both substrates upon removal of its binding domain [12]. Other researchers uncovered more subtle substrate preferences. For example, when CBM1 from *H. jecorina* Cel7A and Cel6A, a CBM 4-1 from the *Cellulomonas fimi* xylanase/exoglucanase, and the CBM 3a from *C. thermocellum* CipA, each was found to have unique binding sites, and additional hydrolysis was accessible by adding a chimeric enzyme with a unique CBM [13]. More recently, other researchers have demonstrated with molecular dynamics simulation that single amino acid changes to Family 1 CBMs can affect the affinity and directional preferences by changing the hydrogen bonding between CBM and cellulose crystal [14].

Clearly, subtle differences in fungal CBMs has produced diversity in cellulose binding. The four corresponding catalytic domains are much more diverse in terms of fold structure, and we would expect a corresponding diversity in both cellulose and lignin interaction. While the fungal CBM1 has been shown to interact with lignin more readily than the binding domain for the *H. jecorina* enzymes Cel7A and Cel5A [15, 16], this information in incomplete. It lacks information on the dynamics of catalytic domain interactions, and it furthermore lacks investigation into the reversibly of the catalytic domain interaction. Recent work from Sammond et al. (2014) presented evidence to support a hypothesis that the size and number of hydrophobic patches on a protein surface controls the enzyme's affinity for lignin [17]. Our comparison of these four cellulases will allow us to test this hypothesis using kinetic parameters extracted from binding data, rather than qualitative comparisons of enzyme binding. We hope that by isolating *H. jecorina* catalytic domains, we can gain new understanding

of the CBM and CD to the adsorption, desorption, and irreversible adsorption of individual cellulases to lignin.

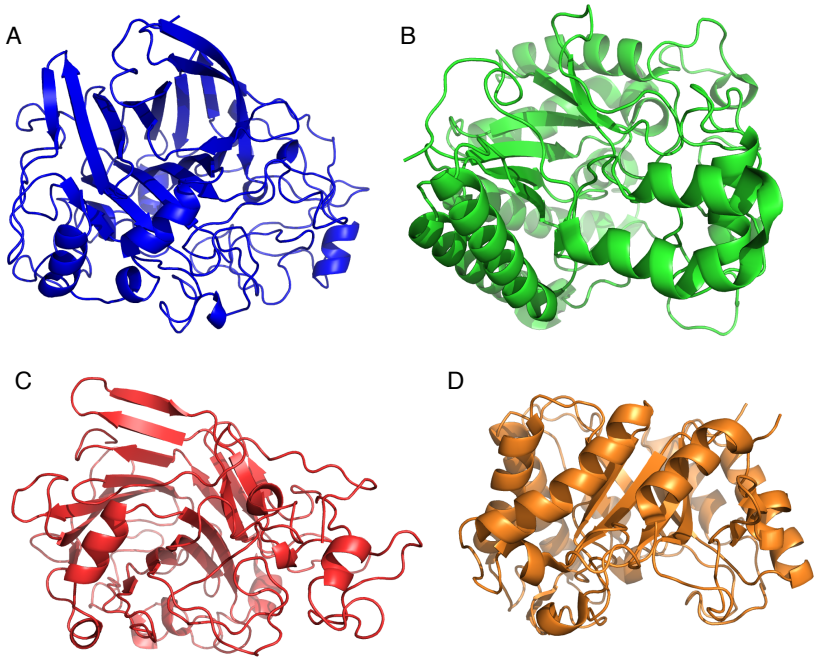

Figure 3.4: Crystal structures of the four main *Hypocrea jecorina* celluase catalytic domains. A: Cel7A, B: Cel6A, C: Cel7B, D: Cel5A.

A dynamic understanding of cellulase-lignin interaction is important in modeling the enzymatic breakdown of biomass. It's been hypothesized that lignin inhibition occurs by competitive adsorption. Any attempts to create a mechanistic model of biomass hydrolysis must account for the adsorption kinetics of cellulase to lignin in order to model biomass accurately and mechanistically.

Finally, we seek to probe the domain contributions of individual cellulases as a method of defining targets for enzyme improvement. Researchers have suggested that CBM engineering for reduced lignin affinity is an opportunity to improve enzyme performance in biomass hydrolysis [18], but other opportunities may exist. For example, *Hj*Cel5A catalytic domain has been shown to have a non-negligible affinity for lignin [15]. Researchers who suggest that removing CBMs may be a viable option for reducing lignin inhibition, especially at high biomass solids concentration [19], may find that further improvements can be found by engineering catalytic domains for reduced affinity for lignin. Understanding the reaction mechanism and kinetics can also provide evolution targets - an enzyme with a slow adsorption rate may also have a relatively high rate of irreversible adsorption - engineering the catalytic

domain for stability may produce better results than engineering for a decreased adsorption rate. Qualitatively, researchers have already shown that enzymes with identical CBMs can be more or less inhibited by lignin by changing the catalytic domain to a more thermostable (and presumably less susceptible to surface denaturation) catalytic domain [20].

3.3 Materials and methods

Trichoderma reesei enzyme purification

Natively produced *H. jecorina* cellobiohydrolase I (Cel7A), cellobiohydrolase II (Cel6A) and endoglucanase II (Cel5A) were purified from Celluclast, a commercially available cellulase mixture produced by *H. jecorina* strain ATCC 26921 (Sigma Aldrich, St. Louis, Missouri). The commercial preparation was dissolved at 7.5 g/L and exchanged into 25mM HEPES Buffer pH 7.35 on a HiPrep desalting column. The eluate was loaded onto a Q Sepharose 16/10 HP column, washed with equilibration buffer, and eluted with a 1M NaCl gradient in 25 mM HEPES pH 7.35. Purified fractions for Cel5A, Cel6A, and Cel7A were pooled from the elution peaks. The fractions were buffer exchanged into acetate buffer, 50 mM, pH 4.85.

Cellobiohydrolase CBM cleavage

H. jecorina cellobiohydrolases Cel7A and Cel6A can be partially proteolyzed to yield isolated catalytic domains. The isolated catalytic domains retain full activity on soluble substrates, but have just 10% and 50% of their respective initial hydrolysis rate on solid crystalline substrates [12].

Cel7A and Cel6A catalytic domains were produced by the limited proteolysis of the full enzyme with papain (Sigma, from papaya latex, >10 units/mg, P4762). The natively expressed enzymes were purified as discussed above. Papain activating buffer, 100 mM sodium phosphate, pH 7.0, 2 mM dithiothreitol, and 2 mM ethylenediaminetetraacetic acid was prepared. Papain, 28 mg/mL, was diluted 11:1 in activating buffer and brought to room temperature for 30 minutes. The activated papain was added to the cellulase solution at a 1:100 protein mass ratio, and left at room temperature overnight.

The reaction products were diluted 2x with water and purified on a MonoQ or Q sepharose FF FPLC column as discussed in the purification of full Cel7A and Cel6A. Purity was verified by running on an SDS-PAGE gel. Enzyme activity was compared with the full enzyme activity on soluble (4-MU-lac) substrate.

Cloning and CD production (Cel7B Δ CBM, Cel5A, Cel5A Δ CBM)

Endoglucanase I (Cel7B) cannot reliably be purified from celluclast, and neither endoglucanase Cel7B nor Cel5A are susceptible to limited proteolysis for the isolation of the catalytic

domain. For this reason, we have pursued the recombinate production of these two endoglucanases in their native host, *Hypocrea jecorina*.

For recombinant expression of production of H. jecorina Cel7B, a Cel7B expression vector was constructed by introducing the Cel7B gene (amplified from genomic DNA) followed by the *N. crassa* Cel7A terminator behind the CNDA1 promoter in pPCDNA1 expression plasmid (generous gift from Bernhard Seiboth) [21].

Cel7B catalytic domain was produced by amplifying around the pCDNA plasmid, forming complementary termini and removing the DNA corresponding to the CBM and linker regions. Primers are listed in Table A.1 . The linear fragment ends were joined using the Gibson isothermal assembly method [22].

Cel5A full enzyme was expressed in a similar manner. Cel5A was PCR amplified from *H. jecorina* genomic DNA using Pfu Ultra II (Agilent, Santa Clara, CA) using primers listed in Appendix A. The pCDNA plasmid was amplified with complementary 3' and 5' termini (A.1C and Gibson assembly was used to assemble the two fragments. The CBM and linker regions were removed by amplifying the plasmid with complementary terminal regions, excluding the CBM and linker, and reassembled by Gibson assembly. Full DNA sequences of the expressed reading frames are available in Appendix A. The Cel5A full enzyme expressed functionally, but although the CD-only and CD/linker constructs were could be amplified from the genomic DNA of transformants, the enzymes failed to express.

Transformation of *H. jecorina* by electroporation was adapted as described previously [23]. Spores of *H. jecorina* QM9414 were harvested from a 90mm potato dextrose agar (PDA) plate and suspended in 1.1 M ice cold sorbitol. Spores were washed twice with 1.1 M ice cold sorbitol, pelleted at 800xg for 4 min, and then re-suspended in 100 uL of 1.1 M ice cold sorbitol. Two micrograms of purified PCR product of the expression cassette using the primers listed in A.1E was added to the spore suspension and incubated on ice for 30min. Electroporation was conducted using a Gene Pulser Xcell (Bio-Rad Laboratories, Inc., Hercules, CA) in 1mm gap cuvette at 1.6kV, $600\Omega and 25°$F. Immediately following electroporation, 900 μL of 1.1 M ice cold sorbitol was added and gently mixed. The spore suspension was then added to 9 mL of YPD (1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) dextrose) and incubated for 12 hr at room temperature. Spores were pelleted at 800xg for 4 min, re-suspended in 1 mL of 1.1 M sorbitol, and plated using the overlay method onto Mandels-Andreotti medium containing 200 μg/mL hygromycin B. Both top and bottom agars contained hygromycin B. Germinated spores were picked onto PDA slants containing hygromycin B. Positive transformants were selected by presence of EGI activity (AZO-CM-Cellulose Assay, Megazyme, Ireland) in the supernatant after growth in Mandels-Andreotti medium supplemented with 3% glucose, in which QM9414 does not secrete native cellulases [21, 24] for 3-4 days. Culture broths were filtered through glass microfiber filters (934-AH, Whatman) followed by 0.22 μm PES filters (Corning). This procedure was done by Dr. Christine Roche in the Clark Lab, as described [25].

QCM-D kinetics

A detailed description of the kinetic model is presented in Chapter 2. The same techniques are used here; the data are fit first to a single-site transition model, and if necessary, a second binding site is added ($k_{A2}\Gamma_{max2}$). The overall strategy for kinetic modeling involves the following steps:

1. Measure initial adsorption rate with varying bulk enzyme concentration. The slope of the resulting line is equal to $k_A\Gamma_{max}$.
2. Measure the total surface capacity by saturating the surface with high-concentration enzyme until the adsorbed mass is maximized.
3. Collect binding and washoff curves for varying bulk enzyme concentrations and contact times. Fit the individual washoff curves to extract the lumped parameter $k_i + k_D$.
4. Summarize the reversibility of binding vs. contact time and bulk enzyme concentration. Numerically solve the mass balance equations to minimize the model's deviation from the reversibility data, extracting individual rate constants k_i and k_D.
5. Using the one-site transition model kinetic parameters from above, fit adsorption and desorption curves at high and low bulk enzyme concentrations to check the model fit. If the one-site model fails to accurately reproduce the adsorption curves, use the lowest concentration adsorption data to fit a new parameter, Γ_{max1}.
6. Using high concentration data and the Γ_{max1} from above, fit k_{A2}.
7. Iterate to solve for Γ_{max1}.

3.4 Results

Cellulase adsorption kinetics measured with QCM-D

The four main *H. jecorina* cellulases, and their isolated catalytic domains (excluding Cel5A CD, which could not be successfully expressed) were purified (either purified from commercial cellulase cocktail or expressed heterologously in *H. jecorina*). In Chapter 2, three plausible models for cellulase adsorption to a lignin surface were presented: a single-site transition model, a two-step model with changing footprint size, and a transition model with multiple binding sites. We will apply the one and two-site transition model to each of the four cellulases in this work, and discuss the model's fit, and any deviations from the model that we observe.

Using the methods described in Chapter 2, the adsorption rates of each full-length cellulase and its catalytic domain to lignin isolated from dilute acid pretreated *Miscanthus* were measured. The results are displayed in Figures 3.5 and 3.6.

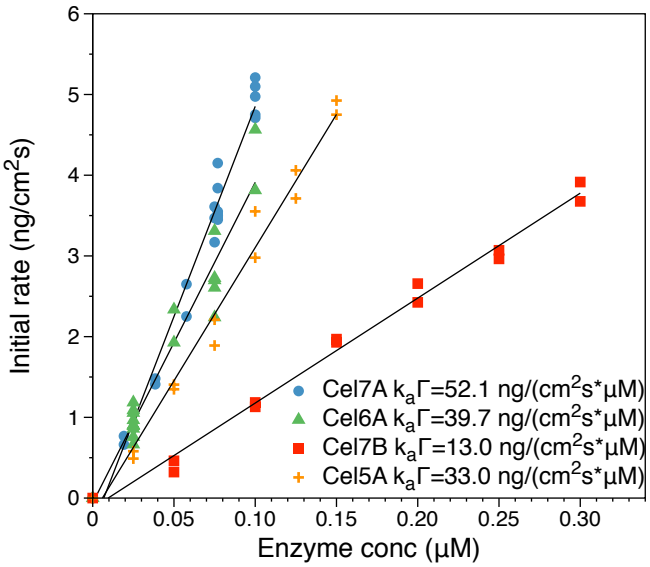

Figure 3.5: Comparison of initial adsorption rate $k_A\Gamma_{max}$ of the four major *H. jecorina* cellulases.

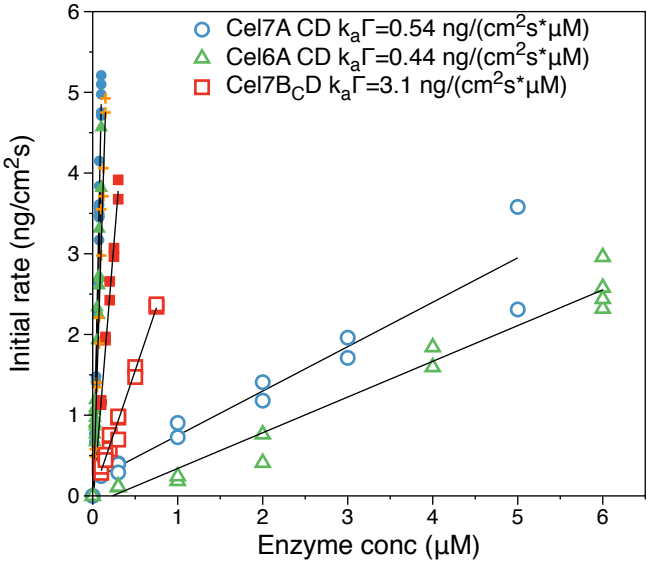

Figure 3.6: Comparison of initial adsorption rate $k_A\Gamma_{max}$ of the four major *H. jecorina* cellulase catalytic domains. Full enzyme data is also displayed for perspective.

Maximum surface capacity

In order to deconvolute the kinetic parameters k_D and k_i, it's first necessary to measure or assume a maximum surface capacity Γ_{max}. We measure this parameter using QCM-D by flowing increasingly concentrated enzyme solutions over the lignin-coated sensor and measuring the surface binding until the adsorption rate is less than the manufacturer's specification for acceptable instrument drift (3 hz/hr), or around 1 hour. Maximum adsorptive capacity ($\Gamma_{max}$is reached when an increase in bulk protein concentration results in no further increase in surface concentration. Figure 3.7 and Table 3.4 summarize the measured values of Γ_{max}.

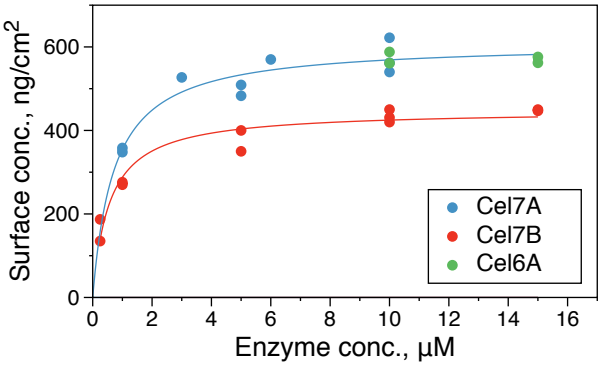

Figure 3.7: Maximum surface concentration of the four main *H. jecorina* cellulases. Low concentration adsorption was not done for Cel6A; high concentration data points were chosen based on the results of Cel7A and Cel7B and achieved surface saturation as expected. Cel5A maximum surface concentration data is still outstanding.

