THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

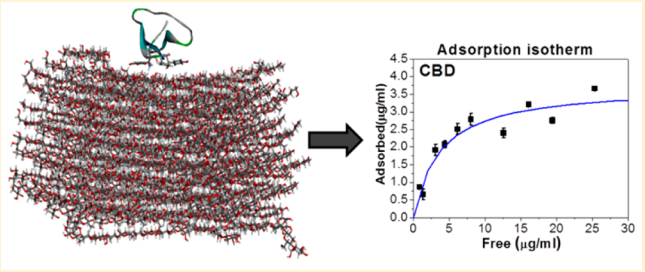
Design of Compact Biomimetic Cellulose Binding Peptides as Carriers for Cellulose Catalytic Degradation

Netaly Khazanov,[†] Taly Iline-Vul,[†] Efrat Noy, Gil Goobes,* and Hanoch Senderowitz*

Department of Chemistry, Bar Ilan University, Ramat Gan 5290002, Israel

Supporting Information

ABSTRACT: The conversion of biomass into biofuels can reduce the strategic vulnerability of petroleum-based systems and at the same time have a positive effect on global climate issues. Lignocellulose is the cheapest and most abundant source of biomass and consequently has been widely considered as a source for liquid fuel. However, despite ongoing efforts, cellulosic biofuels are still far from commercial realization, one of the major bottlenecks being the hydrolysis of cellulose into simpler sugars. Inspired by the structural and functional modularity of cellulases used by many organisms for the breakdown of cellulose, we propose to mimic the cellulose binding domain (CBD) and the catalytic domain of these proteins by small molecular entities. Multiple copies of these mimics could subsequently be tethered together to enhance hydrolytic activity. In this work, we take the first step toward achieving this goal by applying computational approaches to the design of efficient, cost-effective mimetics of the CBD. The design is based on low molecular weight peptides that are amenable to large-scale production. We provide an optimized design of four short (i.e., ∼18 residues) peptide mimetics based on the three-dimensional structure of a known CBD and demonstrate that some of these peptides bind cellulose as well as or better than the full CBD. The structures of these peptides were studied by circular dichroism and their interactions with cellulose by solid phase NMR. Finally, we present a computational strategy for predicting CBD/peptide−cellulose binding free energies and demonstrate its ability to provide values in good agreement with experimental data. Using this computational model, we have also studied the dissociation pathway of the CBDs/peptides from the surface of cellulose.



1. INTRODUCTION

The conversion of biomass into biofuels can reduce the strategic vulnerability of petroleum-based transportation systems and at the same time have a positive effect on the global climate via the reduction of greenhouse gas release into the atmosphere.[1,2] Consequently, many efforts are devoted to extensive research in the field of biofuels as a sustainable and renewable energy source.[3] In particular, lignocellulose (or cellulose) is the cheapest and most abundant source of biomass and consequently has been widely considered as a source for liquid fuel for transportation.[1,4,5]

Lignocellulose can be converted into liquid fuels by three primary routes, namely, the production of syngas by gasification, the production of bio-oil by pyrolysis, or the liquefaction and the hydrolysis of biomass to sugar monomer units followed by fermentation.[1] Multiple efforts focus on the third strategy by utilizing a variety of catalysts both chemical (e.g., liquid and solid acids[2] and nanoparticles[6]) and biological (e.g., enzyme cocktails or cellulosomes secreted/produced by microorganisms; the latter two have recently been shown to work synergistically[5]) for the efficient breakdown of cellulose. Advanced molecular biology approaches are finding their way into this field to achieve cost-effective production of sugars or direct conversion of cellulose into biofuels, such as ethanol or butanol.[7,8] However, despite ongoing efforts along all strategies, cellulosic biofuels are still far from commercial realization.[9−11] A main challenge associated with the efficient conversion of biomass into biofuel common to both chemical and biological catalysis is the highly ordered crystalline structure of cellulose, making it resistant to hydrolysis and often requiring expensive pretreatment methods.[12] Other challenges more unique to biological catalysis involve enzyme inhibition, enzyme costs, and enzyme recycling.[5,13]

An appealing alternative to some of the current strategies can be bioinspired robust and economic catalysts designed using a combination of computational and experimental tools. These may match or even improve upon the catalytic efficiency of the natural cellulose degradation machineries.[7,14] One class of enzymes used by many organisms to break down cellulose are cellulases.[15] These are multidomain proteins consisting of cellulose binding domains (CBD) to bind the polysaccharide and catalytic domains to degrade it. Throughout this work we use the term CBD, although currently, the more common term is carbohydrate binding module (CBM). This latter term is also used by the CAZy database.[16] Inspired by the structural and functional modularity of cellulases, we envision a strategy by

Received: November 11, 2015
Revised: December 22, 2015
Published: December 22, 2015

DOI: 10.1021/acs.jpcb.5b11050
J. Phys. Chem. B 2016, 120, 309−319

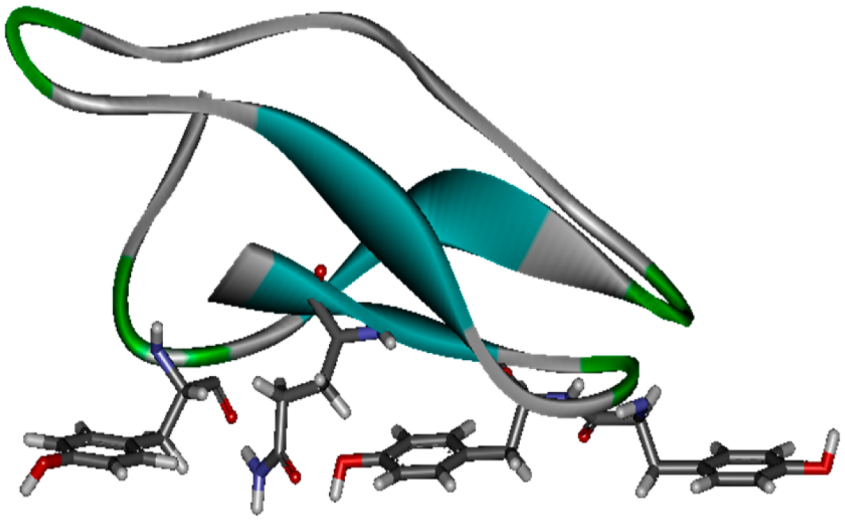


Figure 1. Structure of the CBD domain from *T. reesei* (PDB ID: 1CBH). The domain is presented as a solid ribbon and color coded according to its secondary structure elements (β-sheets in cyan, turns in green). Residues comprising the planar strip are shown as sticks.

which their CBD and their catalytic domain are mimicked by preserving their most essential functional features, yet on small molecular scaffolds. In this design, the role of the CBD mimetic would be to bind cellulose, whereas the role of the catalytic domain mimetic would be to hydrolyze the glycosidic bond between the sugar units. Furthermore, hydrolytic activity could be enhanced by tethering together multiple copies of each domain and/or additional synthetic catalytic centers. A correlation between enhanced cellulose binding and improved catalytic activity for endoglucanase II from *Trichoderma reesei* has been established and lends credit to this design approach.[7] Recently, studies of the carbohydrate binding module structure and dynamics were reported, providing additional insights into interactions of aromatic residues with sugar rings.[17−19]

In this work, we take the first step toward the construction of synthetic cellulose degradation catalysts by applying computational approaches to the design of efficient, cost-effective mimics of the cellulose binding domain (CBD). The design is based on robust, low molecular weight peptides that are amenable to large scale production. We provide an optimized design of four short (i.e., ∼18 residues) peptide mimetics of the CBD based on and inspired by the known three-dimensional (3D) structure of *Trichoderma reesei* CBD. We demonstrate that some of these peptides bind cellulose equally or better than the full CBD. The structures of these peptides were also studies by circular dichroism (CD), and the effect of a selected peptide on the structure of cellulose was studied by solid phase NMR. The interactions of the peptides with cellulose were confirmed experimentally using adsorption isotherm measurements. Finally, we present a computational method for predicting CBD/peptide−cellulose binding free energies and demonstrate its ability to provide values in good agreement with experimental data. This method was also used to study the dissociation pathway of the CBDs/peptides from the surface of cellulose, highlighting the interactions responsible for binding. Our approach could be viewed as an extension to the protein design paradigm, which has been successfully applied for the design of new biocatalysts[20−22] and as a variant of the designer cellulosome approach[12,23] but differs from the two in being more "minimalistic" in its end outcome. This "minimalism" has the potential of cost reduction, which is highly important to meet the challenge of biomass conversion. To the best of our knowledge, this work represents the first successful *rational* design of a peptide-based carbohydrate binding mimetic. Few cellulose binding peptides were identified by phase display technologies.[24−26]

2. RESULTS AND DISCUSSION

To exploit the full potential of molecular modeling approaches in the design of new CBD mimics, one needs to be able to (1) computationally design a series of relevant mimics and (2) predict their cellulose binding affinity prior to their actual synthesis. This second task requires in turn the development of a reliable three-dimensional (3D) cellulose model and the development of a computational method capable of calculating peptide-cellulose binding free energies.

