Bridging the Micro-Macro Gap between Single-Molecular Behavior and Bulk Hydrolysis Properties of Cellulase

Takahiro Ezaki, Katsuhiro Nishinari, Masahiro Samejima, and Kiyohiko Igarashi

Phys. Rev. Lett. **122**, 098102 — Published 7 March 2019

DOI: 10.1103/PhysRevLett.122.098102

Bridging the Micro-Macro Gap between Single-Molecular Behavior and Bulk Hydrolysis Properties of Cellulase

Takahiro Ezaki,[1,2,*] Katsuhiro Nishinari,[1] Masahiro Samejima,[3] and Kiyohiko Igarashi[3,4,†]

[1] *Research Center for Advanced Science and Technology, The University of Tokyo, Meguro-ku, Tokyo, Japan*
[2] *PRESTO, JST, 4-1-8 Honcho, Kawaguchi, Saitama, Japan*
[3] *Department of Biomaterial Sciences, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Bunkyo-ku, Tokyo, Japan*
[4] *VTT Technical Research Centre of Finland, P.O. Box 1000, FI-02044 VTT, Finland*

The microscopic kinetics of enzymes at the single-molecule level often deviate considerably from those expected from bulk biochemical experiments. Here, we propose a coarse-grained-model approach to bridge this gap, focusing on the unexpectedly slow bulk hydrolysis of crystalline cellulose by cellulase, which constitutes a major obstacle to mass production of biofuels and biochemicals. Building on our previous success in tracking the movements of single molecules of cellulase on crystalline cellulose, we develop a mathematical description of the collective motion and function of enzyme molecules hydrolyzing the surface of cellulose. Model simulations robustly explained the experimental findings at both the microscopic and macroscopic levels and revealed a hitherto-unknown mechanism causing a considerable slowdown of the reaction, which we call the crowding-out effect. The size of the cellulase molecule impacted significantly on the collective dynamics, whereas the rate of molecular motion on the surface did not.

Introduction.—Recent observational studies have provided a detailed understanding of enzymatic kinetics at the single-molecule level [1]. But, contrary to expectation, behavior at the single-molecule level is often unable to account for the results of conventional biochemical reactions (involving, for example, $\sim 10^{15}$ molecules in a 1 μM $\times$ 1 mL reaction mixture) [2]. In other words, a theoretical basis for bridging these two very different scales is missing. A well-known example is the enzymatic hydrolysis of cellulose, which is generally very slow, representing a major obstacle to the mass production of biofuels and biochemicals [3].

In 2009, we succeeded for the first time in tracking the movements of individual molecules of a cellulase (*Trichoderma reesei* cellobiohydrolase I: *Tr*Cel7A) during degradation of crystalline cellulose by means of high-speed atomic force microscopy (HS-AFM) [4]. However, contrary to our expectation, the results posed a new enigma: the velocity of individual cellulase molecules on the surface was incredibly fast as compared to the value estimated from bulk experiments (more than 400-fold difference) [5]. To understand this inconsistency, we improved the HS-AFM device to obtain higher spatiotemporal resolution, and found that "traffic jams" of cellulase molecules developed on cellulose microfibrils [6]. We thought that this might account for the decreased hydrolytic rate, but direct evidence that this mechanism might quantitatively explain the results of bulk experiments is still lacking. In fact, the speed of detected motion of a cellulase molecule in this experiment did not agree with the average velocity of cellulase molecules estimated from bulk experiments [5].

