ARTICLE

BIOTECHNOLOGY and BIOENGINEERING

Cellulose Hydrolysis in Evolving Substrate Morphologies II: Numerical Results and Analysis

Wen Zhou,[1,2] Zhiqian Hao,[3] Ying Xu,[1,2] Heinz-Bernd Schüttler[3]

[1]Department of Biochemistry and Molecular Biology, Institute of Bioinformatics, University of Georgia, Athens, Georgia; telephone: 706-542-9779; fax: 706-542-9751; e-mail: xyn@bmb.uga.edu

[2]BioEnergy Science Center (BESC), Oak Ridge, Tennessee

[3]Department of Physics and Astronomy, University of Georgia, Athens, Georgia; telephone: 706-542-3886; fax: 706-542-2492; e-mail: hbs@physast.uga.edu

Received 16 December 2008; revision received 13 March 2009; accepted 27 April 2009

Published online 12 May 2009 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bit.22388

ABSTRACT: Numerical simulation results are presented for a cellulose hydrolysis model which incorporates both the enzymatic glucan chain fragmentation kinetics and the hydrolytic substrate morphology evolution within the general framework of our companion article I. To test the local Poisson (LP) approximation employed in the site number formalism of I, we numerically compare it to the corresponding exact chain number formalism of I. The LP results agree to very high accuracy with the exact chain number kinetics, assuming realistic parameters. From simulations of different types of random and non-random model morphologies, we then show that the details of the random substrate morphology distribution, and its hydrolytic time evolution, profoundly affect the hydrolysis kinetics. Essential, likely very general, experimentally testable features of such morphology-based hydrolysis models are (i) the existence of two distinct time scales, associated with the hydrolysis of the outermost surface-exposed cellulose chains and, respectively, of the entire substrate; (ii) a strongly morphology-dependent hydrolysis slow-down effect, which has also been observed in previous experimental work. Our results also suggest that previously proposed non-morphologic chain fragmentation models can only be applied to describe the hydrolytic short-time behavior in the low enzyme limit. Further experiments to test our modeling framework and its potential applications to the optimization of the hydrolytic conversion process are discussed.

Biotechnol. Bioeng. 2009;xxx: xxx–xxx.

KEYWORDS: cellulose hydrolysis; substrate morphology; mathematical model; site number formalism

Correspondence to: Y Xu and H.-B. Schüttler

Contract grant sponsor: U.S. Department of Energy
Contract grant number: 4000063512
Contract grant sponsor: National Science Foundation
Contract grant number: NSF/DBI-0354771; NSF/ITR-IIS-0407204; NSF/DBI-0542119; NSF/CCF0621700
Additional Supporting Information may be found in the online version of this article.

Introduction

The utilization of cellulosic biomass for bioenergy applications hinges upon the efficient conversion of as-grown solid cellulose substrates into water-soluble, fermentable sugars such as glucose and cellobiose (Himmel et al., 2007). Enzymatic cellulose hydrolysis is presently under extensive investigation as a promising technology to achieve efficient solubilization. The availability of enzyme-accessible surface area is by now recognized as a critical rate-limiting factor in the enzymatic conversion process. Thus, in current engineering-level cellulose solubilization approaches, significant effort must be expended, prior to enzymatic hydrolysis, on various types of physical or chemical biomass pre-treatments whose primary purpose is to produce substrate morphologies with substantially increased enzyme-accessible surface area (Hsu, 1996; Jacobsen and Wyman, 2000; McMillan, 1994; Weil et al., 1994).

Most of the modeling work done so far describes the cellulose hydrolysis kinetics at the level of individual cellulosic glucan chain molecules being hydrolyzed in isolation. While capturing essential features of the enzymatic chain fragmentation process at the molecular level, these "non-interacting chain" approaches, by design, cannot explicitly account for the substrate morphology, that is, for the collective embedding of glucan chains in a dense solid substrate, the resulting steric obstruction of enzyme access, or the hydrolysis-driven evolution of the morphology (Converse and Optekar, 1993; Fenske et al., 1999; Okazaki and Moo-Young, 1978; Suga et al., 1975; Zhang and Lynd, 2006). Some work employed into the model a pre-assumed changing function of the shape and surface area of substrate along the hydrolysis process (Converse and Grethlein, 1987; Converse et al., 1988; Gan et al., 2003; Luo et al., 1997; Movagarnejad et al., 2000; Oh et al., 2001; Philippidis et al., 1992, 1993; Wald et al., 1984). The construction of the changing function is based solely on empirical correlation.

When a substantial fraction of the substrate becomes solubilized during hydrolytic conversion, the resulting changes in substrate morphology will profoundly affect the availability of enzyme-accessible surface area, and hence the hydrolysis kinetics itself. Non-morphologic chain fragmentation modeling approaches are therefore limited to the description of short-time hydrolysis behavior only (Zhang and Lynd, 2006), on time scales where the internal solid surface structure of the substrate remains approximately constant.

In a companion article, hereafter referred to as I (Zhou et al., 2009), we have now developed a general theoretical and computational framework for the modeling of enzymatic cellulose hydrolysis in the presence of a hydrolytically evolving substrate morphology. This modeling formalism combines the fragmentation of surface-exposed enzyme-accessible glucan chains, due a non-complexed hydrolase enzyme system, with an explicit rate equation description of the time-dependent morphology. The morphology evolution is explicitly driven by the hydrolytic solubilization of substrate material. We believe that this new formalism represents a significant step towards the quantitative modeling of the entire hydrolysis process, up to near-complete hydrolytic solubilization of the cellulosic substrate.

In the present article, we will present detailed numerical simulation results obtained with the foregoing surface layer ablation modeling formalism. Our main objectives here are twofold. First, we will carefully test the accuracy of our proposed approximation procedure employed in the site number formalism by direct comparison to numerical results for a corresponding exact chain number fragmentation model, equivalent to the Zhang–Lynd model (Zhang and Lynd, 2006). Secondly, we will present model predictions concerning hydrolytically evolving substrate morphologies, and their effects on the hydrolysis kinetics, which can be tested by future experimental observations and which clearly distinguish our morphology-based models from non-morphologic kinetics approaches. Along the way, we will also demonstrate that a random distribution of evolving substrate morphologies may quite naturally account for a frequently observed, but presently not well understood feature of the hydrolytic conversion process: the phenomenon of hydrolysis slow-down (Lynd et al., 2002). Please note that symbols appearing first in the companion article I are not redefined here, and the reader is referred to the nomenclature of that article. New symbols will first be defined in the context and then summarized in the nomenclature of this article.

Materials and Methods

Five-Site Surface Ablation Model

In our numerical simulations, we model a pure cellulose substrate, such as Avicel, and a non-complexed cellulase system with kinetics and concentration model parameters from Zhang and Lynd (2006) summarized in Table I, unless specified otherwise. Our enzyme concentrations are derived from the natural composition found in *Trichoderma reesei* with mass fractions of 12% EG1, 60% CBH1, and 20% CBH2; and with a remaining 8% of other glycoside hydrolases which we neglect, following Zhang and Lynd (2006). Hence, in Table I, we effectively assume a three-enzyme system with 27.6 mg/L total concentration, consisting only of EG1, CBH1, and CBH2 in a mass ratio EG1/CBH1/CBH2 = 12:60:20. We thus take $k_X = k_Y = 2$, for the chain-end cutting sites of the cellohydro*bi*olases CBH2 and CBH1, respectively. We also assume a minimum insoluble chain length $\ell_S = 7 > k_X + k_Y = 4$ and $\phi_{O,\sigma} = 0$, that is, a pure cellulose substrate without site types Z and O. The chain site distribution models given in I (HDC and CCE) are then equivalent and reduced to a five-site model.

Table I. Simulation parameters.

Parameter	Unit/value	Remarks
M_1	55,000 (g/mol)	Molar mass of EG1
M_3	65,000 (g/mol)	Molar mass of CBH1
M_2	65,000 (g/mol)	Molar mass of CBH2
M_{G_1}	162 (g/mol)	Molar mass of Anhydro-glucose G_1 ($C_6H_{10}O_5$)
$\gamma_{1,N}/M_1$	0.40 (μmol bonds/(mg min))	Specific enzyme activity (by mass) of EG1 on N, X, and Y sites
$\gamma_{3,Y}/M_3$	0.08 (μmol bonds/(mg min))	Specific enzyme activity (by mass) of CBH1 on Y sites
$\gamma_{2,X}/M_2$	0.16 (μmol bonds/(mg min))	Specific enzyme activity (by mass) of CBH2 on X sites
$L_{1,N}$	3.0 (L/mmol)	Adsorption equilibrium coefficient of EG1 to N, X, and Y sites
$L_{3,Y}$	4.0 (L/mmol)	Adsorption equilibrium coefficient of CBH1 to Y sites
$L_{2,X}$	4.0 (L/mmol)	Adsorption equilibrium coefficient of CBH2 to X sites
M_1u_1	0.0036 (g/L)	Mass concentration of EG1 (E1 enzyme system)
M_3u_3	0.0180 (g/L)	Mass concentration of CBH1 (E1 enzyme system)
M_2u_2	0.0060 (g/L)	Mass concentration of CBH2 (E1 enzyme system)
$M_{G_1}x_V^{(o)}$	10.0 (g/L)	Mass concentration of initial anhydro-glucose G_1 in solid
ℓ_s	7	Minimum length ℓ of insoluble chains G_ℓ
k_X, k_Y	2	L-end and R-end exo-cutting lengths (producing cellobiose)

Both the general site number formalism developed in I and the corresponding general chain number formalism, given in Supporting Information section G of I, simplify considerably when applied to the five-site model, as detailed in Supporting Information sections A and B of the present article.