Cellulase	Γ_{max} (ng/cm^2)	±1 StdDev
Cel7A	573	35
Cel6A	572	13
Cel7B	440	13.5
Cel5A	ND	

Table 3.1: Summary of Γ_{max} for each cellulase.

Evidence of surface aggregation at high concentration, 50°C

When attempting to measure maximum surface concentration Γ_{max}, it was experimentally convenient to adsorb the cellulases to the surface at 15°C, as the QCM-D signal is more stable, and the low temperature minimizes bubble formation on the sensor surface. We

assume that Γ_{max} is unchanged at low temperature, and in an effort to verify this assumption we repeated high-concentration, maximum binding experiments at 50°C. The results were surprising - rather than a fast increase to a maximum binding capacity, we saw fast binding followed by slow, continuous surface adsorption, that continued for long periods of time without signs of approaching a maximum, and reached surface concentrations that would not be plausible for a monolayer of protein on the surface.

Surface mass data is acquired by applying the Sauerbrey equation to the frequency shift observed. At low temperature, we observed that the $\Delta D/\Delta f$ ratio is constant, and small (<5%), throughout the adsorption curve. At 50°C, however, the $\Delta D/\Delta f$ undergoes a shift, and a second dissipation regime is observed. Figure 3.8 illustrates this effect.

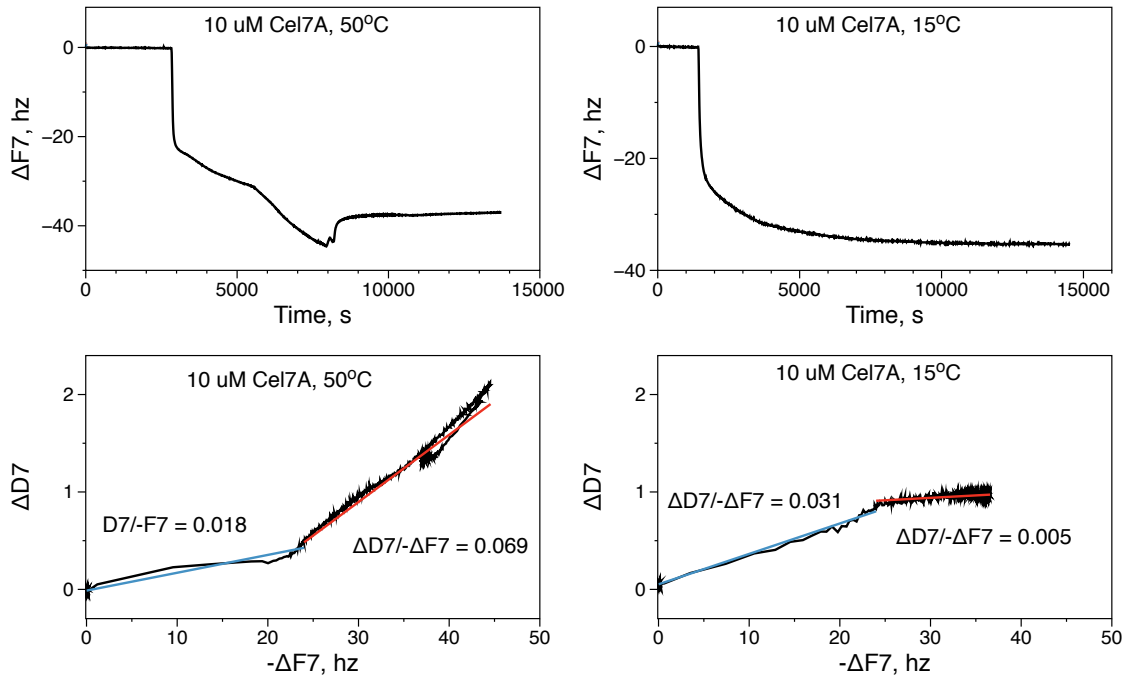

Figure 3.8: Comparison of high and low temperature adsorption of 10 uM Cel7A to lignin surface

At 50°C, the slope of $\Delta D/-\Delta f$ is constant until around -Δf = 24 hz, at which point it begins increasing with a slope of several times that. One hypothesis that could explain this behavior is a shift from cellulase-lignin as the primary binding interaction to a cellulase-cellulase nonspecific interaction, as cellulase on the surface begins to unfold and forms a multilayer film. Large $\Delta D/-\Delta f$ ratios are commonly associated with extended conformations or loose and flexible interactions, as these interactions result in deformation during oscillation and dissipation of mechanical energy [26, 27, 28]. We also note that this associ-

ation is at least partially reversible: when the 10μM Cel7A solution is replaced with buffer, $\Delta D/-\Delta f$ remains constant as some of the adsorbed mass is washed from the surface.

Conversely, a different transition occurs around $-\Delta f = 24$ hz when 10μM Cel7A is flowed over the lignin-coated sensor at lower temperature. Rather than increasing, $\Delta D/-\Delta f$ stabilizes and changes very little as Δf drops further. We hypothesize that as additional mass adsorbs, the dissipation factor increases until a critical surface concentration in reached. Thereafter, enzymes may be crowded on the surface and their ability to deform during oscillation is hindered. Notably, this transition occurs at approximately the same surface concentration as the transition to hypothetical protein multilayers (in the 50°C case).

Small angle light scattering has been used to estimate the size and shape of cellulase enzyme [29, 30, 31]; we can compare these dimensions with the surface mass we observe using QCM-D. *Hj*Cel7A was estimated to have a dimensions of approximately 18nm in maximum length (full enzyme) by 4.4 nm in width (catalytic domain). We can calculate that a perfectly packed monolayer of this 65 kDa enzyme would result in a surface mass of 136 ng/cm^2. We measure a maximum QCM-D mass of 600 ng/cm^2, and measure that approximately 90% of this mass takes the form of adsorbed water. Therefore, we can estimate that the surface is just under 50% covered with Cel7A at its maximum binding capacity.

This measurement compares favorably with the surface coverage expected by a random sequential adsorption model, in which molecules adsorb on a surface at random, and thus do not achieve perfect packing on the surface. When molecules adsorb close to each other, but not perfectly packed, they preclude another molecule from adsorbing between them. The RSA model applied to theoretical hard spherical discs results in a surface coverage of about 55% [32]. Researchers have experimentally determined a maximum surface coverage for fibrinogen, a protein with an aspect ratio of around 7.5, of around 40% [33]; Cel7A's aspect ratio is around 4, thus its RSA surface coverage would fall between that of a globular protein and the high-aspect-ratio fibrinogen.

Desorption and irreversible adsorption kinetics

Transition model validation: kinetic parameters k_D and k_i

Using the methods described in Chapter 2, we can collect adsorption and washoff data for each of the *H. jecorina* cellulase enzymes. The enzymes' kinetic behavior conforms to the model to varying degrees, so each enzyme will be discussed individually in terms of its fit to the transition model (one state or multiple states).

Both the transition model and the multi-site transition model described in Chapter 2 assume that neither the rate constant for desorption k_D nor the rate constant for irreversible adsorption k_i change with time or surface capacity. Because a lumped parameter equal to $k_i + k_D$ is extracted by fitting the washoff curves (see Chapter 2 for more detail), we test the validity of this assumption by comparing the lumped parameter as contact time (t_c) or surface concentration at washoff ($\Gamma_{tot}(t_w = 0)$). If the lumped parameter changes with contact time or surface concentration, we can no longer assume a constant value for these two

kinetic parameters. However, even if the lumped parameter is unchanged as experimental variables are changed, the model is not necessarily validated as the individual parameters may change while their sum remains the same. Comparison of the reversibility data with the model-predicted degree of reversibility can provide additional support for the hypothesis that k_D and k_i are constant.

In Figures 3.9 through 3.12, the sum of k_i and k_D for each individual washoff run is plotted vs. t_c and $\Gamma_{tot}(t_w = 0$. While the data are noisy, there is no clear trend for Cel7A, Cel6A, or Cel7B. Cel5A is an outlier: at short contact times, the Cel5A lumped parameter is larger than at longer contact times (see Figure 3.12). This may be due to changes in rate constants over time, or due to sequential adsorption to sites with varying desorption and irreversible adsorption rates.

We can investigate these two hypotheses by comparing $(k_i + k_D)$ vs. t_c and $(k_i + k_D)$ vs. $\Gamma_{(t_w=0)}$. The comparison of $(k_i + k_D)$ vs. t_c appears to show a correlation between the lumped parameter and contact time. However, at constant bulk concentration, contact time is correlated with the total (reversibly and irreversibly) bound mass at washoff ($\Gamma_{tot,t_w=0}$). When we consider low-concentration data as well, we observe a cluster of high-$(k_i + k_D)$ measurements (corresponding to the $t_c = 1$ minute measurements and a surface concentration of around 80 ng/cm^2). However, when the bulk concentration is decreased, we can observe that although the surface concentration is the same, the lumped parameter decreases. This supports the hypothesis that one or both of k_i or k_d is changing with time, rather than changing depending on the surface coverage. This could occur if Cel5A adsorption occurs first via the CBM, and over time the catalytic domain adsorbs as well, resulting in changing values of the kinetic constants (see Figure 3.12).

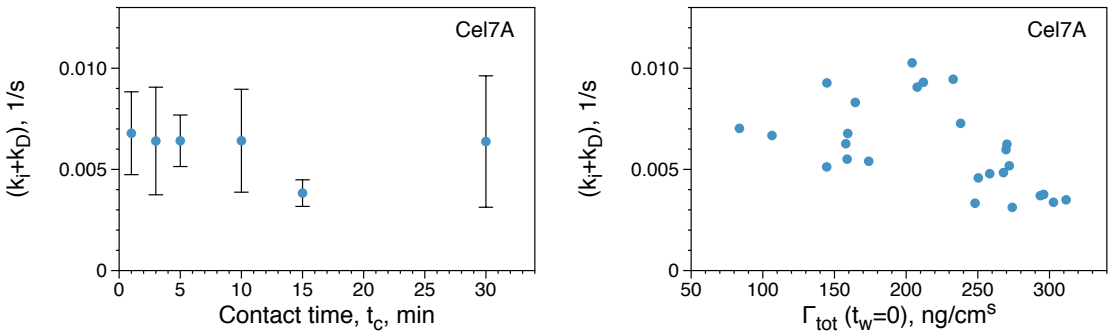

Figure 3.9: Lumped parameter $k_i + k_D$ plotted against experimental variables: Cel7A

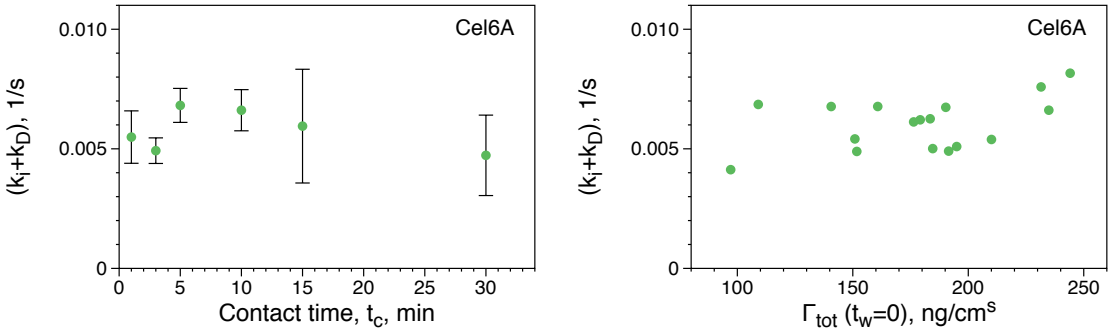

Figure 3.10: Lumped parameter $k_i + k_D$ plotted against experimental variables: Cel6A

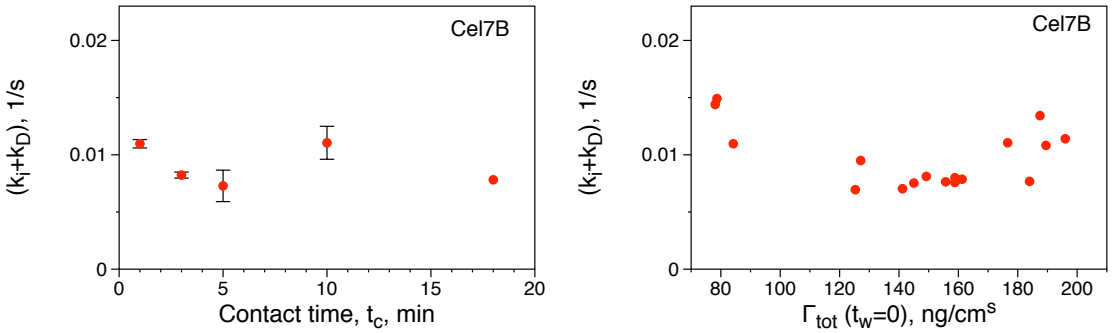

Figure 3.11: Lumped parameter $k_i + k_D$ plotted against experimental variables: Cel7B

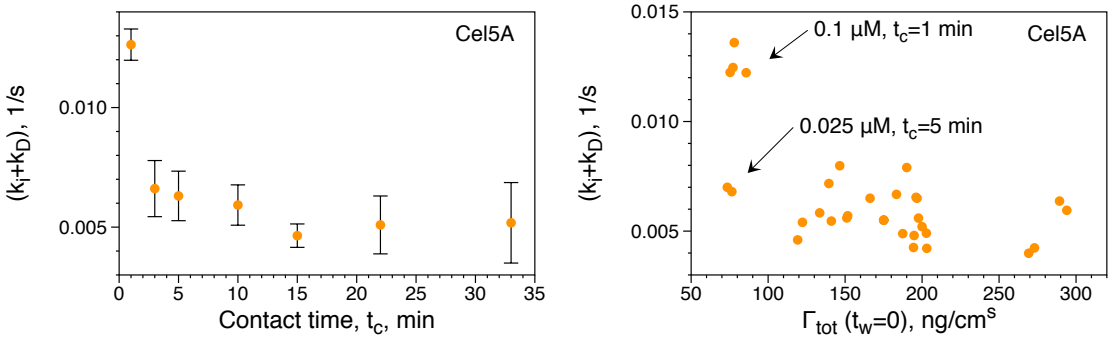

Figure 3.12: Lumped parameter $k_i + k_D$ plotted against experimental variables: Cel5A

The lumped kinetic parameter (k_i+k_D) is then fit to reversibility data (fraction reversibly bound for each bind-and-desorb run) vs. t_c and bulk concentration). Using the mass balance

equations in Chapter 2, we can decolvolute k_i and k_D to minimize the deviation of the predicted reversibility data from the experimentally gathered data. However, as in the Cel7B case detailed in Chapter 2, we observed with each cellulase that application of Γ_{max} values measured in Figure 3.7 resulted in a poor fit to adsorption curves at both high and low concentrations. Therefore, we applied the two-site transition model, and used low-concentration and high-concentration data separately to iteratively solve for Γ_{max1},k_{A1}, and k_{A2}. The model predicted reversibility data is compared to the experimental data in Figure 3.13.