2.1. Design of Mimetic Peptides Based on CBD. CBD mimetics were designed based on the high resolution crystal structures of CBDs available in the PDB.[27] A close inspection of these structures (e.g., Figure 1, the C-terminal 36-residues CBD from *T. reesei* cellobiohydrolase I; PDB ID: 1CBH[28]) reveals that cellulose binding is primarily mediated through a planar strip of aromatic (Y5, Y31, Y32) and nonaromatic (N29, Q34) residues (throughout this work we use the numbering system from the CBD of *T. reesei*) most probably through $\sigma-\pi$ interactions and hydrogen bonds. Designing smaller scaffolds with a similar architecture should therefore preserve the cellulose binding capabilities of the CBD in the new constructs. Following this rationale, we designed 18-mer peptides based on the 36-residue CBD from *T. reesei*.

Two design cycles of CBD mimetics were performed with the second cycle taking advantage of the experimental results obtained for first cycle constructs. In cycle **1**, the original 36-mer CBD from *T. reesei* was shortened by removing residues 9/10−24 and 1−3 and by connecting residue 36 to residue 4 while maintaining the disulfide bond between residues 8 and 25. Our goal was to retain the aromatic strip close to its original 3D conformation albeit on an 18 residue scaffold (Figure 2).

Six constructs, presented in Table 1, were designed that differed in their mutations pattern: L36P introduces a β-turn inducing proline into the newly formed β-turn; H4G introduces a β-turn enhancing glycine into the newly formed β-turn; Q7T and C35T feature mutations to T, which has a higher

DOI: 10.1021/acs.jpcb.5b11050
J. Phys. Chem. B 2016, 120, 309−319

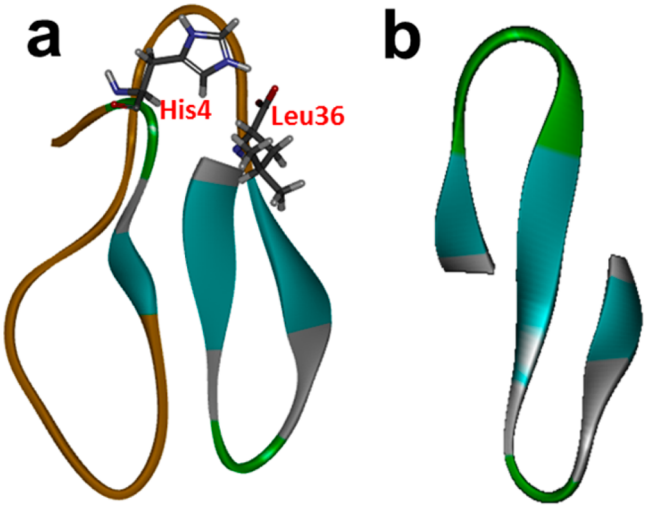


Figure 2. (a) Solid ribbon representation of CBD of *T. reesei* CBHI (PDB ID: 1CBH). Residues that were removed are colored in orange on the ribbon representation. Residues 4 and 36 (stick representation) were linked to form the new peptides. (b) Solid ribbon representation of the 18 residue CBD.

propensity for β-strand conformation; Y5W replaces Y to W to increase the affinity to cellulose;[29] and insT35a inserts T to relieve steric tension in the newly formed β-turn (which is accompanied by removal of G9).

The six CBD mimetics were subjected to 20 ns molecular dynamics (MD) simulations. To assess the ability of the designed constructs to maintain the correct alignment of the aromatic rings, we used two complementary measures, namely, the root mean squared deviation (RMSD) and the root mean squared fluctuation (RMSF) calculated from the MD trajectories. RMSD measures the degree of deviation from a reference structure. Assuming that the reference structure has the aromatic moieties in the "correct" 3D conformation, low RMSD values indicate that the simulated construct maintains a similar conformation. RMSF profiles are calculated on a per-residue (or per-atom) basis and measure the degree of fluctuation of the residue (or atom) from an average structure in the course of the MD simulation. Because experimental and computational studies of protein folding and unfolding have established a connection between early unfolding events and atomic fluctuations,[30−32] low RMSF values are indicative of a more stable construct. Thus, low RMSD values coupled with low RMSF values are the hallmarks of a stable construct with a correct binding strip conformation.

The resulting RMSD and RMSF profiles are presented in Figures 3 and 4, respectively. Constructs 4 and 5 (construct 5

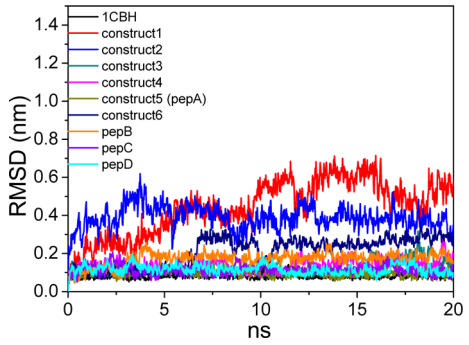


Figure 3. Backbone RMSD plots of the designed 18 residue peptides together with that of the original CBD (PDB ID: 1CBH, 36 residues) obtained from 20 ns MD simulations. Constructs 4, 5 (pepA), and 7 (pepB) from the first round and constructs 8 (pepC) and 9 (pepD) from the second round adopt similar backbone conformations to the original CBD conformation.

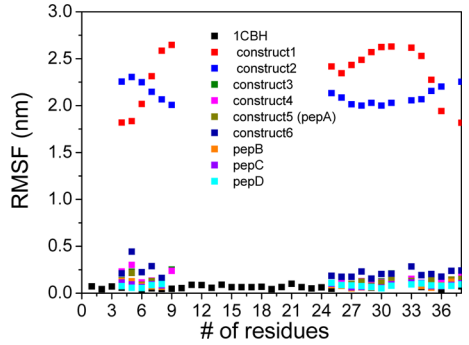


Figure 4. RMSF plots of the designed (18 residue) peptides together with that of the original CBD (PDB ID: 1CBH, 36 residues) calculated from 20 ns MD simulations. Constructs 4, 5 (pepA), and 7 (pepB) from the first round and constructs 8 (pepC) and 9 (pepD) from the second round are stabilized to the same degree as the original CBD. The noncontinuity of the plots for all peptides results from their shorter lengths relative to the full CBD.

hereafter referred to as pepA) demonstrated the smallest RMSD values and the least fluctuating RMSF profiles and were therefore selected for synthesis and functional assays. Because of the high degree of conservation of Q at position 7 in the CBHI family,[33] construct 4 was modified by introducing a T →

Table 1. Designed Constructs

construct	residues	mutations	sequence
		Cycle 1	
1	4−9, 25−36	L36P	CQVLNPYYSQCPHYGQCG
2	4−9, 25−36	H4G, L36P	CQVLNPYYSQCPGYGQCG
3	4−9, 25−36	H4G, Q7T, C35T, L36P	CQVLNPYYSQTPGYGTCG
4	4−9, 25−36	H4G, Y5W, Q7T, C35T, L36P	CQVLNPYYSQTPGWGTCG
5, pepA	4−8, 25−36	H4G, Q7T, C35T, insT35a, L36P	CQVLNPYYSQTTPGYGTC
6	4−8, 25−36	H4G, Y5W, Q7T, C35T, insT35a, L36P	CQVLNPYYSQTTPGWGTC
7, pepB	4−9, 25−36	H4G, Y5W, C35T, L36P	CQVLNPYYSQTPGWGQCG
		Cycle 2	
8, pepC	4−8, 25−36	H4G, Y5W, C35T, insT35a, L36P	CQVLNPYYSQTTPGWGQC
9, pepD	4−8, 25−36	H4G, Y5W, Y31W, C35T, insT35a, L36P	CQVLNPWYSQTTPGWGQC

DOI: 10.1021/acs.jpcb.5b11050
J. Phys. Chem. B 2016, 120, 309−319

Q substitution leading to construct 7 (hereafter referred to as pepB).

