

Stochastic model highlights the impact of crystallinity on saccharification dynamics depending on plant chemotype and pre-treatment

Partho Sakha De, Philipp M Grande, Henrike Heise, Holger Klose, Adélaïde Raguin

Article - Version of Record

Suggested Citation:

De, P. S., Grande, P. M., Heise, H., Klose, H., & Raguin, A. (2025). Stochastic model highlights the impact of crystallinity on saccharification dynamics depending on plant chemotype and pre-treatment. *PLoS ONE*, 20(12), Article e0322367. <https://doi.org/10.1371/journal.pone.0322367>

Wissen, wo das Wissen ist.

This version is available at:

URN: <https://nbn-resolving.org/urn:nbn:de:hbz:061-20260420-124358-2>

Terms of Use:

This work is licensed under the Creative Commons Attribution 4.0 International License.

For more information see: <https://creativecommons.org/licenses/by/4.0>

RESEARCH ARTICLE

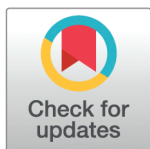
Stochastic model highlights the impact of crystallinity on saccharification dynamics depending on plant chemotype and pre-treatment

Partho Sakha De^{1,2,3}, Philipp M Grande^{4,2}, Henrike Heise^{5,6},
Holger Klose^{4,7,2}, Adélaïde Raguin^{1,2,3,8*}

1 Institute for Computational Cell Biology, Heinrich Heine University, Düsseldorf, Germany, **2** Bioeconomy Science Center (BioSC), c/o Forschungszentrum Jülich, Jülich, Germany, **3** Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich Heine University, Düsseldorf, Germany, **4** Institute of Bio- and Geosciences (IBG) Plant Sciences (IBG-2), Forschungszentrum Jülich GmbH, Jülich, Germany, **5** Institute for Physical Biology, Heinrich Heine University, Düsseldorf, Germany, **6** Institute of Structural Biochemistry (IBI-7), Forschungszentrum Jülich GmbH, Jülich, Germany, **7** RWTH Aachen University, Aachen, Germany, **8** Laboratoire Bordelais de Recherche en Informatique (LaBRI - UMR 5800), CNRS, Univ. Bordeaux, Bordeaux INP, Talence, France

☯ These authors contributed equally.

* adelaide.raguin@hhu.de



OPEN ACCESS

Citation: De PS, Grande PM, Heise H, Klose H, Raguin A (2025) Stochastic model highlights the impact of crystallinity on saccharification dynamics depending on plant chemotype and pre-treatment. *PLoS One* 20(12): e0322367. <https://doi.org/10.1371/journal.pone.0322367>.

Editor: Igor Cesarino, University of Sao Paulo, BRAZIL

Received: July 26, 2024

Accepted: March 20, 2025

Published: December 5, 2025

Copyright: © 2025 De et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: The model code and data, together with the scripts for reproducing the figures, are provided in the GitLab repository accessible via https://gitlab.com/partho9791/code_n_data_for_stochastic_biophysical_model_for_biomass_saccharification.

Funding: This research was done within the BioSC projects PREDIG and OptiCellu that fund the position of PSD. The scientific activities of the Bioeconomy Science Center were financially

Abstract

Enzymatic saccharification of plant-sourced lignocellulosic biomass is a key step in biorefinery approaches. However, these biomasses in their raw form are quite recalcitrant, which invokes the need for pre-treatment processes aimed at not only increasing glucose conversion, but also better valorising non-carbohydrate biopolymers, such as lignin. Here, we use a two-fold computational and experimental approach to investigate enzymatic saccharification time-courses for three cellulosic substrates (i.e. AVICEL, a mixture of AVICEL with Organosolv lignin, and Sigmacell), and four plant-sourced lignocellulosic biomasses following three different conditions for each of them (i.e. untreated, OrganoCat pre-treated with a swelling step, and OrganoCat pre-treated without a swelling step), making a total of fifteen samples. Considering the specific composition of each substrate, the model successfully reproduces the saccharification dynamics for each of the fifteen samples. It additionally provides values for the parameter Crystallinity Fraction that faithfully replicate the substrate Crystallinity Indices experimentally determined by ssNMR. Importantly, we show that the Crystallinity Index of distinct biomasses is differently impacted by swelling, while the sugar release is consistently impacted by pre-treatment across biomasses. Eventually, both artificial cellulosic and plant-sourced lignocellulosic biomasses demonstrate that the sugar release is the result of the combination of the Crystallinity Fraction (the model parameter for experimentally measured ssNMR Crystallinity Index) and the digestibility ratio, the model parameter that represents in a coarse-grained manner complex spatial and structural features. Overall, our results

supported by the Ministry of Culture and Science within the framework of the NRW Strategieprojekt BioSC [005-2012-0107 to AR and HK]. The position of AR was funded by the German federal and state programme Professorinnenprogramm III for female scientists and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) in the framework of EtransColi [470067901 to AR]. The DFG provided additional support in the form of a grant [INST 208/771-1 FUGG to HH]. The Cluster of Excellence on Plant Sciences provided project funding [EXC 2048/1, Project ID: 390686111 to AR].

Competing interests: The authors have declared that no competing interests exist.

stress the need for further experimental investigations that physically explain variations in the digestibility of crystalline bonds across biomasses and pre-treatment conditions. Additionally, we supplemented our work with theoretical investigations on a generic lignocellulosic substrate to highlight the roles of various model parameters in a qualitative manner.

1 Introduction

Lignocellulosic biomass is an abundant source of raw material for the production of biofuels and other valuable chemicals. In biorefinery processes, it is broken down by enzymatic saccharification to release the constituent simple monomeric sugars from the biopolymers. However, its recalcitrance to enzymatic saccharification is a major challenge that has been speculated to be impacted by various factors. Those range from the crystallinity of the substrate, to the inhibitory effect of lignin on the enzymes. For reducing recalcitrance, and hence enhancing saccharification yield, this biomass is subjected to chemical and physical pre-treatments, such as: acidic treatments involving phosphoric, oxalic, acetic, or sulphuric acids [1,2]; alkaline treatments involving slaked lime or ammonia fiber explosion/expansion (AFEX) [3–6], and exposure to steam or liquid hot water [7,8]. Alternative methods for instance rely on innovative enzymatic processes, e.g. the laccase system mediated by 1-hydroxybenzotriazole [9]. In this study, we specifically employ an Organosolv-like OrganoCat process, which involves a diluted acid, e.g. oxalic acid, 2,5-furandicarboxylic acid, or phosphoric acid to hydrolyse hemicelluloses, and *in situ* extract lignin into a second phase of 2-methyltetrahydrofuran (2-MTHF) [10–12]. In addition to experimental approaches, several models have been developed, at multiple scales, to understand the dynamics of enzymatic saccharification and decipher the effect of various structural and compositional properties of lignocellulose on its recalcitrance. Each of these approaches possesses its own set of advantages and disadvantages, which have been nicely summarised in the comprehensive review by Ciesielski et al. [13]. At the smallest scale, density functional theory (DFT), quantum mechanics/molecular mechanics (QM/MM), and molecular dynamics (MD) are typical methods [14–17]. They are used to address problems such as pyrolysis [18,19], the detailed structure and properties of lignocellulosic biomass [20], enzyme mechanisms [21,22], and the effects of lignin binding on cellulose and cellulase enzymes [17]. However, they are very demanding in computational resources, and for instance are unable to depict biopolymers at the scale of seconds. To counter these disadvantages, alternative approaches based on coarse-grained molecular dynamics have been employed with beads or pseudo atoms as elementary units [23]. Instead, considering glucose molecules as elementary units, Kumar and Murthy implemented Monte Carlo simulations to model the digestion of a cellulose bundle under the action of endoglucanase (EG), cellobiohydrolase (CBH), and β -glucosidase (BGL) [24,25]. They considered a substrate of crystalline cellulose with hemicellulose and lignin, but did not study the impact of crystallinity on the saccharification dynamics, and their simulation results displayed significant discrepancies with experimental data. Using Ordinary Differential Equations, Griggs et al. [26,27] developed a mechanistic and kinetic model to simulate the action of a cocktail of cellulase enzymes on a purely cellulosic substrate, while this framework can also be applied at a much higher (i.e. reactor) scale [28]. Their results showed the enzyme synergism, a good agreement with experimental cellulose chain length distributions from literature, and a semi-quantitative agreement with experimental saccharification time-courses. Notably, their model nonetheless neglected all constituents other than cellulose. With agent-based models,

Vetharaniem et al. [29] and Asztalos et al. [30] both investigated enzymatic synergism, while respectively considering: (i) hemicellulosic sugars and crystalline cellulose; and (ii) inter-chain hydrogen bond breaking, hydrolysis of glycosidic bonds, and adsorption and desorption of the cellulases on the substrate. Yet, these studies oversimplify the substrate when neglecting the essential constituent lignin, and considering a two-dimensional and purely cellulosic substrate. To fill these gaps, and build a comprehensive model of lignocellulose and its enzymatic saccharification dynamics, we previously introduced a spatially resolved stochastic biophysical model of a lignocellulose microfibril [31], which has later on been made available as an advanced free, open-source, and user-friendly Web Application <https://predig.cs.hhu.de/> [32]. The model accounted for the detailed biomass structure, the specific action of the enzymes, the crystallinity of cellulose and hemicellulose [33,34], and the role of lignin; and demonstrated an exceptional ability in quantitatively reproducing experimental saccharification time-courses.

Despite our previous work and all the advancements in the investigation of enzymatic degradation of lignocellulosic biomass, it remains unclear how crystallinity in distinct biomasses differently impacts on saccharification recalcitrance, and to which extent experimental measurements of crystallinity can explain saccharification dynamics. To answer these questions, in this study, we develop a two-fold (computational and experimental) approach and focus on the extraction of glucose from various biomasses, with no pre-treatment and two distinct pre-treatment conditions. Our findings suggest that the enzyme kinetics is strongly affected by not only the type of substrate i.e. pure cellulose *vs* lignocellulose, but also the pre-treatment each biomass is subjected to. Moreover, the model shows that even directly accounting for ssNMR measurements of the crystallinity index is insufficient to accurately reproduce the saccharification dynamics, which is however well captured when additionally considering in the model that crystalline bonds in different biomasses have different propensities to be digested, which impact on recalcitrance in a non-trivial way. Overall, we demonstrate the importance of modelling for rationalising experimental data relating to the structure and dynamics of plant biomass, a highly complex complex composite material.

2 Materials and methods

2.1 Extraction and analysis of lignocellulose

All chemicals were purchased from Carl Roth and Sigma-Aldrich (Germany) and used without further purification. The plant biomasses were the same set as published earlier [1]. The extraction and analysis procedures were carried out with slight modifications as outlined in our previous study [35]. In summary, the biomasses were reduced in size before subsequent treatment by grinding to a fine powder using a ball mill M 400 (Retsch, Haan, Germany) in a 50 mL metal beaker (30 s^{-1} , 2 min). For each plant, 1 g of powder was extracted in 50 mL reaction tubes, and the resulting pellet was collected by centrifugation at 3,234 g for 5 minutes. The alcohol-insoluble residue (AIR) was utilised to quantify the content of crystalline cellulose and lignin. Lignin was determined as acetyl bromide soluble lignin (ABSL) and crystalline cellulose content was measured by the Updegraff method, like described by Foster et al. [36,37]. De-starched AIR (d-AIR) was employed to assess the content of matrix polysaccharides (MPS) through High-Performance-Exchange Chromatography with Pulsed Amperometric Detection (HPAEC-PAD) analysis [35]. The composition of each of the samples is summarised in Table 1.