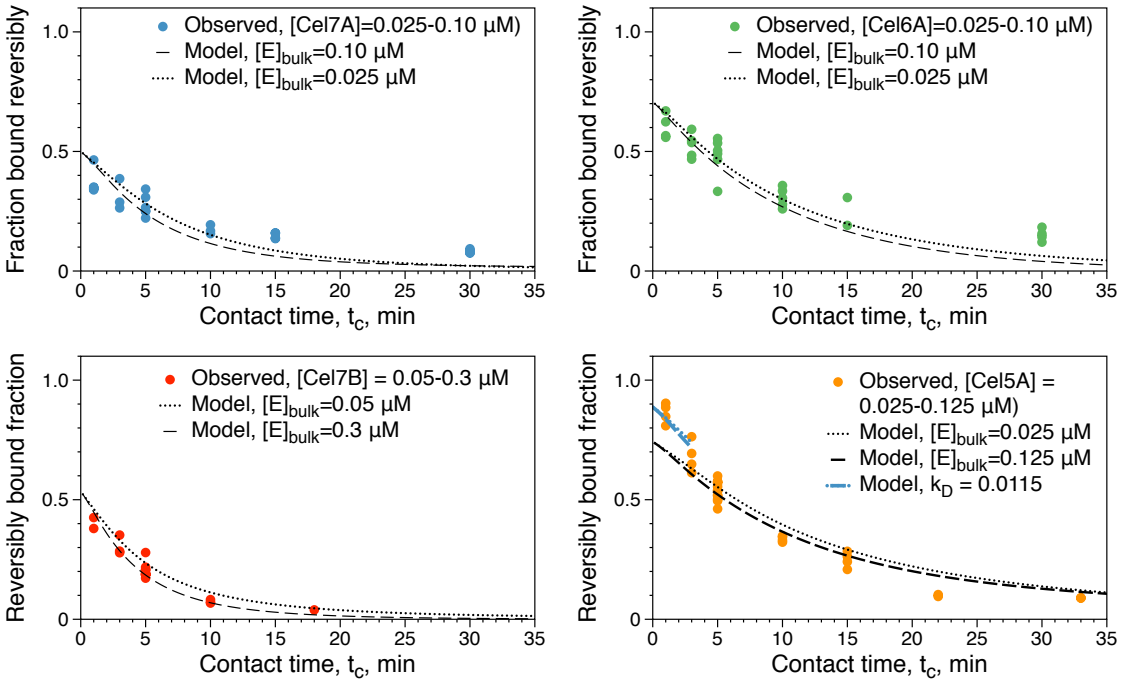

Figure 3.13: Reversiblity data, along with 2-site transition model fits, for each full-length enzyme tested. For each enzyme, high and low concentrations were modeled, since bulk enzyme concentration has an effect on the fraction of enzyme that is ultimately bound irreversibly. For Cel5A, the $k_i + k_D$ parameter is related to t_c. In this case, a second fit is plotted for the early time points, which assumes that k_i is unchanged, and that the increase in $k_i + k_D$ comes from a higher value of k_D. This hypothesis is supported by the data.

Each of the fits uses the two-site transition model to fit the reversibility data. Cel5A, the only cellulase that showed a clearly changing (k_i+k_D) value, has been modeled by fitting the data from t_c3 minute using the average of the (k_i+k_D) values corresponding to the t_c1 minute washoff experiments. We present the fit, along with the experimental reversibility data. For the $t_c = 1$ minute data point, the $(k_i + k_D)$ parameter is about twice that of the longer contact times. This could arise from either an increased value of k_i, k_D, both k_i and k_D, or

an increase in one and a decrease in the other. However, the washoff data at $t_c = 1$ minute (not included in model fit) are higher than predicted. An increased value of k_D at early time points would result in a larger reversible fraction. When the model is plotted using the value of k_i from the t_c3 minutes solution, but setting k_D such that $k_D = (k_i + k_D)_{t_c=1min} - k_i$, we see a good fit of the early-time data.

Kinetics summary

A summary of the kinetic parameters measured and fit using the two-site transition model are presented in Tables 3.2-3.5.

Table 3.2: Cel7A kinetics summary

Parameter	Value (± 1 standard deviation)	Units	R^2
$(k_{a1\|} + k_{a2})\Gamma_{max}$	52.1	ng/(cm^{2}s μM)	0.99 vs. initial rates
$k_{a1}\Gamma_{max1}$	51.2	ng/(cm^{2}s μM)	
$k_{a2}\Gamma_{max1}$	0.9	ng/(cm^{2}s μM)	
Γ_{max1}	230[a]	ng/cm^2	
Γ_{max2}	343[a]	ng/cm^2	
Γ_{max}, measured	573 (35) [a]	ng/cm^2	
$k_i + k_d$	0.0061 (0.0021)	s^{-1}	N/A
k_d	0.0030	s^{-1}	0.64 vs. washoff histories (Fig. 3.13)
ki	0.0031	s^{-1}	0.64 vs. washoff histories (Fig. 3.13)

Table 3.3: Cel6A kinetics summary

Parameter	Value (± 1 standard deviation)	Units	R^2
$(k_{a1\|} + k_{a2})\Gamma_{max}$	39.7	ng/(cm^{2}s μM)	0.97 vs. initial rates
$k_{a1}\Gamma_{max1}$	37.9	ng/(cm^{2}s μM)	
$k_{a2}\Gamma_{max1}$	0.8	ng/(cm^{2}s μM)	
Γ_{max1}	210[a]	ng/cm^2	
Γ_{max}, measured	572 (13) [a]	ng/cm^2	
$k_i + k_d$	0.0059 (0.0013)	s^{-1}	N/A
k_d	.0043	s^{-1}	0.88 vs. washoff histories (Fig. 3.13)
ki	0.00173	s^{-1}	0.88 vs. washoff histories (Fig. 3.13)

[a]QCM-D measured mass (hydrated).

Table 3.4: Cel7B kinetics summary

Parameter	Value (± 1 standard deviation)	Units	R^2
$(k_{a1\|} + k_{a2})\Gamma_{max}$	13.0	ng/(cm^{2}s μM)	0.988 vs. initial rates
$k_{a1}\Gamma_{max1}$	11.55	ng/(cm^{2}s μM)	
$k_{a2}\Gamma_{max1}$	1.45	ng/(cm^{2}s μM)	
Γ_{max1}	98 [a]	ng/cm^2	
Γ_{max2}	172 [a]	ng/cm^2	
Γ_{max}, measured	440 (13)	[a]	ng/cm^2
$k_i + k_d$	0.0098 (0.0026)	s^{-1}	N/A
k_d	0.00515	s^{-1}	0.92 vs. washoff histories (Fig. 3.13)
ki	0.00467	s^{-1}	0.92 vs. washoff histories
QCM mass fraction Cel7B	10.5 ±1.33%	%	N/A

[a]QCM-D measured mass (hydrated).

Table 3.5: Cel5A kinetics summary

Parameter	Value (± 1 standard deviation)	Units	R^2
$(k_{a1\|} + k_{a2})\Gamma_{max}$	33.0	ng/(cm^{2}s μM)	0.99 vs. initial rates
$k_{a1}\Gamma_{max1}$	31.9	ng/(cm^{2}s μM)	
$k_{a2}\Gamma_{max1}$	2.1	ng/(cm^{2}s μM)	
Γ_{max1}	145[a]	ng/cm^2	
Γ_{max2}	405[a]	ng/cm^2	
Γ_{max}, measured	550 [a]	ng/cm^2	
$k_i + k_d$ (t_c1min)	0.0058 (0.0011)	s^{-1}	N/A
$k_i + k_d$ (t_c = 1min)	0.013 (0.0007)[b]	s^{-1}	N/A
ki	0.00149	s^{-1}	0.93 vs. washoff histories (3.13)

[a]QCM-D measured mass (hydrated).

Catalytic domain desorption and irreversibility

The isolated catalytic domains require a much higher protein concentration to produce a QCM-D signal; as a result, a full kinetic analysis is more challenging. We have measured initial adsorption rates, and rather than a full analysis, we do a qualitative comparison of the washoff curves. In Figure 3.14, the data collected for catalytic domain reversibility is plotted against the model results for the full-length enzymes. The CD data closely follows that curves generated for the full-enzyme models. Qualitatively, at least, the catalytic domain appears to follow the same reversibility behavior as the full enzyme. The kinetic parameters determined for the catalytic domains are listed in Table 3.6.

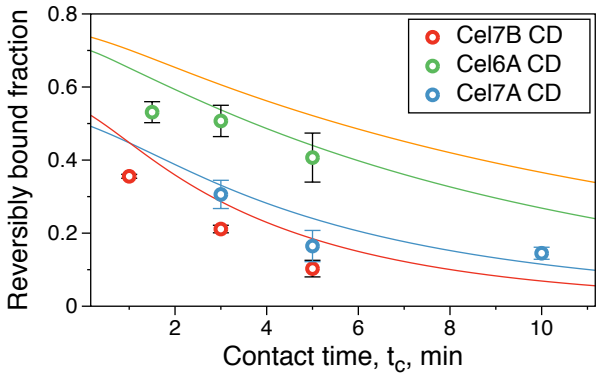

Figure 3.14: Washoff histories of cellulase catalytic domains, compared with the model results for the corresponding full-length enzymes. Although the CDs have a much lower adsorption rate, the irreversible binding to the lignin surface follows the same trends as the full-length constructs.

Table 3.6: Kinetic parameters of *H. jecorina* catalytic domains

Catalytic domain	$k_A\Gamma_max$ (ng/cm^{2}s*μM)	r^2	k_A s^{-1} (assuming Γ_{max} unchanged)
Cel7A	0.54	0.99	$9.4 * 10^{-4}$
Cel6A	0.44	0.96	$7.7 * 10^{-4}$
Cel7B	3.1	0.98	$6.7 * 10^{-3}$
Cel5A	ND		

3.5 Discussion

k_A diversity among *Hj*CBMs and catalytic domains

A summary of the measured adsorption rate constants k_A for the four full-length cellulases examined in this work and the three available catalytic domains are presented in Figure 3.15. The full-length enzymes Cel7A, Cel6A, and Cel5A each have a k_A value within 15% of each other; Cel7B is the outlier. While the four CBMs of these four enzymes are very similar, they have a few differences apparent in the sequence alignment that may explain this behavior. Cel6A and Cel7B both have three disulfide bonds (vs. two for the other CBMs). Cel6A and Cel7B CBM, which may contribute to structural rigidity that renders the CBM less able to interact with lignin. A recent study compared surface hydrophobicity of a diverse set of enzymes and concluded that surface hydrophobicity was a driver of enzyme

affinity for lignin films [17]. Cel7B CBM has several hydrophilic amino acids not present in the other CBMs, such as (referring to the alignment enumeration from Figure 3.2) histidine at position 5, lysine at position 18, threonine at positions 19, 21, and 25, serine at position 29, and aspartic acid at position 31. While some of these charged or polar amino acids are present in some of the other sequences, their combined effect may be to reduce the lignin affinity of the Cel7B CBM.

Both of the endoglucanases have a lower $k_A\Gamma_{max}$ than the exoglucanases. This trend has been observed in the literature: Palonen et al. observed a lower affinity for Cel5A compared to Cel7A [15].

The other major difference between the Cel7B full enzyme construct and the other three studied is its expression. Cel7B is difficult to purify from a cellulase mixture, so it was expressed heterologously in *H. jecorina*. Although the expression host is the same, it's possible that a difference in expression conditions between the native expression and the randomly-inserted promoter and gene, be it folding, glycosylation, or some other parameter, has affected the lignin binding capacity of Cel7B. We would like to heterologously express a Cel5A CD and full-length enzyme to compare with the results we see for Cel7B; unfortunately this work has not yet been completed.

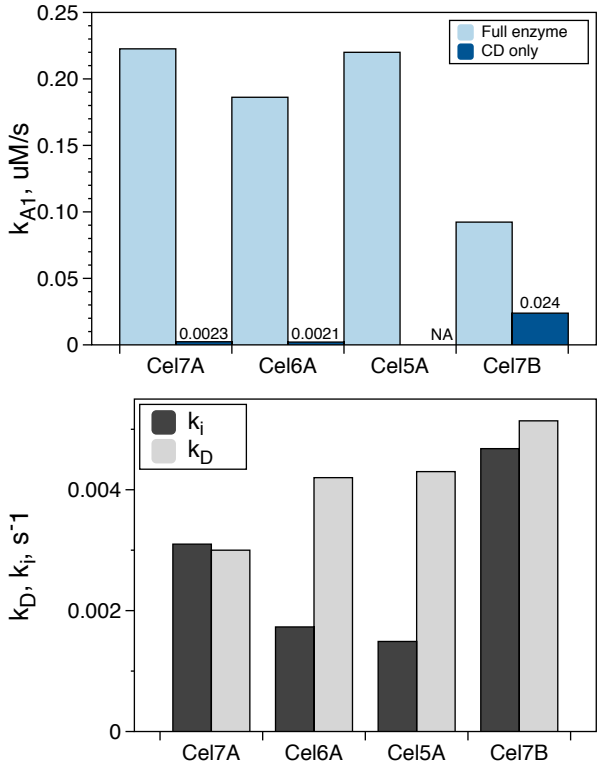

Figure 3.15: Graphical comparison of k_{A1} among the four full-length enzymes and their binding domains

The diversity in k_A seen among the four full-length enzymes is small compared to the difference between the full-length enzymes and their catalytic domains, and between the individual cellulase catalytic domains. The full-length constructs have up to 100-fold higher adsorption rate constants than the isolated catalytic domains. However, catalytic core enzymes are not universal in their slow adsorption rate to lignin. Cel7B CD has an adsorption rate ten-fold higher than the core enzymes of Cel7A and Cel6A, despite its homology with Cel7A. We suggest that this may be due to its open catalytic cleft, in comparison to the closed tunnel exhibited by Cel7A. This allows it to bind to bind in the middle of cellulose chains to produce new chain breaks, which its homolog Cel7A can then thread through its substrate-binding tunnel. An overlay of the catalytic domains of Cel7A and Cel7B is provided in Figure 3.16.

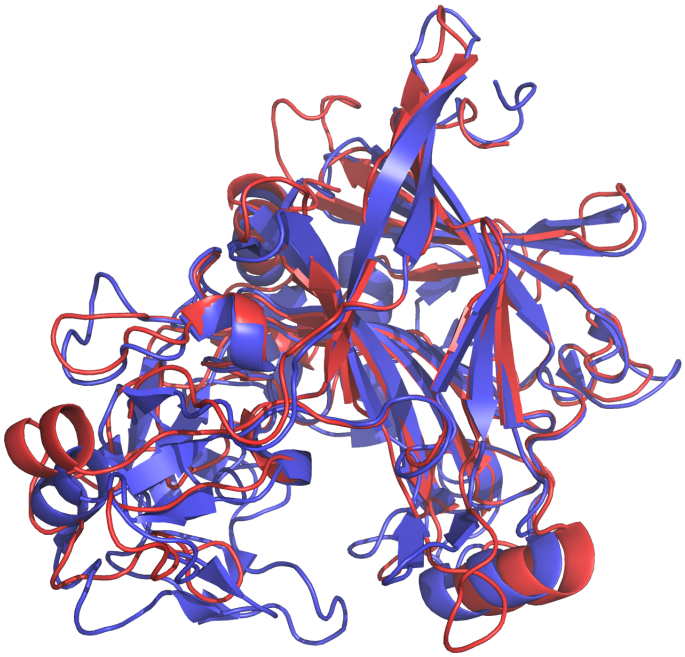

Figure 3.16: Alignment of the catalytic domain crystal structure of *H. jecorina* Cel7A (blue) and Cel7B (red). The binding cleft/tunnel is visible, and several loops that close off the aromatic-residue-rich substrate tunnel in Cel7A are visible.

Does the cellulase catalytic domain determine irreversibility of cellulase-lignin interaction?

The low affinity of cellulase catalytic domains for lignin has led some authors to suggest that under certain circumstances, binding domains may be unnecessary, and impede the recycle and overall conversion of the biomass reaction [19, 34]. However, the high-solids conditions under CBM-free hydrolysis is feasible also presents a challenge in the high concentration of lignin present. Under the long time-scales of a biomass hydrolysis experiment (24-48 hours), even a very low affinity lignin-binder with a significant rate of irreversible binding. This would hinder enzyme recovery and recycle, one of the main advantages of the CBM-free hydrolysis concept.

Both batch binding studies and QCM-D experiments had previously suggested that the CBM is the source of cellulase affinity for lignin. However, our results show that the binding domain might play an important role in the degree of reversibility of the reaction. Figure 3.14 shows reversibility trends for cellulase catalytic domains; qualitatively, the CDs follow the same trends as the full enzyme despite having very different adsorption rates. This could be explained by a surface denaturation mechanism, wherein the CD unfolds on the surface after binding. More information on temperature dependence of the irreversible binding reaction could give us insight into the thermodynamics of this process. This result indicates that successful mitigation of nonproductive lignin binding could involve both the CBM and the catalytic domain - the CBM to decrease the affinity of the enzyme for the surface, and the CD to reduce the rate of irreversible binding to the surface.

3.6 Future work

Evolution and characterization of diverse CBMs and catalytic domains for decreased lignin binding

We've demonstrated that we can measure rate constants for adsorption, desorption, and irreversible adsorption to lignin using QCM-D. Future work in this arena can go in many different directions. Collaborators in the Clark Lab are currently working on mapping the affect of the primary structure of the *Hj*Cel7A binding domain by alanine scanning mutagenesis. Promising mutants that appear to decrease the binding of the protein to lignin can be further characterized using the methods discussed in this work.

We might also consider evolving endoglucanase catalytic domains for decreased lignin affinity. Recent work has shown that at very high solids consistency, hydrolysis of biomass by cellulases without binding domains is as efficient as hydrolysis with intact enzymes, and makes it possible to recycle enzymes, as they are present in high concentration in the hydrolysis supernatant [19, 35, 34]. This approach might be even more successful if the relatively high affinity of endoglucanase catalytic domains were further reduced.