Testing the Local Poisson (LP) Approximation

In the first part of this investigation, the accuracy of the LP approximation in the site number formalism is tested by comparing it to simulation results of the corresponding full chain number formalism. We mainly consider the no-morphology case $\eta_\sigma \equiv 0$, that is, all chains are assumed to be fully exposed at the surface or, equivalently, all SACs contain only one single layer, are therefore identical, and can be described by only a single SAC geometry class ($\sigma = 1 \equiv M_{\mathrm{MD}}$). This is referred to as the "*S*ingle-layer, *S*ingle-geometry" (SS) model hereafter and may be formally regarded as the infinite-dimensional limit, $d_{\mathrm{A}} \to \infty$, of the morphological ($\eta_\sigma \neq 0$) models, as detailed in Supporting Information section A. [A test of the LP approximation for the case of a multi-layer model *with* substrate morphology, $\eta_\sigma \neq 0$, is briefly discussed below and detailed in Supporting Information section D.] As a consequence, $x_{\mathrm{M}} = x_{\mathrm{V}}$ in the SS model, the overall accessibility fraction $\overline{F}_{\mathrm{a}} = x_{\mathrm{M}}/x_{\mathrm{V}} \equiv 1$, the $x_{\mathrm{L},\sigma}$- and $x_{\mathrm{N},\sigma}$-rate equations in the SS model become decoupled from λ_σ, and we can ignore the λ_σ-rate equation altogether. In the site number formalism, we then solve the coupled rate Equations (8) and (9) in Supporting Information section A; and in the corresponding chain number formalism, we solve the rate equation system Equation (30) in Supporting Information section B for *all* chain lengths ℓ, with $\eta_\sigma \equiv 0$ in both. Adopting also the low-enzyme limit, our SS model chain number formalism then becomes equivalent to the Zhang–Lynd chain fragmentation model (Zhang and Lynd, 2006). Two enzyme systems, a mixed endo–exo EG1-CBH1,2 enzyme system with naturally occurring enzyme composition and a purely exo-acting CBH1,2 enzyme system, are used in the simulations. As explained in Supporting Information section A, the LP approximation becomes exact for a purely endo-acting enzyme system and does not require numerical testing in that case. In the chain number formalism, the initial chain length distribution has to be provided as a model input. We consider three such initial distributions: a Delta shape, a (truncated) Gaussian shape and a Global Poisson (GP) shape, as detailed in Supporting Information section B.

Hydrolytic Morphology Evolution by Surface Ablation

In the second part of our investigation, we will consider four specific substrate models: three surface ablation models with fully time-dependent morphology, comprising one "*M*ultiple-layer, *S*ingle-geometry" (MS) model, and two different "*M*ultiple-layer, *M*ultiple-geometry" (MM) models; and, for contrast and comparison, the above-described non-morphologic SS model. The construction and parameterization of these four models is described in detail in Supporting Information section C, and will be only briefly summarized here. The MS model describes a non-random morphology characterized by only one single SAC geometry class: $M_{\mathrm{MD}} = 1$. The two MM models, MM82-1 and MM82-2, describe random substrate morphology distributions, represented by an SAC size (λ_σ-) population comprising $M_{\mathrm{MD}} = 82$ SAC geometry classes. Models MM82-1 and MM82-2 assume, respectively a uniform and a Gaussian distribution of the initial substrate monomer molar fractions $\xi_\sigma^{(o)}$.

Experimentally measured initial (pre-hydrolysis) parameter values for a typical pure cellulose material, Avicel, are employed to constrain our chain fragmentation model and our four substrate morphology models, respectively: the degree of polymerization $\mathrm{DP}^{(0)} = 300$, and the enzymatic surface accessibility fraction $\overline{F}_{\mathrm{a}}^{(o)} = 0.006$ (Zhang and Lynd, 2004). All four morphology models are thus parameterized to represent the same $\mathrm{DP}^{(0)} = 300$ and the same $\overline{F}_{\mathrm{a}}^{(o)} = 0.006$. In the case of the SS model, $\overline{F}_{\mathrm{a}}^{(o)}$ is incorporated as an ad hoc correction factor to account for reduced EG1 enzyme-chain accessibility of the actual solid substrate (Zhang and Lynd, 2006). Except for this accessibility correction in the SS model, all enzyme adsorption and kinetics parameters are taken from Table I. In addition to the mixed enzyme system, referred to as E1 in Table I, a second enzyme system, E200, with 200-fold increased enzyme concentrations will also be studied.

Results and Discussion

Testing the LP Approximation

Mixed EG1-CBH1,2 Enzyme System

Figure 1 shows the results from model calculations for the Zhang–Lynd chain model and for the corresponding site number formalism with the LP approximation (i.e., the SS model), for a mixture of EG1, CBH1 and CBH2, considering different initial DP. It is evident that the site number formalism with approximate LP closure provides an excellent approximation to the exact full chain rate equations when the native (=initial) degree of polymerization $\mathrm{DP}^{(o)}$ exceeds about 20 monomers. For $\mathrm{DP}^{(o)} \geq 20$, the deviations between site number LP and full chain number results are at the 1%-level or less and these deviations decrease rapidly with increasing $\mathrm{DP}^{(o)}$. This is exactly what one would expect, based on the long-chain limit (LCL) arguments discussed in I. We note that typical cellulosic substrates have DP-values well in excess of 20 monomers (Zhang and Lynd, 2004).

Remarkably, this level of accuracy is achieved by the LP approximation for all observable quantities that are relevant

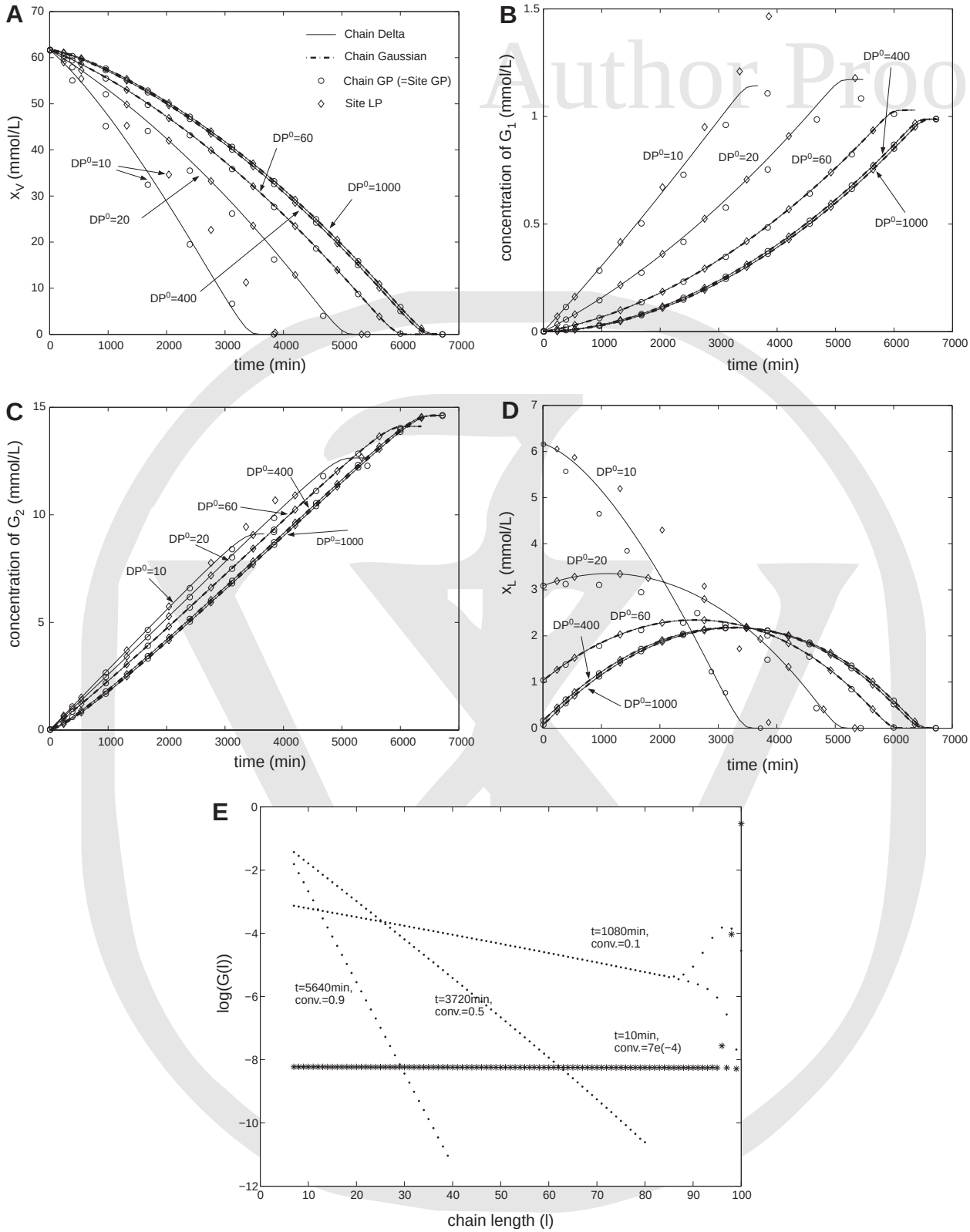

Figure 1. Comparison of the SS model results from chain number formalism (Zhang–Lynd Model) and site number formalism, with different initial chain length distributions, (as detailed in Supporting Information section B, for the case of a GP-shaped initial chain length distribution, the results of site formalism "Site GP" and the chain formalism "Chain GP" are identical.) for the mixed EG1-CBH1,2 enzyme system. In **panels A–D**, full lines, dot-dashed lines and circles are for the chain number formalism with delta-, Gaussian- and global-Poisson-(GP-) shaped initial chain length distributions; diamonds are for the corresponding local Poisson (LP) approximation in the site number formalism. A: total monomer concentration $x_V(t)$ in solid versus time t; (B) concentration of G_1 in solution, $x_S(1, t)$, versus time t; (C) concentrations of G_2 in solution, $x_S(2, t)/2$, versus time t; (D) concentration of non-reducing chain ends, $x_L(t)$, versus time t; (**E**) typical log chain length distribution, $\log\ G(\ell, t)$, versus chain length ℓ at several times t, from SS model chain number formalism with delta-shaped initial distribution from Equation (35) with $\ell_{Wid} \to 0$ and $\ell_{Avg} = 100$.

to hydrolysis, including the total remaining solid substrate monomer concentration $x_V(=x_M)$ in Figure 1A; the soluble oligomeric monomer concentrations $x_S(k, t)$, shown in Figure 1B for oligomer length $k=1$ (=glucose) and in Figure 1C for $k=2$ (=cellobiose); and the total chain (end) concentration $x_L(t)$ shown in Figure 1D. Note that x_L and x_M determine the hydrolytically evolving DP of surface exposed chains by $DP = x_M/x_L$.

