Direct Evidence for Aligned Binding of Cellulase Enzymes to Cellulose Surfaces

Kayla Sprenger,[1] Steven J. Roeters,[2] Sergio Mauri,[3] Rolf Mertig,[2] Yoshiharu Nishiyama,[4] Jim Pfaendtner,[5,*]   Tobias Weidner[2,5,*]

[1]Department of Chemical and Biological Engineering, University of Colorado Boulder, Boulder, CO 80309, United States

[2]Department of Chemistry, Aarhus University, 8000 Aarhus C, Denmark

[3]Max Planck Institute for Polymer Research, 55128 Mainz, Germany

[4]Univ. Grenoble Alpes, CNRS, CERMAV, Grenoble 38000, France

[5]Department of Chemical Engineering, University of Washington, Seattle, WA 98195, United States

Supporting Information Placeholder

ABSTRACT: The conversion of biomass into green fuels and chemicals is of great societal interest. Engineers have been designing new cellulase enzymes for the breakdown of otherwise insoluble cellulose materials. A barrier to the rational design of new enzymes has been our lack of a molecular picture of how cellulase binding occurs. A critical factor is the attachment via the enzyme's carbohydrate binding module (CBM). To elucidate the structural and mechanistic details of cellulase adsorption, we have combined experimental data from sum frequency generation spectroscopy with molecular dynamics simulations to probe the equilibrium structure and surface alignment of a 14-residue peptide mimicking the CBM. The data show that binding is driven by hydrogen bonding and that tyrosine side chains within the CBM align the cellulase with the registry of the cellulose surface. Such an alignment is favorable for the translocation and effective cellulose breakdown and is therefore likely an important parameter for the design of novel enzymes.

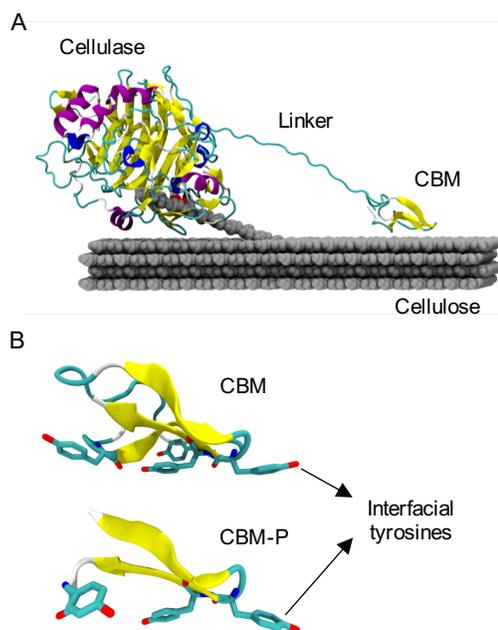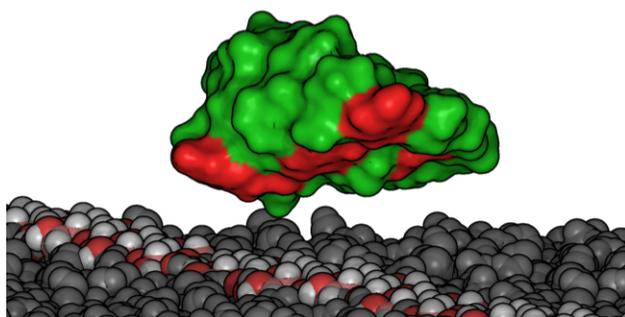## Table of Contents artwork

Figure 1. Model peptide for cellulase binding. (A) Cellulase enzymes bind cellulose surfaces through their carbohydrate-binding modules and break down cellulose. (B) Tyrosine side chains (red) are critical for CBM binding. The CBM-P peptide mimics the tyrosine motif within the binding surface of the CBM. Images were generated in Visual Molecular Dynamics (VMD[1]) using PDB codes 6GRN[2] for cellulase and 2MWK[3] for the CBM. In (B), hydrogens are not pictured for clarity.