The differences that we measured in lignin adsorption rates of four cellulases was somewhat unexpected, considering the 60% sequence identity in the binding domain. In the natural diversity of CBMs, there is likely to be far greater diversity in lignin affinity than we've seen in the small group of homologous domains that we've examined so far. Figure 1.3 shows the crystal structures of just a small subsection of known CBMs, including CBMs targeted to crystalline, amorphous, and lectin-like structures. Screening of non-Family-1 CBMs may uncover binding domains that provide cellulose affinity without nonproductive lignin adsorption. Evidence that this approach might be fruitful exists in the literature. A 2010 study compared the activity of cellulase chimeras with four CBMs of bacterial and fungal origin. The results show that the constructs' specific activity on amorphous Avicel did not always correlate to the activity on AFEX-pretreated corn stover [36]. The source of this discrepancy was not uncovered - it may have been due to the differences in crystallinity between the substrates, or it could have to do with differential lignin binding in the corn stover hydrolysis. Regardless, the exploration of CBM diversity in terms of both substrate and nonproductive binding should be explored.

pH effect on lignin binding and biomass hydrolysis

The four canonical cellulases from *H. jecorina* each have a pH optimum between 4 and 5, when their activity is measured on pure cellulose or soluble substrates. At alkaline pH, their activity drops off sharply around neutral pH (see Table 3.7).

Table 3.7: Optimum pH and relative activities at pH 6 and 7 of the four *H. jecorina* cellulases and *H. jecorina* β-glucosidase

Enzyme	pH optimum	Relative activity, pH 6	Relative activity, pH 7	Expression host	Reference
Cel7A	5	80%	45%	*H. jecorina*	[37]
Cel6A	5	80%	5%	*H. jecorina*	[38]
Cel7B	5	15%	0%	*E. coli*	[39]
Cel5A	4.8	70%	10%	*H. jecorina*	[40]
β-glucosidase	4 - 4.6	N/A	N/A	*H. jecorina*	[41]

The *H. jecorina* secretome is evolved to work optimally at near pH 5. As such, most biomass hydrolysis experiments are conducted between pH 4.6 and 5.0. However, there is some evidence that pretreated biomass could be more efficiently hydrolyzed at a higher pH. A 2013 study showed that optimal hydrolysis pH of pretreated lodgepole pine with *H. jecorina* enzyme (celluclast) was between 5.2 and 6.2. The authors hypothesized that this increased optimum pH was due to deprotonation of moieties on the lignin surface, reducing nonproductive adsorption to lignin [42]. This hypothesis was supported by a study from

another group, who showed that as pH increases from 4 to 7, the degree to which cellulase adsorbed to an isolated lignin substrate decreased [43]. This was true both for a CBM-containing endoglucanase construct, and an isolated endoglucanse catalytic domain (note that endoglucanase catalytic domains have been shown to adsorb to lignin more strongly than the exoglucanases Cel7A and Cel6A). The affinity for microcrystalline cellulose was unchanged. Furthermore, the authors showed experimentally that increasing the pH of the buffer resulted in an increase in negative charges on the lignin surface [43]. A similar effect was shown with β-glucosidase from *H. jecorina*. The enzyme was depleted from the supernatant by lignin adsorption at pH 4-5, but increasing the pH to above 5.5 resulted in a decrease in lignin adsorption [44]. However, when the hydrolysis reaction was run at pH 5.5, the final conversion was only marginally higher, perhaps due to decreased activity of β-glucosidase at elevated pH.

This is unsurprising: biomass pretreatment often results in free carboxyl groups originating from ferulic acid moieties in the biomass; the resulting carboxylic acid pKa is around 4.6. The other main ionizable surface charge on lignin is the phenolic hydroxyl group with a pKa between 7 and 10, depending on the substitution of the aryl ring. These groups are less likely to be deprotonated to a significant extent under hydrolysis conditions. Pareek and colleagues (2013) showed that as the abundance of surface carboxyl groups on lignin increased, the lignin's affinity for enzyme adsorption decreased [45]. This lends further credence to the hypothesis that charged surface groups have the potential to decrease enzyme adsorption.

In addition to nonproductive binding of lignin, several other mechanisms may be at play in this apparent increase in cellulase efficiency at elevated pH. The pretreatment of biomass liberates small amounts of soluble aromatic compounds, present in concentrations ranging from 500-1300 mg/mL [46]. Reports as to the inhibitory effect of these compounds has been mixed, with some research showing a large inhibitory or denaturing effect, especially on β-glucosidase

The effect of pH on nonproductive adsorption of cellulases to lignin, and on their inhibition by lignin-derived inhibitors, should be explored further. If the existing literature that suggests that increasing the pH of pretreated biomass hydrolysis can increase the efficiency and the final conversion of the reaction is verified, and extended to high solids reactions, using pretreated *Miscanthus* as a substrate, the mechanism for this effect should be studied. *H. jecorina* cellulases can be engineered to shift their pH optimum [38, 40, 47], and such a change may result in higher glucose yields and a more cost-effective process.

Bibliography

[1] B. Henrissat, H. Driguez, C. Viet, and M. Schülein. Synergism of cellulases from Trichoderma reesei in the degradation of cellulose. *Nature Biotechnology*, 3(8):722–726, 1985.

[2] L. Meunier-Goddik and M. H. Penner. Enzyme-catalyzed saccharification of model celluloses in the presence of lignacious residues. *J Agric Food Chem*, 47(1):346–51, 1 1999.

[3] Jenni Rahikainen, Saara Mikander, Kaisa Marjamaa, Tarja Tamminen, Angelos Lappas, Liisa Viikari, and Kristiina Kruus. Inhibition of enzymatic hydrolysis by residual lignins from softwood–study of enzyme binding and inactivation on lignin-rich surface. *Biotechnol Bioeng*, 108(12):2823–34, 12 2011.

[4] Alex Berlin, Mikhail Balakshin, Neil Gilkes, John Kadla, Vera Maximenko, Satoshi Kubo, and Jack Saddler. Inhibition of cellulase, xylanase and beta-glucosidase activities by softwood lignin preparations. *J Biotechnol*, 125(2):198–209, 9 2006.

[5] A. O. Converse, H. Ooshima, and D. S. Burns. Kinetics of enzymatic hydrolysis of lignocellulosic materials based on surface area of cellulose accessible to enzyme and enzyme adsorption on lignin and cellulose. *Applied Biochemistry and biotechnology*, 24(1):67–73, 1990.

[6] Samuel A. Maurer, Claire N. Bedbrook, and Clayton J. Radke. Competitive sorption kinetics of inhibited endo-and exoglucanases on a model cellulose substrate. *Langmuir*, 28(41):14598–14608, 2012.

[7] Jerome M. Fox, Seth E. Levine, Douglas S. Clark, and Harvey W. Blanch. Initial- and processive-cut products reveal cellobiohydrolase rate limitations and the role of companion enzymes. *Biochemistry*, 51(1):442–52, 1 2012.

[8] Seth E. Levine, Jerome M. Fox, Harvey W. Blanch, and Douglas S. Clark. A mechanistic model of the enzymatic hydrolysis of cellulose. *Biotechnology and bioengineering*, 107(1):37–51, 2010.

[9] Per J. Kraulis, G. Marius Clore, Michael Nilges, T. Alwyn Jones, Goeran Pettersson, Jonathan Knowles, and Angela M. Gronenborn. Determination of the three-dimensional solution structure of the C-terminal domain of cellobiohydrolase I from Tichoderma reesei. a study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. *Biochemistry*, 28(18):7241–7257, 1989.

[10] M. Linder, G. Lindeberg, T. Reinikainen, T. T. Teeri, and G. Pettersson. The difference in affinity between two fungal cellulose-binding domains is dominated by a single amino acid substitution. *FEBS Lett*, 372(1):96–8, 9 1995.

[11] G. Carrard and M. Linder. Widely different off rates of two closely related cellulose-binding domains from Trichoderma reesei. *Eur J Biochem*, 262(3):637–43, 6 1999.

[12] P. Tomme, H. Van Tilbeurgh, G. Pettersson, J. Van Damme, J. Vandekerckhove, J. Knowles, T. Teeri, and M. Claeyssens. Studies of the cellulolytic system of Trichoderma reesei QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis. *Eur J Biochem*, 170(3):575–81, 1 1988.

[13] G. Carrard, A. Koivula, H. Söderlund, and P. Béguin. Cellulose-binding domains promote hydrolysis of different sites on crystalline cellulose. *Proc Natl Acad Sci U S A*, 97(19):10342–7, 9 2000.

[14] Hirohide Shiiba, Sachio Hayashi, and Toshifumi Yui. Molecular dynamics study of carbohydrate binding module mutants of fungal cellobiohydrolases. *Carbohydr Res*, 374:96–102, 6 2013.

[15] Hetti Palonen, Folke Tjerneld, Guido Zacchi, and Maija Tenkanen. Adsorption of Trichoderma reesei CBH I and EG II and their catalytic domains on steam pretreated softwood and isolated lignin. *Journal of Biotechnology*, 107(1):65–72, 1 2004.

[16] Raquel Martín-Sampedro, Jenni L. Rahikainen, Leena-Sisko S. Johansson, Kaisa Marjamaa, Janne Laine, Kristiina Kruus, and Orlando J. Rojas. Preferential adsorption and activity of monocomponent cellulases on lignocellulose thin films with varying lignin content. *Biomacromolecules*, 14(4):1231–9, 4 2013.

[17] Deanne W. Sammond, John M. Yarbrough, Elisabeth Mansfield, Yannick J. Bomble, Sarah E. Hobdey, Stephen R. Decker, Larry E. Taylor, Michael G. Resch, Joseph J. Bozell, Michael E. Himmel, Todd B. Vinzant, and Michael F. Crowley. Predicting enzyme adsorption to lignin films by calculating enzyme surface hydrophobicity. *J Biol Chem*, 289(30):20960–9, 7 2014.

[18] Alex Berlin, Neil Gilkes, Arwa Kurabi, Renata Bura, Maobing Tu, Douglas Kilburn, and John Saddler. Weak lignin-binding enzymes: a novel approach to improve activity of cellulases for hydrolysis of lignocellulosics. *Appl Biochem Biotechnol*, 121-124:163–70, 2005.

[19] Anikó Várnai, Matti Siika-Aho, and Liisa Viikari. Carbohydrate-binding modules (CBMs) revisited: reduced amount of water counterbalances the need for CBMs. *Biotechnol Biofuels*, 6(1):30, 2013.

[20] Jenni Liisa Rahikainen, Ulla Moilanen, Susanna Nurmi-Rantala, Angelos Lappas, Anu Koivula, Liisa Viikari, and Kristiina Kruus. Effect of temperature on lignin-derived inhibition studied with three structurally different cellobiohydrolases. *Bioresource Technology*, 146(0):118 – 125, 2013.

[21] Fatma Uzbas, Ugur Sezerman, Lukas Hartl, Christian P. Kubicek, and Bernhard Seiboth. A homologous production system for Trichoderma reesei secreted proteins in a cellulase-free background. *Applied microbiology and biotechnology*, 93(4):1601–1608, 2012.

[22] Daniel G. Gibson, Lei Young, Ray-Yuan Y. Chuang, J. Craig Venter, Clyde A. Hutchison, and Hamilton O. Smith. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods*, 6(5):343–5, 5 2009.

[23] André Schuster, Kenneth S. Bruno, James R. Collett, Scott E. Baker, Bernhard Seiboth, Christian P. Kubicek, and Monika Schmoll. A versatile toolkit for high throughput functional genomics with Trichoderma reesei. *Biotechnol Biofuels*, 5(1):1, 2012.

[24] T. Nakari-Setälä and M. Penttilä. Production of Trichoderma reesei cellulases on glucose-containing media. *Appl Environ Microbiol*, 61(10):3650–5, 10 1995.

[25] Harshal Chokhawala, Christine Roche, Tae Wan Kim, Meera Atreya, Neeraja Vegesna, Craig M. Dana, Harvey W. Blanch, and Douglas S. Clark. Mutagenesis of Trichoderma reesei endoglucanase I: Impact of expression host on activity and stability at elevated temperatures. *In preparation*, 2015.

[26] F. Höök, B. Kasemo, T. Nylander, C. Fant, K. Sott, and H. Elwing. Variations in coupled water, viscoelastic properties, and film thickness of a mefp-1 protein film during adsorption and cross-linking: a quartz crystal microbalance with dissipation monitoring, ellipsometry, and surface plasmon resonance study. *Anal Chem*, 73(24):5796–804, 12 2001.

[27] F. Höök, M. Rodahl, B. Kasemo, and P. Brzezinski. Structural changes in hemoglobin during adsorption to solid surfaces: effects of pH, ionic strength, and ligand binding. *Proc Natl Acad Sci U S A*, 95(21):12271–6, 10 1998.

[28] F. Höök, M. Rodahl, P. Brzezinski, and B. Kasemo. Energy dissipation kinetics for protein and antibody-antigen adsorption under shear oscillation on a quartz crystal microbalance. *Langmuir*, 14(4):729–734, 1998.

[29] P. M. Abuja, M. Schmuck, I. Pilz, P. Tomme, M. Claeyssens, and H. Esterbauer. Structural and functional domains of cellobiohydrolase I from Trichoderma reesei. *European Biophysics Journal*, 15(6):339–342, 1988.

[30] P. M. Abuja, I. Pilz, M. Claeyssens, and P. Tomme. Domain structure of cellobiohydrolase II as studied by small angle x-ray scattering: close resemblance to cellobiohydrolase I. *Biochem Biophys Res Commun*, 156(1):180–5, 10 1988.

[31] Véronique Receveur, Mirjam Czjzek, Martin Schülein, Pierre Panine, and Bernard Henrissat. Dimension, shape, and conformational flexibility of a two domain fungal cellulase in solution probed by small angle x-ray scattering. *J Biol Chem*, 277(43):40887–92, 10 2002.

[32] J. Talbot, G. Tarjus, P. R. Van Tassel, and P. Viot. From car parking to protein adsorption: an overview of sequential adsorption processes. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 165(1):287–324, 2000.

[33] P. Schaaf, Ph Déjardin, A. Johner, and A. Schmitt. Characteristic time scales for the adsorption process for fibrinogen on silica. *Langmuir*, 8(2):514–517, 1992.

[34] Annukka Pakarinen, Mai Ostergaard Haven, Demi Tristan Djajadi, Anikó Várnai, Terhi Puranen, and Liisa Viikari. Cellulases without carbohydrate-binding modules in high consistency ethanol production process. *Biotechnol Biofuels*, 7(1):27, 2014.

[35] Dahai Gao, Shishir PS Chundawat, Anurag Sethi, Venkatesh Balan, S. Gnanakaran, and Bruce E. Dale. Increased enzyme binding to substrate is not necessary for more efficient cellulose hydrolysis. *Proceedings of the National Academy of Sciences*, 2013.

[36] Tae-Wan W. Kim, Harshal A. Chokhawala, Dana C. Nadler, Dana Nadler, Harvey W. Blanch, and Douglas S. Clark. Binding modules alter the activity of chimeric cellulases: Effects of biomass pretreatment and enzyme source. *Biotechnol Bioeng*, 107(4):601–11, 11 2010.

[37] Harry Boer and Anu Koivula. The relationship between thermal stability and pH optimum studied with wild-type and mutant Trichoderma reesei cellobiohydrolase Cel7A. *Eur J Biochem*, 270(5):841–8, 3 2003.

[38] Pete Heinzelman, Christopher D. Snow, Indira Wu, Catherine Nguyen, Alan Villalobos, Sridhar Govindarajan, Jeremy Minshull, and Frances H. Arnold. A family of thermostable fungal cellulases created by structure-guided recombination. *Proc Natl Acad Sci U S A*, 106(14):5610–5, 4 2009.

[39] Hikaru Nakazawa, Katsunori Okada, Ryota Kobayashi, Tetsuya Kubota, Tomoko Onodera, Nobuhiro Ochiai, Naoki Omata, Wataru Ogasawara, Hirofumi Okada, and Yasushi Morikawa. Characterization of the catalytic domains of Trichoderma reesei en-

doglucanase I, II, and III, expressed in Escherichia coli. *Appl Microbiol Biotechnol*, 81(4):681–9, 12 2008.

