

Supporting Information

Endoglucanase Peripheral Loops Facilitate Complexation of Glucan Chains on Cellulose via Adaptive Coupling to Emergent Substrate Structures

Yuchun Lin^{†,‡}, Gregg T. Beckham^{§,⊥}, Michael E. Himmel^{||}, Michael F. Crowley^{||},

and Jih-Wei Chu^{†,∇,○,*}

[†]Department of Chemical and Biomolecular Engineering, University of California, Berkeley, Berkeley, California, United States

[‡]State Key Laboratory of Oral Diseases, West China Hospital of Stomatology, Sichuan University, Chengdu, Sichuan, 610041, PRC

[§]National Bioenergy Center, National Renewable Energy Laboratory, Golden, CO, USA

[⊥]Department of Chemical Engineering, Colorado School of Mines, Golden, CO, USA

^{||}Biosciences Center, National Renewable Energy Laboratory, Golden, CO, USA

[∇]Department of Biological Science and Technology, National Chiao Tung University, Hsinchu, Taiwan, ROC

[○]Institute of Bioinformatics and Systems Biology, National Chiao Tung University, Hsinchu, Taiwan, ROC

Keywords: Interfacial biocatalysis; endoglucanase; cellulose; complexation; free-energy simulation; coarse-graining; multiscale simulation

Fig. S1 Segments strongly couple to the 11 glucose residues of the TGC as a function of ξ during the two pathways of complexation. -7: non-reducing end of TGC; +4: reducing end of TGC. Bold:

*Corresponding Author: Jih-Wei Chu. Address: 208 Lab Building 1, 75 Bo-Ai St., National Chiao Tung University, Hsinchu, Taiwan, ROC. Tel: 886-3-5712121 ext: 56996. Email: jwchu@nctu.edu.tw.

strength of mechanical coupling is larger than $2.5 \text{ kcal/mol/\AA}^2$; Light: strength of mechanical coupling is larger than $1.5 \text{ kcal/mol/\AA}^2$. Nearby segments are colored as blue: L1 and HP1; red: L7; green: L2 and $\beta 1$; black: others.