The enzymatic conversion of insoluble polysaccharides at the solid-liquid interface is critically important to designing new systems to convert renewable resources into green fuels and chemicals. Despite this fact, we often lack a molecularly detailed understanding of these interfaces. Such lack of knowledge is

a current barrier to the rational design of improved biocatalytic systems.

Enzymatic biomass conversion is rate-controlled by interfacial phenomena. In the case of biomass conversion, the insoluble cellulose microfibril creates diffusion limitations, as the reaction requires collision and proper alignment/binding. As shown in Fig. 1, cellulase enzymes have evolved precise carbotion of active enzymatic cores near the interface.[4,5]

Some CBMs may also increase rates by beneficially altering the microfibril structure.[6,7]Tight binding of the active core to the surface plays an important role in the enzymatic breakdown of insoluble cellulose. A substantial amount of literature deals with the mechanism by which the CBM recognizes and binds cellulose. Mutagenesis studies have been used to determine key side chains for CBM cellulose affinity and enzymatic activity.[8,9] Efforts have also been made to investigate the structure of the CBM with X-ray diffraction[10] and NMR[3,11,12]. These studies show that the CBM has a flat side containing four tyrosine residues out of which three are essential for surface binding (Fig. 1B). Simulations suggest that the CBM prefers to bind to the hydrophobic face of a cellulose microfibril and that binding to the cellulose's hydrophilic face is transient.[13–18]
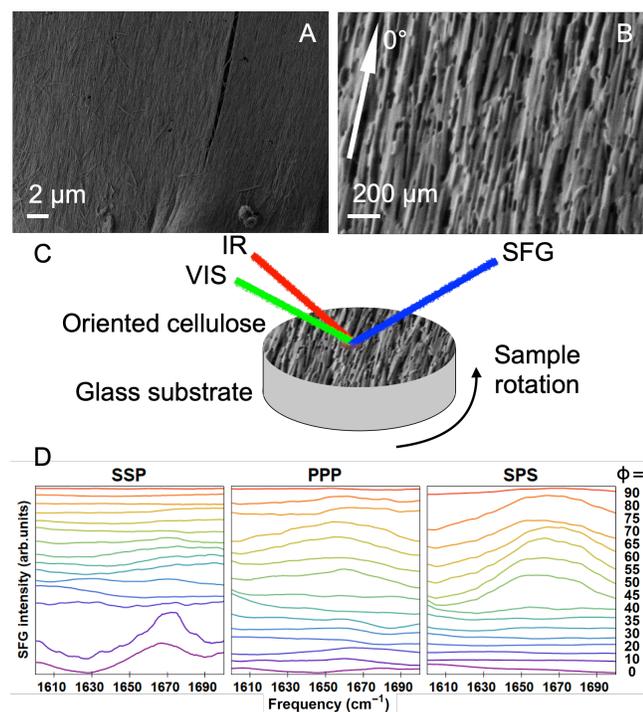

Figure 2. SFG analysis of the CBM-P alignment on cellulose. (A,B) SEM image of the aligned cellulose fibers and definition of the azimuthal angles. (C) Schematic of the SFG experiment with the rotating sample stage. (D) SFG spectra in the amide I region recorded for the CBM-P adsorbed to the cellulose fibers. The spectra show a distinct variance with azimuthal angles, direct evidence for the alignment of the CBM-P relative to the cellulose surface structure.

The importance of the CBM in cellulase action is undisputed. However, recent work has been performed against a backdrop of uncertainty in the structure and orientation of the CBM at the surface/protein interface.

To reduce complexity and focus on the binding of the CBM to the cellulose surface, we have constructed a model peptide, which mimics the binding motive of the CBM of the enzyme cellulase 7A from *T. Reesei*[3]. Fig.1B shows how the flat binding structure of the CBM with four aligned tyrosine rings is naturally suggestive of a model fragment representing the binding surface of the CBM. This 14-residue fragment from region 222-234 of the CBM with an additional tyrosine at the C-terminus (CBM-P, NH2-TCQVLNPYYSQCLY-COOH) is small enough to be synthesized with a peptide synthesizer, thus enabling extensive experimentation.