[40] Yuqi Qin, Xiaomin Wei, Xin Song, and Yinbo Qu. Engineering endoglucanase II from Trichoderma reesei to improve the catalytic efficiency at a higher pH optimum. *Journal of biotechnology*, 135(2):190–195, 2008.

[41] H. Chen, M. Hayn, and H. Esterbauer. Purification and characterization of two extracellular beta-glucosidases from Trichoderma reesei. *Biochim Biophys Acta*, 1121(1-2):54–60, 5 1992.

[42] T. Q. Lan, Hongming Lou, and J. Y. Zhu. Enzymatic saccharification of lignocelluloses should be conducted at elevated pH 5.2–6.2. *BioEnergy Research*, 6(2):476–485, 2013.

[43] Jenni Liisa Rahikainen, James David Evans, Saara Mikander, Anna Kalliola, Terhi Puranen, Tarja Tamminen, Kaisa Marjamaa, and Kristiina Kruus. Cellulase-lignin interactions–the role carbohydrate-binding module and pH in non-productive binding. *Enzyme and Microbial Technology*, 2013.

[44] Ja Kyong Ko, Eduardo Ximenes, Youngmi Kim, and Michael R. Ladisch. Adsorption of enzyme onto lignins of liquid hot water pretreated hardwoods. *Biotechnol Bioeng*, 8 2014.

[45] Nidhi Pareek, Thomas Gillgren, and Leif J. Jönsson. Adsorption of proteins involved in hydrolysis of lignocellulose on lignins and hemicelluloses. *Bioresour Technol*, 148:70–7, 11 2013.

[46] Youngmi Kim, Eduardo Ximenes, Nathan S. Mosier, and Michael R. Ladisch. Soluble inhibitors/deactivators of cellulase enzymes from lignocellulosic biomass. *Enzyme Microb Technol*, 48(4-5):408–15, 4 2011.

[47] Dieter Becker, C. Braet, I. I. I. H. Brumer, Marc Claeyssens, Christina Divne, B. Fagerstrom, Mark Harris, T. Jones, G. Kleywegt, and Anu Koivula. Engineering of a glycosidase family 7 cellobiohydrolase to more alkaline pH optimum: the pH behaviour of Trichoderma reesei Cel7A and its E223S/A224H/L225V/T226A/D262G mutant. *Biochem. J*, 356:19–30, 2001.

Chapter 4

Lignin decomposition: laccase-mediated α-oxidation and base-catalyzed cleavage

4.1 Abstract

The work presented in this chapter presents a process envisioned to incorporate lignin oxidation, cleavage, and extraction into a volatile organic solvent into an ionic liquid pretreatment process. Nonphenoilc β-O-4 linkages, although common in lignin, are not broken under common pretreatment conditions. This work demonstrates that α-oxidation of a model β-O-4 compound renders it labile under basic aqueous conditions, but also under conditions present during ionic liquid (IL) pretreatment of biomass using the basic ionic liquid 1-ethyl 3-methyl imidazolium acetate ([EMIM][OAc]). A non oxidized analog of the same model compound is resistant to both the IL and base-catalyzed breakdown. We present a process by which lignin oxidation could be incorporated into an ionic-liquid based biomass pretreatment process, and begin to map the limitations and requirements of an enzymatic lignin oxidation scheme to fit this process.

The complete oxidation of a non phenolic β-O-4 lignin model dimer requires up to a ten-fold excess mediator concentration for complete α-oxidation. The laccase-mediator system is only slightly effective in oxidizing non phenolic lignin models in the presence of moderate (10-20% w/v) concentrations of [EMIM][OAc]. This is likely due both to decreased stability of the oxidized mediator, and reduced enzyme activity and stability in the presence of ionic liquid.

Under conditions similar to those effective in oxidizing the model dimer, we have attempted to oxidize purified organosolv lignin. The results indicate that while either polymerization or an increase in charged groups on the lignin surface make the lignin less extractable from an ionic liquid solution into a THF phase, there is no change in small lignin-derived fragments detectable by gel permeation chromatography or GC/MS. This may be due to lignin's

low solubility under the reaction conditions, rendering it inaccessible to bulk oxidation.

4.2 Introduction

Lignocellulosic biomass is a potential source of renewable energy for transportation fuels through the conversion of fermentable sugars derived from cellulose and hemicellulose to alcohols. The highly crystalline nature of cellulose, along with the presence of lignin surrounding the cellulose microfibrils, contribute to the slow enzymatic hydrolysis of cellulosic biomass and necessitate pretreatment to increase hydrolysis rates. While mechanical milling, weak and strong acid treatment, ammonia fiber expansion, and other pretreatment methods increase hydrolysis rates, by far the most effective pretreatment devised so far is the dissolution and subsequent regeneration of cellulose in ionic liquids (IL) [1]. This method, which employs polar, imidazolium-based ionic liquids such as 1-ethyl-3-methylimidazolium acetate ([EMIM][OAc]), results in high hydrolysis rates at low enzyme loadings, due to a combination of cellulose decrystallization and physical separation of the lignin and cellulose components of the biomass [2, 3]. In ionic liquid pretreatment, the whole biomass is dissolved in ionic liquid, followed by precipitation by the addition of water or another polar solvent. The solid cellulose fraction is recovered, washed, and hydrolyzed in an aqueous environment by cellulase enzymes; the liquid fraction, composed of soluble lignin, water, and ionic liquid, is dried and recycled to dissolve more biomass [1].

While ionic liquid pretreatment is a promising technology for lignocellulose processing, its commercialization is hindered by the high cost of ILs. The efficient recycling of ionic liquids, as well as the valorization of the lignin byproduct of biomass deconstruction, is essential to the large-scale success of IL pretreatment processes [4]. When polysaccharides are precipitated from an ionic liquid-biomass solution, lignin remains in solution (up to 50% of the initial lignin content is recovered in the IL-water supernatant [1, 5]. If this lignin could be recovered, either as a macromolecule or after its fragmentation, it would not only facilitate IL recycle, but would also be a source of high-quality lignin. Higher value uses for lignin include gasification to mixed alcohols, conversion to aromatic chemicals or high-octane fuel additives, and the production of lignin-based polymers and carbon fiber [6, 7, 8].

The extraction of organosolv lignin and lignin-derived monomers from ionic liquids by organic solvents has been shown to be possible route to recovering lignin and its breakdown products from a biomass pretreatment reaction [9]. [EMIM][OAc] forms a two-phase system with water and the organic solvents tetrahydrafuran, dioxane, and ethyl acetate, and both lignin and monolignols can be extracted from the IL phase, with small molecules extracting more efficiently from the IL solution than large molecular weight lignin [9].

The accumulation of soluble lignin in ionic liquid pretreatment ionic liquid presents an opportunity to break down dissolved lignin, and extract it into an organic solvent for further processing. We envision a process in which biomass is pretreated, solublizing lignin. After the precipitation of the cellulose fraction, the soluble lignin is oxidized. α-oxidized lignin model compounds break down easily in [EMIM][OAc] under pretreatment conditions, thus the

oxidized lignin is recycled with the IL back into the biomass pretreatment vessel. Following pretreatment, the oxidized lignin would be cleaved, and small molecules could be extracted with an organic solvent, and the process repeated. Figure 4.1 gives an overview of a potential process for the oxidation and extraction of lignin.

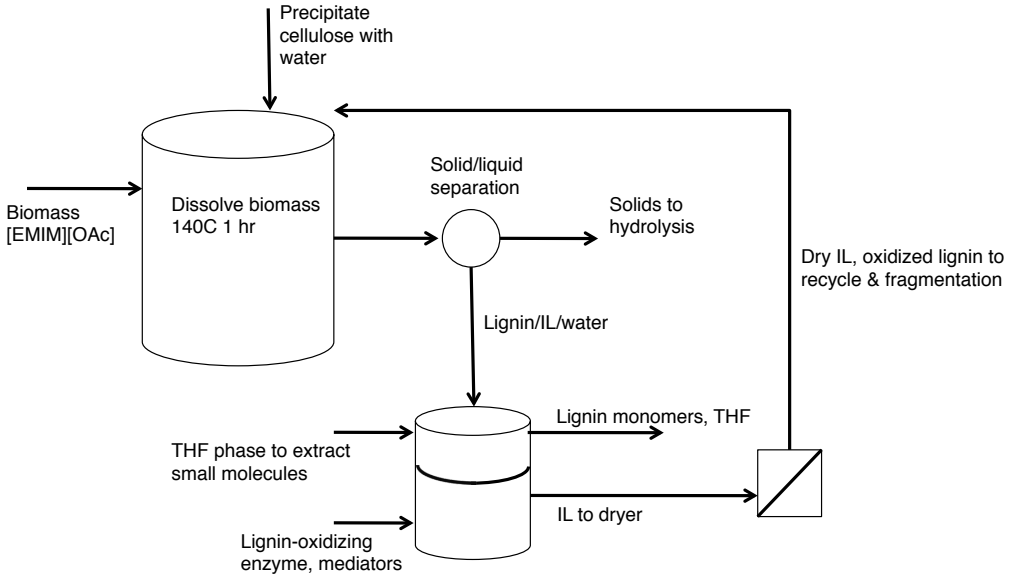

Figure 4.1: A proposed process for the oxidation and extraction of lignin as part of an ionic liquid pretreatment process. This chapter addresses the α-oxidation of lignin model dimers, and the breakdown of oxidized lignin models in the presence of the ionic liquid [EMIM][OAc].

Laccase-mediator system for oxidation of lignin

Fungal lignin deconstruction pathways are only partially understood, but they are known to involve multiple enzymes, including lignin peroxidase (LiP) and manganese peroxidase (MnP), both heme-containing enzyme that use H_2O_2 as a cosubstrate and terminal electron acceptor, as well as the copper-containing monooxygenase, laccase [10]. LiP and MnP are both capable of oxidizing high-redox substrates, including nonphenolic aryl rings like dimethoxybenzene, either directly or via oxidized Mn.

Laccase is a copper-containing oxidase (EC 1.10.3.2) that catalyzes the reduction of molecular oxygen to water, and the single-electron oxidation of their substrates, most notably phenolic hydroxyls. It is found in both lignin synthesis pathways in plants, where it catalyzes the free-radical polymerization of monolignols, and lignin breakdown pathways in fungi and bacteria [11].

Although laccases lack the ability to oxidize non phenolic aromatic substrates directly, fungal laccases can oxidize these high-redox substrates through a redox mediator, which can

be oxidized by the laccase and diffuse to the non phenolic lignin substrate [12]. Figure 4.2 is a schematic for potential reactions between laccase, lignin, and a radical mediator ABTS. On both an industrial and a lab scale, laccase is a more tractable enzyme system than LiP and MnP. It uses O_2 as a cofactor (rather than H_2O_2, as in LiP and MnP, which would be easier to deliver on an industrial scale. Laccase expression might also be less complicated, as it does not require the coordination of a heme group, instead the four copper atoms are coordinated within the protein structure. For these reasons, we chose to focus on laccase and laccase-mediator systems to oxidize lignin.

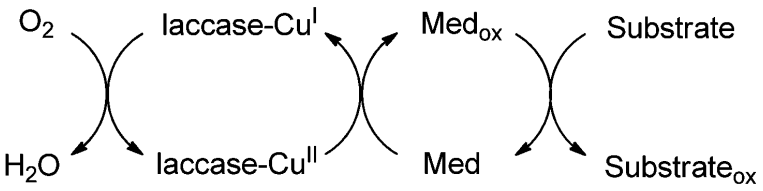

Figure 4.2: Schematic for reactions between lignin, laccase, and a radical mediator. Adapted from [12].

Base-catalyzed breakdown of non-phenolic β-O-4 bonds in lignin

Nonphenolic β-O-4 bonds make up $50 - 60\%$ of the linkages in lignin [13]. As such, the model dimer compound shown in Figure 4.3 is a good small-modecule model for lignin.

Lignin oxidation to render it more vulnerable to base-catalyzed depolymerization and nucleophilic attack has been studied in the pulp and paper industry for pulp delignification since the early 80's [14]. Oxidation of α-hydroxyl to carbonyl groups has been shown to increase the rate of β-O-4 bond breakage during kraft pulping [15].

The work presented in this chapter demonstrates that α-oxidation of a model β-O-4 compound renders it labile under aqueous conditions, but also under conditions present during ionic liquid (IL) pretreatment of biomass using the basic ionic liquid 1-ethyl 3-methyl imidazolium acetate ([EMIM][OAc]). A non-oxidized analog of the same model compound is resistant to both the IL an base-catalyzed breakdown. We present a process by which lignin oxidation could be incorporated into an ionic-liquid based biomass pretreatment process, and begin to understand the limitations and requirements of an enzymatic lignin oxidation scheme to fit this process.

4.3 Methods

Materials

All solvents and chemicals were purchased from Sigma and were analytical grade. Laccase from *Tramedes versicolor* was purchased from Sigma (catalog number 53739, 15U/mg, lyophilized powder) and stored at -20°C. Lignin model compounds were synthesized by Dr. Julian Chan as previously published [16].

Lignin was purified using an organosolv method. Dioxane and ethanol extracted lignin were isolated from *Miscanthus x gianteus* by the following protocol. Dried, ground biomass (2mm screen Retsch mill) was washed with water and solvent (three washes, water followed by dioxane or ethanol) at 100°C for 5 minutes and dried. The dried biomass was milled to 0.5 mM and added at 1:10 mass ratio to 0.2M HCL/95% v/v solvent (dioxane or ethanol), and billed for 1 hour under reflux. Biomass was cooled to room temperature and filtered (grade GF/B glass microfiber, Whatman, Clifton, NJ) and washed with solvent four times. Filtrate was transferred into stirred, acidified water (0.02 M HCl, 20 mL/g starting biomass, let stand overnight, and collected by centrifugation (3200g, 10 min). Precipitate was washed with water and freeze dried. Procedure was carried out by Dr. Stefan Bauer at EBI.

Mediators and substrates for oxidation are shown in Figures 4.3 and 4.4. Activity was assayed in acetate buffer, pH 4.5, at 37°C, with agitation and under oxygen to ensure that the O_2 cofactor did not limit the reaction rate. Where necessary, substrate was dissolved in THF, aliquotted to the correct final concentration in the reaction, and the THF dried in an oven at 37°C. Buffer containing the mediator at the required concentration and the enzyme (0.3 U/mL) was added to the substrate, which was agitated to dissolve the substrate. Samples were taken at the indicated time intervals, and diluted into 50% methanol, 0.2 um filtered, and analyzed by HPLC.

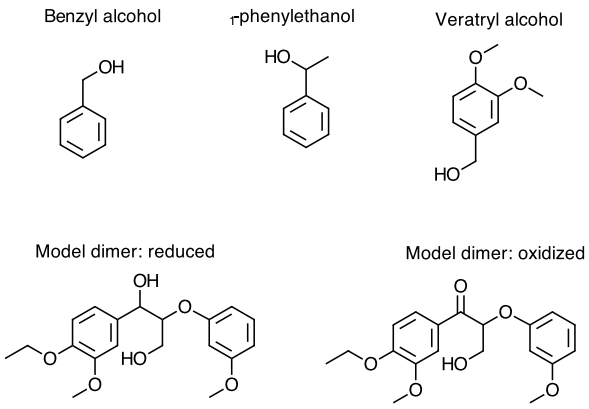

Figure 4.3: Chemical structures for model lignin substrates used in this work.

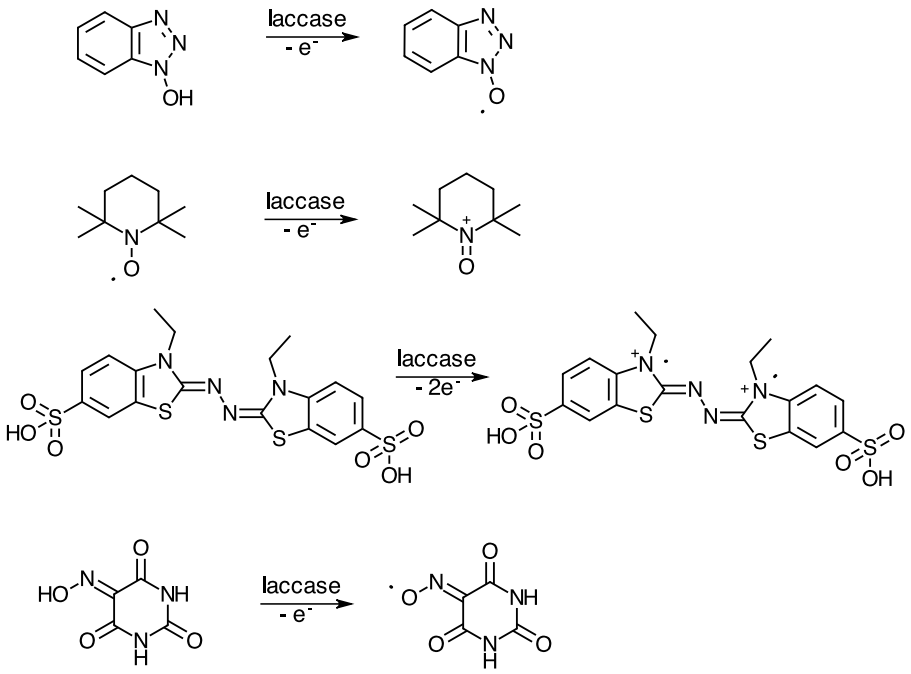

Figure 4.4: Chemical structures for radical mediators used in this work, along with their oxidation products (active radical mediators). From top: hydroxybenzotriazole (HBT), TEMPO, 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) (ABTS), violuric acid (VLA).