Table 1. Composition of cellulosic and lignocellulosic biomasses in various conditions: untreated, OrganoCat treated without swelling step (NS+OCAT), and OrganoCat treated with swelling step (YS+OCAT).

Biomass source	Pre-treatment	cellulose [%]	lignin [%]	MPS [%]	acetate [%]	total [%]
AVICEL	NA	80.6 ± 0.78	11.3 ± 0.75*	3.7 ± 0.17	3.9 ± 0.09	99.5
Sigmacell	NA	82.8 ± 0.82	9.2 ± 0.68*	4.0 ± 0.18	3.6 ± 0.13	99.2
Beech	untreated	45.3 ± 1.7	24.7 ± 2.0	22.8 ± 3.5	5.1 ± 0.5	97.9
	NS+OCAT	71.0 ± 1.5	21.8 ± 1.0	4.2 ± 0.5	0.0 ± 0.1	97.1
	YS+OCAT	68.6 ± 1.5	23.8 ± 2.5	1.9 ± 0.1	0.0 ± 0.0	94.3
Miscanthus	untreated	46.5 ± 4.3	24.4 ± 2.2	24.3 ± 3.6	2.7 ± 0.1	97.8
	NS+OCAT	71.9 ± 2.0	20.9 ± 1.1	5.9 ± 0.8	0.0 ± 0.0	98.7
	YS+OCAT	80.8 ± 2.7	19.7 ± 1.1	1.8 ± 0.1	0.1 ± 0.1	102.4
Sida (Sida hermaphrodita)	untreated	44.7 ± 2.3	25.2 ± 1.2	17.6 ± 2.5	5.1 ± 0.2	92.6
	NS+OCAT	62.1 ± 0.9	18.2 ± 1.5	6.6 ± 0.7	0.1 ± 0.2	87.1
	YS+OCAT	68.5 ± 0.8	23.7 ± 0.5	3.4 ± 0.5	0.4 ± 0.1	96.0
Walnut	untreated	31.4 ± 4.0	28.4 ± 1.9	15.7 ± 3.2	4.6 ± 0.2	80.2
	NS+OCAT	50.9 ± 1.8	28.4 ± 2.5	6.8 ± 0.1	0.2 ± 0.1	86.3
	YS+OCAT	50.0 ± 1.0	29.3 ± 0.3	3.5 ± 0.6	0.2 ± 0.1	83.0

Note: *Positive lignin results in AVICEL and Sigmacell are artifacts of the measurement and should not be considered.

<https://doi.org/10.1371/journal.pone.0322367.t001>

2.2 Lignocellulose fractionation by OrganoCat pulping

Without swelling. In a high-pressure reactor with a volume of 20 mL, 500 mg of lignocellulose biomass, 5 mL of phosphoric acid (0.74 M) and 5 mL of 2-MTHF were introduced. To prevent 2-MTHF evaporation, the stainless-steel high-pressure reactor was sealed and pressurised with 10 bar of argon. The mixture was stirred for 3 hours at 140°C. After cooling the reactor to room temperature, the liquid phases were separated through decantation and the cellulose-enriched solid pulp was filtered. Sugar concentrations were determined using HPAEC-PAD. The solid residue was washed with distilled water until it reached a neutral pH and then dried until a constant weight was achieved. Lignin was obtained by evaporating the 2-MTHF and quantified through gravimetric analysis.

With swelling. In a high-pressure reactor with a volume of 20 mL, 500 mg of lignocellulose biomass, 250 µL of ultra-pure water, and 250 µL of phosphoric acid (85 wt%) were combined. This mixture was then heated to 80°C for 1 hour. Subsequently, 4.5 mL of ultra-pure water and 5 mL of 2-MTHF were introduced. To prevent 2-MTHF evaporation, the stainless-steel high-pressure reactor was sealed and pressurised with 10 bars of argon. The mixture was stirred for 3 hours at 140°C. After cooling the reactor to room temperature, the liquid phases were separated through decantation and the cellulose-enriched solid pulp was filtered. Sugar concentrations in the aqueous phase were determined using HPAEC-PAD. The solid residue was washed with distilled water until it reached a neutral pH and then dried until a constant weight was achieved. Lignin was obtained by evaporating the 2-MTHF and quantified through gravimetric analysis.

The OrganoCat treatment applied to the lignocelluloses did not result in instantaneous expansion and does not have a significant effect on the cellulose fibres, unlike for instance steam explosion, where intense heating leads to the explosive expansion of vapour, causing cell damage.

2.3 Enzymatic hydrolysis of cellulose pulp

Enzymatic hydrolysis and its subsequent analysis were performed following the protocols detailed in our previous study [38]. In brief, each of the cellulosic or lignocellulosic biomasses was suspended at a concentration of 20 g/L in a 0.1 M sodium citrate buffer with a pH of 4.5

(total volume of 1 mL). This suspension was maintained at a temperature of 50°C for specified amount of time. To initiate the enzymatic reaction, Accellerase[®]-1500 (containing 60 FPU/mL and 82 CBU/mL, Genencor, The Netherlands) was added at a volume of 1 vol%, relative to the buffer. Samples (0.3 mL) were withdrawn from the reactions at indicated times, heated to 100°C for 5 minutes to halt the enzymatic process, and subsequently stored at -20°C until the colourimetric analysis was carried out. Sugar release was determined by employing the PAHBAH (4-hydroxybenzoic acid hydrazide) method, following the procedure outlined by Lever et al. [39]. The resulting reaction mixtures were diluted as needed to be in the range of the calibration curve. The absorbance was measured at 410 nm using a BioTek Power Wave HT UV/Vis spectrometer. All experiments were performed as single trials or with replicates, as specified in the text.

2.4 Crystallinity measurement via ssNMR

For solid-state NMR measurements, ca. 30 mg of the sample were packed into Bruker magic-angle spinning (MAS) rotors with an outer diameter of 3.2 mm. ¹³C Cross-Polarisation (CP) MAS spectra were recorded on a 14.1 T (600 MHz ¹H frequency) Bruker Avance wide bore spectrometer equipped with a 3.2 mm MAS triple resonance ¹H, ¹³C, probe and a 14.1 T (600 MHz ¹H frequency) Bruker AVANCE NEO spectrometer equipped with a triple resonance HCN 3.2 mm MAS Efree probe. The CP contact time was 500 μs, the MAS spinning speed was 11 kHz. Spinal-64 ¹H decoupling (RF field of 85 kHz) was applied during acquisition. Spectra were referenced externally to Sodium trimethylsilylpropanesulfonate (DSS) using adamantane as a secondary standard (the low frequency peak was set to 31.4 ppm).

2.5 Stochastic simulation model

Our theoretical model [31] employs a Gillespie algorithm (see Sect 2.6 for details) to stochastically simulate the enzymatic saccharification process of a lignocellulose microfibril. It represents *in silico* the distinct bio-polymers (namely cellulose, hemicellulose, and lignin) constituting the biomass, as well as the three-dimensional configuration of the substrate. The physical structure of the substrate is depicted as a hexagonal-shaped core bundle of cellulose polymer chains of length 200 bonds, surrounded by two layers of randomly positioned hemicellulose and lignin polymers, as seen in Fig 1. The number of cellulose chains that form the core of a microfibril has been strongly debated for decades, and it is widely accepted that they form in multiples of six, i.e. 36, 24, or 18, with a recent preference for the latter one [40,41]. Since this number is known to vary for different species and developmental stages [41–43], and that the conclusions from the model we present here are not impaired by these considerations, we choose to simulate the system for eighteen cellulose chains. The inner polymer bonds are shielded from enzymatic action by the outer chains, until those get digested, giving physical access to the enzymes. Besides, the lignin and hemicellulose layers do not completely surround the cellulose core, but possess some gaps, as seen in Fig 1.

Biomass crystallinity is also known to have a strong impact on enzymatic saccharification. It is well-established that cellulose possesses a crystalline arrangement that negatively impacts saccharification. Even though there is no direct evidence of hemicellulose crystallinity, xylans have been known to partially bind to cellulose microfibrils, adopting a semi-crystalline arrangement [33]. It has also been suggested that hemicellulose adsorption to cellulose interferes with the saccharification process [34]. Furthermore, the hemicelluloses of the biomasses considered in our study are predominantly constituted of xylose. Thus, in our model, we consider both crystalline and amorphous regions for both cellulose and hemicellulose, so that the crystalline bonds are more difficult to enzymatically digest as compared to

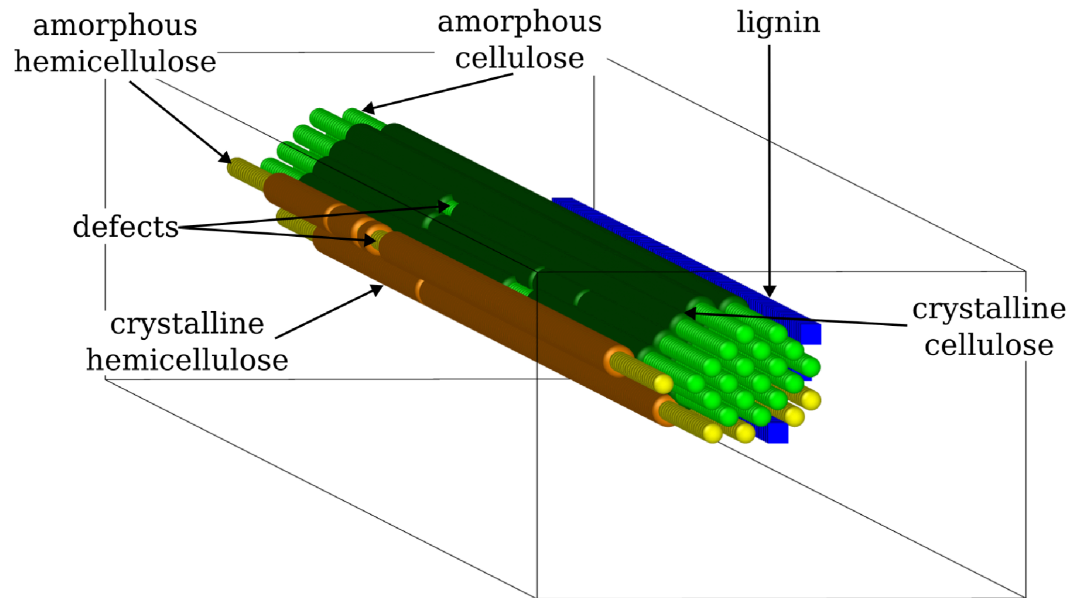


Fig 1. Perspective view of the modelled lignocellulose microfibril, with lignin in blue, hemicellulose in yellow, and cellulose in green. Both hemicellulose and cellulose can be either crystalline (dark colour) or amorphous (light colour), with amorphous regions typically located at the boundaries of the microfibril, or embedded in the crystalline region as 'defect' patches. These defects may for instance arise as a result of pre-treatments.

