A Dynamic Model for Cellulosic Biomass Hydrolysis: a Comprehensive Analysis and Validation of Hydrolysis and Product Inhibition Mechanisms

**Chien-Tai Tsai · Ricardo Morales-Rodriguez ·
Gürkan Sin · Anne S. Meyer**

Abstract The objective of this study is to perform a comprehensive enzyme kinetics analysis in view of validating and consolidating a semimechanistic kinetic model consisting of homogeneous and heterogeneous reactions for enzymatic hydrolysis of lignocellulosic biomass proposed by the U.S. National Renewable Energy Laboratory (Kadam et al., Biotechnol Prog 20(3):698–705, 2004) and its variations proposed in this work. A number of dedicated experiments were carried out under a range of initial conditions (Avicel® versus pretreated barley straw as substrate, different enzyme loadings and different product inhibitors such as glucose, cellobiose and xylose) to test the hydrolysis and product inhibition mechanisms of the model. A nonlinear least squares method was used to identify the model and estimate kinetic parameters based on the experimental data. The suitable mathematical model for industrial application was selected among the proposed models based on statistical information (weighted sum of square errors). The analysis showed that transglycosylation plays a key role at high glucose levels. It also showed that the values of parameters depend on the selected experimental data used for parameter estimation. Therefore, the parameter values are not universal

C.-T. Tsai · A. S. Meyer (✉)
Center for BioProcess Engineering, Department of Chemical and Biochemical Engineering, Technical University of Denmark, 2800 Lyngby, Denmark
e-mail: am@kt.dtu.dk

R. Morales-Rodriguez · G. Sin
Computer-Aided Process-Product Engineering Center, Department of Chemical and Biochemical Engineering, Technical University of Denmark, 2800 Lyngby, Denmark

R. Morales-Rodriguez
e-mail: rmro@xanum.uam.mx

G. Sin
e-mail: gsi@kt.dtu.dk

R. Morales-Rodriguez
Departamento de Ingeniería de Procesos e Hidráulica, Universidad Autónoma Metropolitana-Iztapalapa, Av. San Rafael Atlixco 186, C.P. 09340 Mexico City, Mexico
e-mail: rmro@xanum.uam.mx

and should be used with caution. The model proposed by Kadam et al. (Biotechnol Prog 20(3):698–705, 2004) failed to predict the hydrolysis phenomena at high glucose levels, but when combined with transglycosylation reaction(s), the prediction of cellulose hydrolysis behaviour over a broad range of substrate concentrations (50–150 g/L) and enzyme loadings (15.8–31.6 and 1–5.9 mg protein/g cellulose for Celluclast and Novozyme 188, respectively) was possible. This is the first study introducing transglycosylation into the semimechanistic model. As long as these type of models are used within the boundary of their validity (substrate type, enzyme source and substrate concentration), they can support process design and technology improvement efforts at pilot and full-scale studies.

Nomenclature

BG	β-Glucosidase
CBH	Exo-1,4-β-D-glucanases
Cel	Celluclast 1.5 L
DP	Degree of polymerization
EG	Endo-1,4-β-D-glucanase
E_{iT}	Total enzyme concentration (gram protein per liter) ($i=1$ for Cel; $i=2$ for N188)
E_{iB}	Bound enzyme concentration ($i=1$ for Cel; $i=2$ for N188)
E_{iF}	Concentration of free enzyme in solution ($i=1$ for Cel; $i=2$ for N188)
E_{imax}	Maximum mass of enzyme that adsorbs onto a unit mass of substrate (gram protein per gram substrate)
G_i	Glucose ($i=1$) cellobiose ($i=2$), trisaccharide ($i=3$) and tetrasaccharide ($i=4$) concentration (grams per liter)
$G_{cr,tri}$	Critical glucose concentration of transglycosylation for trisaccharide production (grams per liter)
$G_{cr,tetra}$	Critical glucose concentration of transglycosylation for tetrasaccharide production (grams per liter)
K_{iad}	Dissociation constant for enzyme adsorption/desorption reaction (liters per gram protein) ($i=1$ for Cel; $i=2$ for N188)
K_{3M}	Substrate (cellobiose) saturation constants (grams per liter)
K_{iIG2}	Inhibition constant cellobiose (grams per liter) ($i=1$ for r_1; $i=2$ for r_2)
K_{iIG}	Inhibition constant glucose (grams per liter) ($i=1$ for r_1; $i=2$ for r_2; $i=3$ for r_3)
K_{iIX}	Inhibition constant xylose (grams per liter) ($i=1$ for r_1; $i=2$ for r_2; $i=3$ for r_3)
k_{ir}	Reaction rate constant ($i=1$ and 2 liters per gram per hour; $i=3$ per hour)
k_{G3}	Reaction rate constant of transglycosylation for trisaccharide production
k_{G4}	Reaction rate constant of transglycosylation for tetrasaccharide production
N188	Novozym 188
r_i	Reaction rate (grams per liter per hour) ($i=1$ for cellulose to cellobiose; $i=2$ for cellulose to glucose; $i=3$ for cellobiose to glucose)
r_{tri}	Overall reaction rate (grams per liter per hour) of $3G \leftrightarrow G_3 + 2H_2O$
r_{tri+}	Reaction rate (grams per liter per hour) of $3G \rightarrow G_3 + 2H_2O$
r_{tri-}	Reaction rate (grams per liter per hour) of $3G \leftarrow G_3 + 2H_2O$
r_{tetra}	Overall reaction rate (grams per liter per hour) of $G + G_3 \leftrightarrow G_4 + H_2O$
r_{tetra+}	Reaction rate (grams per liter per hour) of $G + G_3 \rightarrow G_4 + H_2O$
r_{tetra-}	Reaction rate (grams per liter per hour) of $G + G_3 \leftarrow G_4 + H_2O$
R_S	Substrate reactivity

| S | Substrate concentration (grams per liter) (suffix with "0" means initial substrate concentration) |
X	Xylose concentration (grams per liter)
Xbg	A BG other than N188
α	Dimensionless constant for substrate reactivity

Introduction

Ethanol produced from lignocellulose is now recognized as an alternative type of liquid fuel, and industrial large-scale production is gradually becoming a reality [1, 2]. In order to release glucose from biomass for the glucose to be fermented into ethanol, cellulose contained in the lignocellulosic matrix firstly needs to be hydrolysed. Enzymatic hydrolysis besides chemical hydrolysis is a promising method for decomposing the cellulose into small molecules and glucose. In addition to the recently discovered polysaccharide monooxygenase cellulose cleaving type of enzymes [3], the enzymatic hydrolysis reactions on cellulose involve a minimum of three groups of enzymes: endo-1,4-β-D-glucanase (EG) (EC 3.2.1.4), exo-1,4-β-D-glucanases (or cellobiohydrolase, CBH) (EC 3.2.1.91) and β-glucosidase (BG) (EC 3.2.1.21). EG catalyses hydrolysis of cellulose via random endo-attack on the β1,4 bonds in the cellulose backbone. CBH-type enzymes catalyse the hydrolysis of the cellulose from the ends and release mainly cellobiose. Finally, BG hydrolyses cellobiose into glucose [4].

The mechanism and notably the kinetics of converting an insoluble polymeric cellulosic substrate into soluble sugars by the action of cellulase enzymes are not yet completely understood due to the complexity of the involved phenomena (such as adsorption, desorption, enzyme deactivation, accessible area, crystallinity, degree of polymerization, lignin content, enzyme synergism, etc.), which affect the reaction kinetics. Nevertheless, a number of mathematical models for enzymatic hydrolysis have been proposed in the literature [5–9].

However, none of the models proposed in the past were rigorously validated. This lack of vigorous validation partly questions the credibility of the models for engineering applications. Among the factors that can explain the variability are different enzyme sources, substrate type, conceptual framework of the model (model structure), experimental data quality and quantity [10]. The present study was undertaken to validate the conceptual framework of the most recent model (model structure) proposed by the U.S. National Renewable Energy Laboratory published by Kadam et al. [7] based on a focused and dedicated rigorous experimental testing (improved data quality and quantity) in a systematic and iterative manner. The present study was also designed to assess the requirement for expansion of the model structure in view of improving the applicability range of the model. Notably, transglycosylation reactions which convert glucose back to oligosaccharides under high glucose and cellulose concentrations were investigated and introduced into the modelling.

The mathematical model proposed by Kadam et al. [7] describes the conversion of cellulose to cellobiose (r_1), the conversion of cellulose to glucose (r_2) and the conversion of cellobiose to glucose (r_3) (illustrated in the bold dashed-dot square in Fig. 1). The mathematical model was based on a number of assumptions:

- Enzyme adsorption follows a Langmuir-type isotherm with first-order reactions occurring on the cellulose surface.
- Amorphous and crystalline cellulose are lumped and considered uniform in terms of susceptibility to enzymatic attack.
- Enzyme activity remains constant throughout the reaction.
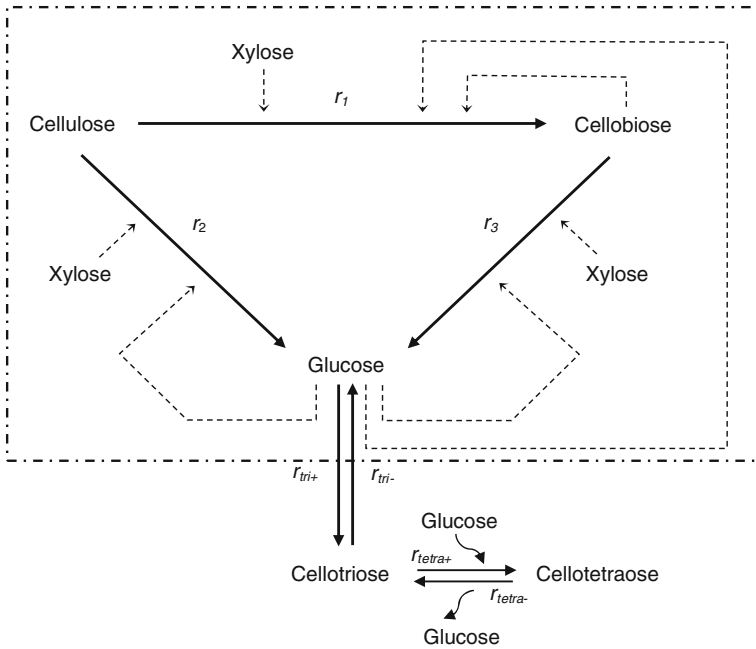
Fig. 1 Reaction scheme for modelling cellulose hydrolysis. *Bold dashed-dot square* illustrates the original Kadam model. *Solid arrows* represent the reaction routes and the *dashed arrows* show the inhibition of sugars on the reactions. Modified from Kadam et al. [7]

- Conversion of cellobiose to glucose occurs in solution and follows the classical Michaelis-Menten kinetics.
- Considers separate cellulase (mainly a mixture of EG and CBH) and BG activities on cellulose breakdown and competitive inhibition by simple sugars.