ξ	Scooping											Threading											
	-7	-6	-5	-4	-3	-2	-1	+1	+2	+3	+4	-7	-6	-5	-4	-3	-2	-1	+1	+2	+3	+4	
0.00							L7																
0.03	HP1						L7															L7	
0.07	HP1						L7									HP1						L7	
0.10	HP1						L7	L7								HP1						L7	
0.13							L7									HP1							
0.17							L7												HP1				
0.20			HP1				L7																
0.23							L7											HP1	HP1				
0.27							L7	L7															
0.30							L7	L7									HP1	HP1					
0.33							L7										HP1						
0.37			HP1				L7										HP1			L7	L7	L7	
0.40			HP1	HP1			L7							HP1						L7	L7	$\beta 4, L7$	
0.43			HP1	HP1										HP1						L7	L7	$\beta 4, L7$	
0.47		HP1	HP1	HP1					L7							L1				L7	L7	$\beta 4$	
0.50		HP1							L7							L1			L1, L2	L2, $\beta 1$	L7	L7	
0.53		HP1	HP1	HP1			L7			$\beta 6$				HP1	L1, HP1	L1, HP1		L1, L2	L2, $\beta 1$		L7	L7	
0.57			HP1				L7			$\beta 6$				HP1	HP1	HP1		L1, L2	L2, $\beta 1$	L7	L7	L7	
0.60			HP1		L1		L7	L7		$\beta 6$				HP1	L1, HP1	HP1	L1, HP1	L1	L7	L7	L7	L7	
0.63			HP1		L1		L7	L7, $\beta 4$	L7					L1, HP1	L1, HP1	L1, HP1	L1	L1	L7	L7	L7		
0.67			HP1		L1	L7	L7	L7	L7					L1, HP1	L1, HP1	L1, HP1	L1	L1, L2	L7	L7	L7		
0.70		HP1	HP1	L1		L7	L7	L7, $\beta 4$	L7					L1, HP1	L1, HP1	L1, HP1	L1, L2	L1, L2		L7	L7		
0.73		HP1	HP1	L1	L1, L2, $\beta 1$	L7	L7	L7	L7					L1, HP1	L1, HP1	L1, HP1	L1, L2	L1, L2, $\beta 1$			L7		
0.77		HP1	HP1	L1	L1, L2, $\beta 1$	L7, $\beta 1$	L7	L7	L5, L7					L1, HP1	L1	L1, HP1	L1	L1, $\beta 1$	L7	L7	L7		
0.80		HP1	HP1	L1	L1, L2, $\beta 1$	L7, $\beta 1$	L7	L7, $\beta 6$	L7					L1, HP1	L1	L1, HP1	L1, L2	L1, $\beta 1$	L7	L7	L7		
0.83		L1, HP1	HP1	L1	L1, L2, $\beta 1$	L7, $\beta 1, \beta 2$	L7	$\beta 4, \beta 6$	L5, $\beta 6$					L1	L1, HP1	L1	L1, L2, $\beta 1$	L7, $\beta 1$	L7	L7	L7		
0.87		L1, HP1	L1, HP1	L1	L1, L2, $\beta 1$	L7, $\beta 1, \beta 2$	L7	$\beta 4, \beta 6$	L5, $\beta 6$	L5				HP1	L1, HP1	L1, HP1	L1, L2	L1, L2, $\beta 1$	L7, $\beta 1$	L7		L7	
0.90	HP1	L1, HP1	L1, HP1	L1	L1, L2, $\beta 1$	L7, $\beta 1, \beta 2$	L7, $\beta 3$	$\beta 4, \beta 6$	L5, $\beta 6$	L5				HP1	HP1	L1, HP1	L1, L2	L7, L2, $\beta 1$	L7, $\beta 1$	$\beta 4, L7$	L5, L7	L7	
0.93	L1, HP1	L1, HP1	L1, HP1	L1	L1, L2, $\beta 1$	L7, $\beta 1, \beta 2$	L7, $\beta 3$	$\beta 4, \beta 6$	L5, $\beta 6$	L5				L1, HP1	HP1	L1	L1, L2	L7, L2, $\beta 1, \beta 2$	L7, $\beta 3$	L7	L5, L7	L7	
0.97	L1, HP1	L1, HP1	L1, HP1	L1	L1, L2, $\beta 1$	L7, $\beta 1, \beta 2$	L7, $\beta 3$	$\beta 4, \beta 6$	L5, $\beta 6$	L5				L1, HP1	L1	L1, HP1	L1	L1, L2, $\beta 1$	L7, $\beta 1, \beta 2$	L7, $\beta 3$	L7, $\beta 4$	L5, L7, $\beta 6$	
1.00	L1, HP1	L1, HP1	L1, HP1	L1	L1, L2, $\beta 1$	L7, $\beta 1, \beta 2$	L7, $\beta 3$	$\beta 4, \beta 6$	$\beta 6$	L5				L1, HP1	L1, HP1	L1, HP1	L1	L1, L2, $\beta 1$	L7, $\beta 1, \beta 2$	L7, $\beta 3$	$\beta 4, \beta 6$	$\beta 6$	L5

Fig. S2 Segments strongly couple to the glucose sites of NNC in parallel to the 11 glucose residues of the TGC as a function of ξ during the two pathways of complexation. Numberings from -7 to +4 represent residues in NNC parallel to the targeted glucan chain. -7: non-reducing end of TGC; +4: reducing end of TGC. Bold: strength of mechanical coupling is larger than 2.5 kcal/mol/Å²; Light: strength of mechanical coupling is larger than 1.5 kcal/mol/Å².