Because in this catalytic reaction, the substrate has a crystalline structure and the size of cellulase molecules is not negligible, we hypothesized that complex nonlinear interactions between molecules dramatically affect the bulk chemical properties. It is often the case that the collective behavior of individuals cannot be trivially understood from that of each individual, especially when the number of individuals involved in the system is large [7]. As demonstrated in our previous study [6], the trafficking of cellulase molecules on a cellulose surface appears to be such a case. In physics, a model-based approach has provided detailed understandings of various transportation phenomena [8, 9], including the flow of human crowds [10, 11] and intracellular transportation of motor proteins [12–14]. Here, we adopt this approach to develop a traffic-flow model tailored to the cellulose-cellulase system, aiming to bridge the micro-macro gap between the single-molecular behavior of cellulase and its properties in the bulk hydrolysis of cellulose. Note that we are not first to construct a course-grained model of the cellulase-cellulose system [15–17]. However, as we discuss below, the previous studies did not incorporate a physical exclusion effect of cellulase, dependent on its form. We find that this factor gives rise to a hitherto-unknown collective effect (i.e., the *crowding-out* effect), which is crucial to account for the results obtained in both single-molecular and bulk-chemical experiments.

Model.—We briefly review the experimentally observed mechanics of *Tr*Cel7A (Fig. 1). The degradation of cellulose crystalline is performed by *Tr*Cel7A on the top surface of the substrate [Fig. 1(a)], which, of course, is altered during the hydrolytic process [18]. *Tr*Cel7A has a two-domain structure composed of the cellulose-binding domain (CBD) and catalytic domain (CD) [19]. It binds to the cellulose surface via the CBD (non-productive binding) and subsequently loads the substrate (i.e., a

single cellulose chain) into its CD (productive binding) [Figs. 1(b) and 1(c)]. After successful loading of the substrate, the CD slides unidirectionally along the chain with successive release of cellobiose units [Fig. 1(c)]. Productively and non-productively bound cellulase molecules are released at different, but constant, rates. The kinetics of these dynamic processes are well documented in the literature [20]. Given these features of enzymatic hydrolysis of cellulose nanofibrils, we constructed a simple coarse-grained stochastic model [Figs. 1(d)–1(g)].

Two modes of binding are assumed, i.e., productive (P) and non-productive (NP) [21] [Figs. 1(b) and 1(c)]. Cellulose fibers are assumed to have 20×20 chains [Fig. 1(g)] [5], each of which contains L cellobiose units. One cellobiose unit has dimensions of 1 nm $\times$ 0.5 nm [22]. The single-molecular kinetics of *Tr*Cel7A is modeled by a set of stochastic transitions characterized by kinetic rates [1]. On the top surface of the cellulose fiber, cellulase molecules can attach and detach themselves with kinetic rates [$k_{\rm on}$ and $k_{\rm off}$, respectively; Figs. 1(d) and 1(e)]. Non-productive binding can occur at any point unless blocked by other cellulase molecules. At the reducing end of each cellulose, a transition from non-productive to productive bindings (complexation) and its reverse transition (decomplexation) occur with rates $k_{\rm c}$ and $k_{\rm d}$, respectively. A productive cellulase molecule hydrolyzes the end of the chain, moving to the next cellobiose unit, with a rate $k_{\rm pr}$ [Fig. 1(f)]. Thus, the flow rate of cellulase molecules coincides with the degradation rate of cellulose. Here, we modeled the exclusion area of each cellulase molecule as a square region for simplicity (Fig. 1). The size of the area is separately defined for the P and NP modes as $R^{\rm P}$ and $R^{\rm NP}$, respectively. The assumption of physical exclusion of cellulase is in line with a previous study, which reported a mutual hindrance effect attributed to steric exclusion of molecules [23]. All processes occur only on the top surface of the cellulose fiber, i.e., the chains lying below unreacted chains do not react. In accordance with the kinetic rates measured in single-molecule experiments [average velocity: $v = 7.1 \pm 3.9$ nm/s [6]; detachment rate (NP): $k_{\rm off} = 0.86 \pm 0.03$ s^{-1}; detachment rate (P) including decomplexation and subsequent NP detachment: $k_{\rm off}^{\rm P} = 0.12 \pm 0.01$ s^{-1} [20]] and the hydrodynamic radius (r) of the *Tr*Cel7A molecule ($R^{\rm P} \approx 2r = 8.4 \pm 1$ nm [24]), we set the parameter values as $k_{\rm pr} = 7$ s^{-1}, $k_{\rm off} = 1$ s^{-1}, $k_{\rm d} = 0.1$ s^{-1}, $R^{\rm P} = 8$ nm, and $R^{\rm NP} = 4$ nm (see also Supplemental Material). The attachment rate, $k_{\rm on}$, is assumed to be proportional to the concentration of cellulase and is a controllable parameter in experiments. Considering the experimental observation that approximately 50% of the cellulase molecules adsorbed on the surface were non-productive in a dilute condition [20], we set $k_{\rm c} (= k_{\rm d}) = 0.1$ s^{-1} to satisfy a balance equation (Supplementary Material). As shown in the Supplemental Material, we confirmed that the conclusions in the main text were not substantially affected by variations in these parameters. Note that the temperature of the environment, which generally affects the kinetic rates, is kept constant throughout the simulations in accordance with the experimental conditions employed in Refs. [5, 6].