<https://doi.org/10.1371/journal.pone.0322367.g001>

their amorphous counterparts. Crystallinity *versus* amorphousness is quantified using two distinct parameters: (i) the Crystallinity Fraction (noted CF) quantifies the number of crystalline bonds over the total number of bonds of a certain type (i.e. cellulose or hemicellulose), and (ii) the digestibility ratio (noted $r_{c,a}$) describes how much harder it is to digest a crystalline bond as compared to its amorphous counterpart, for a certain bond type (i.e. cellulose or hemicellulose). For instance, when $r_{c,a} = 0$, it implies that crystalline bonds are not digestible at all, whereas when $r_{c,a} = 1$, crystalline bonds are digested at the same rate as amorphous ones. We assume that the crystallinity of hemicellulose is induced by its direct vicinity to the highly ordered and crystalline cellulose chains. We also account for the non-productive adsorption of enzymes on lignin, independently of their specific type [44–47].

The microfibril is subjected to digestion by an enzyme cocktail consisting of endoglucanases (EGs), cellobiohydrolases (CBHs), β -glucosidases (BGLs), and hemicellulases (HCs). Endoglucanases can digest any exposed bond in the bulk of a cellulose chain, except the two endmost ones, at each end of the polymer chain [48]. Cellobiohydrolases attach to glucan chains' ends from which they processively cleave-off cellobiose units [49–51]. β -glucosidases complete the saccharification process by splitting cellobiose into two glucose monomers [52, 53]. In our model, we account for the cellulases' specific mode of action, unlike for the hemicellulases that are considered as non-specific for the sake of simplicity, and capable of digesting any exposed digestible hemicellulosic bond. Our model also accounts for end product inhibition of the cellulases, such that free glucose and cellobiose in the reaction mixture have a detrimental effect on their activity. More precisely, these saccharification end products can bind to the active catalytic site of cellulases, and render them ineffective [54]. We assume that free glucose and cellobiose reduce the effective concentration of the enzymes available for saccharification, such that both inhibit EG and CBH, while BGL is only inhibited by glucose, as

follows:

$$[EG] = [EG]_0 - \omega_{EG}^{cbs} \frac{[EG]_0 [cbs]}{[EG]_0 + [CBH]_0 + [cbs]} - \omega_{EG}^{glc} \frac{[EG]_0 [glc]}{[EG]_0 + [CBH]_0 + [BGL]_0 + [glc]}, \quad (1)$$

$$[CBH] = [CBH]_0 - \omega_{CBH}^{cbs} \frac{[CBH]_0 [cbs]}{[EG]_0 + [CBH]_0 + [cbs]} - \omega_{CBH}^{glc} \frac{[CBH]_0 [glc]}{[EG]_0 + [CBH]_0 + [BGL]_0 + [glc]}, \quad (2)$$

$$[BGL] = [BGL]_0 - \omega_{BGL}^{glc} \frac{[BGL]_0 [glc]}{[EG]_0 + [CBH]_0 + [BGL]_0 + [glc]}, \quad (3)$$

where, on the right-hand side of the equations, $[y]_0$ denotes the concentration of the cellulases (i.e. $y = EG, CBH,$ or BGL) if inhibition would not take place, and $[x]$ denotes the time varying concentration of the inhibitors (i.e. $x = glc$ for glucose and cbs for cellobiose). The parameter ω_y^x quantifies the inhibition strength of the inhibitor x on the enzyme y , and can vary between 0 and 1. These equations account for the competition of distinct inhibitors for the same active site of an enzyme, and conversely, for the sharing of common pools of inhibitor molecules by distinct enzymes.

In this study, we expand our model to additionally account for a higher level of detail relating to the crystalline and amorphous regions of the substrate. In the earlier version of the model [31], we assumed a simplistic homogeneous representation of cellulose and hemicellulose crystallinity, with the crystalline bonds being only located in the mid-length region of the microfibril, and the amorphous ones at the periphery. Nonetheless, during material preparation and pre-treatment (e.g. grinding to powder, and OrganoCat) the ordered structure of the crystalline cellulose and hemicellulose polymers may become disrupted, giving rise to random amorphous *defects*. The occurrence of such defects and disruption of crystallinity due to pre-treatments and mechanical stresses have also been reported in earlier studies [55–57]. These *defects* appear in the outer polymer chains of the microfibril, which are most exposed. In this expanded model, two independently adjustable parameters characterise the mean size (μ_{defect}) and the number (N_{defect}) of such *defects*. μ_{defect} can vary in the range 0–0.5, denoting which fraction of the amorphous bonds in a polymer chain is embedded as *defects* in the crystalline region, instead of being located at the ends of the chain. N_{defect} varies in the range 0–1, quantifying which fraction of the outer polymer chains of a microfibril contains *defects*. Fig 1 shows a schematic representation of the entire microfibril, including amorphous *defects* located in the crystalline cellulose and hemicellulose regions. The distinct model parameters and their respective values are listed in the Sect 4.

2.6 Reminder of Gillespie algorithm

The enzymatic saccharification of the lignocellulose microfibril is simulated using a Gillespie algorithm, which is a common technique for implementing stochastic simulations. It mimics the stochastic nature of the system by considering a sequence of randomised events. Here, those can be the binding of enzymes to lignin, adsorption of CBH to a free polymer end, or reactions of enzymatic digestion. In our implementation of the Gillespie algorithm, for improved computational efficacy, we keep track of all and only the possible reactions, by

accounting for the bond accessibility and steric hindrance. At each time step, both the reaction to take place and its duration are randomly selected, in a manner that reflects the reaction's likelihood of occurrence, which can, for instance, depend on the substrate availability, the enzyme availability and the enzyme kinetics. The simulation lasts until the chosen amount of events is attained, or until all of the digestible substrate has been digested.

2.7 Fitting algorithm

In addition to our stochastic simulation model, we also developed a parameter fitting algorithm which enables us to search the parameter space to reproduce experimental saccharification time-course data. It proceeds through generations and sub-generations, with parameters randomly adjusted within specified ranges. The difference between experimental data and simulation curves (averaged over several runs) is recorded. If the difference decreases within a generation, the sub-generation with the lowest variance becomes the basis for the next generation. If the difference increases, the previous generation's parameters are used again as a starting point for random adjustments. This mix of directed and random search ultimately converges to an optimal fit. The minimum difference across all generations defines the best fit for a particular set of experimental data. In this study, the model input parameters for the composition of the substrates are fixed to the experimentally measured values listed in Table 1. Then, in order to reproduce both the saccharification time-courses and the experimentally measured Crystallinity Indices, we fit the same kinetic parameters for all samples under a given pre-treatment condition (i.e. enzyme kinetic rates and end-product inhibition parameters), while the substrate related parameters are substrate specific (i.e. crystallinity fraction and digestibility ratio).

3 Results

3.1 Artificial cellulosic substrates

Fig 2 displays the saccharification time-courses in three cases, i.e. for Sigmacell, AVICEL, and a (1:1) mixture of AVICEL with Organosolv lignin, with the experimental data being represented by points and the simulation results by dashed lines. We observe that the presence of Organosolv lignin in the reaction mixture decreases the overall sugar yield in comparison to pure AVICEL. In contrast to plant biomass, in which lignin is connected to cellulose, and hence physically blocks the access of cellulases to the cellulose polymers, the Organosolv lignin added to the cellulose in this case is not. To capture this, instead of a structured microfibril like shown in Fig 1, we consider polymers freely floating in the solution, where all their bonds can be easily accessed. In this configuration, lignin can still inhibit the cellulases by non-productively adsorbing on their surface, which we also represent in our model. Together panels (a) and (b) in Fig 2 show that the model is able to fairly reproduce the experimental time-courses, for each of the three cases considered.

As can be seen in Table 2, in the three cases considered in Fig 2, the Crystallinity Fraction of the model (averaged over the substrate) also fairly reproduces the Crystallinity Indices obtained by ssNMR. Since crystallinity in the model is quantified by two parameters: (i) the Crystallinity Fraction (noted CF) and (ii) the digestibility ratio (noted $r_{c,a}$) (as described in Sect 2.5), this confirms that ssNMR quantifies the abundance of crystalline bonds in the total substrate, rather than their propensity to be digested. It is also interesting to note that while the Crystallinity Fraction of AVICEL (53%) is much higher than that of Sigmacell (22%), their sugar yields from cellulose at 48 hours are similar (86.6%). This non-trivial observation is well rationalised by the digestibility ratio of cellulose (cellu $r_{c,a}$) determined by the model.

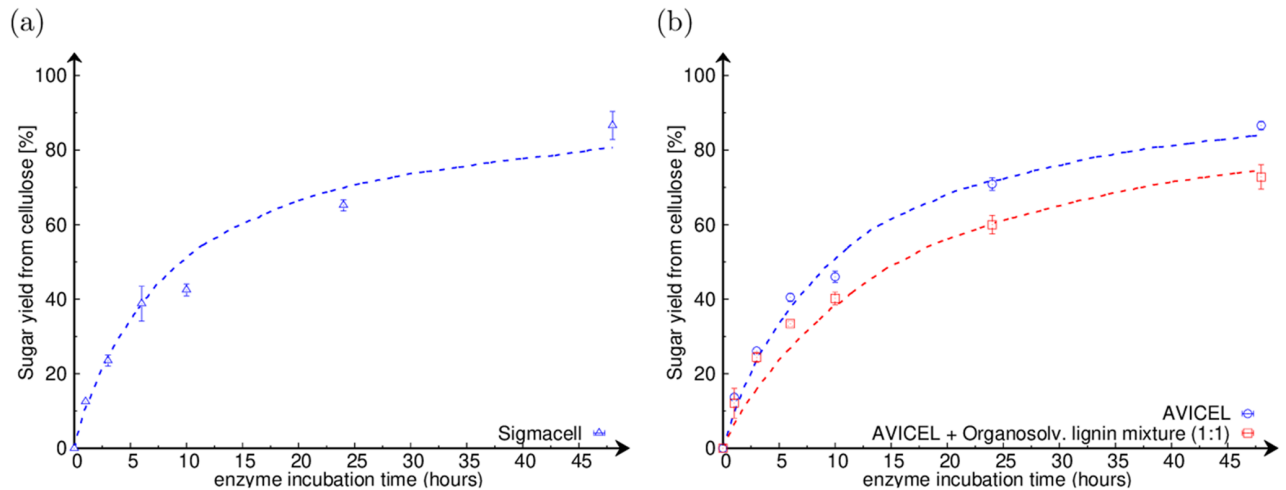


Fig 2. Enzymatic saccharification of cellulosic substrates. (a) Sigmacell; (b) in blue, AVICEL; in red, AVICEL and Organosolv lignin mixture (1:1). Experimental data are shown as points. Model simulation results are shown as dashed lines. The stochastic model fairly reproduces the saccharification dynamics in each case over the entire time-course.