In cycle 2, two new peptide constructs (pepC and pepD; see Table 1) were designed based on the sequences of pepA (CQVLNPYYSQTTPGYGTC, Mw = 1995 g/mol) and pepB CQVLNPYYSQTPGWGQCG, Mw = 2001 g/mol) because both peptides demonstrated favorable properties in terms of binding to and coverage of cellulose (see below). PepC (CQVLNPYYSQTTPGWGQC, Mw = 2045 g/mol) is composed of residues 25−34 of the parent CBD connected to residues 5−8 through residues TTPG with residue 5 mutated to Trp and residue 7 not mutated as in pepB. PepD (CQVLNPWYSQTTPGWGQC, Mw = 2068 g/mol) is similar to pepC with Y31 mutated to W. As before, the Y to W replacement is expected to increase affinity to cellulose. RMSD and RMSF profiles of PepB, pepC, and pepD are presented in Figures 3 and 4, respectively, and are consistent with stable constructs with the "correct" binding strip conformation.

2.2. Computational Evaluation of the Binding Affinity of the Designed Peptides to Cellulose. Predicting the binding affinities of the designed peptides to cellulose required the construction of a reliable 3D model of the polysaccharide and the development of a protocol for binding free energy calculations. This protocol was first evaluated for a set of CBDs for which experimental data are available.

Construction of a 3D Model for Cellulose. A prerequisite of any computational protocol involving cellulose is the ability to reproduce its known structural features in the crystalline state. Several molecular dynamics (MD) simulations of cellulose were reported in the literature using different force fields, simulation protocols, and model systems.[34−36] Recently, Brady et al. have compared the performances of four different force fields using the same simulation protocol and concluded that all resulted in different structures.[34] In general, however, it seems that using the Glycam06 parameters provide the best fit to experimental data.[34]

In this work, we built a 3D crystalline model of cellulose Iβ (see Methods) and simulated it for 100 ns. An analysis of the resulting trajectory (Figure 5) showed that most of the glucose hydroxymethyl groups adopted a trans-gauche conformation except for those located at the surface of the structure, which is in agreement with the conformations found in cellulose crystal structures.[37] The resulting cellulose model was used for all subsequent simulations.

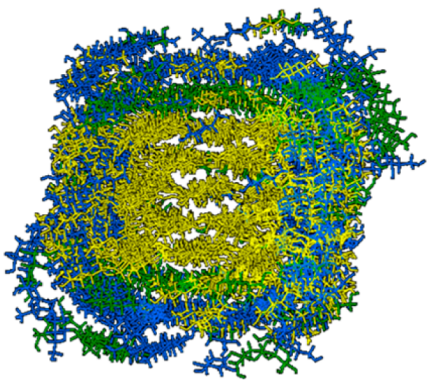


Figure 5. Last snapshot from the 100 ns simulation performed on the cellulose model. Colors indicate conformation of hydroxymethyl groups: Yellow-TG, Blue-GT, and Green-GG (T-trans, G-gauche).

Binding Free Energy Calculations. Peptide-cellulose binding free energies were evaluated through potential of mean force (PMF) calculations. These involved pulling simulations followed by umbrella sampling (see Methods). We first used the method to evaluate the cellulose binding affinity of the CBD from *T. reesei* (CBHI; PDB ID: 1CBH) and some of its mutants known to disrupt cellulose binding (Y32A, Y31A, Y5A).[38] The resulting PMF plots are presented in Figure 6, and

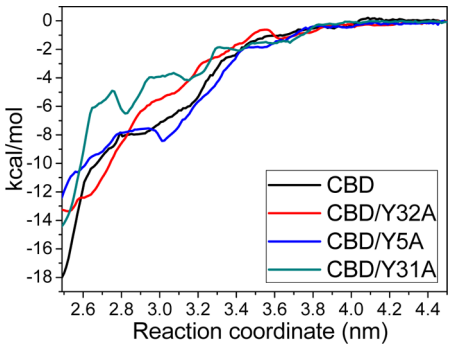


Figure 6. Potential of mean force (PMF) plots describing the binding of wild-type and mutant CBDs to cellulose. The reaction coordinate was taken as the center of mass (COM) distance between CBDs and cellulose.

the calculated $\Delta\Delta G$ values (relative to wt-CBD) are presented in Table 2. Encouragingly, calculated and measured experimental data follow the same trend, but at least for the Y31A mutant, the calculated number is higher. This discrepancy could result from multiple factors, both computational and experimental. Binding of CBD to cellulose is a complex process mediated by multiple factors (e.g., exact solution composition, cellulose concentration, cellulose surface size, and cellulose crystallinity), which may have not been fully accounted for in the computational protocol. Moreover, the starting point of the pulling simulations (e.g., the precise positioning of the CBD with respect to the cellulose) may also affect the results. Nevertheless, the calculations well-reproduced the strong effect on cellulose binding exerted by the three tyrosine residues (5, 31, and 32) located on the flat face of CBD as well as the experimentally observed stronger effect of Y5 and Y32.[39]

Table 2. Experimental[39] and Calculated Free Energy Differences ($\Delta G_{binding}^{mutant\text{-}CBD} - \Delta G_{binding}^{wt\text{-}CBD}$) between wt-CBD and its Mutants to Cellulose in which Positive Numbers Indicate Lower Binding Affinities than wt-CBD

mutation	$\Delta\Delta G_{exp}$ (kcal/mol)[a]	ΔG_{cal} (kcal/mol)[c]	$\Delta\Delta G_{cal}$ (kcal/mol)[c]
Y5A	>1.74[b]	−12.4 (±0.2)	5.6 (±0.1)
Y31A	1.74	−14.3 (±0.4)	3.7 (±0.3)
Y32A	>1.74[b]	−13.3 (±0.1)	4.7 (±0.1)
WT-CBD		−18.0 (±0.03)	

[a]The free energy of binding was calculated from the equation: $\Delta\Delta G = -RT\ln(K_{mutant\text{-}CBD}/K_{wt\text{-}CBD})$, where $K_{mutant\text{-}CBD}$ and $K_{wt\text{-}CBD}$ are the partition coefficients. [b]In ref 39, exact free energy values could not be extrapolated from the adsorption isotherm due to its small slope. These values (>1.74) were estimated through a comparison with other mutations for which exact values were derived. [c]The error associated with energy minima was calculated using a bootstrap method.[40,41]

Having validated the computational protocol, we have applied it to the evaluation of the cellulose binding free

DOI: 10.1021/acs.jpcb.5b11050
J. Phys. Chem. B 2016, 120, 309−319

energies of the CBD mimetics designed in this work (pepA–pepD). The resulting PMF plots are presented in Figure 7, and the comparison with the experimental results is discussed in section 2.6 below.

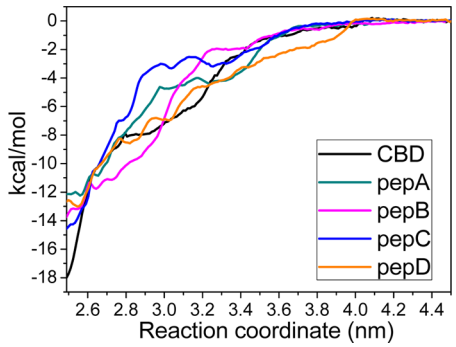


Figure 7. Potential of mean force (PMF) plots describing the binding of wild-type CBD and of the designed peptides to cellulose. The reaction coordinate was taken as the COM distance between CBD/peptides and cellulose.