For comparison, we are also showing results for $DP^{(o)} = 10$. Here the LCL conditions discussed in I, for example, the condition $DP^{(o)} \gg \ell_C$, are not really satisfied, since $\ell_c = 8$. As expected, the deviations between site number LP and chain number formalism become quite noticeable here as hydrolysis progresses.

In comparing the full chain results for these three different chain length distribution shapes, we notice that all three give almost identical results at least for $DP^{(o)} \geq 60$. This is again fully consistent with the general discussion of the LCL in I: as long as the initial chain length distribution satisfies the LCL conditions, the hydrolysis kinetics is very insensitive to the actual initial chain length distribution shape. The only parameter that matters under LCL conditions is the initial average chain length, that is, the $DP^{(o)}$-value; other details of the distribution shape become essentially irrelevant.

Pure CBH1,2 Enzyme System

Figure 2 shows the comparison between the site number formalism with LP approximation and the chain number results for pure CBH1,2 enzymes. All model parameters are from Table I, except that the total EG1 concentration is set to $u_1 = 0$. As seen in Figures 2A–C, respectively, for initial DP-values $DP^{(o)} \geq 60$ the site number LP approximation is again remarkably accurate in reproducing chain number results for total insoluble substrate monomer, $x_V(=x_M)$, and

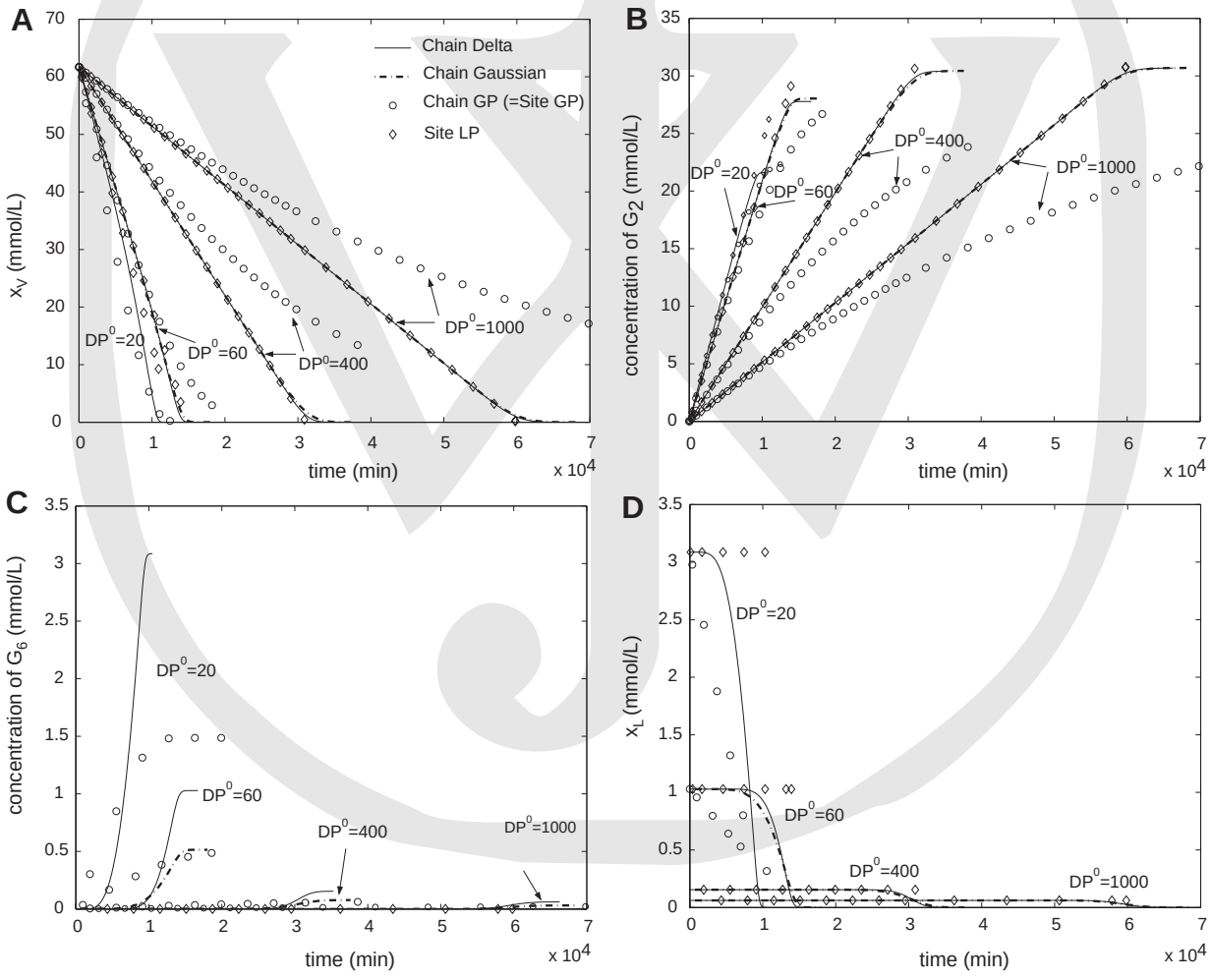

Figure 2. Comparison of the SS model results from chain number formalism (Zhang–Lynd Model) and site number formalism, with different initial chain length distributions, for the pure CBH1,2 enzyme system. In all **panels A–D**, abbreviations, full lines, dot-dashed lines, circles and diamonds are defined as in Figure 1. A and D: see descriptions in Figure 1; (B) concentration of G_2 in solution, $x_S(2, t)/2$, versus time t; (C) concentrations of G_6 in solution, $x_S(6, t)/6$, versus time t.

for dissolved monomer concentrations $x_S(k)$ in soluble oligomers. As expected, the dominant dissolved oligomer species is $k=2$ (cellobiose), since both CBH2 and CBH1 can cut off only oligomers lengths $k_X = k_Y = 2$ from the non-reducing and reducing ends, respectively. In addition, a small amount of $k=5$ (cellopentose) and $k=6$ (cellohexose) oligomers can be produced in the exact full chain calculations, as seen in Figure 2C.

In the limit of a pure CBH1,2 enzyme system, $u_1 = 0$ and hence $\Gamma_{N,\sigma} = 0$. This implies that the $x_{L,\sigma}$-term in Equation (11) does not contribute to the production of short-chains $G_\sigma(\ell)$ with $\ell \le \ell_D$. There is then essentially no source of production of short chains and the LP approximation gives zero short-chain concentrations $G_\sigma(\ell)$ for the entire duration of the hydrolysis process. As a consequence, the (very small) fraction of $k=5$ and $k=6$ oligomers is simply approximated by zero. However, in terms of the overall oligomer distribution, LP is actually a very good approximation to the exact full chain results for delta- or Gaussian-shaped initial distribution for realistic chain lengths $DP^{(0)} \ge 60$, since it reproduces the dominant $k=2$ oligomer very accurately for realistic $DP^{(o)}$.

In addition, the LP approach will then give a zero rate $R_{L,\sigma}$ for the production of chains or chain ends from chain fragmentation processes. In the single-layer limit ($\eta_\sigma(\lambda) = 0$), the chain concentration $x_{L,\sigma}$ thus becomes t-independent in the LP approximation. As shown in Figure 2D, this result agrees poorly with the chain number results in the case of short chains with $DP^{(o)} < 60$. However, under LCL conditions, that is, for larger $DP^{(o)}$-values, the full chain model with delta- or Gaussian initial distribution shape also predicts approximately t-independent x_L for most of the hydrolysis time: in comparing Figure 2A and D for $DP^{(o)} \ge 60$, we note that the full chain x_L remains approximately constant until about 80–95% of the substrate has been converted, which is then followed by a very quick downturn of x_L. Hence, the rather simple approximate LP result for x_L agrees, for most of the hydrolysis time with the full chain number result.

Both the site LP approximation and the exact full chain results for delta- or Gaussian-shaped initial distribution deviate noticeably in Figure 2 from the exact full chain results for the GP-shaped initial distribution. Compared to the other distributions, the GP distribution shows a slower loss of substrate monomers, as seen in Figure 2A, and, at the same time, a faster loss in the total number of chains in Figure 2D. This can be understood by noting that the GP distribution contains a larger fraction of its monomers in longer chains with $\ell > DP^{(o)}$, but larger fraction of its chains (and chain ends) at in shorter chains with $\ell \ll DP^{(o)}$. Recall here that both the site LP approximation and the full chain delta- and Gaussian-shaped initial distributions assume that there are initially negligible short chains, whereas GP assumes that the shortest chains have the largest concentrations right from the start. In the mixed endo–exo-acting enzyme system shown in Figure 1, this difference in the initial chain length distribution does not affect the hydrolysis rate significantly, since endo-cutting processes are very efficient (see Fig. 1E) in quickly producing a large population of short chains, even if short chains are initially absent. However, in the purely exo-acting enzyme system, the difference in the initial chain length distribution has a much more pronounced effect on the hydrolysis and chain loss rate, since it takes comparatively a much longer time for exo-cuts alone to produce short chains from long ones.

The primary pathology of the LP approximation which manifests itself in Figure 2D is that, for purely exo-acting enzyme systems (and only for those!), the LP site number formalism fails to eventually remove the chain ends (x_L) from the substrate (i.e., x_L is constant), even after all substrate monomers (x_M) have been completely dissolved. The persistence of these "phantom chains" in the LP approximation would be of no consequence if single-layer substrates are considered, as shown in Figure 2D. However, in the case of a full multi-layer surface ablation model with hydrolytically evolving substrate morphology and a purely exo-acting enzyme system, this pathology currently still limits the applicability of the site number formalism. Thus in LP-based site number simulations for surface ablation models with purely exo-acting enzyme systems, we must restrict ourselves to a short-time limit where only a few SAC surface layers are solubilized so that the accumulation of surface phantom chain ends remains a negligible artifact.