Surface-specific sum frequency generation (SFG) spectroscopy and atomistic simulations have recently been combined into a versatile tool to determine protein structure at interfaces[19–25]. We use this strategy here to probe the equilibrium structure and binding orientation of the CBM-P cellulase peptide mimic on cellulose. As a model cellulose surface, we deposited highly oriented cellulose fibers on glass substrates[26]. The scanning electron microscopy (SEM) images in Figs. 2A and 2B illustrate the sample geometry. The alignment of the fibers allowed us to probe the alignment of the CBM-P relative to the cellulose registry by recording spectra at different azimuthal sample orientations. The samples were immersed in a CBM-P solution overnight, rinsed, and then introduced to the SFG sample stage (Fig. 2C).

SFG spectra were collected in the amide I region to follow the protein backbone structure at the cellulose surface. Fig. 2D shows SFG spectra recorded as a function of the azimuthal angle in ssp (s-polarized SFG, s-polarized visible, and p-polarized infrared), ppp and sps polarization. All three sets show a pronounced variance with azimuthal angle. At angles close to the fiber direction, there is a broad amide I signal visible in ssp polarization. A spectral dip near 1630 cm^{-1} is accompanied by a peak centered around 1670 cm^{-1}. These modes can be assigned to β-sheet type structures[27,28]. The ppp and sps spectra show no appreciable signal intensity along this direction. At higher angles, more perpendicular to the fiber direction, the ssp intensity drops while ppp and sps spectra grow in signal level. For both polarizations, the spectra are broad and centered just below 1670 cm^{-1}, which is typically assigned to turn structures or β-sheet[29].

Typically, SFG spectra are invariant with the azimuthal sample orientation because of the statistical orientation of molecules at surfaces with respect to the azimuthal angle[30]. The angle variance of the spectra provides direct evidence for a preferential orientation of the CBM-P peptides relative to the cellulose surface structure.

To determine the preferential orientation and conformation of the CBM-P on the cellulose surface, we performed all-atom molecular dynamics (MD) simulations of the CBM-P/cellulose Iβ interface, which we can combine with the SFG data. Classical MD simulations were first performed of the CBM-P bound in four different orientations (rotated 90° apart; Fig. 3A; top left) to the cellulose surface, initiated from the full β-like conformation of the crystal structure. After the 50 ns production simulations (see SI for details), two structures were found to have maintained the full β-like conformation, whereas the other two structures (in purple and blue) adopted a more relaxed β-like conformation (Fig. 3A; bottom left). An energetic analysis of the CBM-P/cellulose interface shows the two structures that adopted the relaxed β-like conformation were able to form more favorable van der Waals interactions with the surface (Fig. 3A; middle).

Furthermore, one structure (in purple) was found to have re-oriented itself at approximately 45° to the cellulose chains, allowing it to achieve slightly more favorable van der Waals interactions with the surface than the other relaxed β-like conformation (in blue). The structure in purple was also found to have formed the most favorable electrostatic interactions with the surface (Fig. 3A; right). While this was the case from the start of the simulation, the electrostatic interactions remained favorable even after the peptide re-oriented itself against the cellulose chains, and perhaps even facilitated this re-orientation given the high energetic barriers typically present in classical MD simulations of surface-adsorbed proteins[31].

Fig. 3B-C show representative snapshots from the simulation of the relaxed β-like conformation of CBM-P in purple in Fig. 3A. The snapshots highlight the observed 45° orientation of the CBM-P's backbone atoms to the cellulose chains, permitting favorable surface alignment of its tyrosine residues with cellulose glucose units. To quantify differences in structure and dynamics between the full and relaxed β-like conformations, we performed a hydrogen-bond analysis in VMD over the last 20 ns of a simulation trajectory of each conformation (relaxed β-like structure: Fig. 3A, purple; full β-like structure: Fig. 3A, green). Compared to the full β-like conformation, we observe fewer *intra*molecular hydrogen bonds in the relaxed β-like conformation (Fig. S1 A-B). However, due to the in-creased flexibility, we observe the relaxed β-like conformation is more often able to form at least three simultaneous hydrogen bonds with the surface (Fig. S1 C-D).