2.3. Experimental Binding Measurements of the Designed Peptides to Cellulose. The affinities of the CBD mimetics (pepA–pepD) and of CBD to cellulose were determined using adsorption isotherm measurements. Figure 8 presents the concentration of adsorbed peptide relative to the concentration of the free peptide for all constructs. Data fitting based on a Langmuir model provided the binding constants (K_{ads}) and maximal surface coverage (N_{max}) (Table 3). Comparing the adsorption characteristics of the full CBD to values reported before for CBD from *Clostridium cellulovorans*[42]

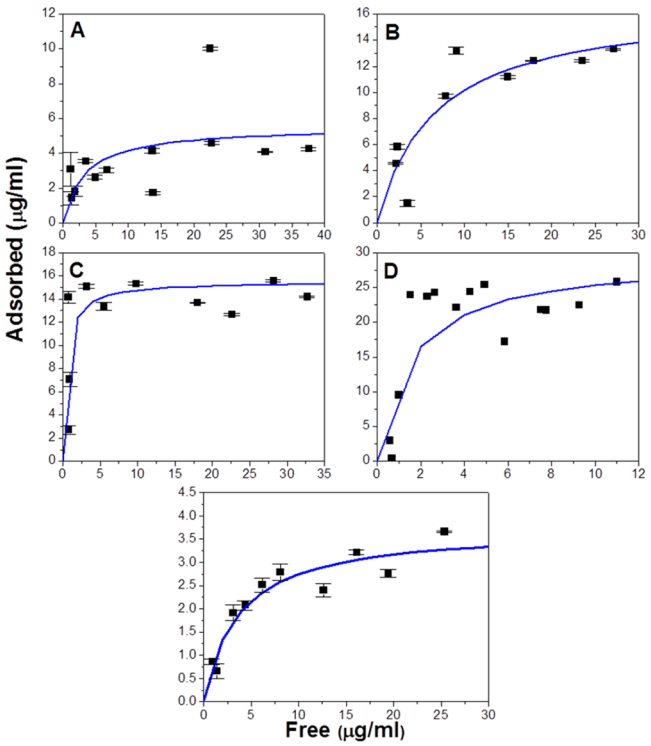


Figure 8. Experimental adsorption isotherm data (squares) of peptides A–D and full-length CBD to cellulose. Langmuir fits (blue line) to a single set of peptide adsorption sites on cellulose are shown.

Table 3. Summary of Adsorption Measurements of the Full CBD and the Design Peptides to Cellulose

construct	K_{ads} (M^{-1})	N_{max} (μmol/gr)
CBD	0.94(±0.15) × 10^6	5.07 (±0.05)
PepA	0.59(±0.15) × 10^6	1.35 (±0.13)
PepB	0.62(±0.15) × 10^6	3.50 (±0.10)
PepC	3.25(±0.30) × 10^6	3.79 (±0.08)
PepD	1.45(±0.30) × 10^6	6.77 (±0.10)

indicates a slightly weaker binding affinity (0.94(±0.15) × 10^6 M^{-1} in the present work vs (1.1–1.6) × 10^6 M^{-1} in previous reports) but 2.4–2.6 times higher coverage (5.07(±0.05) μmol/gr vs ~2 μmol/gr). The first-generation peptides (pepA and pepB) are either similar to (pepA) or weaker than (pepB) the full CBD in terms of their ability to bind cellulose. However, the second-generation peptides, pepC and pepD, exhibit excellent affinity to cellulose that surpasses that of the complete domain and more efficient coverage of the cellulose than the first-generation peptides.

Figure S1 presents the normalized heat release during isothermal titration calorimetry (ITC) of 1 mM peptide B to 0.23 mM suspension of cellulose in sodium acetate buffer at 30 °C. The titration curve exhibits a very low heat release signal and quantitative analysis gives high errors on values obtained. The K_{ads} value derived from the curve was 2.5(±0.9) × 10^5 M^{-1}, and the ΔH_{ads} was −63 (±7) cal/mol. The adsorption constant value is on the same order of magnitude as the one obtained from the adsorption isotherm measurements despite the poor sensitivity of the measurement. Titrations of the other constructs and of CBD to Avicell cellulose produced even lower heat release signatures and could not be analyzed quantitatively. It is clear from the ITC data that weak forces are driving binding of CBD and the designed peptides to the polysaccharide (e.g., the above-mentioned σ–π interactions). These interactions are much weaker than, e.g., electrostatic interactions, and consequently lead to low enthalpy changes upon binding. This in turn suggests that entropy may play the dominant role in the binding process as was previously shown by Georgelis et al.[43]

2.4. Circular Dichroism Measurements and Their Relation to Cellulose Binding Results. To obtain information regarding the secondary structure of the peptides, CD measurements were performed (see Figure 9). The normalized CD spectra of peptides A–D and of the full CBD showed a similar pattern with a dip at 193 nm corresponding to

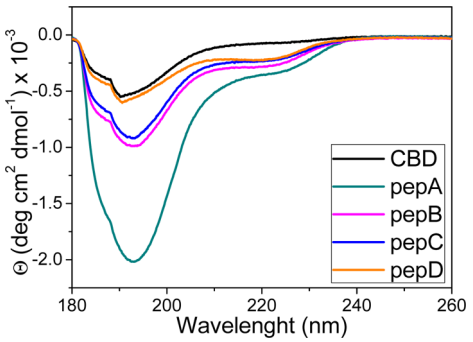


Figure 9. Circular dichroism data of CBD and peptides A–D in 0.1 M sodium acetate buffer solution, similar to the one used in adsorption isotherm measurements.

DOI: 10.1021/acs.jpcb.5b11050
J. Phys. Chem. B 2016, 120, 309–319

a random coil conformation of the peptide backbone and a secondary dip at 184 nm marking the presence of an antiparallel beta (ap-β) motif. The absence of a positive signal at ~200 nm indicates that the random coil dominates the β conformation. The spectrum of the CBD based on its crystal structure was calculated using DichroCalc (data not shown).[44] The predicted spectrum is typical of ap-β structure that is very different from the spectra in Figure 9. This suggests that, under the experimental conditions employed in this study, higher populations of disordered conformations of the CBD (and probably of its mimetic designs) were obtained. On the basis of the intensity of the 193 nm dip, we suggest that the random coil population increases in the order pepD → pepC → pepB → pepA. Interestingly, the two constructs with the higher content of random coil have the lower affinity to cellulose. However, cellulose affinity likely depends on additional factors, such as the identity of the residues comprising the "binding strip".

2.5. Solid State NMR Measurements and Their Relation to Cellulose Ordering. Selective ^{13}C magic angle spinning (MAS) NMR experiments were carried out on pepB, cellulose, and their complex. The sugar carbons in the Avicell cellulose were recently assigned and are shown in Figure 10.[45]

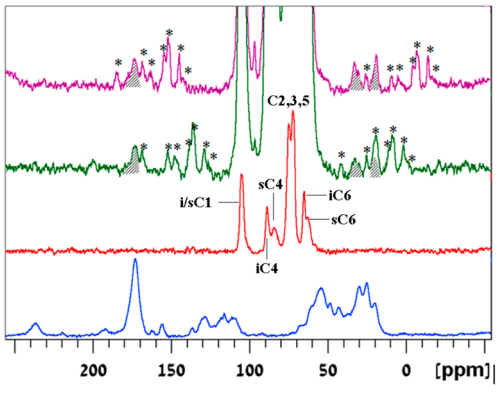


Figure 10. ^{13}C CPMAS spectra of pepB (blue), Avicell cellulose (red), and the pepB−cellulose complex at 8 kHz spinning (green) and at 10 kHz spinning (purple). Peak assignments of sugar carbons are shown.

The spectra of the three materials are compared in Figure 10. The sugar lines of cellulose (green) also dominate the spectra of the pepB−cellulose complex recorded at spinning rates of 10 and 8 kHz in purple and green, respectively. One prominent feature of the pepB−cellulose spectra is the appearance of strong sidebands (marked with asterisks) flanking the cellulose lines. These sidebands indicate that cellulose, in the presence of pepB, acquired large chemical shift anisotropy. This is unlike free cellulose, which exhibits negligible sidebands on its sugar lines. Line width differences of center sugar peaks are also observed as shown in Figure S2. The changes observed for cellulose may be related to a decrease in puckering motion of backbone sugar rings in the presence of the peptide. However, further measurements are required to validate that these changes are not a consequence of measuring the slightly hydrated complex in acetate buffer. The lines from the peptide (marked as shaded regions in the purple and green spectra) are hard to notice under the large cellulose sidebands but confirm the presence of pepB in this sample.