Lignin oxidation and IL-mediated deconstruction

The process detailed in Figure 4.1 relies upon a laccase-mediator system, preferably working under IL or solvent conditions in which lignin is soluble, to oxidize the α-hydroxyl groups on non phenolic β-aryl-ether linkages. We show that oxidation renders model dimers susceptible to cleavage under mild conditions in the presence of NaOH or neat [EMIM][OAc]. Therefore, to test the system on lignin polymer, we first oxidize the lignin (under conditions shown to result in full oxidation of model dimers), dry the lignin samples, and add 100% [EMIM][OAc] to the lignin residue and incubate at 80°C for 8 hours to ensure full cleavage of the oxidized linkages. Then, having shown that both organosolv lignin and small lignin-derived models can be extracted from a 1:1:2 v:v:v two-phase system of [EMIM][OAc], water, and THF, the lignin and small molecules are extracted into the THF phase, where they can be analyzed by gel permeation chromatography for molecular weight, or by GC/MS for the presence of new volatile fragments.

Analytical techniques

HPLC samples were analyzed on a 1200 series Agilent HPLC with a reverse-phase Alltech Prevail C18 3 m 3.0x150 mm column with a 0.4 ml/min flow rate and a gradient of 50 ? 90% methanol over 10 min with a constant 0.05% concentration of formic acid.

Gel permeation chromatography was used to approximate the molecular weight of isolated lignin. PL-GPC 50 from A Varian, Inc. with two MesoPore Columns (each 300mm length, 7.5mm diameter) and a UV-vis detector at 280nm were used. The flow rate was 1.0 mL/min; THF was the sample-containing mobile phase. The column temperature was 30°C. Polystyrene standards were used for mass calibration.

GC/MS analysis was done using a Varian CP-3800 gas chromatograph equipped with Varian 320-MS mass spectrometer, as described previously [9].

4.4 Results

Base-catalyzed breakdown of α-oxidized β-O-4 model compound

Previous research has shown that α-oxidation of a non phenolic β-O-4 model dimer such as that shown in Figure 4.3 is vulnerable to base-catalyzed cleavage [14, 15]. [EMIM][OAc], a room-temperature ionic liquid, absorbs water and results in a solution with a pH between 8 and 10. While pH is difficult to measure at low water concentrations, the high pH results in deprotonation of phenolic hydroxyls and other ionizable groups in lignin, and increases its solubility [17, 5].

When the model compounds shown in Figure 4.3 are incubated for several hours at 80°C, it's clear that breakdown of the oxidized model proceeds quickly, while the α-hydroxyl version of the dimer remains stable in solution. Dimer concentration was measured by HPLC.

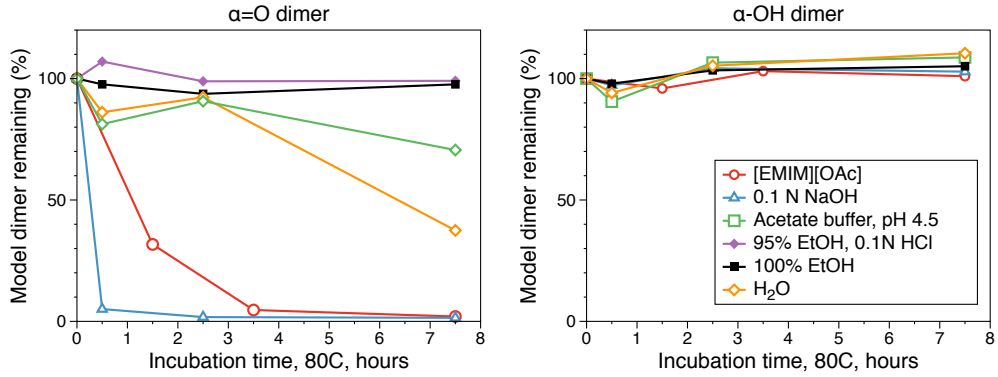

Figure 4.5: Nonphenolic lignin dimers incubated in a variety of solvents, including [EMIM][OAc], dilute base, and water. The α-hydroxyl model is stable in each solvent tested, while the α-carbonyl model degrades in base, IL, and water. The figure legend applies to both sides of the figure.

The products of base-catalyzed hydrolysis of the oxidized dimer shown in Figure 4.3 include guaiacol, according to the reaction mechanisms proposed in the literature [15]. When the products of the ionic liquid decomposition of the non phenolic model dimer are analyzed by HPLC, a peak corresponding to the retention time observed from guaiacol forms concurrently with the disappearance of the model dimer peak. However, when the same reaction takes place in NaOH, very little guaiacol is observed. Instead, a white precipitate was formed that was insoluble in the methanol-water mobile phase of the HPLC. This may be due to secondary reactions that cause liberated fragments to form condensed structures, as occurs during the base-catalyzed deconstruction of lignin during paper pulping [18].

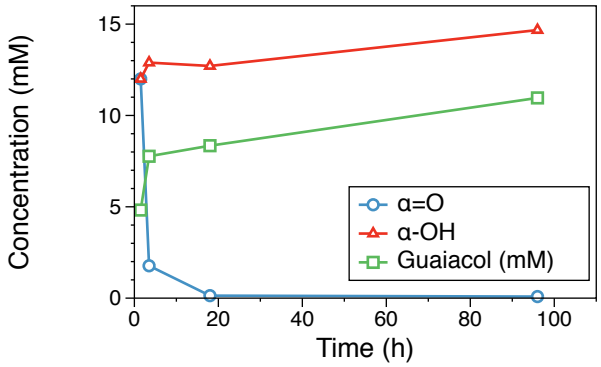

Figure 4.6: The reaction products of the IL incubation of the α-oxidized model dimer include guaiacol, produced in near-stoichiometric amounts.

In order to determine whether lignin fragments could be extracted from an IL/water mixture as suggested in Figure 4.1, we extracted several lignin-like aromatic monomer models, including vanillin, 2-methoxy 4-methylphenol, 4-methyl catechol, and vanillic acid, at a concentration of 15 mM. This concentration was chosen by considering a process with a 5% w/w solids loading in pretreatment, 20% lignin content in biomass, 50% recovery of lignin in the supernatant after cellulose regeneration, and an arbitrary 10% conversion of soluble lignin to the model vanillin. The resulting concentration in this hypothetical process would be about 15 mM monomer. Several concentrations of [EMIM][OAc] and pH were tested. The concentration of the analyte in the extract phase was tested by HPLC. The data show that the monomers extract efficiently, with the exception of the vanillic acid. The pKa of vanillic acid is about 4.5; it is fully deprotonated at pH 8 and thus would have limited solubility in THF.

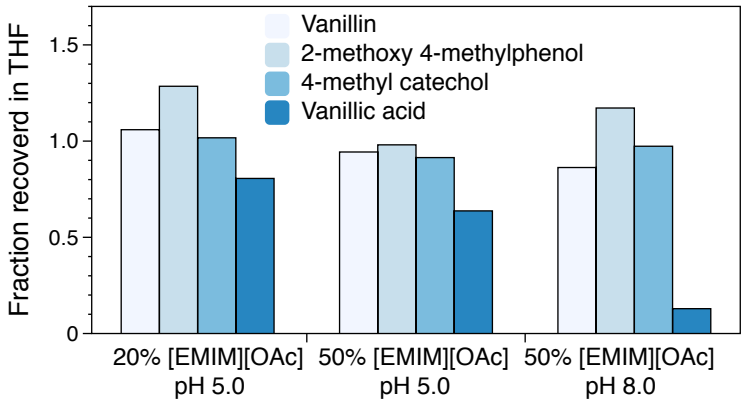

Figure 4.7: Extraction of lignin-like model aromatic monomers to demonstrate that the extraction of lignin breakdown products is possible using the THF/water/IL extraction discussed in the Methods section.

Laccase-catalyzed oxidation of non phenolic lignin models

There are many small-molecule radical mediators upon which laccase is active. We have tested several of them for activity as part of a laccase mediator system to oxidize non phenolic lignin models, and isolated lignin. We describe a process (see Figure 4.1) in which the lignin is dissolved in an ionic liquid/water solution, so we have also tested the activity of the LMS in both in buffer solutions and in solutions containing [EMIM][OAc].

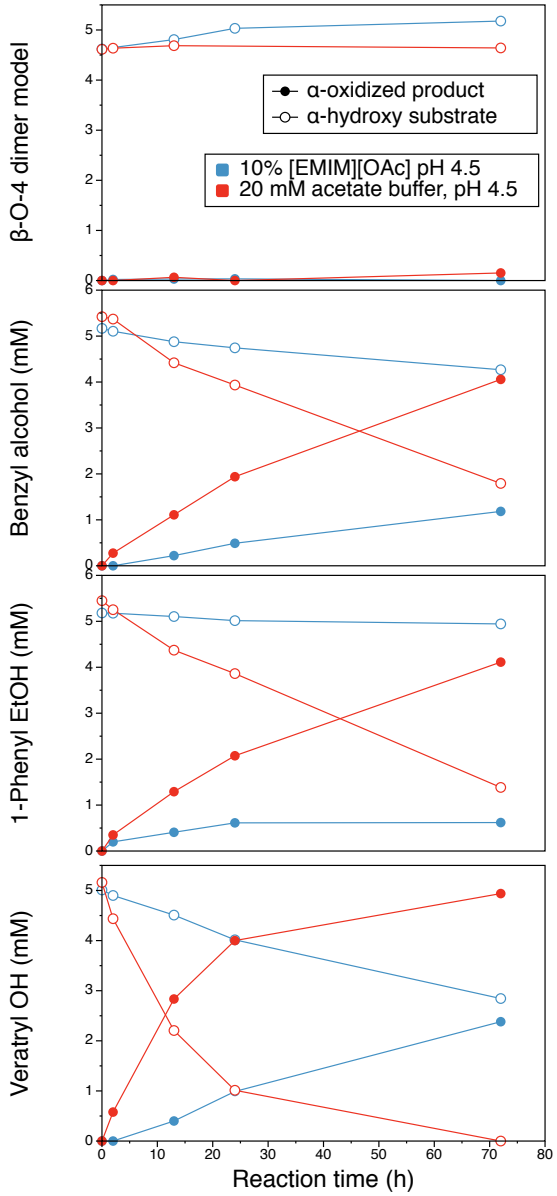

Figure 4.8: Laccase-TEMPO oxidation of four non phenolic lignin model substrates.

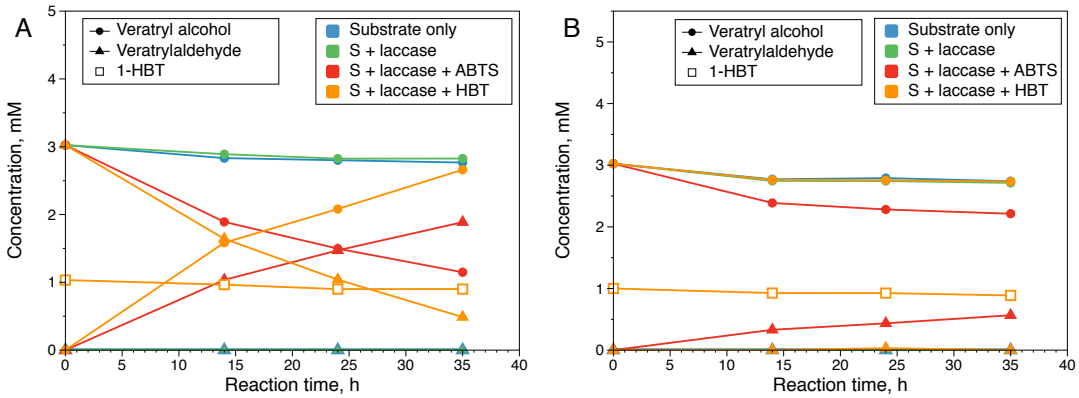

Figure 4.9: Laccase-mediator oxidation of veratryl alcohol, in pH 4.5 acetate buffer (A) and 20% [EMIM][OAc], pH 4.5 (B). Mediators used are 1-HBT, and ABTS (1 mM). 1-HBT concentration is monitored throughout the reaction and its concentration does not change. The LMS is able to catalyze the oxidation of the model monomers, both in the presence of buffer and 20% IL, but the oxidation of the model dimer was only detectible in buffer, and to a very limited extent. The oxidized product peak was detectable by HPLC.

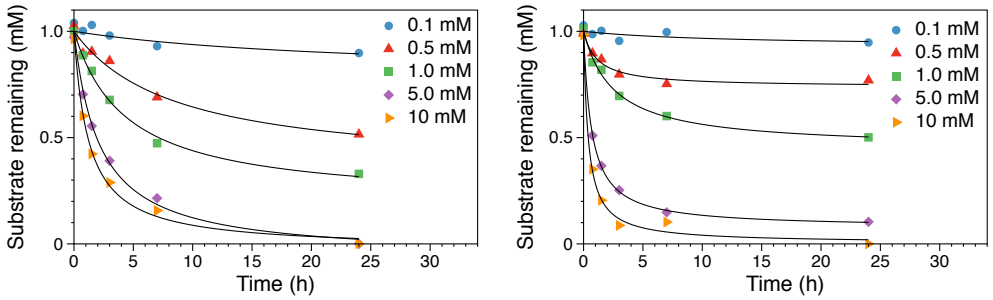

Figure 4.10: The effect of increasing mediator concentrations on the oxidation of non phenolic lignin dimer model. HBT (left) and VLA (right) are used as mediators. Both require greater than 10-fold more mediator than substrate to achieve full oxidation of the substrate.

While all combinations of radical mediators, mediator concentrations, substrates, and solvents were not investigated, the data clearly show that the model dimer is more resistant to oxidation by the LMS than the monomers. Further, the reaction is very limited in the presence of ionic liquids. Full oxidation of the model dimer was achieved, however only in the presence of 5-10x excess mediator concentration, and in a buffer solution. In order to try to demonstrate a proof of concept for the process we described in Figure 4.1, we tried to apply the conditions used in Figure 4.10 to oxidize organosolv lignin, followed by incubation in [EMIM][OAc] to break oxidized linkages.

Laccase-catalyzed oxidation of organosolv lignin

Having found conditions under which, industrially impractical as they may be, α-oxidation of non phenolic β-O-4 model dimer linkages occurs, we attempted to apply these conditions (buffer, pH 4.5, 10:1 mediator to substrate ratio) to model lignin. Organosolv lignin (both ethanol- and dioxane-extracted lignin) at 1 mg/mL was reacted with laccase (3 U/mL). After incubation at 37°C for 24 hours, the lignin is lyophilized, 5 mL [EMIM][OAc] added, incubated for 8 hours at 80°C simulate pretreatment conditions and to break any labile, oxidized β-O-4 bonds, and finally, the lignin is extracted into THF by adding 5 mL water, 10 mL THF, and extracting. The THF extract was filtered and analyzed by GPC to compare the relative molecular weight.