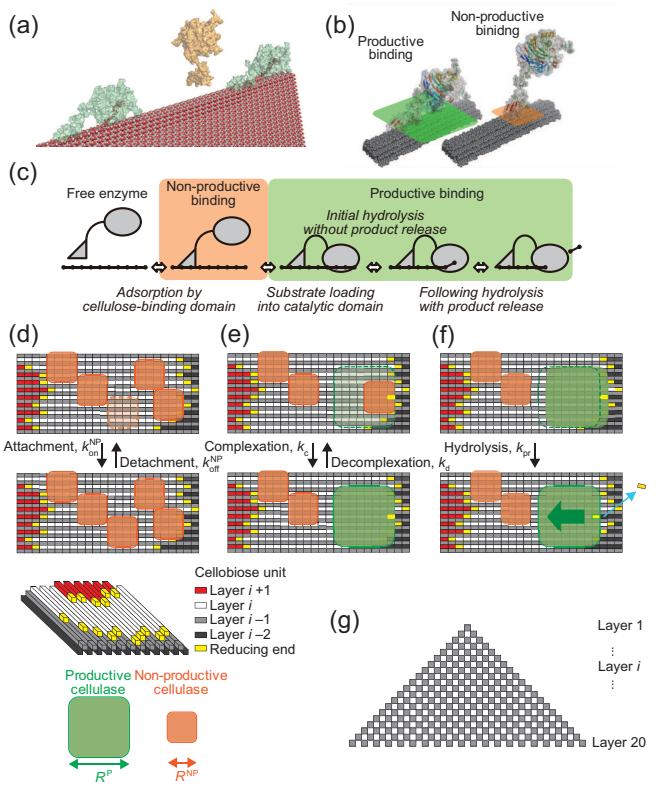

FIG. 1. Overview of the system. (a) Schematic illustration of cellulase molecules on crystalline cellulose. (b) Two types of binding: productive binding (left) and non-productive binding (right). (c) Reaction mechanism of enzymatic hydrolysis of cellulose. (d–g) Definitions of the model. (d) A cellulase attaches to the bulk cellulose surface with a rate of $k_{\rm on}$ in a non-productive manner. A non-productive cellulase leaves the surface with a rate of $k_{\rm off}$ (detachment). (e) At the reducing end of each cellobiose chain (shown in yellow), complexation (decomplexation) occurs with a rate of $k_{\rm c}$ ($k_{\rm d}$) if a non-productive (productive) cellulase is present. (f) A productive cellulase moves forward by one cellobiose unit with a rate of $k_{\rm pr}$, decomposing the site where it was originally located. (g) Conformation of cellulose chains. The upper half of the structure is shown.