The hydrolysis pathways and the inhibition effects by intermediate and final products (cellobiose, glucose and xylose) can be classified into r_1, r_2 and r_3. These can be expressed as Eqs. (1) to (8):

Enzyme Adsorption

$$\text{Langmuir isotherm } E_{i\text{B}} = \frac{E_{i\max}K_{i\text{ad}}E_{i\text{F}}S}{1 + K_{i\text{ad}}E_{i\text{F}}} \tag{1}$$

Cellulose-to-Cellobiose Reaction with Competitive Glucose, Cellobiose and Xylose Inhibition

$$r_1 = \frac{k_{1\text{r}}E_{1\text{B}}R_\text{S}S}{1 + \dfrac{G_2}{K_{1\text{IG2}}} + \dfrac{G}{K_{1\text{IG}}} + \dfrac{X}{K_{1\text{IX}}}} \tag{2}$$

Cellulose-to-Glucose Reaction with Competitive Glucose, Cellobiose and Xylose Inhibition

$$r_2 = \frac{k_{2\text{r}}(E_{1\text{B}} + E_{2\text{B}})R_\text{S}S}{1 + \dfrac{G_2}{K_{2\text{IG2}}} + \dfrac{G}{K_{2\text{IG}}} + \dfrac{X}{K_{2\text{IX}}}} \tag{3}$$

Cellobiose-to-Glucose Reaction with Competitive Glucose and Xylose Inhibition Reaction

$$r_3 = \frac{k_{3r} E_{2F} G_2}{K_{3M} \left(1 + \dfrac{G}{K_{3IG}} + \dfrac{X}{K_{3IX}}\right) + G_2} \tag{4}$$

Mass Balances

$$\text{Cellulose}: \quad \frac{dS}{dt} = -r_1 - r_2 \tag{5}$$

$$\text{Cellobiose}: \quad \frac{dG_2}{dt} = 1.056 r_1 - r_3 \tag{6}$$

$$\text{Glucose}: \quad \frac{dG}{dt} = 1.111 r_2 + 1.053 r_3 \tag{7}$$

$$\text{Enzyme}: \quad E_{iT} = E_{iF} + E_{iB} \tag{8}$$

This model [7] is based on the assumption that enzyme activity remains constant. Factors decreasing the reaction rate are as follows: (1) product inhibition, (2) inactivation of enzymes (BG) after the adsorption to the substrate and (3) change of substrate reactivity (R_S). Substrate reactivity is derived from the secondary hydrolysis rate of the residual substrate at any given time [7, 11, 12] and is expressed as:

$$R_S = \alpha \frac{S}{S_0} \tag{9}$$

S_0 is the initial substrate concentration and S is the substrate concentration at a given time (grams per liter). α is a dimensionless constant derived from experimental data, e.g. the relation between the secondary initial hydrolysis rate and S/S_0. The introduction of R_S in the model is based on the observation/assumption that the substrate may become less susceptible to enzymatic hydrolysis over time. Since the reasons for decreasing substrate reactivity are complicated (change of the substrate structure like crystal structure, degree of polymerization, pore size distribution, etc.), in practice, these properties are difficult to evaluate. Therefore, R_S represents an empirical factor for the correction and consideration of the phenomena mentioned above in the reaction rate equations.

The model proposed by Kadam et al. [7] does not include any considerations of reverse reactions, i.e. transglycosylation reactions. To the best of our knowledge, transglycosylation was not considered in previous studies. However, this reaction transferring glucose back to cello-oligosaccharides by β-glucosidase at high glucose or cellobiose concentrations has been reported or described as a possible event in different studies [13–19]. It might be a relevant mechanism to be considered in kinetic models, because in the bioethanol industry, the hydrolysis process is typically operated above 15 % (w/w) dry matter of biomass to obtain higher glucose concentration in order to obtain higher distillation efficiency [20]. Already 30 years ago, Gusakov et al. [16] proposed a very detailed, relatively complex reaction scheme for transglycosylation.

Therefore, two simplified reactions were proposed and tested in the present research. The first reaction is for trisaccharide production $3G \leftrightarrow G_3 + 2H_2O$. The reaction rate can be expressed as:

$$r_{tri} = r_{tri+} - r_{tri-} \tag{10}$$

Where

$$r_{tri+} = k_{G3+} \left[\frac{1}{1 + e^{(G_{cr,tri} - G)}} \right] G \tag{11}$$

$$r_{tri-} = k_{G3-} \left[\frac{1}{1 + e^{(G_{cr,tri} - G)}} \right] G_3 \tag{12}$$

$G_{cr,tri}$ is the critical glucose concentration, which means that when the glucose level is above this concentration, the transglycosylation for trisaccharide production is significant. The second reaction is for tetrasaccharide production $G + G_3 \leftrightarrow G_4 + H_2O$ (r_{tetra}) and the reaction rate can be expressed as:

$$r_{tetra} = r_{tetra+} - r_{tetra-} \tag{13}$$

Where

$$r_{tetra+} = k_{G4+} \left[\frac{1}{1 + e^{(G_{cr,tetra} - G)}} \right] G \tag{14}$$

$$r_{tetra-} = k_{G4-} \left[\frac{1}{1 + e^{(G_{cr,tetra} - G)}} \right] G_4 \tag{15}$$

$G_{cr,tetra}$ is again the critical glucose concentration, which means that when the glucose level is above this concentration, the transglycosylation reaction for tetrasaccharide production is significant. Figure 1 illustrates all the reaction pathways considered for glucose production from cellulose.

Materials and Methods

Substrate and Enzymes

The substrates employed in this research were Avicel® PH-101 (Sigma-Aldrich, St. Louis, MO, USA) and pretreated barley straw, respectively. The barley straw was grown and harvested in Denmark and then pretreated at DONG Energy (Danish Oil and Natural Gas Energy, Denmark) as described previously [21]; in brief, the pretreatment method consisted of a three-stage heating process, which involved triple heating treatment of the straw at increasing temperatures (15 min at 60 °C; liquids removal; 10 min at 180 °C; 3 min at 195 °C) [21]. After pretreatment, the liquids were removed. The standard procedures for acid hydrolysis and compositional calculation analysis of the dry solid were done according to the protocol of the U.S. National Renewable Energy Laboratory [22]. The cellulose, xylose

and acid-insoluble lignin contents in the pretreated barley straw were 66.3, 3.5 and 26.6 % by weight dry matter, respectively.

The enzymes, Celluclast 1.5 L, Novozyme 188 and a special product of BG (not commercialized, named Xbg in this research), were obtained from Novozymes A/S (Bagsværd, Denmark): Celluclast 1.5 L (Cel, mainly EG + CBH) derived from *Trichoderma reesei*, having an activity of 65 FPU/mL (FPU = filter paper units), 10 CBU/mL (CBU = cellobiose units) and protein concentration of 79 mg/mL. Novozyme 188 (N188, mainly BG) derived from *Aspergillus niger* had an activity of 870 CBU/mL and a protein concentration of 88 mg/mL. Xbg had a protein concentration of 60 mg/mL. The CBU activity was determined by measuring glucose production from hydrolysis of cellobiose at 50 °C, pH 4.8 [23, 24].

Analyses of Protein and Sugars

Protein concentrations of the enzymes were measured by the Quick Start Bradford protein assay (Bio-Rad, Hercules, CA) and γ-globulin was used as standard. Hexokinase (420 U/mL) + glucose-6-phosphate-dehydrogenase (210 U/mL), purchased from Megazyme (Wicklow, Ireland), were used for glucose analysis. The concentration of cellobiose followed a previous research [25] calculated from the increase of glucose after treatment with excess of N188 for 24 h at 50 °C. Xylose contents in the pretreated barley straw were only 3.5 %, even completely released after hydrolysis, and the concentration was only 3.15 g/L, which was ignored and no further measurements were performed.

Experimental Design and Hydrolysis Reaction

Hydrolysis reactions under different enzyme concentrations/combinations, substrate and inhibitor concentrations are shown in Table 1. These data were used for parameter estimation or validation, depending on the modelling strategies. All reactions were conducted in 2 mL Eppendorf tubes, incubated in thermomixers at 50 °C and mixed at 1,000 rpm. Reaction buffer was 50 mM sodium acetate, pH 4.8, with 0.04 % sodium azide.

Determination of Langmuir Adsorption Constants

Ten grams per liter of Avicel or 5 g/L of pretreated barley straw was mixed with different concentrations of enzyme and incubated for 1 h at 50 °C with mixing at 1,000 rpm. Free enzymes were measured as the protein concentrations in the supernatant. Adsorbed enzymes were calculated by subtracting free enzyme concentrations from the initial enzyme concentrations. K_{ad} and E_{max} were determined by Eq. (16), which was rearranged from Eq. (1).

$$\frac{E_F}{(E_B/S)} = \frac{1}{E_{max}K_{ad}} + \frac{E_F}{E_{max}} \tag{16}$$

The obtained Langmuir adsorption constants are shown in Table 2.