<https://doi.org/10.1371/journal.pone.0322367.g002>

Table 2. Sugar yield from cellulose at 48 hours, experimentally measured Crystallinity Index (by ssNMR), and model parameters Crystallinity Fraction and digestibility ratio, from simulations that best fit saccharification time-courses in Fig 3.

Biomass source	Pre-treatment	Sugar yield from cellulose at 48 hrs [%]	Crystallinity Index from ssNMR [%]	Crystallinity Fraction [%]	$cellur_{c,a}$
Sigmacell	N.A	86.6 ± 3.8	22.0 ± 0.8	22.0	0.20
AVICEL	N.A	86.6 ± 1.1	53.0 ± 2.0	55.0	0.60
AVICEL and Organosolv Lignin	N.A	72.8 ± 3.3	53.0 ± 2.0	55.0	0.40
Beech wood	untreated	9.5 ± 0.2	21.0 ± 0.8	22.0	0.005
	NS+OCAT	50.8 ± 0.4	43.0 ± 1.6	45.0	0.04
	YS+OCAT	40.9 ± 0.6	51.0 ± 1.9	45.0	0.05
Miscanthus	untreated	11.5 ± 0.6	29.0 ± 0.8	31.0	0.05
	NS+OCAT	81.9 ± 0.3	45.0 ± 1.7	43.0	0.30
	YS+OCAT	47.0 ± 4.4	30.0 ± 1.1	33.0	0.001
Sida (Sida hermaphrodita)	untreated	12.4 ± 0.5	23.0 ± 0.9	22.0	0.05
	NS+OCAT	87.5 ± 4.3	38.0 ± 1.4	37.0	0.50
	YS+OCAT	40.4 ± 1.4	34.0 ± 1.3	36.0	0.001
Walnut shells	untreated	3.0 ± 0.2	26.0 ± 1.0	29.0	0.0001
	NS+OCAT	41.0 ± 2.5	37.0 ± 1.4	37.0	0.01
	YS+OCAT	22.6 ± 1.0	47.0 ± 1.7	45.0	0.0001

<https://doi.org/10.1371/journal.pone.0322367.t002>

For AVICEL, $r_{c,a} = 0.60$, while for Sigmacell, $r_{c,a} = 0.20$. It means that, the crystalline bonds are much harder to digest than the amorphous ones in Sigmacell as compared to AVICEL. This explains why a substrate with less crystalline bonds that are harder to digest (Sigmacell) releases the same amount of sugar like a substrate with more crystalline bonds that are easier to digest (AVICEL). This signifies that solely the measurement of the Crystallinity Index is not reflective of the amount of sugar being released from the substrate, which is instead influenced by a combination of the two factors: how much of the crystalline bonds are present (experimentally measured by Crystallinity Index and represented by the model parameter Crystallinity Fraction), and how hard they are to digest (represented by the model parameter digestibility ratio $r_{c,a}$).

3.2 Plant biomass lignocellulosic substrates

In Table 2, we also consider various plant biomasses (i.e. beech wood, miscanthus, sida hermaphrodita (sida), and walnut shells), and conditions (i.e. no pre-treatment, OrganoCat pre-treatment without swelling (NS+OCAT), and OrganoCat pre-treatment with swelling (YS+OCAT)). We observe that the measured Crystallinity Indices from ssNMR for the untreated biomasses are always lower than that of the OrganoCat pre-treated samples. This is expected to be induced by the pre-treatment process, which removes a majority of the hemicelluloses from the biomass, and thereby amorphous polymers, leading to an increase in crystallinity of the overall substrate. For miscanthus and sida, the additional swelling step (YS+OCAT compared to NS+OCAT) swells crystalline cellulose and reduces the Crystallinity Indices. For beech wood and walnut shells, instead, during swelling amorphous cellulose is partially hydrolysed, which increases the overall Crystallinity Indices. To evaluate if any correlation can be drawn between the Crystallinity Index measured by ssNMR and the 48-hour sugar release (see Table 2), we rank them as follows:

For untreated biomass:

CI

beech wood (21.0 ± 0.8) < sida (23.0 ± 0.9) < walnut shells (26.0 ± 1.0) < miscanthus (29.0 ± 0.8)

48-hour sugar release

sida (12.4 ± 0.5) \approx miscanthus (11.5 ± 0.6) > beech wood (9.5 ± 0.2) > walnut shells (3.0 ± 0.2)

For NS+OCAT:

CI

walnut shells (37.0 ± 1.4) \approx sida (38.0 ± 1.4) < beech wood (43.0 ± 1.6) \approx miscanthus (45.0 ± 1.7)

48-hour sugar release

sida (87.5 ± 4.3) > miscanthus (81.9 ± 0.3) > beech wood (50.8 ± 0.4) > walnut shells (41.0 ± 2.5)

For YS+OCAT:

CI

miscanthus (30.0 ± 1.1) < sida (34.0 ± 1.3) < walnut shells (47.0 ± 1.7) < beech wood (51.0 ± 1.9)

48-hour sugar release

miscanthus (47.0 ± 4.4) > beech wood (40.9 ± 0.6) \approx sida (40.4 ± 1.4) > walnut shells (22.6 ± 1.0).

It is therefore apparent that, for the different biomasses considered, there is no correlation between the ranking of the measured Crystallinity Indices and their respective saccharification yields, which is another clue that the Crystallinity Index measured *via* ssNMR is insufficient to completely capture the biomass recalcitrance. Instead, one must also account for the digestibility ratio ($r_{c,a}$) of the crystalline cellulose, that varies amongst biomasses.

Fig 3 displays the enzymatic saccharification time-courses of biomass from: (a) beech wood, (b) miscanthus, (c) sida, and (d) walnut shells, with experimental data being shown as points and simulation results as dashed lines. In each case, we observe that the untreated biomasses (in green) have the lowest sugar release; whereas the OrganoCat pre-treated ones without swelling (NS+OCAT, in red) have the highest, and the OrganoCat pre-treated ones with swelling (YS+OCAT, in blue) lie in between. In each case, the model fairly reproduces the experimental saccharification time-courses, while it takes into account the specific composition of the substrate (detailed in Table 1), and predicts a value for the model parameter

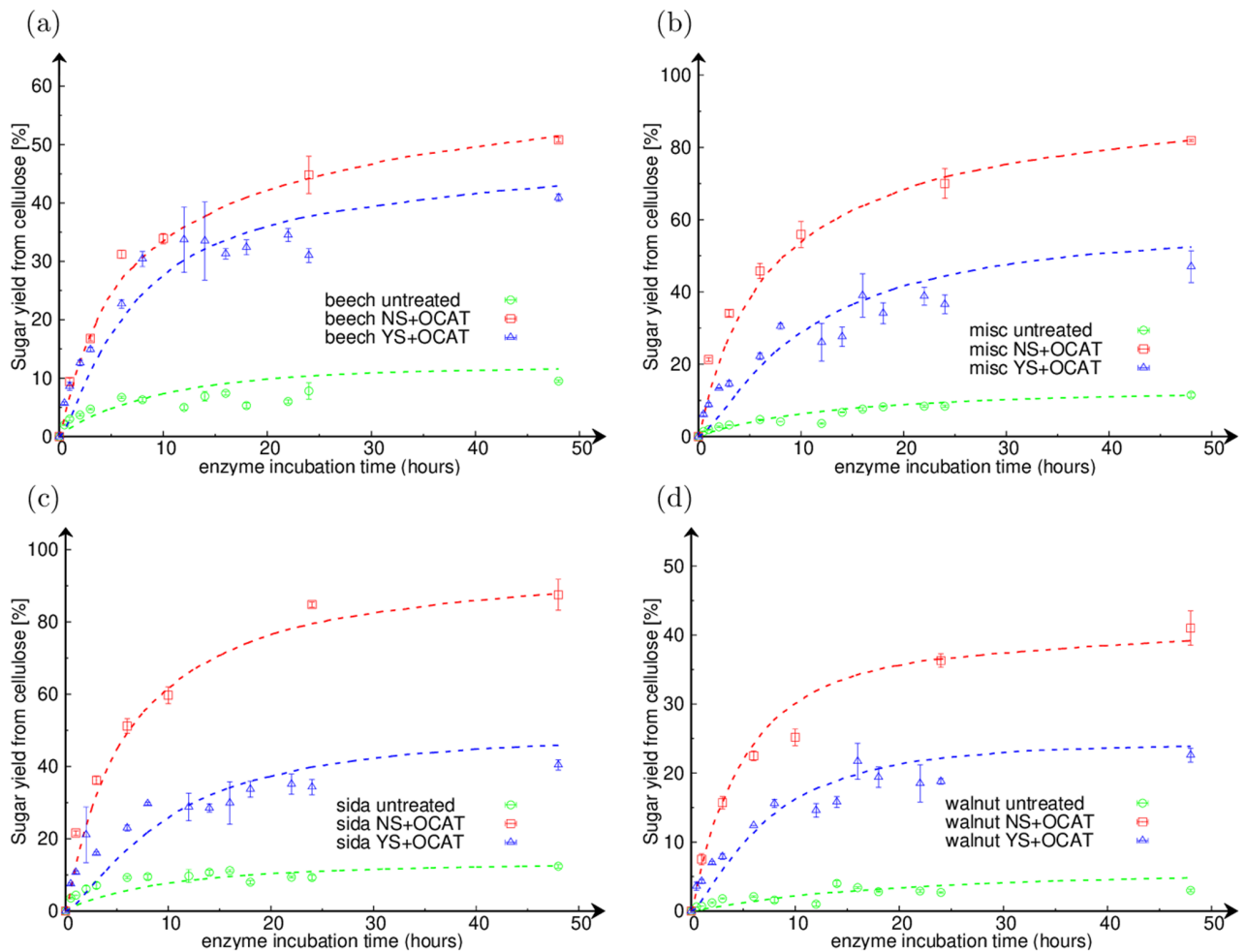


Fig 3. Enzymatic saccharification time-courses. The plots show the sugar release from cellulose *versus* the enzyme incubation time, for (a) beech (beech wood), (b) misc (miscanthus), (c) sida (*Sida hermaphrodita*), and (d) walnut (walnut shell) biomasses. Experimental data are shown as points and simulation results as dashed lines. Untreated biomass is in green, OCAT without swelling in red, and OCAT with swelling in blue.

<https://doi.org/10.1371/journal.pone.0322367.g003>

Crystallinity Fraction that matches well the experimental Crystallinity Index, excepted for OrganoCat pre-treated beech wood with swelling (detailed in Table 2). In this case, the Crystallinity Fraction provided by the model is roughly 6% below the experimental Crystallinity Index. However, one could note that the impact of swelling on OrganoCat pre-treated beech wood is unique as compared to the other biomasses. Beech wood shows the lowest decrease in sugar release from cellulose when comparing NS+OCAT and YS+OCAT (i.e. ca. 9.9%), while this amounts to ca. 34.9% for miscanthus, ca. 47.1% for sida, and ca. 17.4% for walnut shells. The uniqueness of beech wood is also reflected in the fitted value of the cellulose digestibility ratio ($cellu r_{c,a}$), which unlike for other biomasses, is almost the same when comparing NS+OCAT and YS+OCAT conditions.