Results.—Figure 2(a) shows sample trajectories of cellulase molecules. The simulated results were consistent with the HS-AFM recordings in the previous studies [4, 6]. This agreement confirms that the motion of individual cellulase molecules simulated in our model with parameter values based on empirical data is consistent with that observed experimentally. We also confirmed that the time course of total cellobiose production was qualitatively consistent with that obtained in previous experiments [6] [Fig. 2(b)] . Next, we examined the abil-

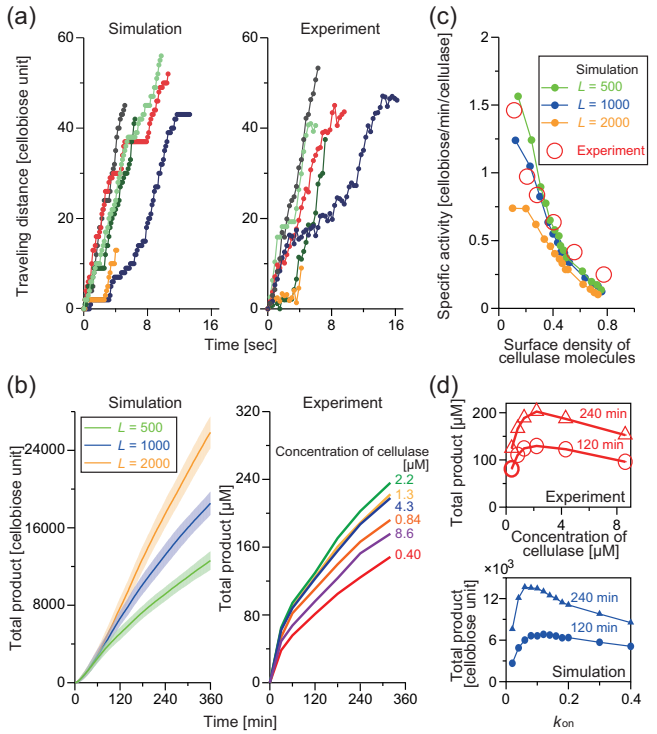

FIG. 2. Comparisons between simulation and experimental results [5, 6]. (a) Sample time-space trajectories of cellulase molecules. We selected typical trajectories that occurred between $t = 0$ and $t = 10\,000$ s. (b) Time course of cellobiose production for $L = 500$, 1000, and 2000. Shaded areas represent s.d. (c) Specific activity as a function of surface density at $t = 120$ min. (d) Relationship between the concentration of free cellulase (i.e., $k_{\rm on}$) and total production at $t = 120$ and 240 min. (a,d) We set $k_{\rm on} = 0.1$ s^{-1}. For more details of simulation conditions and experimental procedures, see Supplemental Material.

ity of our model to explain the considerable slowdown of the bulk degradation rate. We defined the specific activity of the *Tr*Cel7A molecules as the speed of total cellobiose production divided by the number of cellulase molecules on the surface (i.e., production speed per cellulase) [6]. The previous experimental results showed that the specific activity was less than 1 cellobiose unit/min when the total production was maximized, which is quite small considering the degradation rate of an isolated cellulase molecule (i.e., 7 cellobiose unit/s). Our model realistically reproduced this small specific activity for a wide range of attachment rates [i.e., concentrations of *Tr*Cel7A; Fig. 2(c)]. We also computed the total product formation at $t = 120$ and 240 min for various concentration conditions [Fig. 2(d)]. Because the length distribution of cellulose microfibrils used in the experiment is not readily traceable, reliable comparison of the absolute values of total product formation between simulation and experiment (in mol unit) was not possible. The results well explained the decrease in production in

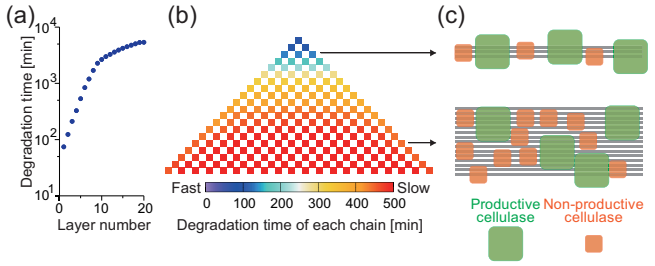

FIG. 3. Relationship between the layer number and degradation time. (a) Degradation time of each layer. The time at which each layer is completely decomposed is shown. (b) Intrinsic degradation time of each chain. (a,b) We set $k_{\rm on} = 0.1$ s^{-1}. (c,d) Schematic representations of configurations of cellulase molecules for (c) layer 3 and (d) layer 15.