Transglycosylation Reaction

N188 and Xbg (0.585 and 0.293 mg/mL) and Celluclast 1.5 L (1.58 and 0.585 mg/mL) were mixed with different glucose concentrations. Reactions were incubated at 50 °C for 48 h. Final glucose concentrations were analyzed and compared with the initial concentrations. The decreased glucose was regarded as being converted to oligosaccharides by transglycosylation.

Table 1 Experimental conditions of hydrolysis for the development and validation of the kinetic model

Substrate (g/L)	Enzyme/substrate[a] (mg protein/g substrate)	Case no.	Initial inhibitor	Inhibitor concentration (g/L)
Avicel (100)	Cel+N188 (15.8+5.9)	A1	No	–
		A2	Glucose	50
		A3	Cellobiose	30
		A4	Xylose	80
	Cel (10.5)	B1	No	–
		B2, B3	Glucose	25 and 50
		B4, B5	Cellobiose	15 and 30
		B6, B7	Xylose	40 and 80
	Cel (21.1)	C1	No	–
		C2, C3	Glucose	25 and 50
		C4, C5	Cellobiose	15 and 30
		C6, C7	Xylose	40 and 80
Cellobiose (37.5)	N188 (1.95)	D1	No	–
		D2, D3	Glucose	25 and 50
		D4, D5	Xylose	40 and 80
	N188 (3.9)	E1	No	–
		E2, E3	Glucose	25 and 50
		E4, E5	Xylose	40 and 80
Avicel (100)	Cel+N188 (15.8+3)	F	Xylose	40
	Cel+N188 (15.8+1)	G	Xylose	40
	Cel+N188 (31.6+5.9)	H	Xylose	40
Avicel (150)	Cel+N188 (15.8+5.9)	I	No	–
Avicel (50)	Cel+N188 (15.8+5.9)	J	No	–
Avicel (100)	Cel+Xbg (15.8+1)	K	Xylose	40
Cellobiose (37.5)	Xbg (3.9)	L	No	–
		M	Glucose	50
Barley straw (90)	Cel+N188 (15.8+5.9)	N	No	–
		O	Glucose	50
		P	Cellobiose	30
		Q	Xylose	80

[a] Enzyme loading: "Cel+N188 (15.8+5.9)" in nos. A1~A5 means the enzyme/substrate in the reaction is 15.8 mg Cel protein/g Avicel+5.9 mg N188 protein/g Avicel

Estimation of Kinetic Parameters and Model Validation Methodology

To achieve the goal of the study, a systematic framework was introduced and used. The framework consisted of the following steps: model construction, experimental data collection, estimation of the parameters and validation (see Fig. 2). Three kinetic parameters (k_{3r} K_{3M}, K_{3IG} and K_{3IX}) of the cellobiose-to-glucose conversion (r_3) were calculated using data set E1–E5 (see Table 1), and glucose and xylose were used as initial inhibitors. Then the parameters were validated by data sets with different enzyme concentrations (data set D1–D5, see Table 1). If the fitting and validation did not seem completely correct, a further analysis of the mathematical model was made, and in addition, the cellotriose production transglycosylation reaction

Table 2 Parameters of the Langmuir adsorption and substrate reactivity

Parameters	Value
Avicel	
K_{1ad} (L/g protein)	1.238
K_{2ad} (L/g protein)	1.865
E_{1max} (g protein/g substrate)	0.03257
E_{2max} (g protein/g substrate)	0.00102
Pretreated barley straw	
K_{1ad} (L/g protein)	18.98
K_{2ad} (L/g protein)	21.71
E_{1max} (g protein/g substrate)	0.02255
E_{2max} (g protein/g substrate)	0.00517
α	1

(r_{tri}) was introduced into the original Kadam model. Two parameters (k_{+G3}, k_{-G3}) involved in transglycosylation were estimated. The next step of the procedure was estimation of eight unknown parameters (k_{1r}, k_{2r}, K_{1IG2}, K_{2IG2}, K_{1IG}, K_{2IG}, K_{1IX} and K_{2IX}) in the kinetic reactions for cellulose conversion to cellobiose (r_1) and glucose (r_2) using the experimental data set A1–A4 (Celluclast + N188) or B1–B7 (Celluclast only) which involved analysis of the hydrolysis and product inhibition mechanism (see Table 3). Different initial concentrations of inhibitory agents, glucose, cellobiose and xylose, were used to quantify their inhibition effects on the cellulosic hydrolysis. R_S of Avicel in the cellulose-to-cellobiose reaction (r_1) and the cellulose-to-glucose reaction (r_2) (Eqs. (2) and (3)) were regarded as constant, i.e. setting R_S as 1, according to [11, 26] assuming that the reactivity of Avicel did not change over time. The validation of the models was performed using data sets with a different enzyme combination and concentration, substrate concentration and pretreated barley straw (data set C–Q, see Table 1). The R_S of pretreated barley straw was not constant and α was set as 1, according to the values derived from other types of lignocellulose used in previous studies (e.g. [7, 12]). If the mathematical model predictions did not fit or predict the experimental data well, then a new return to the modelling step was done and finally another transglycosylation reaction was introduced (production for cellotetraose, r_{tetra}). The expresssion of r_1 and r_2 could also be revised by introducing K_{1M} and K_{2M} (see Table 3, model 3). Finally, the implementation of the framework for parameter estimation and validation for further investigation could be done.

To this end, three different variations of the original mathematical model proposed by Kadam et al. [7] were proposed in this research (Table 3) and evaluated. Model 1 only considered the transglycosylation for cellotriose production. Model 2 and model 3 involved both trisaccharide and tetrasaccharide production. Therefore, Eq. (7) for the mass balance of glucose was modified for model 1 as follows:

$$\frac{dG}{dt} = 1.111r_2 + 1.053r_3 - 1.071r_{tri+} + 1.071r_{tri-} \quad (17)$$

and Eqs. (7) and (10) were modified for models 2 and 3 as shown below, respectively

$$\frac{dG}{dt} = 1.111r_2 + 1.053r_3 - 1.071r_{tri+} + 1.071r_{tri-} - 0.2702r_{tetra+} + 0.2702r_{tetra-} \quad (18)$$

$$\frac{dG_3}{dt} = r_{tri+} + r_{tri-} - 0.7567r_{tetra+} + 0.7567r_{tetra-} \quad (19)$$
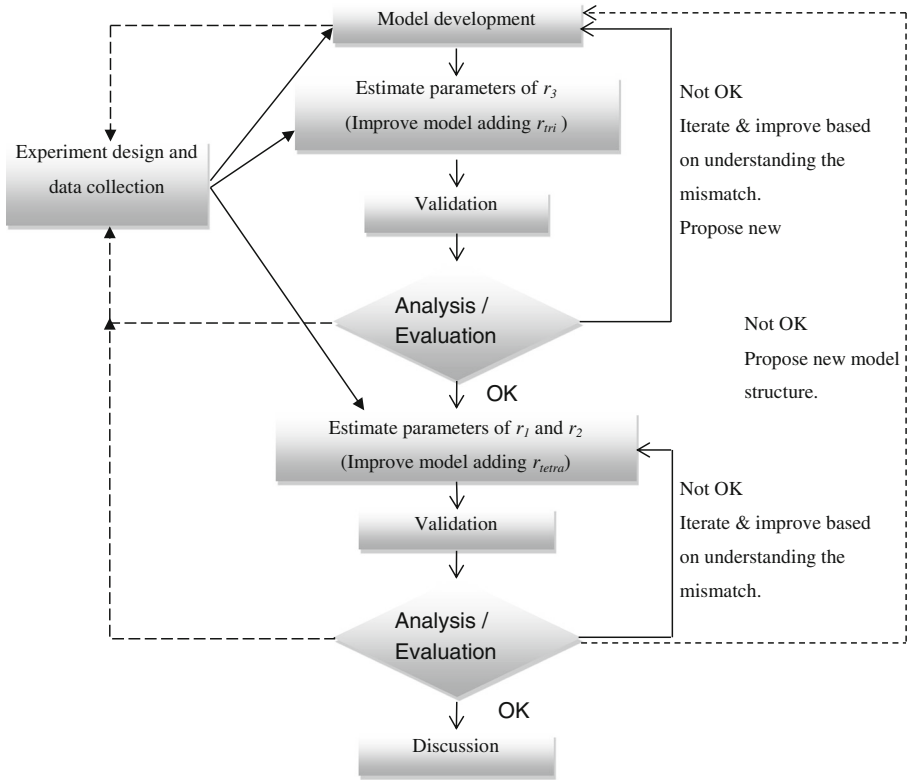
Fig. 2 Framework for construction of the enzymatic cellulose hydrolysis mathematical model and parameter estimation

in order to fulfill the mass balances of the compounds following the reaction pathway proposed in Fig. 1.

The modelling and estimation of the parameters was done using MatLab (The Mathworks, Natick, MA). A nonlinear least squared method was used for the parameter estimation. The initial values for the parameter estimation were taken from [7]. The fminseach function from Matlab and the Levenberg-Marquardt search were used to solve the objective function, $J(\theta)$:

$$\text{argmin} \quad J(\theta) = \sum_{j=1}^{M} \sum_{i=1}^{N} \left(ym_{i,j} - f_{i,j}(\theta) \right) \tag{20}$$

where $J(\theta)$ is the sum of squared errors, $ym_{i,j}$ is the ith measurement in the jth experiment and $f_{i,j}(\theta)$ is the corresponding model prediction for the measurement and θ is the parameter subset used for parameter estimation. This is to mean that for parameter estimation, a set of experiments are typically used in view of identifying the parameter subset in question better. This is explained below.

Experimental data sets under different hydrolysis conditions used for parameter estimation and validation are shown in Table 1. Each data set was given a specific code from A to Q where different initial substrate concentrations, substrate types, enzyme loadings and combinations and inhibitor loadings were related with the different data sets. To make the article concise, these codes will be used repeatedly in this paper.