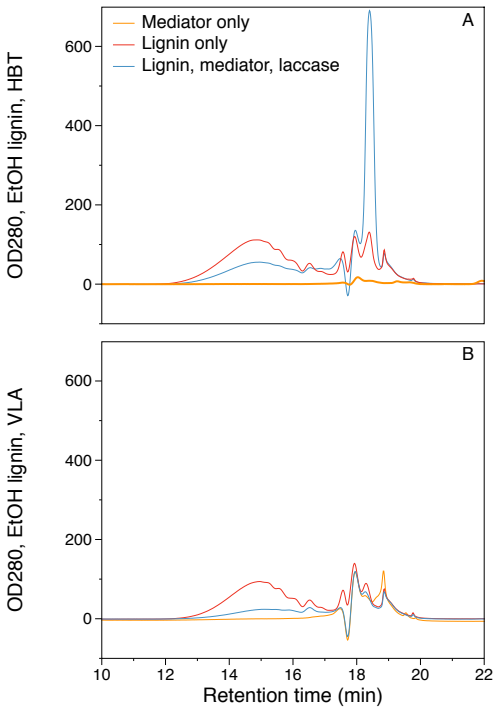

Figure 4.11: Comparison of GPC molecular weight profiles of extracted lignin treated with laccase-mediator system, mediator alone, and lignin alone. A: ethanol-extracted lignin treated with LMS including the mediator hydroxybenzotriazole. B: ethanol-extracted lignin treated with LMS including mediator violuric acid. In A, the HBT mediator sample (orange) does not give the same low-molecular weight peak seen in the LMS-lignin sample (blue). In B. the low molecular weight peaks observed appear to be artifacts of violuric acid addition.

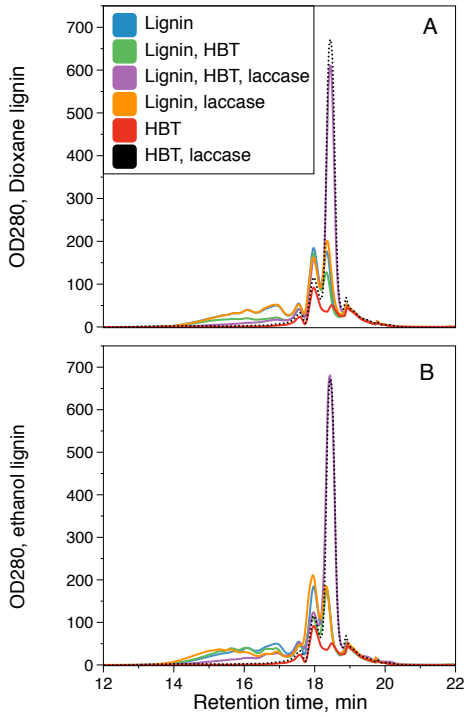

Figure 4.12: Comparison of LMS with the mediator hydroxybenzotriazole on organosolv lignin with oxidized mediator controls: lignin alone (blue), lignin/mediator (orange), lignin/mediator/laccase (purple), lignin/laccase (orange), mediator alone (red), mediator/laccase (black). The comparison shows that unlike the VLA mediator, non oxidized HBT does not give a small-molecule sized OD280 peak on the GPC trace. However, when HBT is oxidized by laccase, it produces the large artifact that is also seen in the lignin/LMS samples. No lignin degradation is detected that cannot be explained by the mediator.

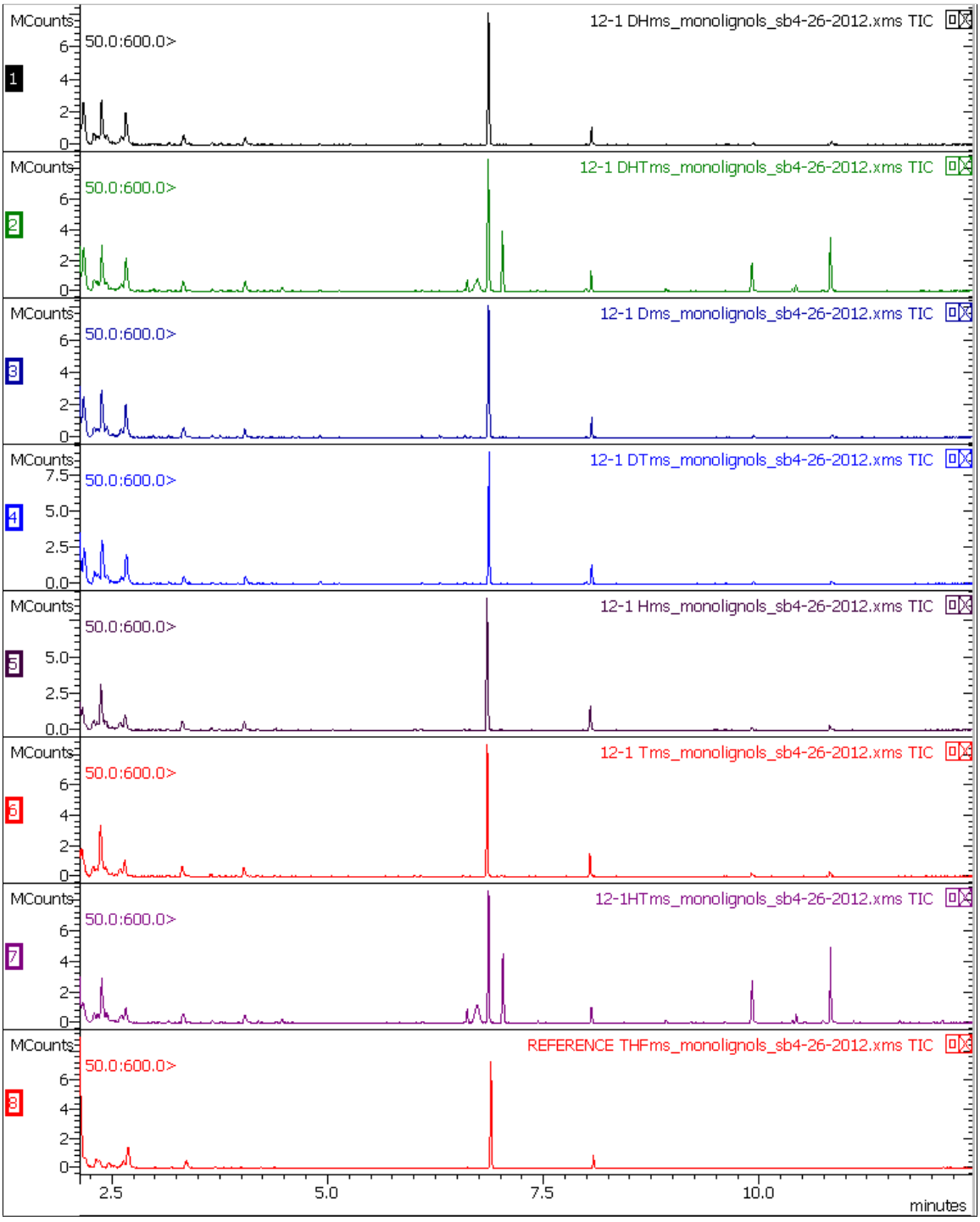

Figure 4.13: GC/MS of the extract of the lignin/LMS reactions and their controls. A: lignin, HBT; B: lignin, HBT, laccase; C: lignin only; D: lignin, laccase; E: HBT alone; F: laccase alone; G: HBT, laccase; H: THF solvent background (extractant). Comparison of each of these controls reveals that no new peaks are produced by the laccase/mediator system that is not produced by the oxidized mediator.

4.5 Discussion

Cleavage of oxidized lignin during ionic liquid pretreatment

Oxidation of non phenolic lignin modes dimers

The oxidation of non phenolic β-O-4 linkages at the α-hydroxyl position makes the dimer vulnerable to cleavage into monomers, which have been partially characterized. In the presence of NaOH, an insoluble precipitate forms, whereas in neat [EMIM][OAc] at moderate temperature, near-stoichiometric amounts gf guaiacol is formed from the cleavage of the oxidized dimer shown in Figure 4.3. This is an expected product of the base-catalyzed hydrolysis of this model [15]. Our goal was to incorporate the ability of [EMIM][OAc] to deconstruct oxidized lignin with a pretreatment process using [EMIM][OAc], as it is a well-known solvent for dissolving and whole biomass, rendering it more susceptible to enzymatic hydrolysis [3, 5].

Using the laccase mediator system with several different mediators, it is possible to oxidize non phenolic β-O-4 linkages at the α-hydroxyl position. However, the data collected indicates that oxidation of a model dimer is more difficult than the model monomers that are often used to model this reaction, such as veratryl alcohol. Oxidation of the non phenolic dimer shown in Figure 4.3 requires very high mediator concentration, in excess of the concentration of the substrate. In order for mediator to be used as a reusable catalyst for substrate oxidation, as is hypothesized in Figure 4.2, it must not be consumed in the reaction. While the work done so far can't show whether the mediator has been consumed or is merely required in high concentration to effectively oxidize the substrate, a mediator concentration in excess of the substrate concentration would be difficult to justify industrially. There are many combinations of mediator, laccase, and substrate, so it's possible that a more thorough investigation into this area would prove fruitful.

Oxidation of organosolv lignin with LMS

When we attempted to oxidize organosolv lignin using the laccase mediator system under conditions that produced full oxidation of the model dimer, we found what appeared to be low-molecular weight compounds when the lignin was LMS-oxidized, [EMIM][OAc] treated, extracted with THF, and analyzed using gel permeation chromatography. However, further controls indicated that the source of the small molecular weight compounds was the mediator extracting into the THF; these results were confirmed when no low molecular weight compounds were identified in the extract using GC/MS.

Another trend that is visible in the GPC results is the decrease of early-eluting, high molecular weight lignin following oxidation by the laccase mediator system. There are several plausible explanations for this effect. It's unlikely that this decrease is due to fragmentation of the lignin into smaller compounds, as all of the low molecular weight peaks appear to be artifacts of the mediator. The decrease in the intensity of the high molecular weight

lignin may be due to polymerization of the lignin that renders it less extractable [9]. It may also be due to an increase in carboxylic acid functionality in the lignin. Oxidation producing carboxylic acids has been detected from laccase mediator systems [19], and we have shown that lignin model monomers containing carboxylic acid functionality are likely to be deprotonated and not extracted in the THF/IL/water system used in this work (see Figure 4.7). However, we also know that both the LMS and laccase alone can oxidize phenolic hydroxyl groups on lignin, and the resulting radical will delocalize and undergo radical coupling to form new lignin linkages (this is the mechanism by which lignin is polymerized in the plant cell wall). It's likely that this process increases the molecular weight of isolated lignin as well, decreasing its extractability.

4.6 Future work

The breakdown of oxidized lignin during ionic liquid pretreatment is a promising addition to IL-pretreatment's many benefits over other mechanical and chemical pretreatment methods. We've shown that in its current form, laccase mediator systems are either unable to catalyze the oxidation of α-hydroxyl groups in β-O-4 linkages, or the breakdown of these oxidized linkages in model dimers don't extend to the lignin polymer.

There is ample reason to believe that the problem with this system is the former. Under conditions in which lignin is soluble in IL (>50% [EMIM][OAc], pH >8), laccase and the laccase mediator system is unable to catalyze the oxidation of even monomer or dimer models. Under conditions in which the LMS can oxidize the models, lignin is insoluble. If conditions can be found under which lignin is soluble, the enzyme is active and stable, and oxidized mediators are stable, there is hope that enzymatic oxidation is a route to successful lignin breakdown. Examples of thermostable, alkaliphilic, and solvent-stable laccases exist in the literature ([20, 21, 22], respectively). Extremophiles might be one route toward oxidizing lignin under industrially relevant conditions.

The ionic liquid used in this work, [EMIM][OAc], has proven very efficient at dissolving biomass and increasing cellulose hydrolysis rates after pretreatment. However, it has not been successfully used as a cosolvent for enzymes at high concentrations. Ionic liquids are a diverse group of solvents, with thousands of combinations of anions and cations possible. Non-aromatic cations may be better solvents for oxidized radical mediators. Aromatic molecules, such as the imidazolium cation, are efficient quenchers of radicals; non aromatic ILs might be better solvents for laccase mediator systems. These IL candidates are already in development: amino acid-based, biomass-derived, and enzyme-stabilizing ionic liquids are in development as new solvents for biomass [23, 24, 25, 26]. These solvents may also prove to be good solvents for the laccase-catalyzed oxidation of lignin, and its eventual breakdown.

Other parts of the lignin-degrading systems of fungi and other organisms might also be better-suited for lignin oxidation. This work failed to explore the ability of lignin peroxidase, manganese peroxidase, both of which have the ability to oxidize higher redox substrates with and without the use of soluble radical mediators. One can even imagine an enzyme with an

active site evolved to specifically oxidize β-O-4 linkages with α hydroxyl groups, providing a route for specific oxidation of lignin that isn't currently known.

Bibliography

[1] Kierston Shill, Sasisanker Padmanabhan, Qin Xin, John M. Prausnitz, Douglas S. Clark, and Harvey W. Blanch. Ionic liquid pretreatment of cellulosic biomass: enzymatic hydrolysis and ionic liquid recycle. *Biotechnol Bioeng*, 108(3):511–20, 3 2011.

[2] Dongbao Fu, Giuseppe Mazza, and Yukihiro Tamaki. Lignin extraction from straw by ionic liquids and enzymatic hydrolysis of the cellulosic residues. *J Agric Food Chem*, 58(5):2915–22, 3 2010.

[3] Anantharam P. Dadi, Sasidhar Varanasi, and Constance A. Schall. Enhancement of cellulose saccharification kinetics using an ionic liquid pretreatment step. *Biotechnol Bioeng*, 95(5):904–10, 12 2006.

[4] Daniel Klein-Marcuschamer, Blake A. Simmons, and Harvey W. Blanch. Techno-economic analysis of a lignocellulosic ethanol biorefinery with ionic liquid pre-treatment. *Biofuels, Bioproducts and Biorefining*, 5(5):562–569, 2011.

[5] Sang Hyun Lee, Thomas V. Doherty, Robert J. Linhardt, and Jonathan S. Dordick. Ionic liquid-mediated selective extraction of lignin from wood leading to enhanced enzymatic cellulose hydrolysis. *Biotechnol Bioeng*, 102(5):1368–76, 4 2009.

[6] J. Shabtai, W. Zmierczak, S. Kadangode, E. Chornet, and D. K. Johnson. Lignin conversion to high-octane fuel additives. In *Biomass, Proceedings of the Fourth Biomass Conference of the Americas, Oakland, CA*, pages 811–818, 1999.

[7] Zuyong Xia, Takashi Yoshida, and Masamitsu Funaoka. Enzymatic synthesis of polyphenols from highly phenolic lignin-based polymers (lignophenols). *Biotechnol Lett*, 25(1):9–12, 1 2003.

[8] Joseph Zakzeski, Pieter C. A. Bruijnincx, Anna L. Jongerius, and Bert M. Weckhuysen. The catalytic valorization of lignin for the production of renewable chemicals. *Chem Rev*, 110(6):3552–99, 6 2010.

[9] Qin Xin, Katie Pfeiffer, John M. Prausnitz, Douglas S. Clark, and Harvey W. Blanch. Extraction of lignins from aqueous-ionic liquid mixtures by organic solvents. *Biotechnol Bioeng*, 109(2):346–52, 2 2012.

[10] Annele Hatakka. Lignin-modifying enzymes from selected white-rot in lignin degradation. *FEMS Microbiology Reviews*, 13:125–135, 1994.

[11] H. P. Call and I. Mücke. History, overview and applications of mediated lignolytic systems, especially laccase-mediator-systems (lignozym®-process). *Journal of Biotechnology*, 53(2):163–202, 1997.

[12] R. Bourbonnais, M. G. Paice, B. Freiermuth, E. Bodie, and S. Borneman. Reactivities of various mediators and laccases with kraft pulp and lignin model compounds. *Appl Environ Microbiol*, 63(12):4627–32, 12 1997.

[13] R. ten Have and P. J. Teunissen. Oxidative mechanisms involved in lignin degradation by white-rot fungi. *Chem Rev*, 101(11):3397–413, 11 2001.

[14] J. Gierer and I. Norén. Oxidative pretreatment of pine wood to facilitate delignification during kraft pulping. *Holzforschung-International Journal of the Biology, Chemistry, Physics and Technology of Wood*, 36(3):123–130, 1982.

[15] D. L. Criss, T. H. Fisher, and T. P. Schultz. Alkaline hydrolysis of nonphenolic α-carbonyl β-0-4 lignin dimers substituted on the leaving phenoxide ring: Comparison with benzylic hydroxyl analogues. *Holzforschung-International Journal of the Biology, Chemistry, Physics and Technology of Wood*, 52(1):57–60, 1998.

[16] Julian MW Chan, Stefan Bauer, Hagit Sorek, Sanil Sreekumar, Kun Wang, and F. Dean Toste. Studies on the vanadium-catalyzed nonoxidative depolymerization of miscanthus giganteus-derived lignin. *ACS Catalysis*, 3(6):1369–1377, 2013.

[17] Yunqiao Pu, Nan Jiang, and Arthur J. Ragauskas. Ionic liquid as a green solvent for lignin. *Journal of Wood Chemistry and Technology*, 27(1):23–33, 1 2007.