2.6. A Comparison between Computational and Experimental Results. Table 4 provides a comparison between experimentally and computationally derived ΔG and $\Delta\Delta G$ values (relative to wt-CBD) for the four CBD mimetics considered in this work. Both the computational and experimental work suggest that, among the designed peptides, pepC is the strongest cellulose binder, and pepA is the weakest binder. PepC features the Y5W mutations, whereas in pepA, none of the "strip" residues are mutated to Trp. These finding are consistent with an important role of Trp as an enhancer of cellulose binding affinity. Predictions also show the correct trend (albeit with no statistical significance) in the cellulose binding potency of pepA and pepD, in agreement with experiments with the two Y to W mutations nearly compensated for by the T to Q mutation. This emphasizes the importance of Q7 in binding cellulose. The largest calculated difference in $\Delta\Delta G$ (between pepA and pepC) is 2.2 kcal/mol, and the largest measured difference is 0.9 kcal/mol (also between pepA and pepC), illustrating nicely that, albeit the discrepancy in absolute values obtained by the two approaches, on a relative scale, they are able to pinpoint surprisingly well the differences between the constructs. They differ markedly however with respect to the CBD. While the computational work suggests it to be the strongest binder to cellulose, the experimental work suggests that both pepC and pepD are better binders. In fact, even if one accounts for the previously measured affinity of CBD, which was reportedly 1.7 times stronger than measured here, it would still be ~2.5 times weaker than that of pepC. The factors leading to the discrepancy between calculated and measured values are likely found both in experiments and computations. As already mentioned, the calculations applied in this work probably did not account for ionic strength or local pH effects. Other sources of errors could come from the specific details of the PMF protocol, such as the pulling velocity, the selected reaction coordinate, and the simulation length in each window. Finally, in this work, we derived ΔG values from multiple windows generated from a single COM pulling simulation. Better results (albeit at the expense of more time-consuming simulations) could perhaps be obtained with Jarzynski's method,[46] which

Table 4. Experimental and Calculated Free Energy Differences ($\Delta G^{peptide}_{binding} - \Delta G^{wt\text{-}CBD}_{binding}$) between wt-CBD and its Mimetics to Cellulose in which Positive Numbers Indicate Lower Binding Affinity than wt-CBD and Negative Numbers Indicate Higher Binding Affinity than wt-CBD

designed construct	ΔG_{exp}(kcal/mol)	$\Delta\Delta G_{exp}$(kcal/mol)	ΔG_{cal} (kcal/mol)[a]	$\Delta\Delta G_{cal}$ (kcal/mol)[a]
pepA	−7.29 (±0.70)	0.25 (±0.08)	−12.2 (±0.2)	5.8 (±0.1)
pepB	−7.32 (±0.21)	0.22 (±0.07)	−13.7 (±0.5)	4.0 (±0.4)
pepC	−8.22 (±0.17)	−0.68 (±0.16)	−14.4 (±0.5)	3.6 (±0.4)
pepD	−7.78 (±0.11)	−0.24 (±0.06)	−12.6 (±0.6)	5.4 (±0.4)
WT-CBD	−7.54 (±0.07)		−18.0 (±0.03)	

[a]The error associated with energy minima was calculated using a bootstrap method.[40,41]

DOI: 10.1021/acs.jpcb.5b11050
J. Phys. Chem. B 2016, 120, 309−319

uses the weighted histogram analysis method (WHAM) algorithm[47] to extract free energies from multiple non-equilibrium COM pulling simulations. The adsorption isotherm measurements are prone to errors in accurate concentration determination as well as sample to sample variability in terms of both the peptides and polysaccharide. This is also manifested in the statistical variance of the experimental results.

2.7. The Dissociation Pathway of CBDs and Their Mimetics. Some insight into the dissociation mechanism of cellulose binders could be obtained by examining the pulling trajectories of the CBDs and peptides considered in this work. Despite pulling the CBD/peptides from a single point located at their COM, we observed that the aromatic strip composed of residues 5, 31, and 32 did not detach from the surface of the cellulose in a parallel manner. To analyze the dissociation pathway, distances between the C_β atoms of residues 5, 31, and 32, and cellulose (C3 atom of the glucose unit closest to residues 5, 31, and 32) were measured using a tcl scripting interface in VMD.[48]

Figure 11 presents the distances between the C_β atoms of residues 5, 31, and 32 and between cellulose as a function of the

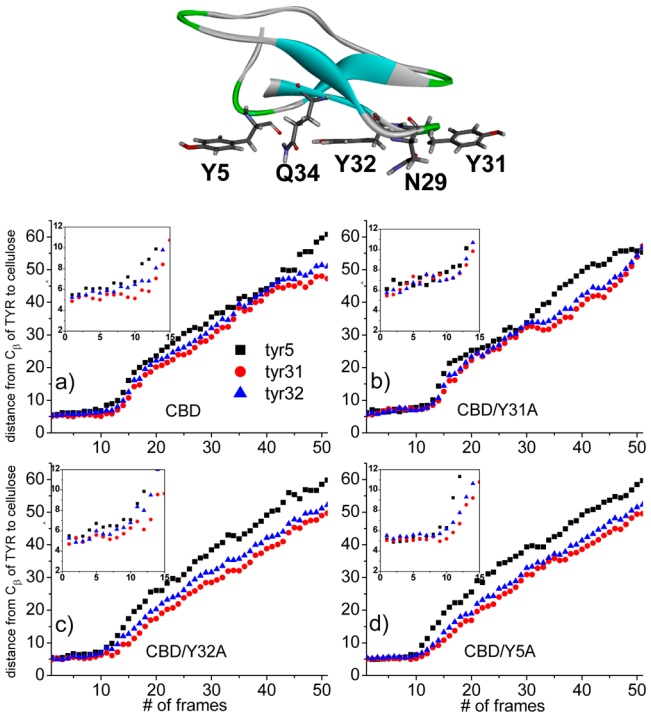


Figure 11. (Top) Ribbon representation of the CBD highlighting (in sticks) the residues participating in hydrogen bonds and hydrophobic interactions with cellulose. (Bottom) Distance from C_β atoms of residues 5, 31, and 32 to cellulose as a function of frame number from the pulling simulations. The insets focus on the initial stage of the simulations.

pulling trajectory frame for the four CBDs considered in this work. As can be seen, residues in position 5 (black squares in Figure 11) are the first to dissociate from the cellulose surface, followed by residues in positions 32 and 31. This behavior could be explained by monitoring the hydrogen bonds during the pulling simulation. Most hydrogen bonds between the CBDs and the sugar hydroxyls of cellulose are formed through the strictly conserved residues Q34 and N29 and to a lesser extent through the hydroxyl moiety of Y32. In CBD, Q34 is located on one side of the domain (close to Y5), whereas N29 and Y32 are located on the other side. Because breaking one hydrogen bond is energetically easier than breaking two bonds, dissociation begins from the Y5 side. However, a similar dissociation pathway is also observed for the Y32A mutant that has the same number of hydrogen bonds on both sides of the domain. Thus, hydrogen bonding alone is unable to fully account for the observed dissociation pathway. However, an analysis of hydrophobic interactions between CBDs and cellulose at a 4.5 Å cutoff suggests that Y31 is more heavily engaged in hydrophobic interactions than either Y5 or Y32 (Figure 12). Hydrophobic interactions affect the dissociation

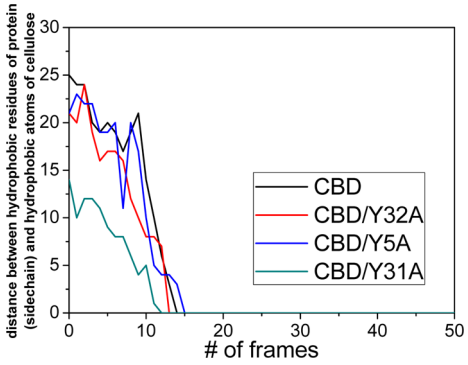


Figure 12. Hydrophobic interactions between CBDs and cellulose as a function of frame number during pulling simulations. Interactions were calculated by measuring the distances between the hydrophobic atoms of the hydrophobic side chain residues A, L, V, I, P, F, M, W, and Y and the hydrophobic atoms of cellulose using the g_mindist analysis tool in Gromacs.

path of CBDs from cellulose to a larger extent than hydrogen bonds. This is evident from the inset of Figure 11(b) that shows a much more balanced detachment pathway for the Y31A mutant in the initial stage of the simulation despite having imbalanced hydrogen bond interactions.