The comparison and selection for the proper mathematical structure among the proposed mathematical models was performed relying on the weighted sum of square error (WSSE) as follows:

Table 3 Summary of the models proposed in this research

	Model 1	Model 2 ($G_{cr,Tetra}$ = 75 or 80)	Model 3 ($G_{cr,Tetra}$ = 75 or 80)
r_3	$r_3 = \dfrac{k_{3r} E_{2F} G_2}{K_{3M}\left(1 + \dfrac{G}{K_{3IG}} + \dfrac{X}{K_{3IX}}\right) + G_2}$		Data E1~E5 were used for parameter estimation
	Transglycosylation for cellotriose production: $3G \rightleftharpoons G_3 + 2H_2O$		
	$r_{Tri} = r_{Tri+} - r_{Tri-} = k_{G3+}\left[\dfrac{1}{1 + e^{(G_{cr,Tri} - G)}}\right]G - k_{G3-}\left[\dfrac{1}{1 + e^{(G_{cr,Tri} - G)}}\right]G_3$		($G_{cr,Tri}$ = 40)
r_1 and r_2	$r_1 = \dfrac{k_{1r} E_{1B} R_S S}{1 + \dfrac{G_2}{K_{1IG2}} + \dfrac{G}{K_{1IG}} + \dfrac{X}{K_{1IX}}}$	$r_1 = \dfrac{k_{1r} E_{1B} R_S S}{1 + \dfrac{G_2}{K_{1IG2}} + \dfrac{G}{K_{1IG}} + \dfrac{X}{K_{1IX}}}$	$r_1 = \dfrac{k_{1r} E_{1B} R_S S}{K_{1M}\left(1 + \dfrac{G_2}{K_{1IG2}} + \dfrac{G}{K_{1IG}} + \dfrac{X}{K_{1IX}}\right) + S}$
	$r_2 = \dfrac{k_{2r}(E_{1B} + E_{2B}) R_S S}{1 + \dfrac{G_2}{K_{2IG2}} + \dfrac{G}{K_{2IG}} + \dfrac{X}{K_{2IX}}}$	$r_2 = \dfrac{k_{2r}(E_{1B} + E_{2B}) R_S S}{1 + \dfrac{G_2}{K_{2IG2}} + \dfrac{G}{K_{2IG}} + \dfrac{X}{K_{2IX}}}$	$r_2 = \dfrac{k_{2r}(E_{1B} + E_{2B}) R_S S}{K_{2M}\left(1 + \dfrac{G_2}{K_{2IG2}} + \dfrac{G}{K_{2IG}} + \dfrac{X}{K_{2IX}}\right) + S}$
	[Strategy 1] Data B1~B7 were used for parameter estimation	Transglycosylation for cellotetraose production	Transglycosylation for cellotetraose production
		$G + G_3 \rightleftharpoons G_4 + H_2O$	$G + G_3 \rightleftharpoons G_4 + H_2O$
		$r_{Tetra} = r_{Tetra+} - r_{Tetra-}$	$r_{Tetra} = r_{Tetra+} - r_{Tetra-}$
	[Strategy 2] Data A1~A4 were used for parameter estimation	$= k_{G4+}\left[\dfrac{1}{1 + e^{(G_{cr,Tetra} - G)}}\right]G$ $- k_{G4-}\left[\dfrac{1}{1 + e^{(G_{cr,Tetra} - G)}}\right]G_4$	$= k_{G4+}\left[\dfrac{1}{1 + e^{(G_{cr,Tetra} - G)}}\right]G$ $- k_{G4-}\left[\dfrac{1}{1 + e^{(G_{cr,Tetra} - G)}}\right]G_4$
		($G_{cr,Tetra}$ = 75 or 80)	($G_{cr,Tetra}$ = 75 or 80)
		Data A1~A4 were used for parameter estimation	Data A1~A4 and I were used for parameter estimation.

$$WSSE = \sum_j \sum_i \left(\frac{y_{i,j} - f(x_i, j.k)}{\sigma_{i,j}}\right)^2 \qquad (21)$$

where $y_{i,j}$ is the experimental value i of the dataset j, $f(x_{i,j})$ is the value of the function evaluated at the same experimental conditions i of the data set j in the model k and σ is the standard deviation of the measurement error (in this study a similar value was assumed for all measurement points which was calculated from triplicate measurements).

Kinetic Parameters Derived from a Standard Enzymatic Procedure

In enzymology, the standard procedure for deriving kinetic properties of enzymes is through the measurement of the initial reaction rate. In order to compare the parameters estimated from a standard procedure and numerical method in this research, kinetic parameters were also obtained in this way. The proper concentration of BG was mixed with different concentrations of cellobiose and inhibitors (glucose or xylose) and incubated at 50 °C with gentle shaking for 8 min. All reactions were stopped by heating at 100 °C for 5 min. Parameters, k_3, K_{3M}, K_{3IG} and K_{3IX}, were derived from Hanes-Woolf plots; values are shown in Table 4.

Results and Discussion

### Parameter Estimation and Validation of Cellobiose-to-Glucose Conversion (r_3)

The parameters of describing the rate of cellobiose hydrolysis, r_3, were obtained using the experimental data from hydrolysis of cellobiose by N188 using data set E1–E5 (see Table 1).

Table 4 Hydrolysis kinetic parameters derived from model simulation and standard assay

		Model 1 (strategy 1)	Model 1 (strategy 2)	Model 2 ($G_{cr,tetr}$=75)	Model 2 ($G_{cr,tetr}$=80)	Model 3 ($G_{cr,tetr}$=75)	Model 3 ($G_{cr,tetr}$=80)	Kadam's value	N188, standard assay	Xbg, standard assay
1	k_{1r}	18.86	24.19	16.73	16.45	23.28	24.16	22.3		
2	K_{1IG2}	0.0042	0.0040	0.0089	0.0058	0.0465	0.0498	0.015		
3	K_{1IG}	0.0947	0.2041	0.2773	0.3720	0.4284	0.6635	0.1		
4	K_{1IX}	0.0859	0.0463	0.0558	0.0998	0.4613	0.2608	0.1		
5	K_{1M}	–	–	–	–	6.7546	6.4797	–		
6	k_{2r}	7.926	7.077	6.593	3.101	7.449	6.768	7.18		
7	K_{2IG2}	147.87	3467.78	118.68	1328.59	129.39	128.97	132.0		
8	K_{2IG}	0.01509	0.0199	0.0286	0.0059	0.0496	0.0323	0.04		
9	K_{2IX}	0.0097	1.2023	1.6950	1.2514	288.9385	39.6367	0.2		
10	K_{2M}	–	–	–	–	6.5795	6.8987	–		
11	k_{3r}	228.264						285.5	91,690	815,497
12	K_{3M}	3.1740						24.3	0.5814	1.4924
13	K_{3IG}	0.9617						3.9	0.6600	1.4061
14	K_{3IX}	12.614						201.0	12.007	9.1762
15	k_{G3+}	0.0075								
16	k_{G3-}	0.1197								
17	k_{G4+}	–	–	0.0083	0.0070	0.0079	0.0102			
18	k_{G4-}	–	–	0.0754	0.0523	0.1249	0.0785			

The order of the calculations are as follows: (1) k_{3r}, K_{3M}, K_{3IG}, k_{+G3}, k_{-G3}; (2) K_{3IX}; (3) k_{1r}, K_{1IG2}, K_{1IG}, K_{1IX}, K_{1M}, k_{2r}, K_{2IG2}, K_{2IG}, K_{2IX}, K_{2M}, k_{+G4}, k_{-G4}; (4) K_{1IX}, K_{2IX}