As a further test of our observations on the preferred structure and orientation of cellulose-adsorbed CBM-P, we performed additional MD simulations using the parallel tempering metadynamics in the well-tempered ensemble (PTMetaD[32]-WTE[33])[34] enhanced sampling method (see SI for details). Subsequently, we clustered the surface-bound structures of the CBM-P together with the cellulose surface. Fig. 3D shows that the second most densely populated cluster consists of states in which the CBM-P adopted a relaxed β-like conformation at an approximately 52° orientation to the cellulose chains, in line with the findings from our classical MD simulations (see SI for further discussion).

For a test of the simulation results we calculated SFG spectra for the relaxed β-like conformation for the different relative orientations of the plane of laser beam incidence and cellulose fibers. The computed spectra in Figure 3E capture the resonance positions, peak width and, most importantly, the azimuthal orientation dependence of the experimental data very well. It should be noted that we calculate the spectra without non-resonant contributions and therefore the phase of the spectral features is not included in the calculated spectra. This is particularly noticeable for the ssp spectra, where the experimental data consists of two features with opposite phases, which results in a dip and a peak in the spectra. In addition, it is difficult to rotate the sample precisely on-axis and slightly different spots on the sample are probed as the sample is rotated. Minor variances in the spectra due to sample inhomogeneity can therefore not be excluded. The qualitative match with experiments strongly supports the simulation results. Evidently, CBM-P aligns with the registry of the cellulose surface structure.
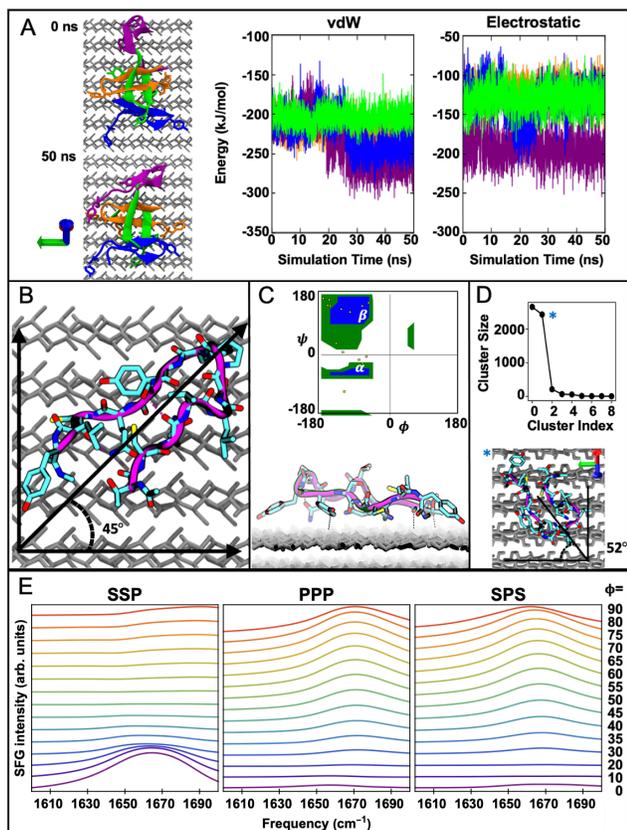
Figure 3. Theoretical results of the alignment and conformation of the CBM-P on cellulose. (A) van der Waals and electrostatic interfacial energy as a function of relative CBM-P orientation. (B) The most stable orientation predicted by classical MD aligns at 45° to cellulose chains. (C) A relaxed β-like structure is predicted to be the most stable surface-bound structure of CBM-P, as it allows for energetically favorable alignment of the CBM-P backbone and side chains with the surface. A Ramachandran plot of the figure below and in (B) highlights the β-sheet character of residues in surface-bound CBM-P. (D) A relaxed β-like structure aligned at 52° to cellulose chains is the second most densely populated cluster from PTMetaD-WTE simulations. Images in (A-D) were generated in VMD[1]. (E) Theoretical SFG spectra calculated for the CBM-P pose shown in (B, C) as a function of the azimuthal angle in ssp, ppp and sps polarization. The spectra are in good agreement with the experimental data.