terms of congestion of cellulase molecules on the cellulose surface [see also Fig. 2(b)]. Figures 2(a)–2(d) provide compelling evidence that our model works at both the microscopic and macroscopic levels, notwithstanding the presence of extensive collective interactions among cellulase molecules. We consider that this level of accuracy of prediction is sufficient considering the simplicity of our model and the fact that our aim here is not to pursue a perfect fit to the experimental data by detailed modeling (see also Supplemental Material for the robustness of the results).

Given the acceptable fidelity of our model, we can now analyze the enzymatic hydrolysis of cellulose in more detail than has previously been possible, because every quantity is measurable in the simulations. Figure 3(a) shows the time required to finish degradation to the i-th layer. The first few layers were decomposed rapidly, but as the layer number increased, the degradation significantly slowed down. This remarkable slowdown is attributed to the increase in the width of the cellulose surface. To study this effect, we measured the time required to decompose each chain when each layer was isolated from the others [Fig. 3(b)]. When the width is small, the configuration of cellulases on the surface is approximately one-dimensional [Fig. 3(c)]. In contrast, when the width is large, the cellulase molecules are distributed on the surface in a two-dimensional manner, and blocking occurs more frequently [Fig. 3(d); see also Supplementary Movie S1]. This intrinsic degradation time of each chain converges to a constant value for an increasing layer number because the effects of the boundaries become negligible. The increase of exclusion effects according to the time-dependent development of surface width can account for the slowdown of the reaction during the time course [Fig. 2(b)], which is orthogonal to previous theories [25, 26].

Finally, taking advantage of the controllability of the parameter values in the model, we assessed the impact of each parameter on the entire degradation process [Figs.

4(a)–4(f)]. Variations in the exclusion size (i.e., R^{P} and R^{NP}) had significant impacts on the degradation speed [Figs. 4(a) and 4(b)], whereas variations in the kinetic rate parameters had only marginal effects [Figs. 4(c)–4(f)]. If $R^{\mathrm{P}} = R^{\mathrm{NP}}$, the production rate was approximately ten times larger and the decrease in production rate due to congestion of cellulase molecules did not appear [Figs. 4(a) and 4(b)]. In other words, it appears that the difference in the exclusion size between the two binding modes (i.e., productive and non-productive) of real *Tr*Cel7A is the primary cause of the slow degradation of cellulose. We also examined the number of productive cellulase molecules on the surface [Figs. 4(g) and 4(h)]. The results suggested that the decrease in the production speed can be simply attributed to the decrease in the number of productive cellulase molecules. This is counterintuitive, because the total number of cellulase molecules on the surface (i.e., productive and non-productive) increases with k_{on} (i.e. with increasing concentration). In fact, however, when R^{NP} is smaller than R^{P}, non-productive cellulases, whose attachment and detachment processes are faster than those of productive cellulases, reduce the space available for productive binding, which requires a larger open space [Fig. 4(i)]. In this way, productive cellulase is crowded out, and the effective speed of degradation is considerably reduced. In contrast, if $R^{\mathrm{NP}} = R^{\mathrm{P}}$ for example, productive binding is not crowded out because the emptied space after detachment of cellulase is sufficiently large for productive binding [Fig. 4(j); see also Supplementary Movie S2]. Therefore, there is a monotonic increase of productive cellulase with increasing k_{on} [Fig. 4(g) and 4(h)]. See also Supplemental Material for further simulation results supporting the relevance of R^{P} and R^{NP} to the system dynamics.