However, the preliminary evaluation showed that the deviations between the fitting/prediction and experimental data were large when the glucose concentration was higher than 70 g/L (Fig. 3a). Since the product inhibition was already taken into account, this deviation was thought to be a result of transglycosylation. This hypothesis was tested and proven by incubating the enzymes in different concentrations of glucose solutions to examine the effect of transglycosylation. The data showed that the higher the initial glucose concentration, the more glucose disappeared after incubation for 2 days (Fig. 3b). The lost glucose was assumed to be converted to trisaccharide. The level of transglycosylation was approximately proportional to the glucose concentration when the glucose concentration was over 40 g/L. Therefore, the original model of Kadam was revised to account for this reaction step, namely by introducing transglycosylation reactions for trisaccharide production (r_{tri}, Eqs. (10)–(12)). The value of the critical glucose concentration ($G_{cr,tri}$) was set to 40 g/L, meaning that when glucose concentration is higher than 40 g/L, the effect of transglycosylation is taken into account. To describe the rate expression mathematically, two terms were used: the first term is a sigmoid function which switches the transglycosylation reaction on or off depending on the critical glucose concentration and the other term is a first-order conversion rate of glucose to trisaccharide (see Eqs. (11)–(12)). With this modification, a better description of the experimental data was obtained as shown in Fig. 4. It should be noticed that under same enzyme loading (protein/substrate ratio), Xbg was less inhibited by glucose; in addition, the Xbg hydrolysis efficiency was much higher than that of N188 as shown in Fig. 3a. This is consistent with the observation that the values derived from the standard assay were different
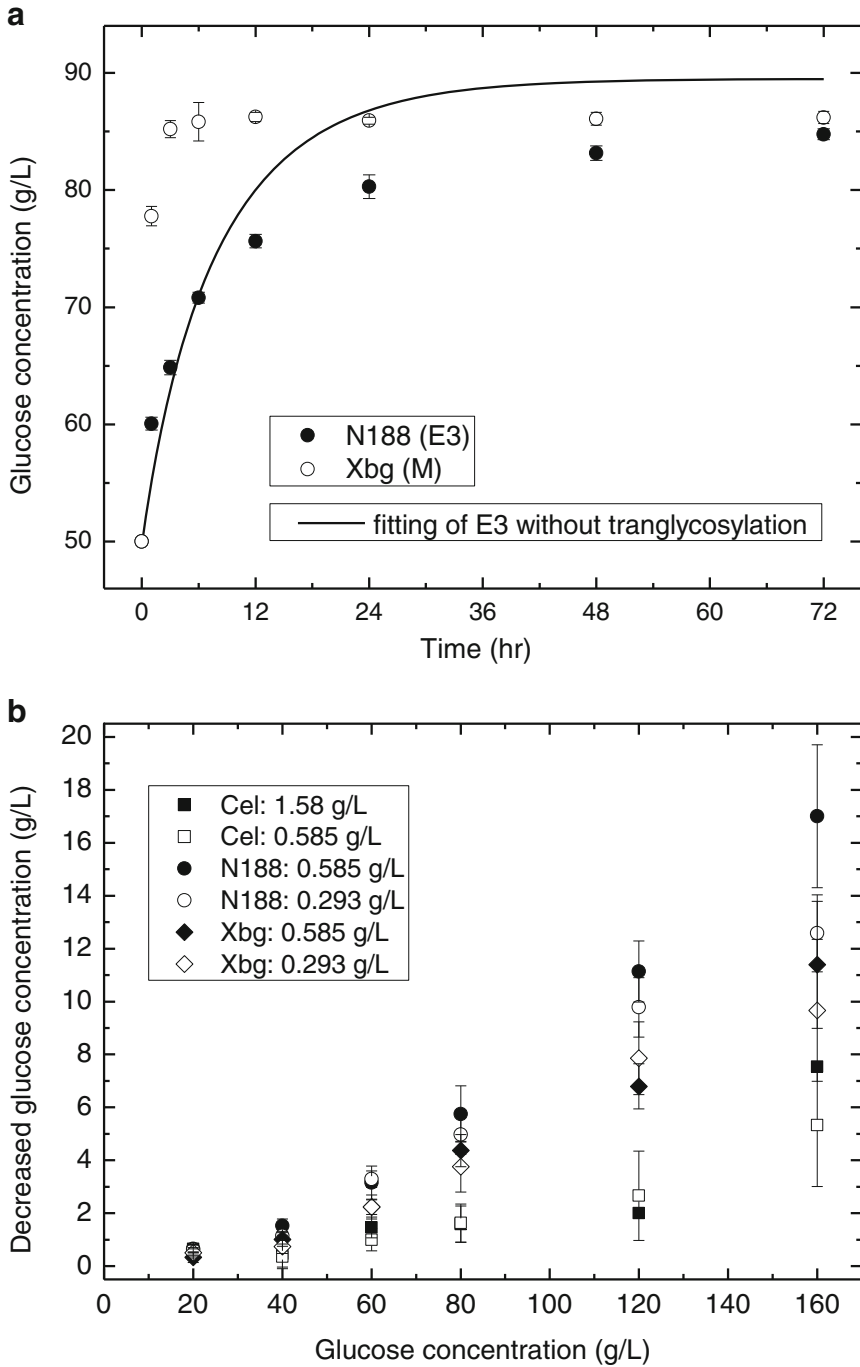
Fig. 3 The effect of transglycosylation. **a** Hydrolysis of 37.5 g/L cellobiose by N188 and Xbg (3.9 mg protein/g cellobiose) with initial background of glucose 50 g/L. *Solid line* indicates the fitting by model without introducing transglycosylation into the model. **b** Transglycosylation effect observed by incubating Celluclast 1.5 L, N188 and Xbg with different glucose concentrations

(Table 4). Therefore, we can expect that the parameters of N188 and Xbg derived from mathematical modelling estimation will be different.

Parameter Estimation and Validation of Cellulose to Glucose (r_1) and Cellulose to Cellobiose (r_2)

In this study, several model structures were proposed based on an iterative analysis of data and model predictions. The models are referred to as models 1, 2 and 3 representing modifications of kinetic expressions of the original Kadam model (see Table 3). Model 1 includes tranglycosylation for trisaccharide production (r_{tri}). Model 2 includes the same model structure as model 1 with the addition of tranglycosylation for tetrasaccharide production (r_{tetra}). Model 3 was based on model 2. But K_{1M} and K_{2M} were introduced into r_1 and r_2 (Eqs. (22) and (23)), that is, adding this type of kinetic for r_1 and r_2 which originally were not included in the model published by Kadam. For the parameter estimation strategy, in addition to data set A1–A4, data set I (150 g/L of Avicel) was also included. Each model candidate was evaluated step by step from model 1 to model 3 to optimise the performance of r_1 and r_2. Therefore, several factors, such as experimental data selected for parameters estimation, transglycosylation for tetrasaccharide (r_{tetra}, Eqs. (13)–(15)), values of critical glucose concentration ($G_{cr,tetra}$) and the mathematical expression of r_1 and r_2, were analyzed. Due to the limited space of this
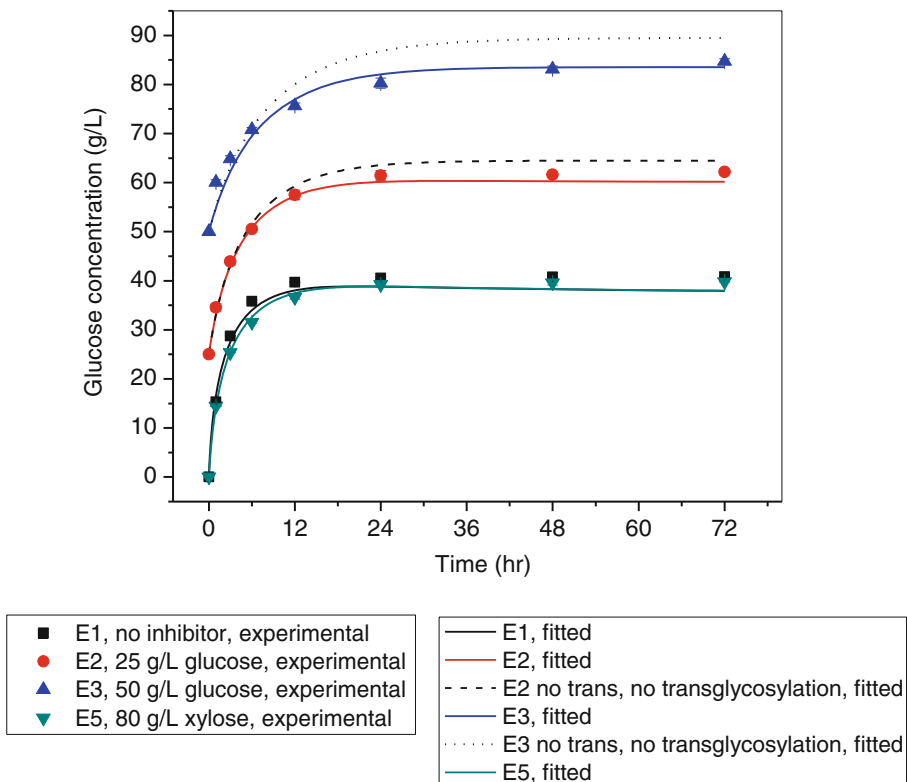
Fig. 4 Parameter estimations of the enzymatic cellobiose hydrolysis reaction; 37.5 g/L cellobiose hydrolysed by N188 (3.9 mg protein/g substrate). Different inhibitor background with 40 g/L xylose (E4) is not shown. Fitting curves without the incorporation of transglycosylation reaction are also shown

article, only selected results of fitting and validation are shown and discussed. Additional results and figures are shown in the Appendix (Supplementary Material). The comparison and evaluation of the models were performed by WSSE analysis (Table 5). It should be noted that in the present work, "fitting" means that the values or curves were derived from the nonlinear least squared method according to the experimental data sets used for parameter estimation, whereas "prediction" means that the values or curves were calculated using the parameters estimated from the fitting and the initial conditions used in the validation experiments.

In model 1, transglycosylation for tetrasaccharide production (r_{tetra}) was not included. Two strategies for parameter estimation were pursued: the first strategy involved using data set B1–B7 for parameter estimation. The fitting of data set B was good (WSSE value is small). The predicted values of data set A2 and A3 were higher than the experimental value(s) when the glucose concentration was higher than around 70~80 g/L (Fig. 5b). The second strategy used data set A1–A4 for parameter estimation instead of data set B1–B7. The fitted values of data set A2 and A3 were still higher than the experimental values obtained under high glucose concentration; besides, the prediction of data set C1–C7 (Fig. 6b) became worse than stategy 1 (Fig. 5a). The WSSE values for prediction of data set C of strategy 1 and strategy 2 were 1,835 and 5,254, respectively (Table 5).

According to the curves in data set A2 and A3 of model 1, the "over-prediction" problem (for strategy 1) and "over-fitting" problem (for strategy 2) were observed when the glucose concentration was higher than around 70~80 g/L (see Figs. 5b and 6a). Hence, the equations r_1, r_2, r_3 and r_{tri}, did not suffice to describe the hydrolysis kinetic behaviour under such high glucose concentrations. In order to address this issue, an additional transglycosylation reaction for tetrasaccharide production, r_{tetra}, was proposed in model 2 (see Fig. 1), of which the critical value for transglycosylation ($G_{cr,tetra}$) was set as 75 and 80 g/L, and both values were evaluated. Data set A1–A4 was used for parameter estimation. In each case, the "over-fitting"

Table 5 WSSE analysis of the performance of the models

	Model 1 (strategy 1)	Model 1 (strategy 2)	Model 2 ($G_{cr,Tetra}$=75)	Model 2 ($G_{cr,Tetra}$=80)	Model 3 ($G_{cr,Tetra}$=75)	Model 3 ($G_{cr,Tetra}$=80)
A	1,696	2,874	1,081	1,190	1,502	1,300
B	622	2,417	4,085	1,533	1,253	921
C	1,835	5,254	6,981	1,811	4,556	2,230
D	336	336	336	336	336	336
E	263	263	263	263	263	263
F	446	68	64	203	273	530
G	350	34	70	167	101	236
H	193	103	165	183	164	261
I	1,138	3,928	2,397	1,518	314	647
J	550	217	139	260	667	412
K[a]	–	–	–	–	–	–
N	331	407	372	567	339	262
O	169	363	50	72	136	25
P	385	368	126	507	389	131
Q	337	398	411	428	285	273
WSSE sum	8,650	17,032	16,540	9,039	10,579	7,827