3.3 The digestibility ratio ($r_{c,a}$)

In Table 2, considering all the samples investigated in this study, the $cellu r_{c,a}$ values indicate that the crystalline bonds in sida (NS+OCAT) are the easiest to enzymatically digest, while

those in walnut shells (untreated and YS+OCAT) are the hardest. Interestingly, this correlates with highest sugar conversion at 48 hours for sida and lowest one for walnut shells. Additionally, all the samples showing more than 70% of sugar yield from cellulose at 48 hours have a $cellu r_{c,a}$ value in the range 0.2–0.6, while in all other cases $cellu r_{c,a}$ is at least one order of magnitude lower. It is therefore apparent that distinct samples (considering biomass and pre-treatment) have crystalline bonds with largely different digestibility ratios ($r_{c,a}$), which strongly impacts, together with the Crystallinity Fraction, on the saccharification dynamics. Moreover, in the case of plant-sourced lignocellulosic biomasses, the digestibility ratios $cellu r_{c,a}$ follow the trend: NS+OCAT > untreated \geq YS+OCAT, which suggests that $cellu r_{c,a}$ is a signature of the impact of the pre-treatment on the biophysical properties of the crystalline polymers.

It must be made explicit that the physical meaning of $r_{c,a}$ is undoubtedly complex. Indeed, even though in the same chemical environment the same bonds need the same energy to be cleaved, there might be other factors such as the detailed spatial and structural arrangements of the molecules in and around the crystalline cellulose bonds which contribute to the biomass recalcitrance. For instance, the polysaccharide orientation, lignin composition and spatial arrangement of aromatic rings, lignin-carbohydrate-complexes, hemicellulosic decorations, or *kinks* in the linear arrangement of the cellulose chains may contribute to uneasiness for the enzymes to access and digest crystalline cellulose bonds. These features are neither accounted for in the model nor measured experimentally, and so, they might be captured, in a coarse-grained manner, by $r_{c,a}$.

Noticeably, the model can also reproduce the experimental saccharification time-courses for the different biomasses and pre-treatment conditions using the same value of the digestibility ratio or the same value of the enzyme kinetics for all samples. However, then, the predicted values of the model parameter Crystallinity Fraction are very far from the experimental Crystallinity Indices. Thus, we instead have digestibility ratios that are specific for each biomass and pre-treatment, and enzyme kinetic rates that are common to all samples under a given pre-treatment condition. This confirms the impact of pre-treatments on the material properties and the consequent ability of an enzyme cocktail to digest it.

3.4 Further insights from the model

Apart from the crystallinity parameters (Crystallinity Fraction and digestibility ratio), other model parameters can have an impact on the saccharification dynamics [58]. For instance, below we illustrate the role of the end product inhibition and enzymes' size, before summarising the overall phenomenology.

3.4.1 Impact of cellulase inhibition. Fig 4 shows the effect of the inhibition of cellulases by their end products cellobiose and glucose on the simulated saccharification dynamics. All the inhibition factors ω_y^x (see Sect 2.5) are kept equal to each other, and simultaneously varied. All the other parameters are kept fixed to the best-fitted values for the case of miscanthus OrganoCat pre-treated without swelling (which we arbitrarily picked). In Fig 4a, as one might intuitively expect, we observe that, as the value of the inhibition factor increases, the sugar yield from cellulose decreases. In Fig 4b, we study the effect of the end-product inhibition factors ω_y^x on the sugar release at different points of the time-course. At earlier times, an increase in the inhibition causes a pseudo-linear decrease in the sugar release. At later times, two regimes are observed: first a very slow decrease of the sugar release with increasing ω_y^x , second, a sharp decrease of the sugar release with increasing ω_y^x . Additionally, when comparing the sugar yield from cellulose at $\omega_y^x = 0$ and $\omega_y^x = 1$, we clearly observe that this difference is stronger for late time points than for early time points, which highlights the cumulative

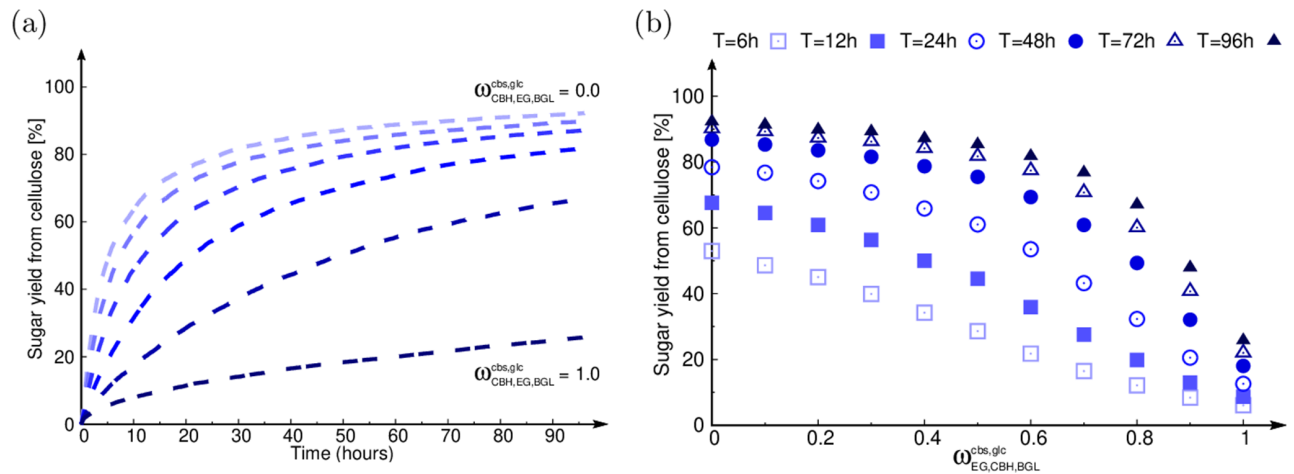


Fig 4. Simulated effect of end product inhibition on saccharification. (a) Saccharification time-courses with inhibition by both cellobiose and glucose. Each dashed line corresponds to a specific value for the inhibition factors, that are kept equal to one another. The inhibition factors are varied from $\omega_y^x = 0$ to $\omega_y^x = 1$ in steps of 0.2, where $x = cbs, glc$ and $y = EG, CBH, BGL$. (b) Sugar released from cellulose versus ω_y^x at different time points along the simulated saccharification time-course. Each point type corresponds to a specific time.

<https://doi.org/10.1371/journal.pone.0322367.g004>

impact of inhibition throughout the saccharification time-course. We also tested the cases where the only inhibitor was either cellobiose or glucose (data not shown). We observed that inhibition by only cellobiose has a negligible effect on the saccharification dynamics, since it does not accumulate and is instead readily digested into glucose by BGL. A very similar, but weaker, profile like in Fig 4 is observed when only glucose inhibits, hence, we deduce that inhibition by cellobiose reinforces that by glucose.

3.4.2 Effect of enzymes' size. Fig 5 shows the effect of varying enzymes' size on the simulated saccharification dynamics. All the other model parameters are fixed to the best-fitted values for miscanthus OrganoCat pre-treated without swelling (which we arbitrarily picked). The enzymes are approximated as hard spheres, with a radius of r_{enzyme} , such that they interact among themselves and with the substrate through excluded-volume interactions. Due to their intrinsic volume, both cellulases and hemicellulases may have limited access to their respective substrate because it is being shielded by surrounding non-substrate polymers. Furthermore, processive enzymes (i.e. CBH) that remain attached to the microfibril for a while, prevent other enzymes from approaching and digesting bonds. As a consequence, we observe that steric hindrance increases with the enzyme radius, which directly slows down the saccharification process. The sugar yield from cellulose at 48 hours can be fitted with the right side of a Gaussian distribution, using the parameters stated in the caption of Fig 5.

3.4.3 Overall phenomenology. Fig 6 illustrates how the various model parameters affect the saccharification time-courses for a generic lignocellulose substrate. We assume, for each line, the variation of only the marked parameter, with respect to the control, while all the other parameters, including the substrate composition, are considered to be constant. The saccharification yield increases with the increase in the enzyme reaction rates (K_{enzyme}), or with the decrease in the digestibility ratio ($r_{c,a}$). Conversely, the saccharification yield decreases with the increase of the end product inhibition (ω_y^x), or with that of the Crystallinity Fraction (CF), or with that of the enzymes' radius (r_{enzyme}). These generic considerations are presented for a control sample, whose substrate composition is kept fixed. In order to fully rationalise experimental saccharification time-courses (such as in Fig 3), one should be reminded that

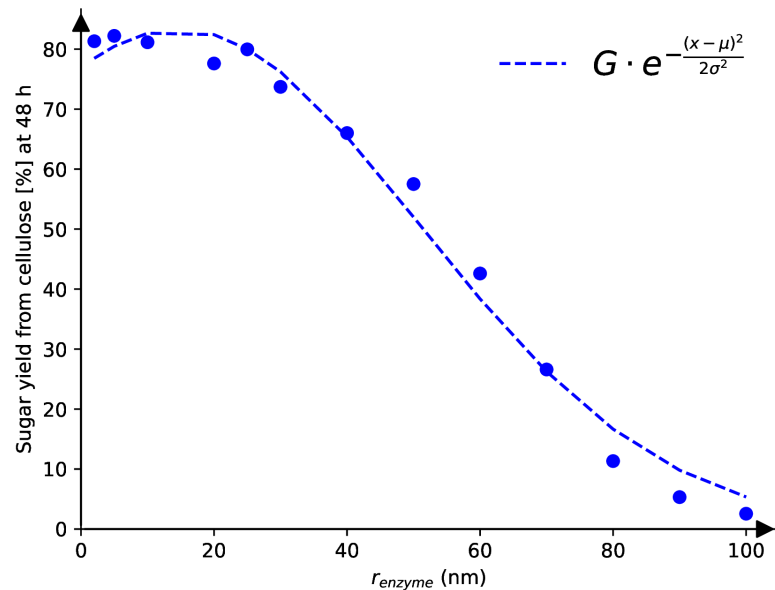


Fig 5. Simulated effect of the enzymes' size on the saccharification dynamics. Sugar yield from cellulose at 48 hours decreases with increasing enzymes' size, following a Gaussian trend ($G = 83.32 \pm 1.65$, $\mu = 14.62 \pm 2.25$, $\sigma = 36.43 \pm 2.06$).