Discussions.—Our model, developed from a microscopic level, was nevertheless able to explain experimental results obtained at a macroscopic level qualitatively over a wide range of concentration conditions, including the disputed slow reaction rate. Recent developments in experimental techniques have made it possible in this work to move beyond extant theories [25] and models [15–17]. First, sophisticated single-molecular observations have provided detailed information on enzymatic kinetics [20], which can be used to enable plausible modeling. In fact, the kinetic rates estimated by using an assumed kinetic model in previous studies significantly deviated from those directly measured in single-molecular experiments [20]. Second, experimental data obtained by using almost pure crystalline cellulose without amorphous structure [27] allowed us to make quantitative comparisons between simulation and experimental results. Thus, our modeling is better founded on available experimental evidence. Indeed, we succeeded in validating the model dynamics in various ways (Fig. 2), which significantly extended the previous studies that performed validations

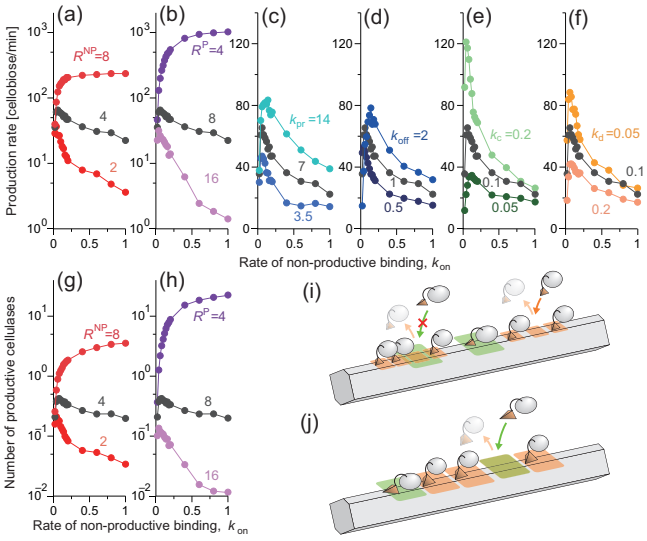

FIG. 4. Impacts of cellulase properties on production rate at t = 120 min. The results for the original condition (gray circles) and two other conditions varying a single parameter (a) R^{NP} = 2 nm and 8 nm, (b) R^{P} = 4 nm and 16 nm, (c) k_{pr} = 3.5 s^{-1} and 14 s^{-1}, (d) k_{off} = 0.5 s^{-1} and 2 s^{-1}, (e) k_{c} = 0.05 s^{-1} and 0.2 s^{-1}, and (f) k_{d} = 0.05 s^{-1} and 0.2 s^{-1} are shown. The other parameters not explicitly shown in each panel were left unchanged. (g,h) Number of productive cellulase molecules on the surface for different values of (g) R^{NP} and (h) R^{P}. For more details of simulation conditions, see Supplemental Material. (i,j) Schematic representations of the crowding-out effect. If R^{NP} is smaller than R^{P} (i), the productive cellulase is crowded out; otherwise (j), the effect is not observed. The colors of the areas and arrows, i.e. green and orange, represent productive and non-productive bindings, respectively.

using only the time course of cellulose production for a limited number of concentrations [15–17]. The improved fidelity of our model led to new findings.

Specifically, we claim that the central cause of the reaction slowdown is the crowding-out effect due to the two-domain structure of the cellulase. Although there are some similarities between the cellulase dynamics and vehicular traffic, as mentioned in Ref. [6], traffic-jamming of cellulase molecules cannot fully account for the suppression of the reaction rate (see also Supplemental Material). It is worth mentioning that in the context of intracellular molecular dynamics, macromolecular crowding is known to impact significantly on the speed of enzymatic reactions [28]. Thus, our results might also provide insight into such problems.

Our findings here suggest that we still have opportunities to enhance the degradability of cellulose not only by means of pretreatment to increase the susceptibility of the substrate [29], but also by engineering more processive cellulases with an increased size of the CBD. Thus, we believe our model opens up opportunities for completely new approaches to the design of next-generation enzymes for cellulosic biomass utilization.