[a] Data set K involved Xbg, not evaluated by WSSE analysis
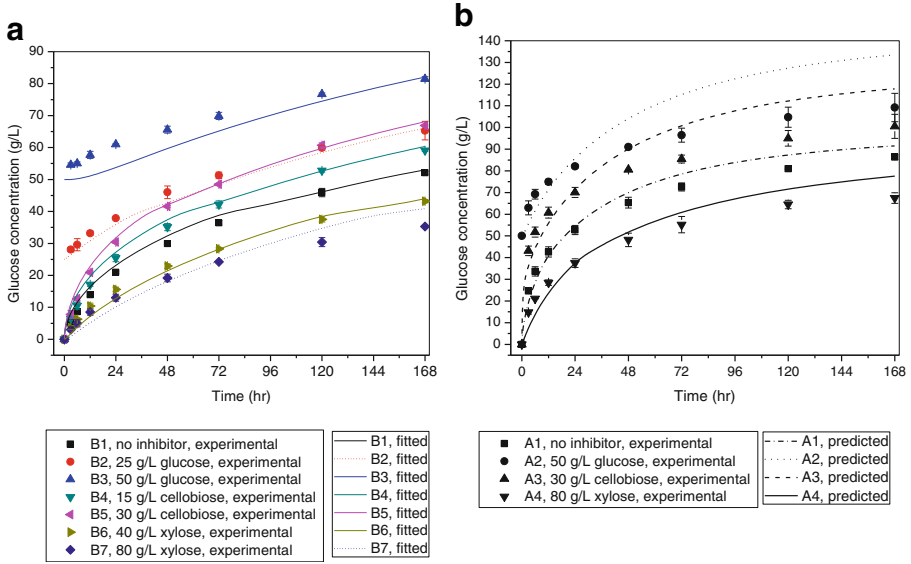
Fig. 5 Evaluation of strategy 1 of model 1 by enzymatic hydrolysis of 100 g/L Avicel. **a** Parameter estimation by data sets hydrolysed by Cel (10.5 mg protein/g substrate) with different initial background inhibitors: B1, no inhibitor; B2, 25 g/L glucose; B3, 50 g/L glucose; B4, 15 g/L cellobiose; B5, 30 g/L cell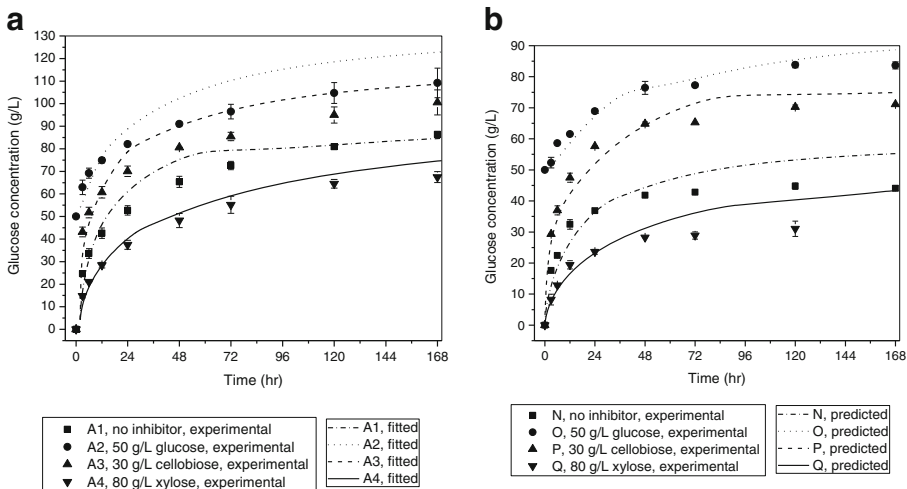obiose; B6, 40 g/L xylose; B7, 80 g/L xylose. **b** Validation of the model by data sets hydrolysed by Cel (15.8 mg protein/g substrate) and N188 (5.9 mg protein/g substrate) with different background inhibitors: A1, no background; A2, 50 g/L glucose; A3, 30 g/L cellobiose; A4, 80 g/L xylose
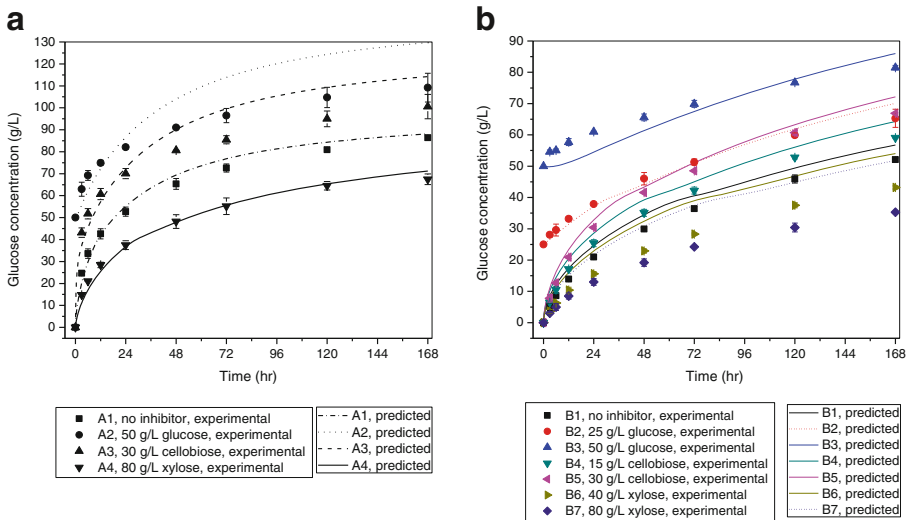
Fig. 6 Evaluation of strategy 2 of model 1 by enzymatic hydrolysis of 100 g/L Avicel. **a** Parameter estimation by data sets hydrolysed by Cel (15.8 mg protein/g substrate) and N188 (5.9 mg protein/g substrate) with different background inhibitors: A1, no background; A2, 50 g/L glucose; A3, 30 g/L cellobiose; A4, 80 g/L xylose. **b** Validation by data sets hydrolysed by Cel (10.5 mg protein/g substrate) with different initial background inhibitors: B1, no inhibitor; B2, 25 g/L glucose; B3, 50 g/L glucose; B4, 15 g/L cellobiose; B5, 30 g/L cellobiose; B6, 40 g/L xylose; B7, 80 g/L xylose

problem under high glucose concentration in data set A2 and A3 (Fig. 7a) and the "over-prediction" problem in data set O and P were reduced (Fig. 7b).

However, when the Avicel concentration was up to 150 g/L (data set I), the "over-prediction" problem was not solved by model 2. Therefore, model 2 was modified further by changing the expression of cellulose-to-cellobiose (r_1) (Eq. (22)) and cellulose-to-glucose (r_2) (Eq. (23)). The modified model was called model 3:

$$r_1 = \frac{k_{1r}E_{1B}R_S S}{K_{1M}\left(1 + \dfrac{G_2}{K_{1IG2}} + \dfrac{G}{K_{1IG}} + \dfrac{X}{K_{1IX}}\right) + S} \tag{22}$$

$$r_2 = \frac{k_{2r}(E_{1B} + E_{2B})R_S S}{K_{2M}\left(1 + \dfrac{G_2}{K_{2IG2}} + \dfrac{G}{K_{2IG}} + \dfrac{X}{K_{2IX}}\right) + S} \tag{23}$$

Of which the denominators of both equations were espressed according to the Michaelis-Menten equation with competitive inhibition. In addition to data set A1–A4, data set I (150 g/L Avicel concentration) was also used for parameter estimation. The critical value of transglycosylation for tetrasaccharide production ($G_{cr,tetra}$) was set as 75 and 80 g/L and both were evaluated. The fitting curve of data set I came closer to the experimental data (Fig. 8). Among the WSSE values of data set I, the model 3 data were also smaller than the others (Table 5). For hydrolysis of barley straw (data set N, O, P and Q), the predicted values in all models were smaller than the experimental values during the initial phase (<24 h) but then produced an "over-shoot" (Fig. 7b). This can be ascribed to the difference of physical and chemical properties between Avicel and pretreated barley straw.

Fig. 7 Evaluation of model 2 ($G_{cr,tetr}$=75). **a** Enzymatic hydrolysis of 100 g/L Avicel. Parameter estimation by data sets hydrolysed by Cel (15.8 mg protein/g substrate) and N188 (5.9 mg protein/g substrate) with different background inhibitors: A1, no background; A2, 50 g/L glucose; A3, 30 g/L cellobiose; A4, 80 g/L xylose. **b** Validation of the model by 90 g/L pretreated barley straw hydrolysed by Cel (15.8 mg protein/g substrate) and N188 (5.9 mg protein/g substrate) with different background inhibitors: N, no background; O, 50 g/L glucose; P, 30 g/L cellobiose; Q, 80 g/L xylose
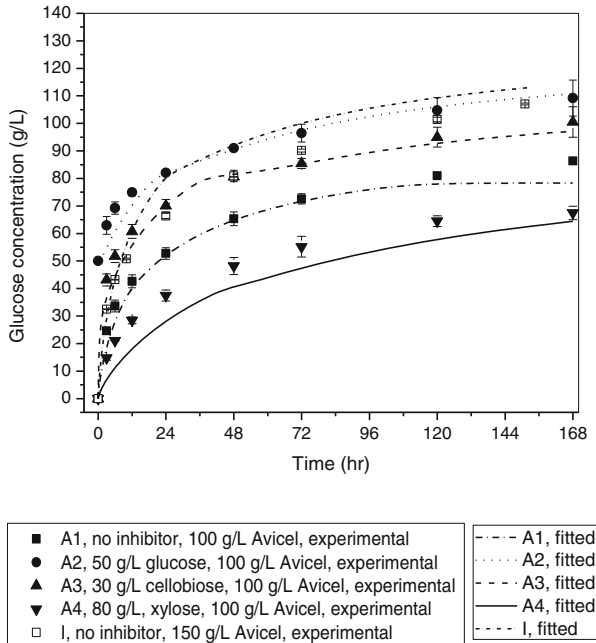
Fig. 8 Evaluation of model 3 by enzymatic hydrolysis of 100 and 150 g/L Avicel. Parameter estimation by data sets hydrolysed by Cel (15.8 mg protein/g substrate) and N188 (5.9 mg protein/g substrate) with different background inhibitors: A1, no background; A2, 50 g/L glucose; A3, 30 g/L cellobiose; A4, 80 g/L xylose; I, no background