The location of the hydrogen bonds forming residues in pepA−pepD is similar to that in wt-CBD with Q34 located on one side (close to Y38) and N29, Y31, and Y32 located on the other side of the peptides. This explains the earlier cellulose dissociation of the "Y38 side" upon pulling. For pepC, Y31 also participates in hydrogen bonding to cellulose in accordance with the more pronounced imbalanced dissociation (Figure 13c). In pepB, the Y to W mutation increases the hydrophobic interactions between the peptide and the cellulose leading to a more balance dissociation pathway (Figure 13b).

3. CONCLUSIONS

In this work, we present a combined computational/experimental study aimed at designing new CBD mimetics capable of binding to cellulose. This is a first step toward the development of synthetic cellulose degradation machinery. The designed 18-residue peptides are able to bind cellulose roughly to the same degree or even stronger than the complete wt-CBD. Furthermore, the overall good agreement between computational and experimental results suggests that computational approaches can be useful in the future design of additional cellulose binders. Finally, the dissociation pathways of the CBDs and their mimetics from cellulose suggest that enhanced cellulose binding can be obtained by increasing the

DOI: 10.1021/acs.jpcb.5b11050
J. Phys. Chem. B 2016, 120, 309−319

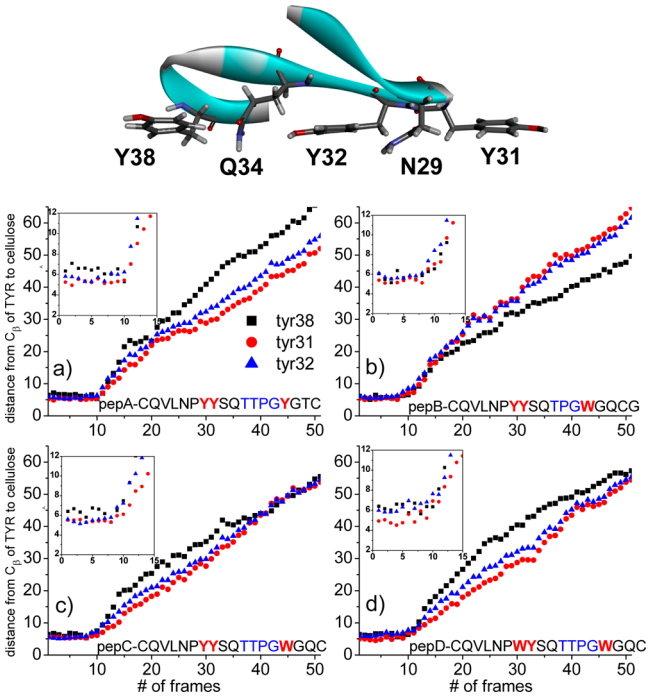


Figure 13. (Top) Ribbon representation of pepA–pepD highlighting (in sticks) the residues participating in hydrogen bonds and hydrophobic interactions with the cellulose. (Bottom) Distance from C_β atoms of residues in position 38, 31, and 32 to cellulose as a function of frame number from the pulling simulations. The insets focus on the initial stage of the simulations.

hydrogen bond and hydrophobic interactions between the two binding partners.

4. METHODS

4.1. Computational Methods. MD simulations were performed using the Gromacs Molecular Dynamics package (version 4.5.5)[49,50] with different force fields and water models as indicated below. The systems were submerged in a cubic box with an extra extension along each axis of the system of 10 Å. When needed, ions were added to make the system electrically neutral. The simulations were performed at 300 K typically with a time step of 2 fs using the leapfrog algorithm.[51,52] The cutoff for van der Waals and Coulomb interactions was set to 10 Å. Long-range electrostatic interactions were computed using Particle Mesh Ewald Summation.[53] Periodic boundary conditions were applied in all directions.

Mimetics Design. The structure of the 36-residue CBD from *T. reesei* cellobiohydrolase I was retrieved from the PDB (PDB ID: 1CBH) and used for the design of pepA–pepD as indicated in the text. Prior to MD simulations, all constructs were prepared using the prepare protein protocol as implemented in Discovery Studio (DS) version 3.5 to assign the correct protonation states to titrateable residues.[54] Following preparation, structures were minimized, equilibrated (first under NVT conditions for 100 ps and then under NPT conditions for an additional 100 ps), and finally simulated under NPT conditions for 20 ns. All simulations utilized the OPLS/AA force field[55] and TIP4P water.[56]

Cellulose Model. A 3D crystalline model of cellulose Iβ composed of 720 repeating glucose units (40 chains each consisting of 18 glucose units)[37] was built based on a glucose structure obtained from the Cambridge Structural Database (JINROO01.cif).[57] The structure was minimized, equilibrated (first under NVT conditions for 100 ps and then under NPT conditions for an additional 100 ps), and finally simulated under NPT conditions for 100 ns. Simulations utilized the Amber 99 SB-ILDN forcefield[58] and the Glycam06 carbohydrate parameter set.[59] For this system, a time step of 1 fs was used.

Binding Free Energy Simulations. Binding free energies were evaluated from potential of mean force (PMF) graphs. These were calculated using pulling simulations followed by umbrella sampling. According to this method, a series of configurations is generated along a reaction coordinate, ξ, between two interacting partners. In the present case, the cellulose served as a reference point, whereas its binding partner (i.e., CBDs or CBD-like peptides), initially placed on its surface, was pulled away and placed at increasing center-of-mass (COM) distances fixed by a biasing potential. These COM distances represented "sampling windows", wherein independent simulations were conducted to generate an ensemble of structures along the reaction coordinate. For assembling a PMF curve as a function of the entire reaction coordinate, energy values in adjacent windows were reassembled into a continuous function.

For generating the starting conformation for the simulations, wt-CBD was manually located on an already equilibrated cellulose model according to the CBD binding mode suggested by Beckham et al.[47,60] The starting conformation was selected by first running a "scanning simulation" spanning a distance of 50 Å from the manually located conformation along the *z*-axis followed by nine simulations initiated from equally spaced snapshots each spanning 10 Å along the *y*-axis. The lowest energy structure identified through this scanning procedure was selected as the starting point for the pulling procedure. The Y32A, Y31Y, and Y5A mutations were introduced on the best conformation found for the native CBD. Following minimization and NPT equilibration for 100 ps, CBDs were pulled away from cellulose along the *x*-axis over 500 ps using a spring constant of 1000 kJ mol^{-1} nm^{-2} and a pull rate of 0.01 nm/ps. A final center of mass (COM) distance between CBDs and cellulose of approximately 50 Å was reached.

From these trajectories, snapshots were taken to generate the starting configurations for the umbrella sampling windows. An asymmetric distribution of sampling windows was used and resulted in 32 windows for each of the CBDs. In each window, the system was first equilibrated for 100 ps under NPT conditions and finally simulated for 10 ns under NPT conditions. Distances to cellulose were restrained within a defined window using a spring constant of 3000 kJ mol^{-1} nm^{-2}. Good overlap between windows was obtained (data not shown). The final PMF was constructed using the weighted histogram analysis method (WHAM).[47] In this work, we attribute no significance to the particular shape of the PMF curve but only the two end points. The energy difference between these points is the binding free energy. The uncertainty of the PMF was estimated using Bayesian bootstrapping analysis of complete WHAM histograms, which provides a reliable estimated error without the requirement to carry out new independent simulations.[40]

COM pulling and umbrella sampling simulations were performed in explicit solvent. Parameters from the Amber 99SB-ILDN force field[58] and the Glycam06 carbohydrate parameter set[59] were applied to all molecules in the simulated

DOI: 10.1021/acs.jpcb.5b11050
J. Phys. Chem. B 2016, 120, 309–319

system. Short-range nonbonded interactions were cut off at 14 Å.

4.2. Experimental Methods. Cellulose (Avicell PH101) was purchased from Sigma-Aldrich and used as is. Full CBD from *T. reesei* was purchased from Selleck.

Peptide Synthesis. Peptides were synthesized in 100 μmol quantities using standard solid phase peptide synthesis (SPPS) techniques based on 9-fluorenylmethoxycarbonyl (FMOC) chemistry on Rink Amide resin. Products were purified on a Waters HPLC using a VYDAC C18 reverse phase column for separation based on the hydrophobic index of the products. Mass spectrometry analysis of crude and purified product composition was performed on a matrix-assisted laser desorption/ionization (MALDI) mass spectrometer equipped with an Autoflex III smartbeam by Bruker.