<https://doi.org/10.1371/journal.pone.0322367.g005>

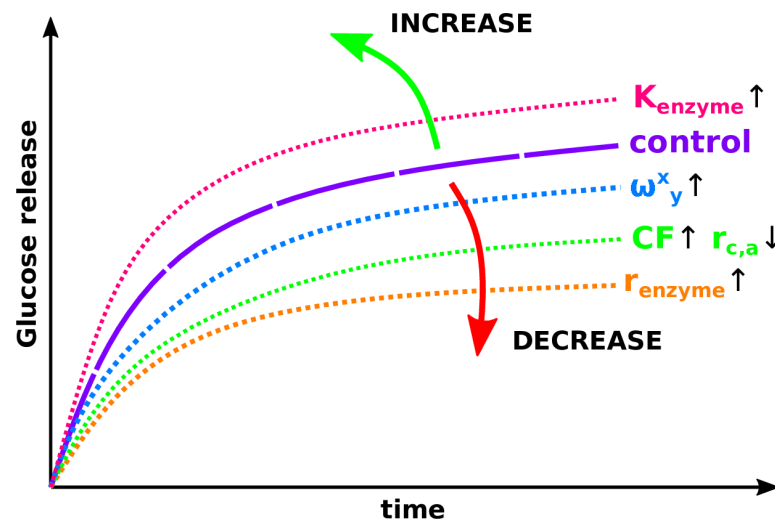


Fig 6. Representative plot showing the impact of several model parameters on the simulated saccharification time-courses, in comparison to a control case (purple line), considering a generic lignocellulose substrate. The depicted parameters are: the enzyme reaction rates (noted K_{enzyme} , pink line), the end product inhibition parameters (noted ω_y^x , blue line), the enzymes' size (noted r_{enzyme} , orange line), and the substrate crystallinity parameters (noted CF and $r_{c,a}$, green line).

<https://doi.org/10.1371/journal.pone.0322367.g006>

the variation in substrate composition can also play a significant role on its recalcitrance, in particular when considering biomass from different sources.

4 Discussion

In this study, we use a two-fold computational and experimental approach to investigate the enzymatic saccharification of not only purely cellulosic substrates (i.e. AVICEL and Sigma-cell) but also lignocellulosic biomasses from four different plant sources (i.e. beech wood, miscanthus, sida, and walnut shells), that we additionally examine under three distinct conditions (i.e. untreated, OrganoCat pre-treated without swelling, and OrganoCat pre-treated with swelling). We employ an upgraded version of our earlier published stochastic biophysical model [31] to reproduce the experimental saccharification time-courses of the fifteen samples while taking into account their specific composition. Our best-fits provide values for the model parameter Crystallinity Fraction that match well the Crystallinity Indices experimentally measured by ssNMR. The enzymatic saccharification of AVICEL in solution with Organosolv lignin highlights the inhibitory effect of lignin on saccharification due to its adsorption on the enzymes, even though it is not physically connected to the cellulosic substrate in that case. For the four plant-sourced biomasses, the pre-treatment consistently impacts on the sugar release. Compared to untreated biomasses, OrganoCat pre-treatment brings about a many-fold increase in the saccharification yield. Untreated biomasses have the lowest yield, OrganoCat pre-treated with swelling the intermediate one, and OrganoCat pre-treated without swelling the highest one.

The Crystallinity Index of distinct plant-sourced biomasses can be differently affected by swelling, that can either increase or decrease it. Additionally, with both plant-sourced biomasses and purely cellulosic substrates (i.e. AVICEL and Sigmacell), we highlight that the Crystallinity Index is not correlated to the sugar release that is instead determined by a combination of the Crystallinity Fraction (that is the model parameter for the experimentally determined Crystallinity Index) and the digestibility ratio. This clearly suggests that the Crystallinity Index measured by ssNMR is important but insufficient to predict the saccharification yield of a given biomass. Therefore, the digestibility ratio appears as an essential parameter to rationalise the saccharification dynamics, and while it can be different for distinct biomasses, it emerges as a footprint of the impact of pre-treatments on the material. Yet, its exact physical meaning is undoubtedly complex, and can be considered as a coarse-grained representation of various spatial and structural features of the material that explain why crystalline bonds are more or less easily digested in distinct biomasses. Other model parameters also impact on the saccharification dynamics, in particular the end product inhibition and the enzyme size. Although resulting from very distinct mechanisms, based on the simulations, we show that either an increase in end product inhibition or enzyme size consistently reduces the saccharification yield in a non-trivial fashion.

Overall, our study illustrates the benefit of modelling to rationalise complex experimental data related to the dynamics of lignocellulose biomass saccharification. Importantly, it also stresses the need for further experimental investigations towards deciphering the fine structural properties of the material that explain why the digestibility of crystalline regions vary that much across biomasses and pre-treatment conditions. In this perspective, complementary approaches could be based on lower scale modelling, such as atomistic-level Molecular Dynamics or Density Field Theory.

Acknowledgements

PSD: wrote the manuscript and its revised version equally with AR, on the basis of a draft prepared by himself, with some contribution of PG, primarily on the Material and Methods section. He performed the entire computational work, analysed the results, and produced all the figures.

PG: wrote the majority of the experimental Material and Methods section, extensively commented on the manuscript and its revised version, collected the experimental data for biomass preparation, composition, and saccharification, and analysed these results together with those from ssNMR spectra.

HH: provided access to the Jülich-Düsseldorf Biomolecular NMR Center, supervised the related experiments, and wrote the corresponding Materials and Methods paragraph.

HK: supervised the experimental work related to the biomass selection, preparation, composition, and saccharification. He contributed to the design of the study and obtained the PREDIG funds together with AR.

AR: designed the study, provided computational resources, supervised the computational work, and obtained the funding for the projects PREDIG, OptiCellu, and EtransColi that she all coordinates. Equally with PSD, she wrote the content of the manuscript and its revised version, on the basis of a draft prepared by PSD, with some contribution of PG, primarily on the Material and Methods section.

The authors express their gratitude to Benedict Ohrem for his essential contribution to conducting the experimental work. They also thank Dagmar Drobiez, Andrea Neuwohner, Michael Gitzen, and Stefan Gelderie for their assistance in preparing the materials and performing the wet chemical analysis. They also thank Luis Gardon and Nina Becker for their essential contribution to the ssNMR measurements of biomass Crystallinity Indices. Access to the Jülich-Düsseldorf Biomolecular NMR Center jointly run by Forschungszentrum Jülich and HHU is acknowledged.

Author Contributions

Conceptualization: Holger Klose, Adélaïde Raguin.

Data curation: Partho Sakha De, Philipp M Grande.

Formal analysis: Partho Sakha De, Philipp M Grande, Adélaïde Raguin.

Funding acquisition: Henrike Heise, Holger Klose, Adélaïde Raguin.

Investigation: Partho Sakha De, Philipp M Grande, Henrike Heise.

Methodology: Partho Sakha De, Philipp M Grande, Holger Klose, Adélaïde Raguin.

Project administration: Adélaïde Raguin.

Resources: Henrike Heise, Holger Klose, Adélaïde Raguin.

Software: Partho Sakha De.

Supervision: Henrike Heise, Holger Klose, Adélaïde Raguin.

Validation: Partho Sakha De, Philipp M Grande, Adélaïde Raguin.

Visualization: Partho Sakha De, Philipp M Grande.

Writing – original draft: Partho Sakha De, Philipp M Grande, Adélaïde Raguin.

Writing – review & editing: Partho Sakha De, Philipp M Grande, Adélaïde Raguin.

Appendix

Glossary

List of symbols

$r_{c,a}$	Ratio of digestibility of the crystalline to the amorphous bonds
CF	Crystallinity fraction
[EG]	Concentration of endoglucanase
[EG] ₀	Concentration of endoglucanase as if no inhibition would take place
[CBH]	Concentration of cellobiohydrolase
[CBH] ₀	Concentration of cellobiohydrolase as if no inhibition would take place
[BGL]	Concentration of β -glucosidase
[BGL] ₀	Concentration of β -glucosidase as if no inhibition would take place
ω_{EG}^{cbs}	Inhibition factor for cellobiose on endoglucanase
ω_{CBH}^{cbs}	Inhibition factor for cellobiose on cellobiohydrolase
ω_{EG}^{glc}	Inhibition factor for glucose on endoglucanase
ω_{CBH}^{glc}	Inhibition factor for glucose on cellobiohydrolase
ω_{BGL}^{glc}	Inhibition factor for glucose on β -glucosidase
[cbs]	Concentration of free cellobiose in the system
[glc]	Concentration of free glucose in the system
r_{enzyme}	Radius of the enzymes (hard-spheres)
μ_{defect}	Mean size of amorphous defects in crystalline cellulose and hemicellulose
N_{defect}	Number of amorphous defects in crystalline cellulose and hemicellulose

List of abbreviations

DP: degree of polymerisation, EG: endoglucanase, CBH: cellobiohydrolase, BGL: β -glucosidase, HC: hemicellulase, MLG: mixed-linkage glucan, DFT: density functional theory, MD: molecular dynamics, MPS: matrix polysaccharide, AIR: alcohol insoluble residue, ssNMR: solid-state nuclear magnetic resonance, OCAT: OrganoCat, OCAT+NS: OrganoCat No-Swelling, OCAT+YS: OrganoCat Yes-swelling.

Input parameters and files of the stochastic biophysical model

The model source code available in the [GitLab repository](#) is organised such that input parameters are read as text files, which are in the folder 'Params'. These files are named 'kinetic_parameters.txt' and 'initial configuration_parameters_*.txt'. The parameters are listed as comments at the bottom of the respective files for ease of the user. We also list them below.

Parameters listed in the input file 'kinetic_parameters.txt':

- ⇒ **EG K_{cat}**
parameter #1
Endoglucanase catalytic constant.
Units: 1/s. Typical Range: 0.001-1000.
- ⇒ **EG K_m**
parameter #2
Endoglucanase Michaelis-Menten constant.
Units: mM. Typical Range: 0.001-1000.
- ⇒ **CBH processive digestion rate**

- parameter #3
Reaction rate of cellobiohydrolase processive action.
Units: reactions/hour. Typical Range: 10-10,000.
- ⇒ **BGL K_{cat}**
parameter #4
 β -glucosidase catalytic constant.
Units: 1/s. Typical Range: 0.001-1000.
- ⇒ **BGL K_m**
parameter #5
 β -glucosidase Michaelis-Menten constant.
Units: mM. Typical Range: 0.001-1000.
- ⇒ **XYL K_{cat}**
parameter #6
Xylanase catalytic constant.
Units: 1/s. Typical Range: 0.001-1000.
- ⇒ **XYL K_m**
parameter #7
Xylanase Michaelis-Menten constant.
Units: mM. Typical Range: 0.001-1000.
- ⇒ **CBH K_{cat} (attachment reaction)**
parameter #8
Cellobiohydrolase catalytic constant.
Units: 1/s. Typical Range: 0.001-1000.
- ⇒ **CBH K_m (attachment reaction)**
parameter #9
Cellobiohydrolase Michaelis-Menten constant.
Units: mM. Typical Range: 0.001-1000.
- ⇒ **inhibition binding affinity of cellobiose to EG**
parameter #10
 ω_{EG}^{cbs}
Units: none. Range: 0-1.
- ⇒ **inhibition binding affinity of cellobiose to CBH**
parameter #11
 ω_{CBH}^{cbs}
Units: none. Range: 0-1.
- ⇒ **inhibition binding affinity of glucose to EG**
parameter #12
 ω_{EG}^{glc}
Units: none. Range: 0-1.
- ⇒ **inhibition binding affinity of glucose to CBH**
parameter #13
 ω_{CBH}^{glc}
Units: none. Range: 0-1.
- ⇒ **inhibition binding affinity of glucose to BGL**
parameter #14
 ω_{BGL}^{glc}
Units: none. Range: 0-1.
- ⇒ **Enzyme size: radius**
parameter #15

r_{enzyme}

Units: nm. Value used: 2.5 (based on formula 2.2 in [59] and enzyme molecular masses from BRENDA.)