A comparison of the two strategies of model 1 (strategy 1: data set B1–B7 for parameter estimation; strategy 2: data set A1–A4 for parameter estimation) indicated that in strategy 2 the values were over-predicted on the xylose background (data set B6 and B7) (Fig. 6b). In contrast, in strategy 1, the fitting curves of data set B1–B7 were quite good and the prediction curves of data set A1–A4 were similar to the fitting curves of data set A1–A4 in strategy 2. The reason for this is not completely clear because a synergistic effect cannot explain this observation. In strategy 1, the parameters of r_1 and r_2 were derived from the effect of Celluclast addition alone (data set B1–B7); therefore, the synergism effects contributed from the cooperation of BG were not included in the paramers. That means, when the parameters derived from strategy 1 are used to describe the kinetic behaviour under the conditions of data set A1–A7 (Celluclast + BG), the predicted values should be lower than the experimental data. The experimental values may include the effect of synergism, but the prediction values do not account for synergism. However, this hypothesis is not consistent with the observed results. The predicted curves of data set A1 and A4 fit the experimental data very well by strategy 1 (not lower, as we expected). This discrepancy needs further investigation.

The reason why model 2 failed to predict the kinetics under Avicel concentration 150 g/L is not known. However, Fig. 3b reveals that the degree of transglycosylation not only depends on glucose concentration but also on enzyme concentration, i.e. the higher the enzyme concentration, the more transglycosylation. In this research, enzyme concentration was not considered in the equation of transglycosylation (either r_{tri} or r_{tetra}). Therefore, although the ratios of enzyme to cellulose in data set A1–A4 and data set I were the same (Table 1), the cellulose concentration of 100 g/L (data set A1–A4) and 150 g/L (data set I), respectively, means that the enzyme concentration of the latter reaction in total is 1.5 times higher than the former. It is
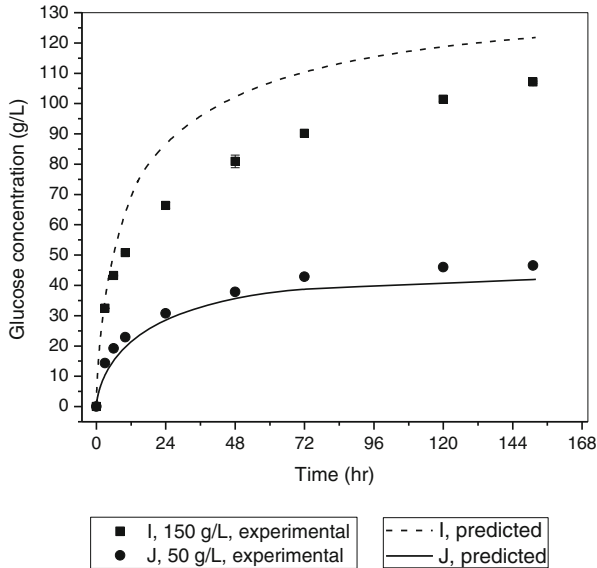
possible that the level of transglycosylation reaction under the condition of data set I is higher than model 2 predicted. The same concept is supported by the fact that parameters of transglycosylation for trisaccharide production (r_3) were estimated from data set E1–E5, but the enzyme concentration of BG in those experiments were only 3.9 mg protein/g cellobiose. Therefore, the BG concentration in the reaction was 3.9×37.5=146 mg/L (enzyme to substrate ratio×substrate concentration = enzyme concentration), compared to the BG concentration in data set I, which was 5.9×150=885, i.e. six times that of data set E. In turn, the parameters of r_3 estimated from data set E1–E5 "hypo-estimated" the transglycosylation reaction of data set I. The detailed relation between enzyme concentration and transglycosylation needs further investigation.

Another explanation could be that the parameters were derived from data fitting by numerical analysis based on kinetic principles (enzyme adsorption, product inhibition and transglycosylation). Thus, when data set I was included in the parameter estimation, certain unknown factors not considered in the semimechanistic model were lumped into the other parameters. That would explain why, e.g. data set I fitted by model 3 was better than the predictions obtained by model 1 and model 2.

In all models, the validation of hydrolysis of Avicel under different ratios of Celluclast and N188 with 40 g/L xylose background (data F, G and H, Fig. 9) was good, meaning that the variation of the enzyme combination within a typically used range (Celluclast, 15.8–31.6 and N188, 1–5.9 mg protein/g cellulose), the models can describe the reaction kinetics precisely. The predictions of the hydrolysis kinetics at lower cellulose concentration (50 g/L Avicel, data set J) were also good (Fig. 10). From the sum of WSSE values in Table 5, the performance of the models was not only influenced by the reaction equations selected in the models, but also affected by the critical glucose concentration ($G_{cr,tetra}$) of transglycosylation. If the reaction is conducted at a cellulose concentration up to 150 g/L, model 3 with $G_{cr,tetra}=80$ is the
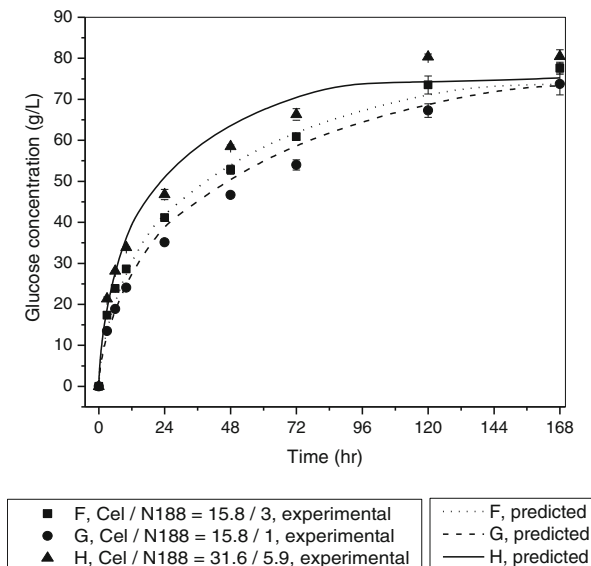
Fig. 9 Validation of model 2 ($G_{cr,tetr}=75$) by enzymatic hydrolysis of 100 g/L Avicel under different enzyme loadings and combinations. F, Cel/N188=15.8/3 mg protein/g substrate; G, Cel/N188=15.8/1 mg protein/g substrate; H, Cel/N188=31.6/5.9 mg protein/g substrate. Background inhibitor is 40 g/L xylose

Fig. 10 Validation of model 2 ($G_{cr,tetr}$=75) by enzymatic hydrolysis of different Avicel concentrations by Cel (15.8 mg protein/g substrate) and N188 (5.9 mg protein/g substrate). I, 150 g/L; J, 50 g/L

best choice (sum of WSSE=7,827). However, when the cellulose concentration is below 100 g/L, model 1 suffices to predict reactions without considering an initial inhibitor background.

Limitations of the Models Proposed in This Research

The models proposed in this study can predict enzymatic cellulose hydrolysis kinetics well within certain ranges of conditions, including various enzyme-substrate ratios and substrate dry matter. However, there are two limitations, namely conditions under which the model does not predict the glucose release very well:

(1) Substrate with different properties. The composition and physical structures of Avicel and pretreated barley straw are different. Most pretreated biomass contains more than 15 % of lignin, which non-specifically adsorbs the enzymes and then the adsorbed enzymes lose their activity. However, in the present work, all bound enzymes were regarded as active; thus, non-productive adsorption was not considered. The lack of this consideration may explain why the model over-estimated the released glucose concentration after ~48 h of reaction (Fig. 7b). The other explanation is that the crystallinity of Avicel is higher than the cellulose in pretreated barley straw. This may be why the hydrolysis curve of pretreated barley straw bends earlier than Avicel. In addition, the lower crystallinty of pretreated barley straw may allow adsorbtion of more water. Hence, at the same dry matter, more free water may exist in the Avicel system. If so, this results in different physical environments in the two systems.

(2) Enzyme from different sources. The intrinsic properties of N188 and Xbg derived from the Hanes-Woolf plot (Table 4) were different, especially k_{3r}, the value for Xbg being more than 8 times higher than that of N188. Xbg moreover triggered less transglycosylation than N188 (Fig. 3b). The influence of these differences on the hydrolysis kinetics is shown in Fig. 3a.

The Significance of the Parameters Derived from This Research and Real Reactions

The semimechanistic kinetic model proposed by Kadam et al. [7] and the modified models presented in the present work provide a passable way to predict the kinetics of enzymatic cellulose hydrolysis because all the variables are easily obtained by regular laboratory measurements. However, the reaction pathways of r_1 and r_2 are oversimplified. As far as Celluclast 1.5 L is concerned, this is a mixture of different enzymes, mainly EG I, EG II, CBH I and CBH II, and the hydrolysis mechanisms of each type of enzyme differ. The main products of CBH I are cellobiose but may be accompanied by some glucose and cellotriose [27, 28], but CBH II produces cellobiose and cellotriose and only trace amounts of glucose may be considered [28]. The products of EG II are generally cellobiose, glucose and cellotriose. CBH I and EG II may interact negatively with respect to substrate adsorption but can exert positive synergistic effects on hydrolysis [27]. Thus, the parameters of r_1 and r_2 derived here are "lumped properties" of EG I, EG II, CBH I and CBH II.