Chain Length Distributions

Figure 1E shows a typical chain length distribution evolution from the full chain model result, in common logarithm, with mixed EG1-CBH1,2 enzymes, and Delta shape initial chain distribution and $DP^{(o)} = 100$. The different time points for the plots were selected to cover most of the hydrolysis time from almost zero up to 90% hydrolytic conversion. The crucial point to note in Figure 1E is that $\log[G(\ell, t)]$ is an almost perfectly linear function of ℓ, at least for short chain lengths up to $\ell \le 80 - 90$, for all times t shown. That is precisely the ℓ-dependence assumed in our LP extrapolation. We have verified that such short-chain Poisson behavior is indeed almost universally realized for mixed EG1-CBH1,2 as well as for pure EG1 systems, under widely varying parameter conditions, including different rate and adsorption coefficients, different initial chain length distributions $Q(\ell)$ and initial DP (within LCL: $DP^{(o)} \gg \ell_S$), and different enzyme concentrations and mixing ratios. For the purely exo-acting CBH1,2 enzyme system, a non-Poissonian initial distribution shape $Q(\ell)$ will generally *not* evolve towards a Poisson shape, while a global Poissonian initial shape leads to exactly preserved Poisson shape thereafter, for any enzyme composition, as proven in Supporting Information section H of I.

In the presence of non-negligible amounts of endo-activity, a population of insoluble chains of *all* fragment chain lengths $\ell < \ell_{Max}$, down to short chains with $\ell \sim \ell_S$, gets produced immediately by the endo-cuts and this population very quickly evolves a Poisson distribution

shape, regardless of the initial distribution shape $Q(\ell)$. Hence, in LCL the fragmentation kinetics becomes "universal," that is, independent of initial distribution shape, since a Poisson distribution shape is established, especially at short chain lengths, long before even a small fraction of the substrate has been hydrolyzed. This also explains why the mixed enzyme system is much less sensitive to both initial DP in the LCL regime and to initial chain length distribution shape than the pure exo-enzyme system.

Hydrolysis Controlled by Morphology

While all results discussed below are based on the LP approximation in the site number formalism, we have also carried out calculations of the hydrolysis kinetics, *with* surface morphology term $\eta_\sigma > 0$, in the exact chain number formalism and compared to the corresponding site number LP results, for the case of the MS model. The results, discussed in Supporting Information section D, show again excellent agreement between the LP-based site formalism and the exact chain formalism.

Figure 3 shows results of three surface ablation models (i.e., MS, MM82-1, and MM82-2) and the SS model with $\overline{F}_a^{(o)} = 0.006$ applied, for the complete hydrolytic conversion process, using the low-concentration enzyme system E1. As illustrated in Figure 3A by the total monomer concentration x_V in solid substrate, the overall hydrolytic conversion in the three surface ablation models is significantly slower than in the pure chain fragmentation single-layer model SS. Furthermore, there are significant differences in hydrolytic conversion times between the three surface ablation models: MS, representing a zero-width Gaussian, hydrolyzes faster than MM82-2 with a finite-width Gaussian initial-SAC-size distribution; and MM82-2

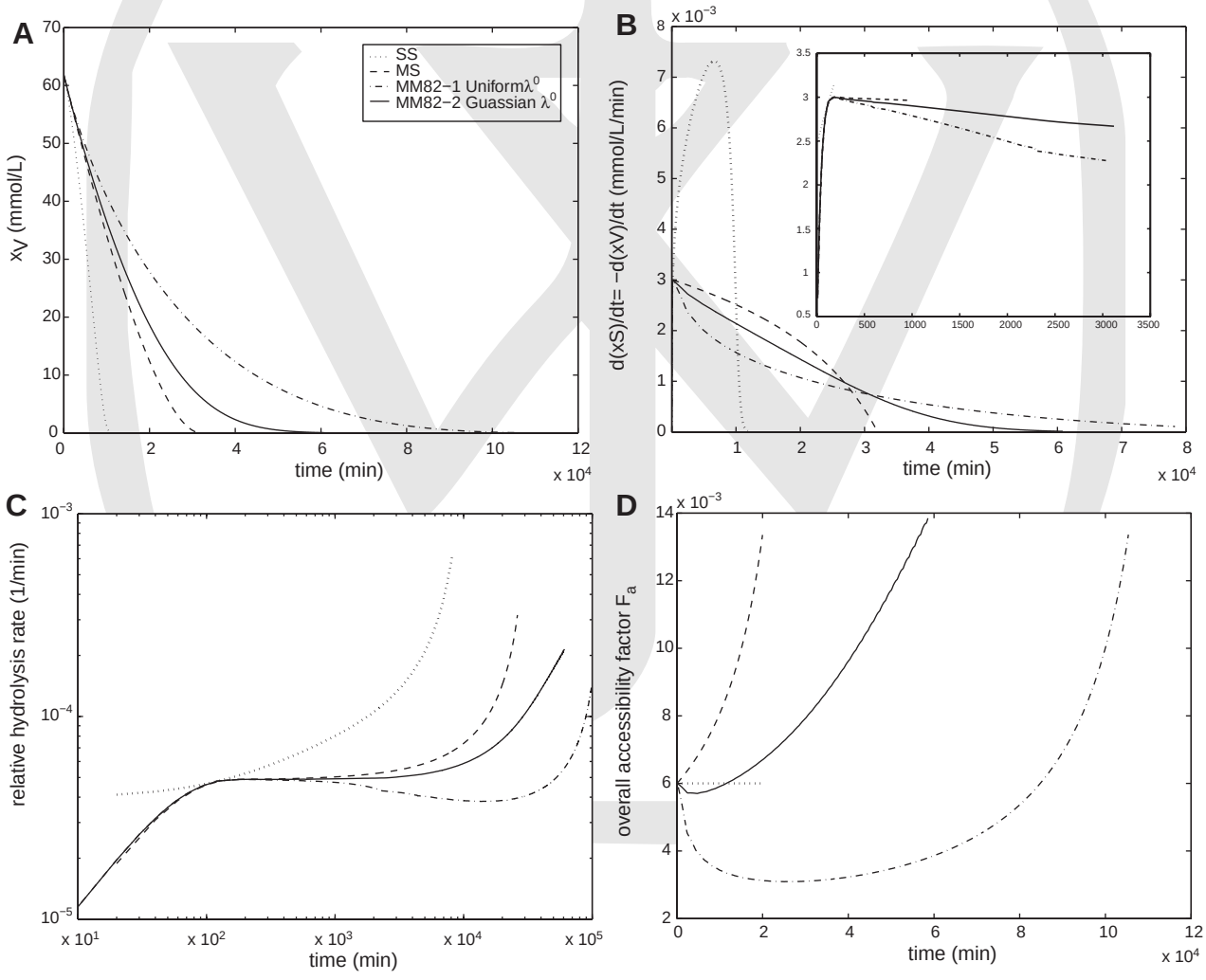

Figure 3. Simulation results of the MM82-1, MM82-2, MS, and SS models, with $DP^{(0)} = 300$ and $\overline{F}_a^{(o)} = 0.006$, for the E1 enzyme system. Plotted as functions of time t are (**A**) total monomer concentration in solid, x_V; (**B**) hydrolysis rate $|dx_V/dt|$; (**C**) relative hydrolysis rate Γ_{rel}; (**D**) overall accessibility fraction $\overline{F}_a$.

in turn, is faster than MM82-1 representing the much wider uniform initial-SAC-size distribution: the hydrolytic conversion time increases with the width of the initial SAC size ($\lambda_\sigma^{(o)}$) distribution.

The crucial point to emphasize here is that, in *all* four models, we have assumed the *same* chain fragmentation mechanism, with the *same* kinetic rate coefficient and enzyme parameters, and the *same* macroscopic substrate parameters, that is, the same initial molar amount of substrate $x_V(t^{(o)})$, the initial degree of polymerization $DP^{(o)}$ and the same initial enzyme surface accessibility fraction, $\overline{F}_a^{(o)}$, respectively. Clearly, the hydrolysis kinetics is very substantially dependent upon "other" factors, beyond the rate coefficient, enzyme adsorption or macroscopic substrate ($DP^{(o)}$ and $\overline{F}_a^{(o)}$) parameters used in single-layer chain fragmentation models (Okazaki and Moo-Young, 1978; Zhang and Lynd, 2006). The substrate morphology is one such critically important factor determining the overall hydrolytic conversion time.

It is evident from Figure 3 that the simulation results within the morphologic surface ablation models (MS and MM) are similar to each other, but quite different from the non-morphologic SS model. From the inset of Figure 3B and C, we can see for very short hydrolysis times (up to ~180 min) the behavior of the three morphologic surface ablation models is almost identical. As explained later, this characteristic time scale of ~180 min corresponds to the hydrolytic fragmentation of the initially accessible fraction of substrate material, residing in the outermost SAC layers. While the three morphologic models diverge from each other thereafter, this divergence is much less pronounced than their profound differences from the non-morphologic SS model at longer time scales. In particular, the non-morphologic model predicts a much higher solubilization rate, which can be understood as a consequence of the fundamental neglect of the obstruction of enzyme access to the chain ends.

The corresponding results for the E200 enzyme system shown in Figure 4 are qualitatively very similar to results from the E1 system: the three morphologic surface ablation models have much longer hydrolytic conversion time than the SS model. There are also significant differences in hydrolytic conversion times between the three surface ablation models. Due to the 200-fold increase in enzyme concentration, the reaction rates are scaled up, and the overall times scales are scaled down, by a factor of order 100. From the inset of Figure 4B and C, it can be seen that, even on very short hydrolysis time scales, the non-morphologic SS model is not a good approximation to the morphologic surface ablation models for the E200 enzyme system.

Hence, the hydrolytic conversion of cellulose substrate is crucially impacted by the substrate morphology. The above results also demonstrate very clearly that non-morphologic models can only be relied upon for the low-enzyme limit regime and only for very short time scales, up to the hydrolysis of the initial accessible fraction of substrate material. On time scales required to achieve substantial or near-complete hydrolytic conversion, or at higher enzyme loading, non-morphologic models are likely to fail.