This research was supported by a Grant-in-Aid for Innovative Areas from the Japanese Ministry of Education, Culture, Sports, and Technology (MEXT) (No. 24114001, 24114008, and 18H05494), an Impulsing Paradigm Change through Disruptive Technologies (ImPACT) from the Japan Science and Technology Agency (JST), and partially supported by the Asahi Glass Foundation to KI. K.I. thanks the Finnish Funding Agency for Innovation (TEKES) for the support of the Finland Distinguished Professor (FiDiPro) Program "Advanced approaches for enzymatic biomass utilisation and modification (BioAD)".

* ezaki@jamology.rcast.u-tokyo.ac.jp
† aquarius@mail.ecc.u-tokyo.ac.jp

[1] R. D. Smiley, G. G. Hammes, Chem. Rev. **106**, 3080–3094 (2006).
[2] A. N. Kapanidis, T. Strick, Trends Biochem. Sci. **34**, 234–243 (2009).
[3] M. E. Himmel et al., Science. **315**, 804–807 (2007).
[4] K. Igarashi et al., J. Biol. Chem. **284**, 36186–36190 (2009).
[5] K. Igarashi, M. Wada, R. Hori, M. Samejima, FEBS J. **273**, 2869–2878 (2006).
[6] K. Igarashi et al., Science. **333**, 1279–1282 (2011).
[7] M. Mitchell, *Complexity: A Guided Tour* (Oxford University Press, New York, 2009).
[8] D. Chowdhury, L. Santen, A. Schadschneider, Phys. Rep. **329**, 199–329 (2000).
[9] D. Helbing, Rev. Mod. Phys. **73**, 1067–1141 (2001).
[10] D. Helbing, I. J. Farkas, T. Vicsek, Nature. **407**, 487–490 (2000).
[11] C. Burstedde, K. Klauck, A. Schadschneider, J. Zittartz, Physica A. **295**, 507–525 (2001).
[12] K. Nishinari, Y. Okada, A. Schadschneider, D. Chowdhury, Phys. Rev. Lett. **95**, 118101 (2005).
[13] A. B. Kolomeisky, M. E. Fisher, Annu. Rev. Phys. Chem. **58**, 675–695 (2007).
[14] C. Leduc et al., Proc. Natl. Acad. Sci. U. S. A. **109**, 6100 (2012).
[15] A.C. Warden, B. A. Little, V. S. Haritos, Biotechnol. Biofuels **4**, 39 (2011).
[16] A. Asztalos, et. al., Biotechnol. Biofuels **5**, 55 (2012).
[17] B. Z. Shang, R. Chang, J. W. Chu, J. Bio. Chem. **288**, 29081–29089 (2013).
[18] J. Lehtiö et al., Proc. Natl. Acad. Sci. U. S. A. **100**, 484–489 (2003).
[19] P. M. Abuja et al., Eur. Biophys. J. **15**, 339–342 (1988).
[20] Y. Shibafuji et al., J. Biol. Chem. **289**, 14056–14065 (2014).
[21] J. Ståhlberg, G. Johansson, G. Pettersson, Nat. Biotechnol. **9**, 286–290 (1991).
[22] Y. Nishiyama, J. Sugiyama, H. Chanzy, P. Langan, J. Am. Chem. Soc. **125**, 14300–14306 (2003).
[23] N. Sugimoto, K. Igarashi, M. Wada, M. Samejima, Langmuir. **28**, 14323–14329 (2012).
[24] M. Hall, J. Rubin, S.H. Behrens, A.S. Bommarius, J. Biotechnol. **155**, 370–376 (2011).
[25] P. Bansal, M. Hall, M. J. Realff, J. H. Lee, A. S. Bommarius, Biotechnol. Adv. **27**, 833–848 (2009).
[26] S. J. Park, et al., J. Phys. Chem. Lett. **8**, 3152–3158 (2017).
[27] J. Araki, M. Wada, S. Kuga, T. Okano, Colloids Surf. A **142**, 75–82 (1998).
[28] I. M. Kuznetsova, K. K. Turoverov, V. N. Uversky, Int. J. Mol. Sci. **15**, 23090–23140(2014)
[29] K. Igarashi, M. Wada, M. Samejima, FEBS J. **274**, 1785–1792 (2007).