[18] S. A. Medvedeva, G. P. Aleksandrova, L. V. Kanitskaya, S. G. D'yachkova, and V. A. Babkin. Chemical transformations of lignin during the biodegradation and organosolv pulping of deciduous wood. *Chemistry of Natural Compounds*, 31(4):503–510, 1995.

[19] Paola Galletti, Matteo Pori, Federica Funiciello, Roberto Soldati, Alberto Ballardini, and Daria Giacomini. Laccase-mediator system for alcohol oxidation to carbonyls or carboxylic acids: toward a sustainable synthesis of profens. *ChemSusChem*, 7(9):2684–9, 9 2014.

[20] Kentaro Miyazaki. A hyperthermophilic laccase from Thermus thermophilus HB27. *Extremophiles*, 9(6):415–25, 12 2005.

[21] Woro Triarsi Sulistyaningdyah, Jun Ogawa, Hiromi Tanaka, Chiharu Maeda, and Sakayu Shimizu. Characterization of alkaliphilic laccase activity in the culture supernatant of Myrothecium verrucaria 24G-4 in comparison with bilirubin oxidase. *FEMS Microbiol Lett*, 230(2):209–14, 1 2004.

[22] G. Elegir, S. Daina, L. Zoia, G. Bestetti, and M. Orlandi. Laccase mediator system: Oxidation of recalcitrant lignin model structures present in residual kraft lignin. *Enzyme and Microbial Technology*, 37(3):340–346, 8 2005.

[23] Yukinobu Fukaya, Yoshiki Iizuka, Kenta Sekikawa, and Hiroyuki Ohno. Bio ionic liquids: room temperature ionic liquids composed wholly of biomaterials. *Green Chemistry*, 9(11):1155–1157, 2007.

[24] Kazuaki Ninomiya, Takashi Yamauchi, Masafumi Kobayashi, Chiaki Ogino, Nobuaki Shimizu, and Kenji Takahashi. Cholinium carboxylate ionic liquids for pretreatment of lignocellulosic materials to enhance subsequent enzymatic saccharification. *Biochemical Engineering Journal*, 71:25–29, 2013.

[25] Xue-Dan D. Hou, Ning Li, and Min-Hua H. Zong. Significantly enhancing enzymatic hydrolysis of rice straw after pretreatment using renewable ionic liquid-water mixtures. *Bioresour Technol*, 136:469–74, 5 2013.

[26] Ning Sun, Ramakrishnan Parthasarathi, Aaron M. Socha, Jian Shi, Sonny Zhang, Vitalie Stavila, Kenneth L. Sale, Blake A. Simmons, and Seema Singh. Understanding pretreatment efficacy of four cholinium and imidazolium ionic liquids by chemistry and computation. *Green Chemistry*, 16(5):2546–2557, 2014.

Appendix A

Primers and sequences for heterologously expressed *H. jecorina* cellulases and truncated cellulases

The following primers were used to create the CBM-free Cel7B catalytic domain, the Cel5A full enzyme (expressed heterologously in the same organism) and the Cel5A catalytic domain. A was used to remove the CBM from an existing construct of Cel7B in the pCDNA plasmid (described in Chapter 3). The primers describe in Table A.1B are used to amplify Cel5A genomic DNA, with ends complimentary to the pCDNA plasmid fragment produced by amplification of the pCDNA plasmid with the primers from A.1C. The fragments were assembled using Gibson assembly. Finally, the CBM was removed from the assembled plasmid by amplifying and recombining the plasmid using A.1D.

The gene sequences for Cel7B and Cel5A constructs are listed in Table A.

Table A.1: Primers used for construction of pCDNA plasmid for heterologous expression in *H. jecorina*

	Construct	Direction	Sequence
A	Cel7B ΔCBM	forward	5'-AACCACTTTCCAAACCATCTCAAC-3'
		reverse	5'-GCGAAGGTCACAAGACAAAGG
B	Cel5A gDNA	forward	5'-CAATCCAACAACTTCTCTCATCGATACTAGTATGAACAAGTCCGTGGCTCCATTGCTGC-3'
		reverse	5'-CAGGCTAGTAATCCTGCAGGTTAATTAACTACTTTCTTGCGAGACACGAGCTGACCAAGG-3'
C	pCDNA plasmid amplification	forward	5'-GTCAGCTCGTGTCTCGCAAGAAAGTAGTTAATTAACCTGCAGGATTACTAGCCTGCATGG-3'
		reverse	5'-GGAGCCACGGACTTGTTCATACTAGTATCGATGAGAGAAGTTGTTGGATTGATCAAAAAG-3'
D	Cel5A ΔCBM	forward	5'-CGTCGCAGGGGTCCGATTTGCC-3'
		reverse	5'-TCGGACCCCTGCGACGGCGCCG-3'
E	Cassette amplification	forward	5'-GGATCCGAGAGCTACCTTACATC-3'
		reverse	5'-CGAACTACCTCGCGAAACTCG-3'

Table A.2: Cel5A and Cel7B gene sequences

Gene	Sequence
Cel7B full enzyme	ATGGCGCCCTCAGTTACACTGCCGTTGACCACGGCCATCCTG GCCATTGCCCGGCTCGTCGCCGCCCAGCAACCGGGTACCAGC ACCCCCGAGGTCCATCCCAAGTTGACAACCTACAAGTGTACA AAGTCCGGGGGGGTGCGTGGCCCAGGACACCTCGGTGGTCCT TGACTGGAACTACCGCTGGATGCACGACGCAAACTACAACTC GTGCACCGTCAACGGCGGCGTCAACACCACGCTCTGCCCTGA CGAGGCGACCTGTGGCAAGAACTGCTTCATCGAGGGCGTCG ACTACGCCGCCTCGGGCGTCACGACCTCGGGCAGCAGCCTCA CCATGAACCAGTACATGCCCAGCAGCTCTGGCGGCTACAGCA GCGTCTCTCCTCGGCTGTATCTCCTGGACTCTGACGGTGAGT ACGTGATGCTGAAGCTCAACGGCCAGGAGCTGAGCTTCGAC GTCGACCTCTCTGCTCTGCCGTGTGGAGAGAACGGCTCGCTC TACCTGTCTCAGATGGACGAGAACGGGGGCGCCAACCAGTAT AACACGGCCGGTGCCAACTACGGGAGCGGCTACTGCGATGCT CAGTGCCCCGTCCAGACATGGAGGAACGGCACCCTCAACACT AGCCACCAGGGCTTCTGCTGCAACGAGATGGATATCCTGGAG GGCAACTCGAGGGCGAATGCCTTGACCCCTCACTCTTGCACG GCCACGGCCTGCGACTCTGCCGGTTGCGGCTTCAACCCCTAT GGCAGCGGCTACAAAAGGTGAGCCTGATGCCACTACTACCCC TTTCCTGGCGCTCTCGCGGTTTTCCATGCTGACATGGTTTTC CAGCTACTACGGCCCCGGAGATACCGTTGACACCTCCAAGAC CTTCACCATCATCACCCAGTTCAACACGGACAACGGCTCGCC CTCGGGCAACCTTGTGAGCATCACCCGCAAGTACCAGCAAAA CGGCGTCGACATCCCCAGCGCCCAGCCCGGCGGCGACACCAT CTCGTCCTGCCCGTCCGCCTCAGCCTACGGCGGCCTCGCCAC CATGGGCAAGGCCCTGAGCAGCGGCATGGTGCTCGTGTTCA GCATTTGGAACGACAACAGCCAGTACATGAACTGGCTCGACA GCGGCAACGCCGGCCCCTGCAGCAGCACCGAGGGCAACCCAT CCAACATCCTGGCCAACAACCCCAACACGCACGTCGTCTTCT CCAACATCCGCTGGGGAGACATTGGGTCTACTACGAACTCGA CTGCG

Continued on next page

Table A.2 – *Continued from previous page*

Gene	Sequence
Cel7B CD	ATGGCGCCCTCAGTTACACTGCCGTTGACCACGGCCATCCTG GCCATTGCCCGGCTCGTCGCCGCCCAGCAACCGGGTACCAGC ACCCCCGAGGTCCATCCCAAGTTGACAACCTACAAGTGTACA AAGTCCGGGGGGGTGCGTGGCCCAGGACACCTCGGTGGTCCT TGACTGGAACTACCGCTGGATGCACGACGCAAACTACAACTC GTGCACCGTCAACGGCGGCGTCAACACCACGCTCTGCCCTGA CGAGGCGACCTGTGGCAAGAACTGCTTCATCGAGGGCGTCG ACTACGCCGCCTCGGGCGTCACGACCTCGGGCAGCAGCCTCA CCATGAACCAGTACATGCCCAGCAGCTCTGGCGGCTACAGCA GCGTCTCTCCTCGGCTGTATCTCCTGGACTCTGACGGTGAGT ACGTGATGCTGAAGCTCAACGGCCAGGAGCTGAGCTTCGAC GTCGACCTCTCTGCTCTGCCGTGTGGAGAGAACGGCTCGCTC TACCTGTCTCAGATGGACGAGAACGGGGGCGCCAACCAGTAT AACACGGCCGGTGCCAACTACGGGAGCGGCTACTGCGATGCT CAGTGCCCCGTCCAGACATGGAGGAACGGCACCCTCAACACT AGCCACCAGGGCTTCTGCTGCAACGAGATGGATATCCTGGAG GGCAACTCGAGGGCGAATGCCTTGACCCCTCACTCTTGCACG GCCACGGCCTGCGACTCTGCCGGTTGCGGCTTCAACCCCTAT GGCAGCGGCTACAAAAGGTGAGCCTGATGCCACTACTACCCC TTTCCTGGCGCTCTCGCGGTTTTCCATGCTGACATGGTTTTC CAGCTACTACGGCCCCGGAGATACCGTTGACACCTCCAAGAC CTTCACCATCATCACCCAGTTCAACACGGACAACGGCTCGCC CTCGGGCAACCTTGTGAGCATCACCCGCAAGTACCAGCAAAA CGGCGTCGACATCCCCAGCGCCCAGCCCGGCGGCGACACCAT CTCGTCCTGCCCGTCCGCCTCAGCCTACGGCGGCCTCGCCAC CATGGGCAAGGCCCTGAGCAGCGGCATGGTGCTCGTGTTCA GCATTTGGAACGACAACAGCCAGTACATGAACTGGCTCGACA GCGGCAACGCCGGCCCCTGCAGCAGCACCGAGGGCAACCCAT CCAACATCCTGGCCAACAACCCCAACACGCACGTCGTCTTCT CCAACATCCGCTGGGGAGACATTGGGTCTACTACGAACTCGA CTGCG

Continued on next page

Table A.2 – *Continued from previous page*

Gene	Sequence
Cel5A full enzyme	ATGAACAAGTCCGTGGCTCCATTGCTGCTTGCAGCGTCCATA CTATATGGCGGCGCCGTCGCACAGCAGACTGTCTGGGGCCAG TGTGGAGGTATTGGTTGGAGCGGACCTACGAATTGTGCTCCT GGCTCAGCTTGTTCGACCCTCAATCCTTATTATGCGCAATGT ATTCCGGGAGCCACTACTATCACCACTTCGACCCGGCCACCA TCCGGTCCAACCACCACCACCAGGGCTACCTCAACAAGCTCA TCAACTCCACCCACGAGCTCTGGGGTCCGATTTGCCGGCGTT AACATCGCGGGTTTTGACTTTGGCTGTACCACAGAGTGAGTA CCCTTGTTTCCTGGTGTTGCTGGCTGGTTGGGCGGGTATACA GCGAAGCGGACGCAAGAACACCGCCGGTCCGCCACCATCAAG ATGTGGGTGGTAAGCGGCGGTGTTTTGTACAACTACCTGACA GCTCACTCAGGAAATGAGAATTAATGGAAGTCTTGTTACAGT GGCACTTGCGTTACCTCGAAGGTTTATCCTCCGTTGAAGAAC TTCACCGGCTCAAACAACTACCCCGATGGCATCGGCCAGATG CAGCACTTCGTCAACGAGGACGGGATGACTATTTTCCGCTTA CCTGTCGGATGGCAGTACCTCGTCAACAACAATTTGGGCGGC AATCTTGATTCCACGAGCATTTCCAAGTATGATCAGCTTGTT CAGGGGTGCCTGTCTCTGGGCGCATACTGCATCGTCGACATC CACAATTATGCTCGATGGAACGGTGGGATCATTGGTCAGGG CGGCCCTACTAATGCTCAATTCACGAGCCTTTGGTCGCAGTT GGCATCAAAGTACGCATCTCAGTCGAGGGTGTGGTTCGGCAT CATGAATGAGCCCCACGACGTGAACATCAACACCTGGGCTGC CACGGTCCAAGAGGTTGTAACCGCAATCCGCAACGCTGGTGC TACGTCGCAATTCATCTCTTTGCCTGGAAATGATTGGCAATC TGCTGGGGCTTTCATATCCGATGGCAGTGCAGCCGCCCTGTC TCAAGTCACGAACCCGGATGGGTCAACAACGAATCTGATTTT TGACGTGCACAAATACTTGGACTCAGACAACTCCGGTACTCA CGCCGAATGTACTACAAATAACATTGACGGCGCCTTTTCTCC GCTTGCCACTTGGCTCCGACAGAACAATCGCCAGGCTATCCT GACAGAAACCGGTGGTGGCAACGTTCAGTCCTGCATACAAGA CATGTGCCAGCAAATCCAATATCTCAACCAGAACTCAGATGT CTATCTTGGCTATGTTGGTTGGGGTGCCGGATCATTTGATAG CACGTATGTCCTGACGGAAACACCGACTAGCAGTGGTAACTC ATGGACGGACACATCCTTGGTCAGCTCGTGTCTCGCAAGAAA GTAG

Continued on next page

Table A.2 – *Continued from previous page*

Gene	Sequence
Cel5A CD	ATGAACAAGTCCGTGGCTCCATTGCTGCTTGCAGCGTCCATA
	CTATATGGCGGCGCCGTCGCAGGGGTCCGATTTGCCGGCGTT
	AACATCGCGGGTTTTGACTTTGGCTGTACCACAGAGTGAGTA
	CCCTTGTTTCCTGGTGTTGCTGGCTGGTTGGGCGGGTATACA
	GCGAAGCGGACGCAAGAACACCGCCGGTCCGCCACCATCAAG
	ATGTGGGTGGTAAGCGGCGGTGTTTTGTACAACTACCTGACA
	GCTCACTCAGGAAATGAGAATTAATGGAAGTCTTGTTACAGT
	GGCACTTGCGTTACCTCGAAGGTTTATCCTCCGTTGAAGAAC
	TTCACCGGCTCAAACAACTACCCCGATGGCATCGGCCAGATG
	CAGCACTTCGTCAACGAGGACGGGATGACTATTTTCCGCTTA
	CCTGTCGGATGGCAGTACCTCGTCAACAACAATTTGGGCGGC
	AATCTTGATTCCACGAGCATTTCCAAGTATGATCAGCTTGTT
	CAGGGGTGCCTGTCTCTGGGCGCATACTGCATCGTCGACATC
	CACAATTATGCTCGATGGAACGGTGGGATCATTGGTCAGGG
	CGGCCCTACTAATGCTCAATTCACGAGCCTTTGGTCGCAGTT
	GGCATCAAAGTACGCATCTCAGTCGAGGGTGTGGTTCGGCAT
	CATGAATGAGCCCCACGACGTGAACATCAACACCTGGGCTGC
	CACGGTCCAAGAGGTTGTAACCGCAATCCGCAACGCTGGTGC
	TACGTCGCAATTCATCTCTTTGCCTGGAAATGATTGGCAATC
	TGCTGGGGCTTTCATATCCGATGGCAGTGCAGCCGCCCTGTC
	TCAAGTCACGAACCCGGATGGGTCAACAACGAATCTGATTTT
	TGACGTGCACAAATACTTGGACTCAGACAACTCCGGTACTCA
	CGCCGAATGTACTACAAATAACATTGACGGCGCCTTTTCTCC
	GCTTGCCACTTGGCTCCGACAGAACAATCGCCAGGCTATCCT
	GACAGAAACCGGTGGTGGCAACGTTCAGTCCTGCATACAAGA
	CATGTGCCAGCAAATCCAATATCTCAACCAGAACTCAGATGT
	CTATCTTGGCTATGTTGGTTGGGGTGCCGGATCATTTGATAG
	CACGTATGTCCTGACGGAAACACCGACTAGCAGTGGTAACTC
	ATGGACGGACACATCCTTGGTCAGCTCGTGTCTCGCAAGAAA
	GTAG