Two-Time Scale Behavior

Figure 3B shows the conversion rate $|dx_V/dt|$ as a function of hydrolysis time for the E1 enzyme system. All four models show a very rapid rise in their initial conversion rate at very early times. However, in the SS model, this rise continues unabated until about $t \sim 6{,}600$ min where a maximum rate is reached, followed by a decline on a similar time scale, through completion of hydrolysis until about 11,000 min. By contrast, in the three surface ablation models, the early rapid rise is abruptly arrested and a much lower maximum rate is reached already at a much earlier time, $t \sim 180$ min, followed by a very slow drop-off for about 25,000–60,000 min, consistent with the overall much longer conversion times in the surface layer ablation models. These results strongly suggest that the hydrolysis kinetics in the surface layer ablation models exhibits two quite distinct characteristic time scales: the very short, early arrest time scale, and the much longer hydrolysis completion time scale, indicated, for example, by the 90%-conversion times.

This two-time scale behavior is also clearly seen in Figure 4B, for the E200 enzyme system, in all three morphologic surface ablation models, indicating that this is a common feature of the morphologic models, regardless of enzyme concentrations. The result in Figure 4B is qualitatively very similar to that of Figure 3B, with approximately 100-fold reduction in time, as mentioned before. The early arrest of the reaction rate in the surface ablation models occurs at ~1.7 min.

In the surface layer ablation models, the steric obstruction of enzyme accessibility is not "mimicked" by the ad hoc correction factor as was done in previous modeling studies (Okazaki and Moo-Young, 1978; Zhang and Lynd, 2006). Rather, reduced accessibility results naturally from the actual substrate morphology, that is, from the fact that only surface-exposed sites are available for enzyme adsorption. Inspection of the early arrest and downturn of the ablation rate, near ~180 min for the E1 enzyme system and ~1.7 min for the E200 enzyme system, in simulation results Figures 3B and 4B of the three surface ablation models reveals that this time corresponds to a 0.6% conversion of total substrate, for both enzyme systems; and this 0.6%-fraction is exactly equal to the initial fraction $\overline{F}_a = 0.006$ of substrate material exposed in the outermost SACs layers at the start of hydrolysis. Hence, the early arrest time scale, for both enzyme systems is clearly associated with the hydrolytic chain fragmentation and ablation of the outermost SAC layer.

Up to this "outermost layer ablation time," enzymatic cuts of endo-acting EG1 generates a large number of new chain ends, compared to initially existing native chain ends, which stimulates the activities of exo-acting CBH1 and CBH2. This cooperative work between endo–exo-enzymes

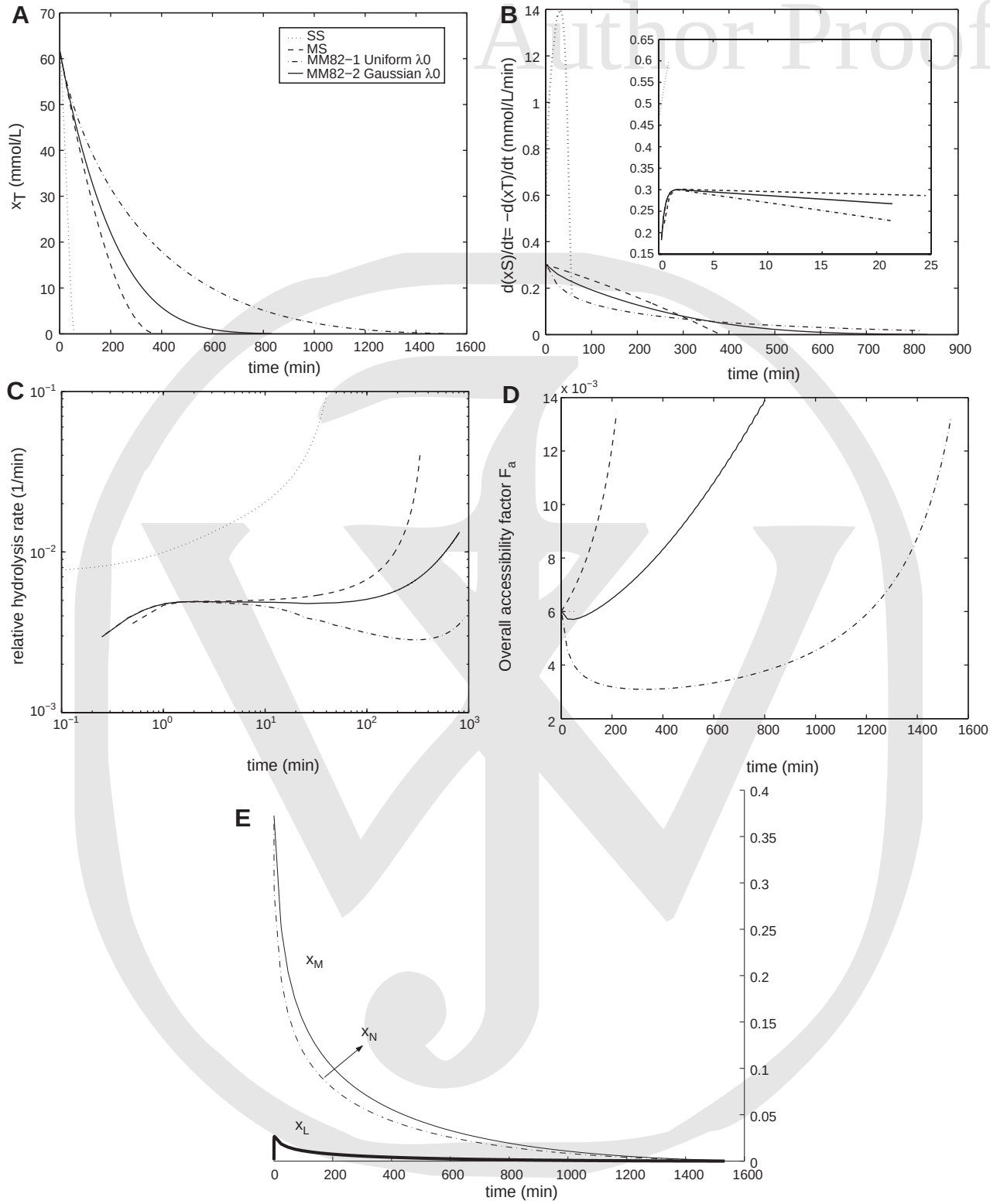

Figure 4. Simulation results of the MM82-1, MM82-2, MS, and SS models, with $DP^{(0)} = 300$ and $\overline{F}_a^{(o)} = 0.006$, for the E200 enzyme system. Plotted as functions of time t are (**A**–**D**): see descriptions of Figure 3; (**E**) concentrations of total surface exposed N sites, $x_N(t) \equiv \sum_\sigma x_{N,\sigma}(t)$, total surface exposed non-reducing ends, $x_L(t) \equiv \sum_\sigma x_{L,\sigma}(t)$, and total surface exposed monomers, $x_M(t)$.

causes the rapid increase in the production rate of soluble oligomers seen at early times in the insets of Figures 3B and 4B. This is also clearly shown in Figure 4E, where x_M and x_N decrease monotonically, and x_L shows a rapid rise during the early hydrolytic stage, with a time scale equal to that of the early arrest time scale seen in $|dx_V/dt|$. After that early arrest time, the steric obstruction by only partially ablated overlaying material affects and persists for all subsequent layers being ablated and hence controls the ablation rate for the entire remaining hydrolytic conversion time. Consequently, the rate of new surface exposure, that is the $-\overline{R}_\sigma \eta_\sigma(\lambda_\sigma) g_{N,\sigma}(\lambda_\sigma)$-term in Supporting Information Equation (8), *not* the enzymatic chain fragmentation, is the rate limiting factor for most of the remaining hydrolytic conversion time. This result clearly has technological implications: to substantially improve the performance of hydrolytic conversion, one may have to consider not only a re-engineering of the available enzyme systems, but also a re-engineering of the substrate morphology.

For ablation of the outermost SAC layer, only the total surface area, or surface site concentration, and the total ablation rate from all SACs of all geometry classes combined are relevant. Consequently, the MM and MS models of identical initial $\overline{F}_a$-, x_V-, and DP-values should exhibit the same early arrest short-time behavior arising from the outermost layer ablation. However, on the much longer overall hydrolysis time scales the three surface ablation models are evidently very different from each other, since the replenishment rate of digested substrate material at the SAC surfaces in these models is quite different because of the effects of the different morphology distributions and their evolution under hydrolysis. As shown in Figures 3A and B, the overall hydrolysis time scales show several fold differences between the three morphologic surface ablation models. This can be easily understood since the overall conversion time is controlled by hydrolysis of the large-size SACs. Hence, even though the initial accessible surface x_M and $\overline{F}_a$, is the same in all three models, the uniform MM82-1 model contains a larger fraction of its substrate in large SACs than the Gaussian MM82-2 and the (Delta-function!) MS model.

Hydrolysis Slow-Down and Morphology Evolution

Rapid decline in cellulose hydrolysis rate, and in the corresponding production rate of soluble glucose equivalent, as shown in Figures 3B and 4B, is a feature that has been frequently observed in real hydrolysis experiments, and is generally referred to as hydrolysis slow-down (Desai and Converse, 1997; Hong et al., 2007; Lynd et al., 2002; Yang et al., 2006; Zhang and Lynd, 2004). Some part of this effect found in real substrates has been attributed to a loss of enzyme activity, either due to enzyme degradation/inactivation or due to enzyme inhibition by the hydrolysis-generated soluble monomer and oligomer products. However, experiments, in which neither enzyme degradation/inactivation nor product inhibition appears operative, suggest that a significant part of the effect could in fact be due to hydrolysis-induced changes in the substrate itself that can not be explained by loss of enzymatic activity or product inhibition (Valjamae et al., 1998; Zhang et al., 1999). Zhang et al. (1999) explained the slow-down effect by declining substrate reactivity caused by substrate heterogeneity, whereby more easily degradable substrate was depleted at a faster rate early during hydrolysis. Valjamae et al. (1998) tried to explain the rate decline in terms of steric hindrance due to nonproductive cellulase adsorption, as well as surface erosion after extended hydrolysis. In some studies (Desai and Converse, 1997; Yang et al., 2006), however, partially hydrolyzed substrate was found to be as reactive as in its pre-hydrolysis state, implying that changes in substrate reactivity is not the cause of the slow-down in hydrolysis. Therefore, the cause of hydrolysis slow-down is uncertain, and whether or not there is a change in substrate reactivity is still very much under debate.