Parameters listed in the input file 'initial_configuration_parameters_*.txt':

- ⇒ **mode_code**
parameter #1
Determines the shape of the microfibril and the number of cellulose polymers present.
Range: 1-5 (integers only). Mode code = 1 or 2, 24 polymers; Mode code = 3 or 4 (used here), 18 polymers; Mode code = 5, 36 polymers.
- ⇒ **pct_EG**
parameter #2
Relative fraction of endoglucanase in the enzyme cocktail.
Units: none. Value used: 0.135 (based on the most efficient enzyme cocktail design from literature [60,61] and calculations in [31]). Range: 0-1.
- ⇒ **pct_CBH**
parameter #3
Relative fraction of cellobiohydrolase in the enzyme cocktail.
Units: none. Range: 0-1. Value used: 0.353 (based on the most efficient enzyme cocktail design from literature [60,61] and calculations in [31]).
- ⇒ **pct_BGL**
parameter #4
Relative fraction of β -glucosidase in the enzyme cocktail.
Units: none. Range: 0-1. Value used: 0.100 (based on the most efficient enzyme cocktail design from literature [60,61] and calculations in [31]).
- ⇒ **pct_XYL**
parameter #5
Relative fraction of xylanase in the enzyme cocktail.
Units: none. Range: 0-1. Value used: 0.412 (based on the most efficient enzyme cocktail design from literature [60,61] and calculations in [31]).
- ⇒ **total_enz_molecules**
parameter #6
Total number of enzyme molecules involved in the digestion of the single simulated microfibril.
Units: none. Value used: 50.
- ⇒ **length_fibril**
parameter #7
Length of the simulated microfibril in terms of β -1-4 bonds in the cellulose chain.
Units: none. Value used: 200.
- ⇒ **boolean_Xyl_or_MLG**
parameter #8
It determines whether the hemicellulose contains either xylose(1) or mixed-linkage glucans(0).
Units: none. Value used: 1.
- ⇒ **pct_xyl**
parameter #9
Percentage of xylose in the hemicellulose.

Value: 1 (Currently unused, kept as a placeholder for future upgrades to accommodate multiple significant hemicellulose types at the same time).

- ⇒ **pct_cellu**
parameter #10
Relative fraction of cellulose in the substrate composition.
Unit: none. Range: 0.01-1 (value as user input from [Table 1](#)).
- ⇒ **pct_hemi**
parameter #11
Relative fraction of hemicellulose in the substrate composition.
Unit: none. Range: 0.01-1 (value as user input from [Table 1](#)).
- ⇒ **pct_lign**
parameter #12
Relative fraction of lignin in the substrate composition.
Unit: none. Range: 0.01-1 (value as user input from [Table 1](#)).
- ⇒ **pct_acetyl_hemi**
parameter #13
Currently unused. Value: 0
- ⇒ **pct_crystalline_cellu**
parameter #14
CF cellulose.
Units: none. Range: 0-1. (Fitted uniquely per sample to reproduce saccharification time-courses and CI of total biomass)
- ⇒ **pct_crystalline_hemi**
parameter #15
CF hemicellulose.
Units: none. Range: 0-1. (Fitted uniquely per sample to reproduce saccharification time-courses and CI of total biomass)
- ⇒ **Mean_defect_size**
parameter #16
 μ_{defect}
Units: none. Range: 0-0.5.
- ⇒ **Nbr_of_defects**
parameter #17
 N_{defect}
Units: none. Range: 0-1.
- ⇒ **r_monomer**
parameter #18
Radius of a single monomer of cellulose and hemicellulose.
Units: nm. Value fixed: 0.6.
- ⇒ **Lignin_adhesion_rate**
parameter #19
The number of monolignols that undergo non-productive adsorption on a single enzyme molecule. A smaller number indicates higher non-productive adsorption.
Units: none. Typical Range: 100-350.
- ⇒ **digestibility_ratio_cellu**
parameter #20
 $r_{c,a}$ cellulose.
Units: none. Range: 0.00001-1. (Fitted uniquely per sample to reproduce saccharification timecourses and CI of total biomass)

⇒ **digestibility_ratio hemi**

parameter #21

 $r_{c,a}$ hemicellulose.

Units: none. Range: 0.00001-1. (Fitted uniquely per sample to reproduce saccharification timecourses and CI of total biomass)

References

- Weidener D, Dama M, Dietrich SK, Ohrem B, Pauly M, Leitner W, et al. Multiscale analysis of lignocellulose recalcitrance towards OrganoCat pretreatment and fractionation. *Biotechnol Biofuels*. 2020;13:155. <https://doi.org/10.1186/s13068-020-01796-8> PMID: 32944071
- Damm T, Grande PM, Jablonowski ND, Thiele B, Disko U, Mann U, et al. OrganoCat pretreatment of perennial plants: Synergies between a biogenic fractionation and valuable feedstocks. *Bioresour Technol*. 2017;244(Pt 1):889–96. <https://doi.org/10.1016/j.biortech.2017.08.027> PMID: 28847077
- Hassan SS, Williams GA, Jaiswal AK. Emerging technologies for the pretreatment of lignocellulosic biomass. *Bioresour Technol*. 2018;262:310–8. <https://doi.org/10.1016/j.biortech.2018.04.099> PMID: 29729930
- Hu Z, Wen Z. Enhancing enzymatic digestibility of switchgrass by microwave-assisted alkali pretreatment. *Biochem Eng J*. 2008;38(3):369–78. <https://doi.org/10.1016/j.bej.2007.08.001>
- Mankar AR, Pandey A, Modak A, Pant KK. Pretreatment of lignocellulosic biomass: A review on recent advances. *Bioresour Technol*. 2021;334:125235. <https://doi.org/10.1016/j.biortech.2021.125235> PMID: 33957458
- Jørgensen H, Kristensen JB, Felby C. Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities. *Biofuels Bioprod Bioref*. 2007;1(2):119–34. <https://doi.org/10.1002/bbb.4>
- Bura R, Chandra R, Saddler J. Influence of xylan on the enzymatic hydrolysis of steam-pretreated corn stover and hybrid poplar. *Biotechnol Prog*. 2009;25(2):315–22. <https://doi.org/10.1002/btpr.98> PMID: 19266561
- Yao K, Wu Q, An R, Meng W, Ding M, Li B, et al. Hydrothermal pretreatment for deconstruction of plant cell wall: Part II. Effect on cellulose structure and bioconversion. *AIChE J*. 2018;64(6):1954–64. <https://doi.org/10.1002/aic.16106>
- Deng Z, Xia A, Liao Q, Zhu X, Huang Y, Fu Q. Laccase pretreatment of wheat straw: effects of the physicochemical characteristics and the kinetics of enzymatic hydrolysis. *Biotechnol Biofuels*. 2019;12:159. <https://doi.org/10.1186/s13068-019-1499-3> PMID: 31249622
- Grande PM, Viell J, Theyssen N, Marquardt W, Domínguez de María P, Leitner W. Fractionation of lignocellulosic biomass using the OrganoCat process. *Green Chem*. 2015;17(6):3533–9. <https://doi.org/10.1039/c4gc02534b>
- Weidener D, Klose H, Leitner W, Schurr U, Usadel B, Domínguez de María P, et al. One-step lignocellulose fractionation by using 2,5-furandicarboxylic acid as a biogenic and recyclable catalyst. *ChemSusChem*. 2018;11(13):2051–6. <https://doi.org/10.1002/cssc.201800653> PMID: 29791080
- Weidener D, Leitner W, Domínguez de María P, Klose H, Grande PM. Lignocellulose fractionation using recyclable phosphoric acid: Lignin, cellulose, and furfural production. *ChemSusChem*. 2021;14(3):909–16. <https://doi.org/10.1002/cssc.202002383> PMID: 33244874
- Ciesielski PN, Pecha MB, Lattanzi AM, Bharadwaj VS, Crowley MF, Bu L, et al. Advances in multiscale modeling of lignocellulosic biomass. *ACS Sustain Chem Eng*. 2020;8(9):3512–31. <https://doi.org/10.1021/acssuschemeng.9b07415>
- Hardy BJ, Sarko A. Molecular dynamics simulations and diffraction-based analysis of the native cellulose fibre: structural modelling of the I- α and I- β phases and their interconversion. *Polymer*. 1996;37(10):1833–9. [https://doi.org/10.1016/0032-3861\(96\)87299-5](https://doi.org/10.1016/0032-3861(96)87299-5)
- Gani TZH, Orella MJ, Anderson EM, Stone ML, Brushett FR, Beckham GT, et al. Computational evidence for kinetically controlled radical coupling during lignification. *ACS Sustain Chem Eng*. 2019;7(15):13270–7. <https://doi.org/10.1021/acssuschemeng.9b02506>
- Zhao Z, Crespi VH, Kubicki JD, Cosgrove DJ, Zhong L. Erratum to: Molecular dynamics simulation study of xyloglucan adsorption on cellulose surfaces: effects of surface hydrophobicity and side-chain variation. *Cellulose*. 2013;21(2):1041–2. <https://doi.org/10.1007/s10570-013-0074-5>
- Vermaas JV, Petridis L, Qi X, Schulz R, Lindner B, Smith JC. Mechanism of lignin inhibition of enzymatic biomass deconstruction. *Biotechnol Biofuels*. 2015;8:217. <https://doi.org/10.1186/s13068-015-0379-8> PMID: 26697106
- Zhou X, Li W, Mabon R, Broadbelt LJ. A mechanistic model of fast pyrolysis of hemicellulose. *Energy Environ Sci*. 2018;11(5):1240–60. <https://doi.org/10.1039/c7ee03208k>