Adsorption of Peptides and CBD to Cellulose. All adsorption isotherm experiments were performed at controlled temperature. Lyophilized peptides were dissolved in 0.1 M sodium acetate buffer at pH 5. Suspensions of Avicel cellulose with a concentration of 2 mg/ml in the presence of different concentrations of peptides were vortexed and incubated for 20−25 h at 4 °C. The initial concentrations of peptides were determined by measuring the UV absorbance at 280 nm. All dilutions of peptides were performed in the same buffer. Then, suspensions were diluted and centrifuged at 8000 rpm for 20 min, and the final concentrations of peptides were measured using the micro BCA protein assay at 562 nm. Each measurement was performed in triplicate.

Isothermal Titration Calorimetry of Peptide B to Cellulose. Calorimetric measurements of isothermal titration of peptide B into a cell containing suspended Avicell cellulose were carried out at 300 K using 5 mL injections of 1 mM peptide solution into 0.23 mM of cellulose in acetate buffer. Experiments were carried out on a Microcal, GE Healthcare VP-ITC calorimeter using high gain mode. Data fitting was performed using a single set of binding sites in Origin Lab 7.0 employing Microcal fitting macros.

CD Analysis. Circular dichroism (CD) measurements were carried out using a Chirascan spectrometer (Applied Photophysics, UK). Measurements were performed at room temperature in a 1 mm optical path length cell, and the spectra were recorded from 260 to 180 nm with a step size of 0.1 nm and a bandwidth of 0.5 nm. Time per point was 0.5 s and approximate scan-time was ∼11 min. The concentration of all samples was 0.1 mg/ml of 0.1 M sodium acetate buffer at pH 5.

Solid State NMR. NMR measurements were performed on a Bruker 11.7T Avance^{III} spectrometer equipped with a 4 mm VTN CPMAS probe at spinning rates of 8 and 10 kHz. ^{13}C CP experiments employed a ^{1}H 90° pulse of 2.4 μs, followed by a 2.5 ms CP contact time using ramped field on ^{1}H (40 to 80 kHz) and 51 kHz field on ^{13}C and composite-pulse ^{1}H decoupling using the SPINAL64 sequence with RF field of 94 kHz during acquisition. ^{13}C CPMAS spectrum was collected with 20150 scans on pepB−cellulose, 2048 scans on cellulose, and 4096 scans on pepB with a recycle delay of 3 s.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.5b11050.

Representative isothermal titration calorimetry plot for titration of cellulose with peptide B (Figure S1) and ^{13}C NMR spectrum of cellolose and pepB−cellulose (Figure S2) (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: gil.goobes@biu.ac.il.

*E-mail: hsenderowitz@gmail.com.

Author Contributions

†N.K. and T.I.-V. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Huber, G. W.; Iborra, S.; Corma, A. Synthesis of Transportation Fuels from Biomass: Chemistry, Catalysts, and Engineering. *Chem. Rev.* **2006**, *106*, 4044−4098.

(2) Wang, J. J.; Xi, J. X.; Wang, Y. Q. Recent Advances in the Catalytic Production of Glucose from Lignocellulosic Biomass. *Green Chem.* **2015**, *17*, 737−751.

(3) Himmel, M. E.; Ding, S. Y.; Johnson, D. K.; Adney, W. S.; Nimlos, M. R.; Brady, J. W.; Foust, T. D. Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production. *Science* **2007**, *315*, 804−807.

(4) Hu, J.; Arantes, V.; Pribowo, A.; Gourlay, K.; Saddler, J. N. Substrate Factors That Influence the Synergistic Interaction of Aa9 and Cellulases During the Enzymatic Hydrolysis of Biomass. *Energy Environ. Sci.* **2014**, *7*, 2308−2315.

(5) Resch, M. G.; Donohoe, B. S.; Baker, J. O.; Decker, S. R.; Bayer, E. A.; Beckham, G. T.; Himmel, M. E. Fungal Cellulases and Complexed Cellulosomal Enzymes Exhibit Synergistic Mechanisms in Cellulose Deconstruction. *Energy Environ. Sci.* **2013**, *6*, 1858−1867.

(6) Akia, M.; Yazdani, F.; Motaee, E.; Han, D.; Arandiyan, H. A Review on Conversion of Biomass to Biofuel by Nanocatalysts. *Biofuel Research Journal* **2014**, *1*, 16−25.

(7) Fukuda, T.; Ishikawa, T.; Ogawa, M.; Shiraga, S.; Kato, M.; Suye, S.-i.; Ueda, M. Enhancement of Cellulase Activity by Clones Selected from the Combinatorial Library of the Cellulose-Binding Domain by Cell Surface Engineering. *Biotechnol. Prog.* **2006**, *22*, 933−938.

(8) Park, S.-H.; Ransom, C.; Mei, C.; Sabzikar, R.; Qi, C.; Chundawat, S.; Dale, B.; Sticklen, M. The Quest for Alternatives to Microbial Cellulase Mix Production: Corn Stover-Produced Heterologous Multi-Cellulases Readily Deconstruct Lignocellulosic Biomass into Fermentable Sugars. *J. Chem. Technol. Biotechnol.* **2011**, *86*, 633−641.

(9) Biello, D. The False Promise of Biofuels (Vol 305, Pg 58, 2011). *Sci. Am.* **2011**, *305*, 10−10.

(10) Dale, B. E. Cellulosic Biofuels and the Road to Energy Security. *Environ. Sci. Technol.* **2011**, *45*, 9823−9823.

(11) Schnoor, J. L. Cellulosic Biofuels Disappoint. *Environ. Sci. Technol.* **2011**, *45*, 7099−7099.

(12) Singhvi, M. S.; Chaudhari, S.; Gokhale, D. V. Lignocellulose Processing: A Current Challenge. *RSC Adv.* **2014**, *4*, 8271−8277.

(13) Dutta, S.; Wu, K. C. W. Enzymatic Breakdown of Biomass: Enzyme Active Sites, Immobilization, and Biofuel Production. *Green Chem.* **2014**, *16*, 4615−4626.

(14) Fox, R. J.; Huisman, G. W. Enzyme Optimization: Moving from Blind Evolution to Statistical Exploration of Sequence−Function Space. *Trends Biotechnol.* **2008**, *26*, 132−138.

(15) Medie, F.; Davies, G. J.; Drancourt, M.; Henrissat, B. Genome Analyses Highlight the Different Biological Roles of Cellulases. *Nat. Rev. Microbiol.* **2012**, *10*, 227−234.

(16) Lombard, V.; Ramulu, H. G.; Drula, E.; Coutinho, P. M.; Henrissat, B. The Carbohydrate-Active Enzymes Database (Cazy) in 2013. *Nucleic Acids Res.* **2014**, *42*, D490−D495.

(17) Ivanir-Dabora, H.; Nimerovsky, E.; Madhu, P. K.; Goldbourt, A. Site-Resolved Backbone and Side-Chain Intermediate Dynamics in a