Zhang and Lynd (2004) stated that "It is widely observed that the heterogeneous structure of cellulose gives rise to a rapid decrease in rate as hydrolysis proceeds, . . . ," and "... it would seem logical to expect that the declining reactivity of residual cellulose during enzymatic hydrolysis is a result of factors such as less surface area and fewer accessible chain ends . . . ". Indeed, our present simulation results of morphology-evolving surface ablation models confirms these expectations. As shown in Figure 4E, both x_M (which is a measure of exposed accessible surface area) and x_L decrease as hydrolysis proceeds (after the short initial rise during the outermost surface ablation time). Please note that in our simulation, neither enzyme degradation/inactivation nor product inhibition are present (i.e., intentionally ignored for purpose of isolating the substrate effects), thus substrate effects contribute to hydrolysis slow-down, as shown by the results of our model.

The heterogeneous structure arising naturally from the solidity of cellulose, that is, steric obstruction of access to the inner, below-surface chains in the solid, is one of the main causes of the slow-down. From the Equations (6) and (5), the rate of solubilization $|dx_{V,\sigma}(\lambda_\sigma)/dt|$ for SACs of size λ_σ is roughly proportional to their surface area, $x_{M,\sigma}(\lambda_\sigma)$. The exposed surface of every individual SAC will shrink as the hydrolysis proceeds. Thus the total solubilization rate $|dx_V/dt|$, will begin to decrease very early on, as soon as the outermost surface layer, but long before any sizeable fraction (<1%) of substrate, has been hydrolyzed. By contrast, the SS model lacks the steric obstruction of a solid substrate and its $|dx_V/dt|$ in Figures 3B and 4B continues to *increase* until well over 80% of the substrate has been hydrolyzed.

However, the foregoing "solid structural heterogeneity" effect alone can not explain the whole picture. As we can see from Figures 3B and 4B, the extent of the hydrolysis slow-down is quite different among the three morphology models. The MM82-1 model exhibits the deepest depression

while the MS model exhibits the least amount of hydrolysis slow-down. As discussed before, the MS and MM82-1 models represent two limiting cases of the Gaussian MM model, namely the "zero-width" limit and the "infinite-width" limit respectively. A real morphology size distribution would likely fall in between these two extremes. In the MS model, all SAC units of the substrate shrink at the same rate and thus are and remain of equal size at all times during the hydrolysis. By contrast, in the MM models, smaller SACs are fully hydrolyzed at earlier times than bigger ones, and this apparently causes the relatively steeper decline in cellulose hydrolysis rate in the MM82-1 model. We will refer to this as the "morphological heterogeneity" effect in the following.

To further analyze these differences between the three multi-layer surface ablation models, we have plotted in Figures 3C and 4C the relative substrate hydrolysis rate Γ_{rel}, for E1 and E200 enzyme system respectively,

$$\Gamma_{rel} \equiv -\frac{1}{x_V}\frac{dx_V}{dt}. \quad (1)$$

In Figures 3C and 4C, the early rapid rise is again arrested in all three substrate ablation models at the outermost layer ablation time. Beyond that point, Γ_{rel} in the MM82-1 model drops noticeably below its early arrest value; Γ_{rel} in the MS models continues to rise, albeit with a markedly slower growth rate; and Γ_{rel} in the MM82-2 model falls between MS and MM82-1.

The decline of Γ_{rel} seems to indicate a decrease in the effective substrate reactivity for the uniform initial-λ distribution model MM82-1: the hydrolysis rate $|x_V/dt|$ declines faster than the remaining substrate concentration x_V itself. This is the model with the widest, distribution of initial SAC sizes $\lambda_\sigma^{(o)}$, extending with uniform weight from $\lambda_1^{(o)} = 20$ to $\lambda_{82}^{(o)} = 1640$. By contrast, in the MS and MM82-2 model, where all SAC units of the substrate have either the same initial size or a narrower, Gaussian size distribution, there is no, or only a very weak Γ_{rel} depression.

In Figures 3D and 4D, we show the hydrolytic evolution of the overall accessibility fraction $\overline{F}_a$, for E1 and E200 enzyme system respectively. For the zero-width distribution MS model, $\overline{F}_a$ increases monotonically; for the finite-width Gaussian distribution MM82-2 model, $\overline{F}_a$ at first declines very slightly for a short time and then increases; and for the widest uniform distribution MM82-1 model, $\overline{F}_a$ declines most strongly and it has the longest duration of decline. Hence, the depression of $\overline{F}_a$ increases with increasing width of the morphology distribution. The proportionality of solubilization rate and surface area also implies, by Equation (1), that the relative hydrolysis rate Γ_{rel} is proportional to the accessibility fraction $\overline{F}_a$. Hence, a decline of $\overline{F}_a$ during early hydrolysis implies a corresponding decline in Γ_{rel}.

To understand why MM82-1 and MM82-2 models exhibit a decline in the accessibility fraction $\overline{F}_a$ and thus a depression in the relative hydrolysis rate Γ_{rel}, we show, for the E200 enzyme system, time evolution of the weighted SAC size (λ_σ)-density distributions: (see Supporting Information section C for their definitions) $Px_M(\lambda, t)$ in Figures 5A and C, $Dx_V(\lambda, t)$ in Figures 5B and D, for the models MM82-1 and MM82-2 respectively. The analogous results for the E1 system are very similar and therefore not shown here. Please note that at $t = t^{(o)}$ both models have the same total number of monomers per SAC volume $x_V^{(o)}$; and the same total surface area $x_M^{(o)}$.

In the uniform model MM82-1, a large fraction of the initial surface area $x_{M,\sigma}$ resides on very small SACs, say at $\lambda < 200$, while most of the volume $x_{V,\sigma}$ is contained in SACs of very large sizes, with $\lambda > 200$. This is seen in Figure 5A and B, respectively. Consequently, most of the early hydrolysis rate will result from the dissolution of small SACs, since those are the ones carrying most of the surface: the large SACs, carrying much smaller surface area will initially be hydrolyzed at a slower rate than the small ones, as seen in Figure 5B. The early fast hydrolysis of small SACs causes a large initial loss of surface area per volume: the initial *relative* surface area loss in Figure 5A is greater than the corresponding *relative* volume loss in Figure 5B, since, as a fraction of total, the small SACs contain far more surface than volume. Consequently, the surface-to-volume ratio, that is, the accessibility fraction $\overline{F}_a$, will initially decline in the uniform model, as is indeed seen in Figures 3D and 4D. On the other hand, in the MM82-2 model a substantially larger fraction of initial surface area resides on intermediate-size SACs, say with $200 < \lambda < 600$; and that is also where most of the SAC volume resides. Hence, early fast hydrolysis of small SACs, with a monomer fraction much less than that of the MM82-1 model, leads to a much smaller decline in $\overline{F}_a$.

Smaller SACs always present more accessible surface area compared to larger SACs, given the same amount of substrate volume contained. As a result, the hydrolysis rate, which is roughly proportional to accessible surface area as we discussed before, of smaller SACs is higher than that of larger SACs. In other words, smaller SACs are easier to be hydrolyzed. The above analysis suggests that the early fast hydrolysis of easily solubilizable small SACs leads to the decline in the overall accessibility fraction, and is thus likely the main cause for an additional morphology-dependent decline in cellulose hydrolysis rate (in addition to that caused by steric accessibility obstruction of inner cellulose material). The random morphology SAC size distribution essentially determines the existence and fraction of these small SACs, and thus the extent of this further hydrolysis rate decline.

It has been proposed in previous studies that cellulose material contains two types of cellulose fractions that differ distinctly in their susceptibility to cellulase enzymatic attack. The basic idea here is that some types of, for example, amorphous, cellulose are easier to hydrolyze and other types, of, say, highly crystalline cellulose, are harder to hydrolyze