The enzyme-cellulose interaction here is described by a Langmuir isotherm. Based on this theory, only enzymes adsorbing onto insoluble cellulose are defined as bound enzymes and regarded as active. However, EG II, CBH I and CBH II were reported to be able to hydrolyse soluble oligosaccharides with DP<8 [27, 28]. Enzymes acting on soluble substrates are classified as free enzymes rather than bound enzymes. In the equations presented in this work, free Celluclast (EG + CBH) was regarded as not contributing to the reactions. The role of BG for r_2 is also ambiguous. BG works on small soluble oligosaccharides [19] rather than on cellulose. Thus, bound BG may not contribute to r_2. This hypothesis has been corroborated by Zheng et al. [12] who assumed that BG only adsorbs onto lignin, not cellulose. Therefore, in their model, BG did not exist in the cellulose-to-glucose reaction (r_2). The values of the parameters were very close to those reported by Kadam et al. [7].

The parameters here were derived from numerical estimation from the time course reaction curves rather than from initial reaction rates as reported by others (e.g. [25]). However, the kinetic parameters of N188 derived from the models and a standard enzymatic assay are quite different (Table 4). Thus, the parameters estimated from the models may not represent the real properties of the enzymes. Furthermore, parameters of r_3 proposed by [7] and those found in the work presented here differ. As far as our understanding is concerned, in the study of Kadam et al. [7], all values were derived from the hydrolysis of lignocellulose with a mixture of cellulase and BG. In this research, parameters of r_3 were obtained from hydrolysis of cellobiose by BG without the participation of Celluclast. It was also found that the two strategies in model 1 used different experimental data for parameter estimation and resulted in different parameter values of r_1 and r_2. Thus, it can be concluded that the parameters derived from these models cannot be regarded as "universal values". However, for engineering applications, the semimechanistic models can be used to describe and predict the observed/measured enzyme-substrate conversion kinetics and product inhibition mechanisms. Therefore, as long as these type of models are used under the range of their validity (substrate type, enzyme source and substrate concentration), they can support process design and technology improvement efforts at pilot and full-scale studies [29] and optimisation studies [20].

Conclusions

The semimechanistic model proposed by Kadam et al. [7] was modified and validated by experimental data with different enzyme loadings/combinations and substrate loading/type.

Factors such as experimental data used for parameter estimation and transglycosylation were evaluated step by step during the modification of the models. Model 3 can describe the hydrolysis behaviour under high cellulose concentration up to 150 g/L, except when (1) the substrate properties differ and (2) different enzyme sources are employed. The parameters are not universal, and the variations depend on the experimental conditions used for the parameter estimation. However, the model provides a useful description of the dynamics of cellulose hydrolysis. As long as these types of models are used, bearing in mind their limitations and staying within the range of their validity (e.g. use a proper data set to estimate parameter values), they can provide useful tools for simulations and support process design, optimisation and scale-up efforts.

References

1. Gnansounou, E. (2010). Production and use of lignocellulosic bioethanol in Europe: current situation and perspectives. *Bioresource Technology, 101*(13), 4842–4850.
2. Larsen, J., Haven, Ø. M., & Thirup, L. (2012). Inbicon makes lignocellulosic ethanol a commercial reality. *Biomass and Bioenergy, 46*, 36–45.
3. Quinlan, R. J., Sweeney, M. D., Lo Leggio, L., Otten, H., Poulsen, J. C. N., Johansen, K. S., et al. (2011). Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *PNAS, 108*, 15079–15084.
4. Zhang, Y.-H. P., & Lynd, L. R. (2004). Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. *Biotechnology and Bioengineering, 88*(7), 797–824.
5. Gan, Q., Allen, S. J., & Taylor, G. (2003). Kinetic dynamics in heterogeneous enzymatic hydrolysis of cellulose: an overview, an experimental study and mathematical modelling. *Process Biochemistry, 38*(7), 1003–1018.
6. Gharpuray, M. M., Lee, Y.-H., & Fan, L. T. (1983). Structural modification of lignocellulosics by pretreatments to enhance enzymatic hydrolysis. *Biotechnology and Bioengineering, 25*(1), 157–172.
7. Kadam, K. L., Rydholm, E. C., & McMillan, J. D. (2004). Development and validation of a kinetic model for enzymatic saccharification of lignocellulosic biomass. *Biotechnology Progress, 20*(3), 698–705.
8. Movagarnejad, K., Sohrabi, M., Kaghazchi, T., & Vahabzadeh, F. (2000). A model for the rate of enzymatic hydrolysis of cellulose in heterogeneous solid-liquid systems. *Biochemical Engineering Journal, 4*(3), 197–206.
9. Philippidis, G. P., Smith, T. K., & Wyman, C. E. (1993). Study of the enzymatic hydrolysis of cellulose for production of fuel ethanol by the simultaneous saccharification and fermentation process. *Biotechnology and Bioengineering, 41*(9), 846–853.
10. Sin, G., Meyer, A. S., & Gernaey, K. V. (2010). Assessing reliability of cellulose hydrolysis models to support biofuel process design—identifiability and uncertainty analysis. *Computers & Chemical Engineering, 34*(9), 1385–1392.
11. Yang, B., Willies, D. M., & Wyman, C. E. (2006). Changes in the enzymatic hydrolysis rate of Avicel cellulose with conversion. *Biotechnology and Bioengineering, 94*(6), 1122–1128.
12. Zheng, Y., Pan, Z., Zhang, R., & Jenkins, B. M. (2009). Kinetic modeling for enzymatic hydrolysis of pretreated creeping wild ryegrass. *Biotechnology and Bioengineering, 102*(6), 1558–1569.
13. Andrić, P., Meyer, A. S., Jensen, P. A., & Dam-Johansen, K. (2010). Effect and modeling of glucose inhibition and in situ glucose removal during enzymatic hydrolysis of pretreated wheat straw. *Applied Biochemistry and Biotechnology, 160*(1), 280–297.
14. Andrić, P., Meyer, A. S., Jensen, P. A., & Dam-Johansen, K. (2010). Reactor design for minimizing product inhibition during enzymatic lignocellulose hydrolysis: I. Significance and mechanism of cellobiose and glucose inhibition on cellulolytic enzymes. *Biotechnology Advances, 28*, 308–324.
15. Bhiri, F., Chaabouni, S., Limam, F., Ghrir, R., & Marzouki, N. (2008). Purification and biochemical characterization of extracellular β-glucosidases from the hypercellulolytic Pol6 mutant of *Penicillium occitanis*. *Applied Biochemistry and Biotechnology, 149*(2), 169–182.

16. Gusakov, A. V., Sinitsyn, A. P., Goldsteins, G. H., & Klyosov, A. A. (1984). Kinetics and mathematical model of hydrolysis and transglycosylation catalysed by cellobiase. *Enzyme and Microbial Technology, 6*(6), 275–282.

17. Kono, H., Waelchli, M. R., Fujiwara, M., Erata, T., & Takai, M. (1999). Transglycosylation of cellobiose by partially purified *Trichoderma viride* cellulase. *Carbohydrate Research, 319*(1–4), 29–37.

18. Pal, S., Banik, S. P., Ghorai, S., Chowdhury, S., & Khowala, S. (2010). Purification and characterization of a thermostable intra-cellular β-glucosidase with transglycosylation properties from filamentous fungus *Termitomyces clypeatus*. *Bioresource Technology, 101*(7), 2412–2420.

19. Watanabe, T., Sato, T., Yoshioka, S., Koshijima, T., & Kuwahara, M. (1992). Purification and properties of *Aspergillus niger* β-glucosidase. *European Journal of Biochemistry, 209*(2), 651–659.

20. Morales-Rodriguez, R., Meyer, A. S., Gernaey, K. V., & Sin, G. (2012). A framework for model-based optimization of bioprocesses under uncertainty: lignocellulosic ethanol production case. *Computers & Chemical Engineering, 42*, 115–129.

21. Rosgaard, L., Pedersen, S., & Meyer, A. S. (2007). Comparison of different pretreatment strategies for enzymatic hydrolysis of wheat and barley straw. *Applied Biochemistry and Biotechnology, 143*(3), 284–296.

22. Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, J., Templeton, D., & Crocker, D. (2008). Determination of structural carbohydrate and lignin in biomass. Laboratory analytical procedure. NREL/TP-510–42618. Updated 2010.

23. Ghose, T. K. (1987). Measurement of cellulase activities. *Pure and Applied Chemistry, 59*(2), 257–268.

24. Sternberg, D., Vijayakumar, P., & Reese, E. T. (1977). β-Glucosidase: microbial production and effect on enzymatic hydrolysis of cellulose. *Canadian Journal of Microbiology, 23*(2), 139–147.

25. Drissen, R. E. T., Maas, R. H. W., Van Dermaarel, M. J. E. C., Kabel, M. K., Schols, H. A., Tramper, J., et al. (2007). A generic model for glucose production from various cellulose sources by a commercial cellulase complex. *Biocatalysis and Biotransformation, 25*(6), 419–429.

26. Ooshima, H., Kurakake, M., Kato, J., & Harano, Y. (1991). Enzymatic activity of cellulase adsorbed on cellulose and its change during hydrolysis. *Applied Biochemistry and Biotechnology, 31*(3), 253–266.

27. Medve, J., Karlsson, J., Lee, D., & Tjerneld, F. (1998). Hydrolysis of microcrystalline cellulose by cellobiohydrolase I and endoglucanase II from *Trichoderma reesei*: adsorption, sugar production pattern, and synergism of the enzymes. *Biotechnology and Bioengineering, 59*(5), 621–634.

28. Nidetzky, B., Zachariae, W., Gercken, G., Hayn, M., & Steiner, W. (1994). Hydrolysis of cellooligosaccharides by *Trichoderma reesei* cellobiohydrolases: experimental data and kinetic modeling. *Enzyme and Microbial Technology, 16*(1), 43–52.

29. Morales-Rodriguez, R., Meyer, A. S., Gernaey, K. V., & Sin, G. (2011). Dynamic model-based evaluation of process configurations for integrated operation of hydrolysis and co-fermentation for bioethanol production from lignocellulose. *Bioresource Technology, 102*(2), 1174–1184.