DOI: 10.1021/acs.jpcb.5b11050
J. Phys. Chem. B 2016, 120, 309−319

Carbohydrate-Binding Module Protein Studied by Magic-Angle Spinning Nmr Spectroscopy. *Chem. - Eur. J.* **2015**, *21*, 10778−10785.
(18) Ivanir, H.; Goldbourt, A. Solid State Nmr Chemical Shift Assignment and Conformational Analysis of a Cellulose Binding Protein Facilitated by Optimized Glycerol Enrichment. *J. Biomol. NMR* **2014**, *59*, 185−197.
(19) Nawrocki, G.; Cazade, P. A.; Thompson, D.; Cieplak, M. Peptide Recognition Capabilities of Cellulose in Molecular Dynamics Simulations. *J. Phys. Chem. C* **2015**, *119*, 24404−24416.
(20) Bommarius, A. S.; Blum, J. K.; Abrahamson, M. J. Status of Protein Engineering for Biocatalysts: How to Design an Industrially Useful Biocatalyst. *Curr. Opin. Chem. Biol.* **2011**, *15*, 194−200.
(21) Privett, H. K.; Kiss, G.; Lee, T. M.; Blomberg, R.; Chica, R. A.; Thomas, L. M.; Hilvert, D.; Houk, K. N.; Mayo, S. L. Iterative Approach to Computational Enzyme Design. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 3790.
(22) Dalby, P. A. Engineering Enzymes for Biocatalysis. *Recent Pat. Biotechnol.* **2007**, *1*, 1−9.
(23) Morais, S.; Morag, E.; Barak, Y.; Goldman, D.; Hadar, Y.; Lamed, R.; Shoham, Y.; Wilson, D. B.; Bayer, E. A. Deconstruction of Lignocellulose into Soluble Sugars by Native and Designer Cellulosomes. *mBio* **2013**, *3*, e00508-12.
(24) Nakazawa, H.; Ikeuchi, A.; Kim, D. M.; Ishigaki, Y.; Asano, H.; Kouda, K.; Kumagai, I.; Umetsu, M. Biomass-Binding Peptides Designed by Molecular Evolution for Efficient Degradation of Cellulose in Biomass by Cellulase. *Green Chem.* **2013**, *15*, 365−369.
(25) Guo, J.; Catchmark, J. M.; Mohamed, M. N. A.; Benesi, A. J.; Tien, M.; Kao, T. H.; Watts, H. D.; Kubicki, J. D. Identification and Characterization of a Cellulose Binding Heptapeptide Revealed by Phage Display. *Biomacromolecules* **2013**, *14*, 1795−1805.
(26) Serizawa, T.; Iida, K.; Matsuno, H.; Kurita, K. Cellulose-Binding Heptapeptides Identified by Phage Display Methods. *Chem. Lett.* **2007**, *36*, 988−989.
(27) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. The Protein Data Bank: A Computer-Based Archival File for Macromolecular Structures. *J. Mol. Biol.* **1977**, *112*, 535−542.
(28) Kraulis, J.; Clore, G. M.; Nilges, M.; Jones, T. A.; Pettersson, G.; Knowles, J.; Gronenborn, A. M. Determination of the Three-Dimensional Solution Structure of the C-Terminal Domain of Cellobiohydrolase I from Trichoderma Reesei. A Study Using Nuclear Magnetic Resonance and Hybrid Distance Geometry-Dynamical Simulated Annealing. *Biochemistry* **1989**, *28*, 7241−7257.
(29) Linder, M.; Lindeberg, G.; Reinikainen, T.; Teeri, T. T.; Pettersson, G. The Difference in Affinity between Two Fungal Cellulose-Binding Domains Is Dominated by a Single Amino Acid Substitution. *FEBS Lett.* **1995**, *372*, 96−98.
(30) Schaeffer, R. D.; Daggett, V. Protein Folds and Protein Folding. *Protein Eng., Des. Sel.* **2011**, *24*, 11−19.
(31) van der Kamp, M. W.; Schaeffer, R. D.; Jonsson, A. L.; Scouras, A. D.; Simms, A. M.; Toofanny, R. D.; Benson, N. C.; Anderson, P. C.; Merkley, E. D.; Rysavy, S.; Bromley, D.; Beck, D. A. C.; Daggett, V. Dynameomics: A Comprehensive Database of Protein Dynamics. *Structure* **2010**, *18*, 423−435.
(32) Benson, N. C.; Daggett, V. Dynameomics: Large-Scale Assessment of Native Protein Flexibility. *Protein Sci.* **2008**, *17*, 2038−2050.
(33) Hoffren, A. M.; Teeri, T. T.; Teleman, O. Molecular Dynamics Simulation of Fungal Cellulose-Binding Domains: Differences in Molecular Rigidity but a Preserved Cellulose Binding Surface. *Protein Eng., Des. Sel.* **1995**, *8*, 443−450.
(34) Bergenstrahle, M.; Matthews, J. F.; Crowley, M.; Brady, J. W. Cellulose Crystal Structure and Force Fields. In *International Conference on Nanotechnology for the Forest Products Industry*; Otaniemi: Espoo, Finland, **2010**.
(35) Matthews, J. F.; Skopec, C. E.; Mason, P. E.; Zuccato, P.; Torget, R. W.; Sugiyama, J.; Himmel, M. E.; Brady, J. W. Computer Simulation Studies of Microcrystalline Cellulose Ibeta. *Carbohydr. Res.* **2006**, *341*, 138−152.
(36) Yui, T.; Nishimura, S.; Akiba, S.; Hayashi, S. Swelling Behavior of the Cellulose Ibeta Crystal Models by Molecular Dynamics. *Carbohydr. Res.* **2006**, *341*, 2521−2530.
(37) Nishiyama, Y.; Langan, P.; Chanzy, H. Crystal Structure and Hydrogen-Bonding System in Cellulose Ibeta from Synchrotron X-Ray and Neutron Fiber Diffraction. *J. Am. Chem. Soc.* **2002**, *124*, 9074−9082.
(38) Mattinen, M. L.; Kontteli, M.; Kerovuo, J.; Linder, M.; Annila, A.; Lindeberg, G.; Reinikainen, T.; Drakenberg, T. Three-Dimensional Structures of Three Engineered Cellulose-Binding Domains of Cellobiohydrolase I from Trichoderma Reesei. *Protein Sci.* **1997**, *6*, 294−303.
(39) Linder, M.; Mattinen, M. L.; Kontteli, M.; Lindeberg, G.; Stahlberg, J.; Drakenberg, T.; Reinikainen, T.; Pettersson, G.; Annila, A. Identification of Functionally Important Amino-Acids in the Cellulose-Binding Domain of Trichoderma-Reesei Cellobiohydrolase-I. *Protein Sci.* **1995**, *4*, 1056−1064.
(40) Hub, J. S.; de Groot, B. L.; van der Spoel, D. G_Wham-a Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates. *J. Chem. Theory Comput.* **2010**, *6*, 3713−3720.
(41) Hub, J. S.; de Groot, B. L. Does Co2 Permeate through Aquaporin-1? *Biophys. J.* **2006**, *91*, 842−848.
(42) Goldstein, M. A.; Takagi, M.; Hashida, S.; Shoseyov, O.; Doi, R. H.; Segel, I. H. Characterization of the Cellulose-Binding Domain of the Clostridium Cellulovorans Cellulose-Binding Protein A. *J. Bacteriol.* **1993**, *175*, 5762−5768.
(43) Georgelis, N.; Yennawar, N. H.; Cosgrove, D. J. Structural Basis for Entropy-Driven Cellulose Binding by a Type-a Cellulose-Binding Module (Cbm) and Bacterial Expansin. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 14830−14835.
(44) Bulheller, B. M.; Hirst, J. D. Dichrocalc-Circular and Linear Dichroism Online. *Bioinformatics* **2009**, *25*, 539−540.
(45) Wang, T.; Salazar, A.; Zabotina, O. A.; Hong, M. Structure and Dynamics of Brachypodium Primary Cell Wall Polysaccharides from Two-Dimensional C-13 Solid-State Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **2014**, *53*, 2840−2854.
(46) Jarzynski, C. Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett.* **1997**, *78*, 2690−2693.
(47) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules 0.1. The Method. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(48) Humphrey, W.; Dalke, A.; Schulten, K. Vmd: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(49) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. Gromacs 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(50) Berendsen, H. J. C.; Vanderspoel, D.; Vandrunen, R. Gromacs - a Message-Passing Parallel Molecular-Dynamics Implementation. *Comput. Phys. Commun.* **1995**, *91*, 43−56.
(51) Hockney, R. W.; Goel, S. P.; Eastwood, J. W. Quiet High-Resolution Computer Models of a Plasma. *J. Comput. Phys.* **1974**, *14*, 148−158.
(52) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(53) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(54) *Discovery Studio*, Accelrys Inc.: San Diego, CA, USA, 2012.
(55) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and Reparametrization of the Opls-Aa Force Field for Proteins Via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.
(56) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.

DOI: 10.1021/acs.jpcb.5b11050
J. Phys. Chem. B 2016, 120, 309−319

(57) Allen, F. The Cambridge Structural Database: A Quarter of a Million Crystal Structures and Rising. *Acta Crystallogr., Sect. B: Struct. Sci.* **2002**, *58*, 380−388.
(58) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber Ff99sb Protein Force Field. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1950−1958.
(59) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. Glycam06: A Generalizable Biomolecular Force Field. Carbohydrates. *J. Comput. Chem.* **2008**, *29*, 622−655.
(60) Beckham, G. T.; Matthews, J. F.; Bomble, Y. J.; Bu, L. T.; 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.

DOI: 10.1021/acs.jpcb.5b11050
J. Phys. Chem. B 2016, 120, 309−319