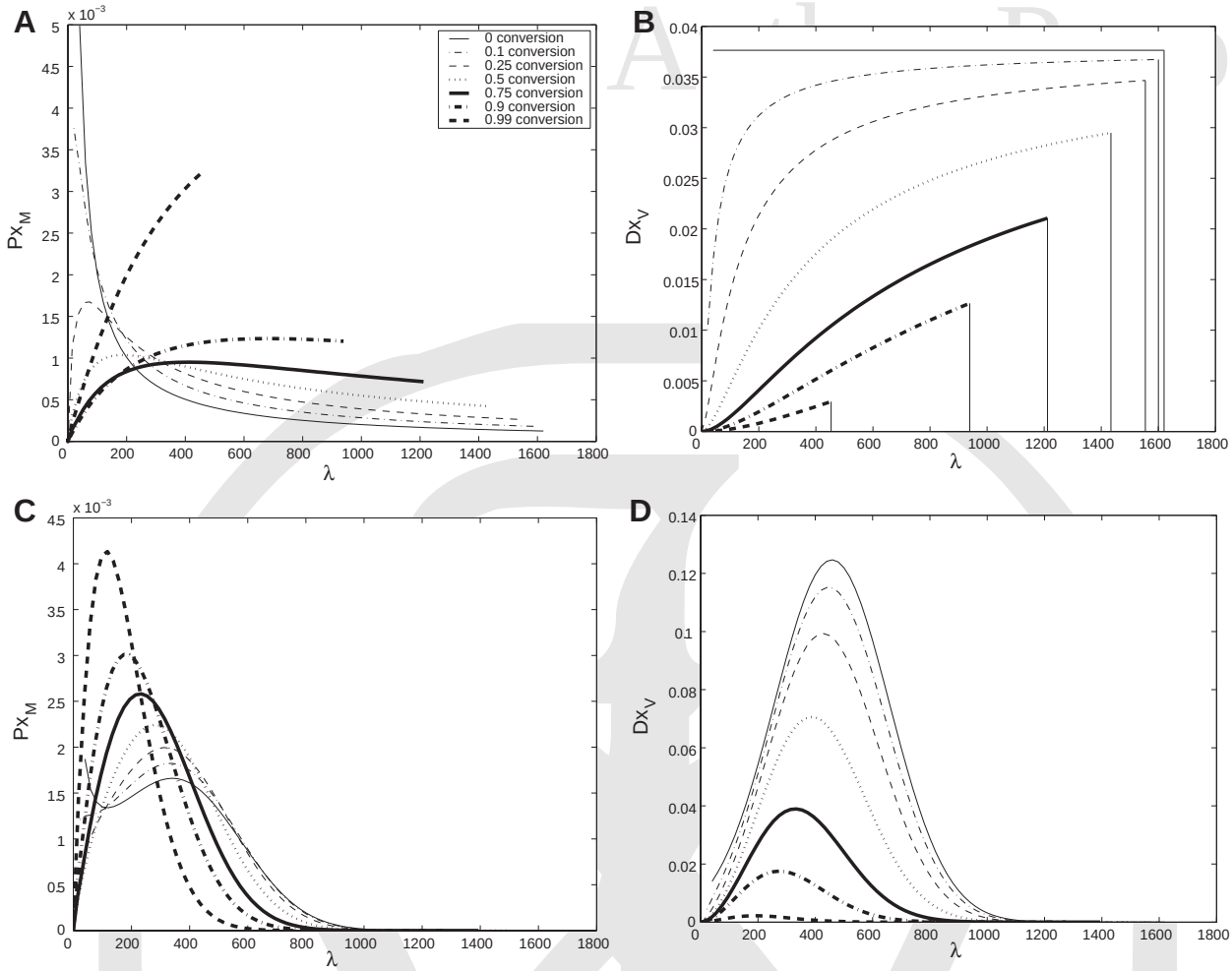

Figure 5. The weighted SAC size (l_o)-density distributions $Px_M(l, t)$ (**panels A** and **C**) and $Dx_V(\lambda, t)$ (**panels B** and **D**) are plotted versus SAC size λ at several times t during hydrolysis, for the models MM82-1 (panel A and B) and MM82-2 (C and D), using the E200 enzyme system.

(Gonzalez et al., 1989; Nidetzky and Steiner, 1993; Scheiding et al., 1984). Thus, if a (hypothetical) material contains both a substantial "fast-hydrolyzing" and a substantial "slow-hydrolyzing" substrate fraction, the fast early hydrolysis of the "fast" substrate results in the decline in hydrolysis rate at later times when only the "slow" substrate fraction remains. This "two-substrate" hypothesis attributes the difference of substrate reactivity specifically to the differences in the crystallinity of the two hypothesized fractions (Gonzalez et al., 1989; Nidetzky and Steiner, 1993; Scheiding et al., 1984). However, this picture has not yet been experimentally supported (Lynd et al., 2002). Although the existence of two (or multiple) substrate fractions of different reactivity within real pre-hydrolysis materials is presently uncertain, our simulation results imply that there do exist different hydrolysis rates among different substrate fractions, which are differentiated simply by volume, surface size and surface-to-volume ratios of their respective accessible substrate compartments. The fact that most of the substrate surface in real cellulosic materials is indeed comprised by internal surfaces (Zhang and Lynd, 2004), strongly suggests that this proposed "fractionation of substrate reactivity by geometry" may in fact be a ubiquitous feature of these materials.

Enzyme Concentration Scale-Up

The E1 enzyme set used above corresponds to a low-enzyme limit. It is speculated that the amount of cellulase required to achieve reasonable hydrolysis rate for real applications can be substantial (Lynd et al., 2002). Specifically, Mandels (1985) estimated that for *T. Reesei* cellulase system, 3% by mass of the initial amount of cellulose is required. Here, we have examined the cellulose enzymatic hydrolysis process in a mimic industrial environment by using the E200 enzyme set, with a 200-fold increase in concentrations in relative to the E1 set.

The corresponding results for the E200 enzyme system shown in Figure 4 are qualitatively very similar to those shown in Figure 3, except for the overall, approximately 100-fold reduction in time scales already discussed. Note that the reaction speeds should scale exactly linearly with enzyme concentrations as long as the enzyme-substrate system remains in the low-enzyme limit. However, the E200 system is already in the intermediate-to-high enzyme regime where the enzymes compete for available substrate sites, rather than substrate sites competing for enzymes. Consequently, the scale up in the reaction speed in going from E1 to E200 is neither exactly linear nor is it the same in all the four model on all time scales.

An analysis of the initial enzyme adsorption equilibrium shows that the free N and L site concentrations decrease from 99% in E1 to around 50% in E200, that is, almost twofold, relative to the total N and L site concentrations. On the other hand, the adsorbed enzyme fractions do not change much for the exo-acting enzymes, CBH1 and CBH2, and decrease from 53% in E1 to 37% in E200 for the endo-acting EG1. As a result, the initial concentration of ES complex, and hence the initial enzymatic cutting rates, show a 200-fold scale-up for CBH1,2 and a 139-fold scale-up for EG1 in E200 in relative to in E1, provided that the reaction rate coefficients remain unchanged.

The depression of Γ_{rel} in the MM82-1 model is somewhat more pronounced in the E200 system. A weak and brief Γ_{rel}-depression is now also seen in the Gaussian-distributed MM82-2 model in Figure 4C. Both of these results suggests that higher enzyme concentrations tend to favor hydrolysis slow-down behavior. This is probably due to greater cooperativity between endo- and exo-enzymes under high enzyme concentrations, consistent with experimental studies (Nidetzky et al., 1994; Woodward et al., 1988). As we can see, both the absolute and relative hydrolysis rate are much higher in the E200 system after the initial rapid rise, compared to the E1 system, as shown in Figures 4B and C and 3B and C. Thus, a much faster consumption of the relatively smaller sized SACs can be expected in the E200 system. This leads to a steeper decline in the supply of accessible substrate sites during the initial high-rate hydrolysis stage, and consequently the deeper decline in the hydrolysis rate.

For industrial applications, there is of course always a trade-off between the cost of enzyme added and the benefit from better performance of the operation using more enzymes. To explore possible performance optimizing applications of our modeling approach, we have also simulated, in addition to the E1 and E200 systems, an enzyme system, labeled E200/50, where only EG1 is increased 200-fold, but CBH1,2 are increased *only* 50-fold. The 90% conversion time in this E200/50 system differs by less than 1% from that in the E200 system. Thus from an economic point of view, one should never use the E200 enzyme system in an industrial application, since lower enzyme usage in E200/50 gives the same performance.

While the E200/50 system's enzyme composition deviates from the naturally occurring composition found in living microbial cells, this does not necessarily mean that the natural composition is not at optimum under in vivo conditions. It is possible that the enzymatic activities exhibited in vivo are different from technologically relevant in vitro environments, and that they may be subject to regulation by the cells. Thus, it is quite possible that substantial improvements of hydrolysis cost/performance under technologically relevant in vitro conditions can be achieved by our modeling approach, even for enzyme systems that have already been optimized, by nature, for in vivo performance. If process operation and economic parameters are available our modeling framework can provide a useful tool for a more detailed process optimization and design, by allowing us to perform systematic computational searches of parameter space for optimal processes and enzyme utilization. This will be the focus of future work.

Conclusions

We have simulated the enzymatic hydrolysis of solid cellulosic biomass within a general functional-based modeling framework which incorporates both the hydrolytic evolution of the substrate morphology and the effect of the morphology on the hydrolysis kinetics into a consistent surface layer ablation rate equation formalism. This surface layer ablation approach explicitly describes the interplay between hydrolysis kinetics and substrate morphology by treating kinetics and morphology on an equal footing.

An essential feature of the surface layer ablation formalism is its ability to capture the effect of random spatial inhomogeneity in the substrate morphology, which is invariably present in all as-grown or pre-treated cellulosic substrates. Substrate randomness includes, but is not limited to, random distributions of enzyme-accessible internal surface areas associated with randomly sized SACs in the substrate morphology; random spatial distributions of non-cellulosic contaminants within SACs; and/or random spatial distributions of the glucan chain degree of polymerization; random spatial distributions of the degree of chain ordering; and random distributions of hydrolysis time scales resulting from all the foregoing random spatial inhomogeneities of the substrate. In the present work, we have limited ourselves to an exploratory study of random geometry distributions and, specifically, random distributions of the initial (pre-hydrolysis) SAC sizes. Other types random-substrate effects can readily be incorporated into the surface layer ablation model.

The surface layer ablation model makes several robust, experimentally testable predictions for the hydrolysis kinetics. First, in all surface ablation models there exist two distinct hydrolysis time scales: the short single-

outermost-layer ablation time and the much longer overall hydrolysis time, required to achieve near-complete hydrolytic conversion. The short single-layer ablation time scale should be observable during the early stages of hydrolytic conversion: $|dx_V/dt|$ is predicted to exhibit an initial rapid rise which is suddenly arrested when the outermost SAC layer has been ablated. The existence of a much longer hydrolytic conversion time scale is a direct, inevitable consequence of the fact that the cellulosic substrate is a *solid*. Thus access to the interior is obstructed by overlaying material at the surface. At any given time, only a small fraction of all chains, namely those exposed at the SAC surfaces, are accessible. As a result, the full hydrolytic conversion process cannot be described by single-layer chain fragmentation models.

A second experimentally testable prediction of the surface layer ablation models is a strong dependence of the overall hydrolysis on the random substrate morphology. We have shown that changes in the width and shape of a random substrate morphology distribution will quantitatively and qualitatively alter the overall hydrolysis kinetics. Two substrates, with the same specific accessible internal surface area and the same degree of polymerization, subjected to the same hydrolytic enzyme system with the same rate coefficients, can exhibit vastly different hydrolytic conversion times, depending upon the widths and shapes of their respective morphology distributions. Cellulosic substrates are differentiated, and should therefore be classified, not only according to their "average" macroscopic substrate parameters, such as internal surface area, enzyme accessibility fraction and degree of polymerization (Zhang and Lynd, 2004); but also, and equally importantly, according to their random morphology distribution characteristics. We have proposed here that substrate morphology heterogeneity (i.e., a sufficiently wide SAC size distribution) can contribute, in addition to substrate solid structural heterogeneity (i.e., steric obstruction of enzyme accessibility), to a hydrolysis slow-down. In real substrates, these substrate effects could be operational *in combination* with others, such as product- and/or substrate-induced enzyme deactivation, in bringing about the observed hydrolysis slow-down phenomena.