SFG and MD draw a detailed picture of CBM binding to cellulose surfaces. Binding is driven by hydrogen bonding between protein side chains and backbone sites with the polysaccharide surface. Evidently, the binding tyrosine side chains align the CBM with the surface registry of the cellulose glucose units. Enzymatic breakdown of cellulose requires cellulase to move across the surface for even and efficient operation. At high cellulase surface coverages, random motion across the surface would likely lead to collisions and steric hindrance. It is conceivable that binding aligned with the macroscopic registry of the cellulose crystal will promote aligned motion, which would allow unhindered, concerted surface motion of cellulase, even at high surface coverages. Future designs of optimized cellulase mutants for tailored green energy applications should therefore focus not only on optimizing attachment and binding. Cellulase alignment and directed motion will likely be important design parameters for enzymes engineered for biomass conversion.

ASSOCIATED CONTENT

Supporting Information. The Supporting Information is available free of charge on the ACS Publications website:

Details of sample preparation, SFG spectroscopy, MD simulations and theoretical spectra calculations.

AUTHOR INFORMATION

Corresponding Authors:

Jim Pfaendtner: jpfaendt@uw.edu

Tobias Weidner: weidner@chem.au.dk

REFERENCES

(1)   Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.

(2)   Hamark, C.; Pendrill, R.; Landström, J.; Dotson Fagerström, A.; Sandgren, M.; Ståhlberg, J.; Widmalm, G. Enantioselective Binding of Propranolol and Analogues Thereof to Cellobiohydrolase Cel7A. *Chem. - A Eur. J.* **2018**, *24*, 17975–17985.

(3)   Happs, R. M.; Guan, X.; Resch, M. G.; Davis, M. F.; Beckham, G. T.; Tan, Z.; Crowley, M. F. O-Glycosylation Effects on Family 1 Carbohydrate-Binding Module Solution Structures. *FEBS J.* **2015**, *282*, 4341–4356.

(4)   Kim, T. W.; Chokhawala, H. A.; Nadler, D.; Blanch, H. W.; Clark, D. S. Binding Modules Alter the Activity of Chimeric Cellulases: Effects of Biomass Pretreatment and Enzyme Source. *Biotechnol. Bioeng.* **2010**, *107*, 601–611.

(5)   Takashima, S.; Ohno, M.; Hidaka, M.; Nakamura, A.; Masaki, H.; Uozumi, T. Correlation between Cellulose Binding and Activity of Cellulose-Binding Domain Mutants of Humicola Grisea Cellobiohydrolase 1. *FEBS Lett.* **2007**, *581*, 5891–5896.

(6)   Arantes, V.; Saddler, J. N. Access to Cellulose Limits the Efficiency of Enzymatic Hydrolysis: The Role of Amorphogenesis. *Biotechnol. Biofuels* **2010**, *3:4*, 1–11.

(7)   Hall, M.; Bansal, P.; Lee, J. H.; Realff, M. J.; Bommarius, A. S. Biological Pretreatment of Cellulose: Enhancing Enzymatic Hydrolysis Rate Using Cellulose-Binding Domains from

Cellulases. *Bioresour. Technol.* **2011**, *102*, 2910–2915.

(8) Urbanowicz, B. R.; Catalá, C.; Irwin, D.; Wilson, D. B.; Ripoll, D. R.; Rose, J. K. C. A Tomato Endo-β-1,4-Glucanase, SlCel9C1, Represents a Distinct Subclass with a New Family of Carbohydrate Binding Modules (CBM49). *J. Biol. Chem.* **2007**, *282*, 12066–12074.