19. Elder T, Beste A. Density functional theory study of the concerted pyrolysis mechanism for lignin models. *Energy Fuels*. 2014;28(8):5229–35. <https://doi.org/10.1021/ef5013648>
20. Zhou S, Jin K, Buehler MJ. Understanding plant biomass via computational modeling. *Adv Mater*. 2021;33(28):e2003206. <https://doi.org/10.1002/adma.202003206> PMID: 32945027
21. Taylor LE 2nd, Knott BC, Baker JO, Alahuhta PM, Hobdey SE, Linger JG, et al. Engineering enhanced cellobiohydrolase activity. *Nat Commun*. 2018;9(1):1186. <https://doi.org/10.1038/s41467-018-03501-8> PMID: 29567941
22. Knott BC, Haddad Momeni M, Crowley MF, Mackenzie LF, Götz AW, Sandgren M, et al. The mechanism of cellulose hydrolysis by a two-step, retaining cellobiohydrolase elucidated by structural and transition path sampling studies. *J Am Chem Soc*. 2014;136(1):321–9. <https://doi.org/10.1021/ja410291u> PMID: 24341799
23. Ingólfsson HI, Lopez CA, Uusitalo JJ, de Jong DH, Gopal SM, Periole X, et al. The power of coarse graining in biomolecular simulations. *Wiley Interdiscip Rev Comput Mol Sci*. 2014;4(3):225–48. <https://doi.org/10.1002/wcms.1169> PMID: 25309628
24. Kumar D, Murthy GS. Stochastic molecular model of enzymatic hydrolysis of cellulose for ethanol production. *Biotechnol Biofuels*. 2013;6(1):63. <https://doi.org/10.1186/1754-6834-6-63> PMID: 23638989
25. Kumar D, Murthy GS. Development and validation of a stochastic molecular model of cellulose hydrolysis by action of multiple cellulase enzymes. *Bioresour Bioprocess*. 2017;4(1). <https://doi.org/10.1186/s40643-017-0184-2>
26. Griggs AJ, Stickel JJ, Lischeske JJ. A mechanistic model for enzymatic saccharification of cellulose using continuous distribution kinetics I: depolymerization by EGI and CBHI. *Biotechnol Bioeng*. 2012;109(3):665–75. <https://doi.org/10.1002/bit.23355> PMID: 22034153
27. Griggs AJ, Stickel JJ, Lischeske JJ. A mechanistic model for enzymatic saccharification of cellulose using continuous distribution kinetics II: cooperative enzyme action, solution kinetics, and product inhibition. *Biotechnol Bioeng*. 2012;109(3):676–85. <https://doi.org/10.1002/bit.23354> PMID: 22034106
28. Lischeske JJ, Stickel JJ. A two-phase substrate model for enzymatic hydrolysis of lignocellulose: application to batch and continuous reactors. *Biotechnol Biofuels*. 2019;12:299. <https://doi.org/10.1186/s13068-019-1633-2> PMID: 31890027
29. Vetharanim I, Kelly W, Attwood G, Harris P. A 3-D model of a perennial ryegrass primary cell wall and its enzymatic degradation. *Computation*. 2014;2(2):23–46. <https://doi.org/10.3390/computation2020023>
30. Asztalos A, Daniels M, Sethi A, Shen T, Langan P, Redondo A, et al. A coarse-grained model for synergistic action of multiple enzymes on cellulose. *Biotechnol Biofuels*. 2012;5(1):55. <https://doi.org/10.1186/1754-6834-5-55> PMID: 22853643
31. Behle E, Raguin A. Stochastic model of lignocellulosic material saccharification. *PLoS Comput Biol*. 2021;17(9):e1009262. <https://doi.org/10.1371/journal.pcbi.1009262> PMID: 34516546
32. De PS, Glass T, Stein M, Spitzlei T, Raguin A. PREDIG: Web application to model and predict the enzymatic saccharification of plant cell wall. *Comput Struct Biotechnol J*. 2023;21:5463–75. <https://doi.org/10.1016/j.csbj.2023.09.026> PMID: 38022701
33. Simmons TJ, Mortimer JC, Bernardinelli OD, Pöppler A-C, Brown SP, deAzevedo ER, et al. Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nat Commun*. 2016;7:13902. <https://doi.org/10.1038/ncomms13902> PMID: 28000667
34. Malgas S, Kwanya Minghe VM, Pletschke BI. The effect of hemicellulose on the binding and activity of cellobiohydrolase I, Cel7A, from *Trichoderma reesei* to cellulose. *Cellulose*. 2019;27(2):781–97. <https://doi.org/10.1007/s10570-019-02848-5>
35. Jablonowski ND, Kollmann T, Nabel M, Damm T, Klose H, Müller M, et al. Valorization of *Sida* (*Sida hermaphrodita*) biomass for multiple energy purposes. *GCB Bioenergy*. 2016;9(1):202–14. <https://doi.org/10.1111/gcbb.12346>
36. Foster CE, Martin TM, Pauly M. Comprehensive compositional analysis of plant cell walls (Lignocellulosic biomass) part I: lignin. *J Vis Exp*. 2010;(37):1745. <https://doi.org/10.3791/1745> PMID: 20224547
37. Foster CE, Martin TM, Pauly M. Comprehensive compositional analysis of plant cell walls (lignocellulosic biomass) part II: carbohydrates. *J Vis Exp*. 2010;(37):1837. <https://doi.org/10.3791/1837> PMID: 20228730
38. Grande PM, de Maria PD. Enzymatic hydrolysis of microcrystalline cellulose in concentrated seawater. *Bioresour Technol*. 2012;104:799–802. <https://doi.org/10.1016/j.biortech.2011.10.071> PMID: 22101072
39. Lever M. Carbohydrate determination with 4-hydroxybenzoic acid hydrazide (PAHBAH): effect of bismuth on the reaction. *Anal Biochem*. 1977;81(1):21–7.

- [https://doi.org/10.1016/0003-2697\(77\)90594-2](https://doi.org/10.1016/0003-2697(77)90594-2) PMID: 900479
40. Ding S-Y, Himmel ME. The maize primary cell wall microfibril: a new model derived from direct visualization. *J Agric Food Chem*. 2006;54(3):597–606. <https://doi.org/10.1021/jf051851z> PMID: 16448156
 41. Song B, Zhao S, Shen W, Collings C, Ding S-Y. Direct measurement of plant cellulose microfibril and bundles in native cell walls. *Front Plant Sci*. 2020;11:479. <https://doi.org/10.3389/fpls.2020.00479> PMID: 32391038
 42. Turner S, Kumar M. Cellulose synthase complex organization and cellulose microfibril structure. *Philos Trans A Math Phys Eng Sci*. 2018;376(2112):20170048. <https://doi.org/10.1098/rsta.2017.0048> PMID: 29277745
 43. Jarvis MC. Structure of native cellulose microfibrils, the starting point for nanocellulose manufacture. *Philos Trans A Math Phys Eng Sci*. 2018;376(2112):20170045. <https://doi.org/10.1098/rsta.2017.0045> PMID: 29277742
 44. Vermaas JV, Petridis L, Qi X, Schulz R, Lindner B, Smith JC. Mechanism of lignin inhibition of enzymatic biomass deconstruction. *Biotechnol Biofuels*. 2015;8:217. <https://doi.org/10.1186/s13068-015-0379-8> PMID: 26697106
 45. van der Weijde T, Torres AF, Dolstra O, Dechesne A, Visser RGF, Trindade LM. Impact of different lignin fractions on saccharification efficiency in diverse species of the bioenergy crop miscanthus. *Bioenerg Res*. 2015;9(1):146–56. <https://doi.org/10.1007/s12155-015-9669-z>
 46. Ying W, Shi Z, Yang H, Xu G, Zheng Z, Yang J. Effect of alkaline lignin modification on cellulase-lignin interactions and enzymatic saccharification yield. *Biotechnol Biofuels*. 2018;11:214. <https://doi.org/10.1186/s13068-018-1217-6> PMID: 30083227
 47. Vermaas JV, Crowley MF, Beckham GT. A quantitative molecular atlas for interactions between lignin and cellulose. *ACS Sustain Chem Eng*. 2019;7(24):19570–83. <https://doi.org/10.1021/acssuschemeng.9b04648>
 48. Scapin SMN, Souza FHM, Zanphorlin LM, de Almeida TS, Sade YB, Cardoso AM, et al. Structure and function of a novel GH8 endoglucanase from the bacterial cellulose synthase complex of *Raoultella ornithinolytica*. *PLoS One*. 2017;12(4):e0176550. <https://doi.org/10.1371/journal.pone.0176550> PMID: 28448629
 49. Brady SK, Sreelatha S, Feng Y, Chundawat SPS, Lang MJ. Cellobiohydrolase 1 from *Trichoderma reesei* degrades cellulose in single cellobiose steps. *Nat Commun*. 2015;6:10149. <https://doi.org/10.1038/ncomms10149> PMID: 26657780
 50. Beldman G, Voragen AG, Rombouts FM, Searle-van Leeuwen MF, Pilnik W. Adsorption and kinetic behavior of purified endoglucanases and exoglucanases from *Trichoderma viride*. *Biotechnol Bioeng*. 1987;30(2):251–7. <https://doi.org/10.1002/bit.260300215> PMID: 18581306
 51. Straathof AJJ. Transformation of biomass into commodity chemicals using enzymes or cells. *Chem Rev*. 2014;114(3):1871–908. <https://doi.org/10.1021/cr400309c> PMID: 23987659
 52. Jørgensen H, Kristensen JB, Felby C. Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities. *Biofuels Bioprod Bioref*. 2007;1(2):119–34. <https://doi.org/10.1002/bbb.4>
 53. Chirico WJ, Brown RD. Purification and characterization of a beta-glucosidase from *Trichoderma reesei*. *Eur J Biochem*. 1987;165(2):333–41.
 54. Gusakov AV, Sinitsyn AP, Gerasimas VB, Savitskene RYu, Steponavichus YuYu. A product inhibition study of cellulases from *Trichoderma longibrachiatum* using dyed cellulose. *J Biotechnol*. 1985;3(3):167–74. [https://doi.org/10.1016/0168-1656\(85\)90017-3](https://doi.org/10.1016/0168-1656(85)90017-3)
 55. Ciesielski PN, Wagner R, Bharadwaj VS, Killgore J, Mittal A, Beckham GT, et al. Nanomechanics of cellulose deformation reveal molecular defects that facilitate natural deconstruction. *Proc Natl Acad Sci USA*. 2019;116(20):9825–30. <https://doi.org/10.1073/pnas.1900161116> PMID: 31036649
 56. Wu Z, Peng K, Zhang Y, Wang M, Yong C, Chen L, et al. Lignocellulose dissociation with biological pretreatment towards the biochemical platform: A review. *Mater Today Bio*. 2022;16:100445. <https://doi.org/10.1016/j.mtbio.2022.100445> PMID: 36212906
 57. Lan L, Chen H, Lee D, Xu S, Skillen N, Tedstone A, et al. Effect of ball-milling pretreatment of cellulose on its photoreforming for H₂ production. *ACS Sustain Chem Eng*. 2022;10(15):4862–71. <https://doi.org/10.1021/acssuschemeng.1c07301> PMID: 35574430
 58. De PS, Theilmann J, Raguin A. A detailed sensitivity analysis identifies the key factors influencing the enzymatic saccharification of lignocellulosic biomass. *Comput Struct Biotechnol J*. 2024;23:1005–15. <https://doi.org/10.1016/j.csbj.2024.01.006> PMID: 38420218
 59. Erickson HP. Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. *Biol Proced Online*. 2009;11:32–51. <https://doi.org/10.1007/s12575-009-9008-x> PMID: 19495910

60. Sinitsyn AP, Korotkova OG, Sinitsyna OA, Rozhkova AM, Dotsenko GS, Proskurina OV, et al. Optimizing the composition of cellulase enzyme complex from *Penicillium verruculosum*: Enhancing hydrolytic capabilities via genetic engineering. *Catal Ind.* 2016;8(1):101–6.
<https://doi.org/10.1134/s2070050416010128>
61. Agrawal R, Semwal S, Kumar R, Mathur A, Gupta RP, Tuli DK, et al. Synergistic enzyme cocktail to enhance hydrolysis of steam exploded wheat straw at pilot scale. *Front Energy Res.* 2018;6.
<https://doi.org/10.3389/fenrg.2018.00122>