There are several experimental approaches which could both provide a more detailed quantitative validation of our proposed morphology-based hydrolysis modeling approach; and help correlate observed hydrolysis kinetics with mesoscopic substrate morphology. Firstly, it would be of considerable interest to perform systematic experiments to measure the hydrolytic evolution of critical substrate and hydrolysis parameters, such as the accessibility fraction $\overline{F}_a$ and the surfaces-exposed (or, failing that, sample-averaged) degree of polymerization, as functions of time on well-characterized cellulosic substrates, while the substrate is undergoing hydrolytic conversion. Such time-dependent results for the accessibility fraction $\overline{F}_a$ and DP could be directly compared to model predictions, such as those shown in Figures 3 and 4, and they would provide a more stringent test of the model, as well as further constraints on the model parameterization.

Secondly, as discussed in our companion article I, the kinetically relevant morphology characteristics of cellulosic substrates, the SAC sizes, are defined on mesoscopic, that is, 10–1,000 nm length scales. Yet, little is presently known experimentally about the mesoscopic structure of these materials. It would therefore be of considerable interest to actually image and map out the SAC surfaces whose existence we have postulated here on geometrical grounds. As explained in I (in the sub-section on *Substrate Morphology and Enzyme Accessibility*), this might be achieved by decorating SAC surfaces or SAC-bounding SAVs, with a microscopically detectable marker (including, for example, gold-conjugated cellulases for electron microscopy; or fluorescently tagged cellulase carbon-binding domains or cellulase-sized fluorescent nanodots for confocal microscopy). If such decoration imaging experiments could be performed on substrate samples extracted at different stages of hydrolytic conversion, they would allow us to track and relate the morphology evolution to the macroscopically observed hydrolysis kinetics.

We cannot assess the technical feasibility of such a decoration imaging approach at the present time. However, if feasible, such experiments, in combination with morphology-based modeling would result in a more detailed understanding of the hydrolytic conversion process at mesoscopic length scales.

Nomenclature

$B_{V,\sigma}$	molar volume prefactor, $\equiv C_\sigma c_{V,\sigma}$
Dx_V	λ_σ-density distribution function of $x_{V,\sigma}$
$G_\sigma(\ell)$	concentration of G_ℓ exposed on class-σ SAC surfaces, mM, $\equiv C_\sigma H_\sigma(\ell)$
Px_M	λ_σ-density distribution function of $x_{M,\sigma}/x_M$
$R_{\nu,\sigma}$	production rate of type-ν site, mM/min, $\equiv C_\sigma V_{\nu,\sigma}$
$\overline{R}_\sigma$	negative rate of monomer loss ($\overline{R}_\sigma < 0$) into solution, mM/min
$x_{S,\sigma}$	concentration of dissolved G_1 monomers from class-σ SACs, mM
$x_S(k)$	concentration of dissolved G_1 monomers contained in dissolved oligomer G_k from all class SACs, mM

Greek Symbols

$\Gamma_{\mu,\sigma}$	enzyme cutting rate factors defined by Equations (19)–(21)
Γ_{rel}	relative hydrolysis rate, $\equiv -dx_V/dt/x_V$

Abbreviations

MM	"Multiple-layer, Multiple-geometry" model
MM82-1	the MM model with uniform distribution of monomer concentration per geometry class, (i.e., $\xi_\sigma^{(o)} = 1/82\ \forall\ \sigma$)
MM82-2	the MM model with Gaussian distribution of monomer concentration per geometry class through Equation (40)

MS "Multiple-layer, *S*ingle-geometry" model

SS "Single-layer, *S*ingle-geometry" model

We acknowledge the U.S. Department of Energy (grant # 4000063512) and the National Science Foundation (grant #: NSF/DBI-0354771, NSF/ITR-IIS-0407204, NSF/DBI-0542119, and NSF/CCF0621700) for financial support. We also acknowledge computing support from the University of Georgia Research Computing Center. The BioEnergy Science Center is supported by the Office of Biological and Environmental Research in the DOE Office of Science.

References

Converse AO, Grethlein HE. 1987. On the use of an adsorption model to represent the effect of steam explosion pretreatment on the enzymatic hydrolysis of lignocellulosic substances. Enzyme Microb Technol 9:79–82.

Converse AO, Optekar JD. 1993. A synergistic kinetics model for enzymatic cellulose hydrolysis compared to degree-of-synergism experimental results. Biotechnol Bioeng 42:145–148.

Converse AO, Matsuno R, Tanaka M, Taniguchi M. 1988. A model of enzyme adsorption and hydrolysis of microcrystalline cellulose with slow deactivation of the adsorbed enzyme. Biotechnol Bioeng 32:38–45.

Desai SG, Converse O. 1997 Substrate reactivity as a function of the extent of reaction in the enzymatic hydrolysis of lignocellulose. Biotechnol Bioeng 56:650–655.

Fenske JJ, Penner MH, Bolte JP. 1999. A simple individual-based model of insoluble polysaccharide hydrolysis: The potential for auto-synergism with dual-activity glycosidases. J Theor Biol 199:113–118.

Gan Q, Allen SJ, Taylor G. 2003. Kinetic dynamics in heterogeneous enzymatic hydrolysis of cellulose: An overview, an experimental study and mathematical modeling. Proc Biochem 38:1003–1018.

Gonzalez G, Caminal G, de Mas C, Lopez-Santin J. 1989. A kinetic model pretreated wheat straw saccharification by cellulose. J Chem Technol Biotechnol 44:275–288.

Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD. 2007. Biomass recalcitrance: Engineering plants and enzymes for biofuels production. Science 315:804–807.

Hong J, Ye X, Zhang YHP. 2007. Quantitative determination of cellulose accessibility to cellulase based on adsorption of a nonhydrolytic fusion protein containing CBM and GFP with its applications. Langmuir 23: 12535–12540.

Hsu T-A. 1996. Pretreatment of biomass. In: Wyman CE, editor. Handbook on biomass ethanol. Taylor and Francis: Washington, DC, pp. 179–212.

Jacobsen SE, Wyman CE. 2000. Cellulose and hemicellulose hydrolysis models for application to current and novel pretreatment processes. Appl Biochem Biotechnol 84:81–96.

Luo J, Xia L, Lin J, Cen P. 1997. Kinetics of simultaneous saccharification and lactic acid fermentation process. Biotechnol Prog 13:762–767.

Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS. 2002. Microbial cellulose utilization: Fundamentals and biotechnology. Microbiol Mol Biol Rev 66:506–577.

1985. Applications of cellulase. Biochem Soc Trans 13:414–416.

McMillan JD. 1994. Pretreatment of lignocellulosic biomass. ACS Symp Ser 566:292–324.

Movagarnejad K, Sohrabi M, Kaghazchi T, Vahabzadeh F. 2000. A model for the rate of enzymatic hydrolysis of cellulose in heterogeneous solid-liquid systems. Biochem Eng J 4:197–206.

Nidetzky B, Steiner W. 1993. A new approach for modeling cellulasecellulose adsorption and the kinetics of the enzymatic hydrolysis of microcrystalline cellulose. Biotechnol Bioeng 42:469–479.

Nidetzky B, Sterner W, Hayn M, Claeyssens M. 1994. Cellulose hydrolysis by the cellulase from Trichoderma reesei: A new model for synergistic interaction. Biochem J 298:705–710.

Oh K-K, Kim S-W, Jeong Y-S, Hong S-I. 2001. Bioconversion of cellulose into ethanol by nonisothermal simultaneous saccharification and fermentation. Appl Biochem Biotechnol 89:15–30.

Okazaki M, Moo-Young M. 1978. Kinetics of enzymatic hydrolysis of cellulose: Analytical description of a mechanistic model. Biotechnol Bioeng 20:637–663.

Philippidis GP, Spindler DD, Wyman CE. 1992. Mathematical modeling of cellulose conversion to ethanol by the simultaneous saccharification and fermentation process. Appl Biochem Biotechnol 34/35:543–556.

Philippidis GP, Smith TK, Wyman CE. 1993. Study of the enzymatic hydrolysis of cellulose for production of fuel ethanol by the simultaneous saccharification and fermentation process. Biotechnol Bioeng 41:846–853.

Scheiding W, Thoma M, Ross A, Schugerl K. 1984. Modeling of the enzymatic hydrolysis of cellobiose and cellulose by a complex enzyme mixture of Trichoderma reesei QM9414. Appl Microbiol Biotechnol 20:176–182.

Suga K, van Dedem G, Moo-Young M. 1975. Degradation of polysaccharides by endo and exo enzymes: A theoretical analysis. Biotechnol Bioeng 17:433–439.

Valjamae P, Sild V, Pettersson G, Johansson G. 1998. The initial kinetics of hydrolysis by cellobiohydrolases I and II is consistent with a cellulose surface—Erosion model. Eur J Biochem 253:469–475.

Wald S, Wilke CR, Blanch HW. 1984. Kinetics of the enzymatic hydrolysis of cellulose. Biotechnol Bioeng 26:221–230.

Weil J, Westgate PJ, Kohlmann K, Ladisch MR. 1994. Cellulose pretreatments of lignocellulosic substrates. Enzyme Microb Technol 16:1002–1004.

Woodward J, Lima M, Lee NE. 1988. The role of cellulase concentration in determining the degree of synergism in the hydrolysis of microcrystalline cellulose. Biochem J 255:895–899.

Yang B, Willies DM, Wyman CE. 2006. Changes in the enzymatic hydrolysis rate of Avicel cellulose with conversion. Biotechnol Bioeng 94:1122–1128.

Zhang Y-HP, Lynd LR. 2004. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: Noncomplexed cellulase systems. Biotechnol Bioeng 88:797–824.

Zhang Y-HP, Lynd LR. 2006. A functionally based model for hydrolysis of cellulose by fungal cellulase. Biotechnol Bioeng 94:888–898.

Zhang S, Wolfgang DE, Wilson DB. 1999. Substrate heterogeneity causes the nonlinear kinetics of insoluble cellulose hydrolysis. Biotechnol Bioeng 66:35–41.

Zhou W, Schüttler HB, Hao Z, Xu Y. 2009. Cellulose hydrolysis in evolving substrate morphologies I: A general modeling formalism. Biotechnol Bioeng xxx:xxx–xxx.