(9) Najmudin, S.; Guerreiro, C. I. P. D.; Carvalho, A. L.; Prates, J. A. M.; Correia, M. A. S.; Alves, V. D.; Ferreira, L. M. A.; Romão, M. J.; Gilbert, H. J.; Bolam, D. N.; Fontes, C. M. G. A. Xyloglucan Is Recognized by Carbohydrate-Binding Modules That Interact with β-Glucan Chains. *J. Biol. Chem.* **2006**, *281*, 8815–8828.

(10) Jamal, S.; Nurizzo, D.; Boraston, A. B.; Davies, G. J. X-Ray Crystal Structure of a Non-Crystalline Cellulose-Specific Carbohydrate-Binding Module: CBM28. *J. Mol. Biol.* **2004**, *339*, 253–258.

(11) Raghothama, S.; Simpson, P. J.; Szabó, L.; Nagy, T.; Gilbert, H. J.; Williamson, M. P. Solution Structure of the CBM10 Cellulose Binding Module from Pseudomonas Xylanase A. *Biochemistry* **2000**, *39*, 978–984.

(12) Courtade, G.; Forsberg, Z.; Heggset, E. B.; Eijsink, V. G. H.; Aachmann, F. L. The Carbohydrate-Binding Module and Linker of a Modular Lytic Polysaccharide Monooxygenase Promote Localized Cellulose Oxidation. *J. Biol. Chem.* **2018**, *293*, 13006–13015.

(13) Nimlos, M. R.; Matthews, J. F.; Crowley, M. F.; Walker, R. C.; Chukkapalli, G.; Brady, J. W.; Adney, W. S.; Cleary, J. M.; Zhong, L.; Himmel, M. E. Molecular Modeling Suggests Induced Fit of Family I Carbohydrate-Binding Modules with a Broken-Chain Cellulose Surface. *Protein Eng. Des. Sel.* **2007**, *20*, 179–187.

(14) Bu, L.; Beckham, G. T.; Crowley, M. F.; Chang, C. H.; Matthews, J. F.; Bomble, Y. J.; Adney, W. S.; Himmel, M. E.; Nimlos, M. R. The Energy Landscape for the Interaction of the Family 1 Carbohydrate-Binding Module and the Cellulose Surface Is Altered by Hydrolyzed Glycosidic Bonds. *J. Phys. Chem. B* **2009**, *113*, 10994–11002.

(15) Beckham, G. T.; Matthews, J. F.; Bomble, Y. J.; Bu, L.; Adney, W. S.; Himmel, M. E.; Nimlos, M. R.; Crowley, M. F. Identification of Amino Acids Responsible for Processivity in a Family 1 Carbohydrate-Binding Module from a Fungal Cellulase. *J. Phys. Chem. B* **2010**, *114*, 1447–1453.

(16) Beckham, G. T.; Bomble, Y. J.; Bayer, E. A.; Himmel, M. E.; Crowley, M. F. Applications of Computational Science for Understanding Enzymatic Deconstruction of Cellulose. *Curr. Opin. Biotechnol.* **2011**, *22*, 231–238.

(17) Shiiba, H.; Hayashi, S.; Yui, T. Molecular Simulation Study with Complex Models of the Carbohydrate Binding Module of Cel6A and the Cellulose Iα Crystal. *Cellulose* **2012**, *19*, 635–645.

(18) Taylor, C. B.; Talib, M. F.; McCabe, C.; Bu, L.; Adney, W. S.; Himmel, M. E.; Crowley, M. F.; Beckham, G. T. Computational Investigation of Glycosylation Effects on a Family 1 Carbohydrate-Binding Module. *J. Biol. Chem.* **2012**, *287*, 3147–3155.

(19) Alamdari, S.; Roeters, S. J.; Golbek, T. W.; Schmüser, L.; Weidner, T.; Pfaendtner, J. Orientation and Conformation of Proteins at the Air-Water Interface Determined from Integrative Molecular Dynamics Simulations and Sum

(20) Frequency Generation Spectroscopy. *Langmuir* **2020**, *36*, 11855–11865.

(20) Lutz, H.; Jaeger, V.; Schmüser, L.; Bonn, M.; Pfaendtner, J.; Weidner, T. The Structure of the Diatom Silaffin Peptide R5 within Freestanding Two-Dimensional Biosilica Sheets. *Angew. Chemie - Int. Ed.* **2017**, *56*, 8277–8280.

(21) Donovan, M. A.; Lutz, H.; Yimer, Y. Y.; Pfaendtner, J.; Bonn, M.; Weidner, T. LK Peptide Side Chain Dynamics at Interfaces Are Independent of Secondary Structure. *Phys. Chem. Chem. Phys.* **2017**, *19*, 28507–28511.

(22) Verreault, D.; Alamdari, S.; Roeters, S. J.; Pandey, R.; Pfaendtner, J.; Weidner, T. Ice-Binding Site of Surface-Bound Type III Antifreeze Protein Partially Decoupled from Water. *Phys. Chem. Chem. Phys.* **2018**, *20*, 26926–26933.

(23) Donovan, M. A.; Yimer, Y. Y.; Pfaendtner, J.; Backus, E. H. G.; Bonn, M.; Weidner, T. Ultrafast Reorientational Dynamics of Leucine at the Air-Water Interface. *J. Am. Chem. Soc.* **2016**, *138*, 5226–5229.

(24) Lutz, H.; Jaeger, V.; Berger, R.; Bonn, M.; Pfaendtner, J.; Weidner, T. Biomimetic Growth of Ultrathin Silica Sheets Using Artificial Amphiphilic Peptides. *Adv. Mater. Interfaces* **2015**, *2*, 1500282.

(25) Hosseinpour, S.; Roeters, S. J.; Bonn, M.; Peukert, W.; Woutersen, S.; Weidner, T. Structure and Dynamics of Interfacial Peptides and Proteins from Vibrational Sum-Frequency Generation Spectroscopy. *Chem. Rev.* **2020**, *120*, 3420–3465.

(26) Yoshiharu, N.; Shigenori, K.; Masahisa, W.; Takeshi, O. Cellulose Microcrystal Film of High Uniaxial Orientation. *Macromolecules* **1997**, *30*, 6395–6397.

(27) Nguyen, K. T.; King, J. T.; Chen, Z. Orientation Determination of Interfacial β-Sheet Structures in Situ. *J. Phys. Chem. B* **2010**, *114*, 8291–8300.

(28) Weidner, T.; Apte, J. S.; Gamble, L. J.; Castner, D. G. Probing the Orientation and Conformation of A-Helix and β-Strand Model Peptides on Self-Assembled Monolayers Using Sum Frequency Generation and NEXAFS Spectroscopy. *Langmuir* **2010**, *26*, 3433–3440.

(29) Singh, B. R. Basic Aspects of the Technique and Applications of Infrared Spectroscopy of Peptides and Proteins. *ACS Symp. Ser.* **2000**, *750*, 2–37.

(30) Shen, Y. R. *Fundamentals of Sum-Frequency Spectroscopy*; Cambridge University Press, 2016.

(31) Prakash, A.; Sprenger, K. G.; Pfaendtner, J. Essential Slow Degrees of Freedom in Protein-Surface Simulations: A Metadynamics Investigation. *Biochem. Biophys. Res. Commun.* **2018**, *498*, 274–281.

(32) Bussi, G.; Gervasio, F. L.; Laio, A.; Parrinello, M. Free-Energy Landscape for Hairpin Folding from Combined Parallel Tempering and Metadynamics. *J. Am. Chem. Soc.* **2006**, *128*, 13435–13441.

(33) Bonomi, M.; Parrinello, M. Enhanced Sampling in the Well-Tempered Ensemble. *Phys. Rev. Lett.* **2010**, *104*, 190601.

(34) Deighan, M.; Bonomi, M.; Pfaendtner, J. Efficient Simulation of Explicitly Solvated Proteins in the Well-Tempered Ensemble. *J. Chem. Theory Comput.* **2012**, *8*, 2189–2192.