ξ	Scooping											Threading										
	-7	-6	-5	-4	-3	-2	-1	+1	+2	+3	+4	-7	-6	-5	-4	-3	-2	-1	+1	+2	+3	+4
0.00	HP1			HP1		L7																
0.03	HP1		HP1	HP1		L7	L7															L7
0.07	HP1		HP1			L7	L7	L7										L2		L7	L7	
0.10						L7	L7	L7										HP1, L2		L7	L7	
0.13						L7	L7	L7										HP1, L2		L7	L7	
0.17						L7	L7															L7
0.20						L7	L7															L7
0.23				HP1		L7	L7										HP1		HP1			
0.27				HP1		L7	L7										HP1	HP1	HP1			L7
0.30				HP1		L7	L7										HP1	HP1	HP1		L7	L7
0.33						L7	L7										HP1	HP1			L7	L7
0.37			HP1	HP1		L7	L7	L7							HP1		HP1	HP1		L7	L7	L7, L8
0.40			HP1	HP1		L7	L7	L7							HP1		HP1	HP1			L7	L7, L8
0.43		HP1	HP1	HP1		L7	L7								HP1	HP1	HP1	HP1			L7	L7, L8
0.47		HP1	HP1	HP1		L7	L7	L7							HP1		HP1	HP1		L7	L7	
0.50		HP1	HP1	HP1		L7	L7	L7	L8						HP1	HP1		HP1	HP1		L7	L7
0.53			HP1	HP1		L7	L7	L7	L8						HP1	HP1	HP1	HP1	HP1	L7	L7	L7
0.57			HP1			L7	L7	L7	L8						HP1	HP1	HP1	HP1		L7	L7	L7
0.60			HP1	HP1	L7	L7	L7	L7	L8						HP1	HP1	HP1	HP1	L7	L7	L7	L7
0.63			HP1	HP1	L7	L7	L7	L7	L8					HP1		HP1	HP1	HP1	L7	L7	L7	L8
0.67		HP1	HP1	HP1	L7	L7	L7	L7	L8					HP1	HP1	HP1	HP1	HP1	L7	L7	L7	L8
0.70		HP1	HP1	HP1		L7	L7	L7	L8					HP1		HP1	HP1	HP1	L7	L7	L7	
0.73		HP1	HP1			L7	L7								HP1	HP1		L7	L7	L7		
0.77		HP1	HP1			L7	L7								HP1	HP1		L7	L7	L7		
0.80		HP1	HP1			L7	L7								HP1	HP1			L7	L7		
0.83		HP1	HP1	HP1		L7	L7								HP1	HP1			L7	L7		
0.87		HP1	HP1	HP1		L7	L7								HP1	HP1	HP1		L7	L7		
0.90		HP1	HP1	HP1		L7	L7								HP1	HP1	HP1		L7	L7	L7	
0.93		HP1	HP1			L7	L7					HP1	HP1	HP1	HP1		L7	L7	L7			
0.97		HP1	HP1	HP1		L7	L7	L7				HP1	HP1	HP1	HP1		L7					
1.00		HP1	HP1	HP1		L7	L7	L7					HP1	HP1	HP1	HP1		L7	L7	L7		

Fig. S3 Segments strongly couple to the glucose sites of RNC in parallel to the 11 glucose residues of the TGC as a function of ξ during the two pathways of complexation. Numberings from -7 to +4 represent residues in RNC parallel to the targeted glucan chain. -7: non-reducing end of TGC; +4: reducing end of TGC. Bold: strength of mechanical coupling is larger than 2.5 kcal/mol/Å²; Light: strength of mechanical coupling is larger than 1.5 kcal/mol/Å².

ξ	Scooping											Threading										
	-7	-6	-5	-4	-3	-2	-1	+1	+2	+3	+4	-7	-6	-5	-4	-3	-2	-1	+1	+2	+3	+4
0.00						H3	H3															H3
0.03			L2			H3	H3														H3	H3
0.07						H3	H3												L2		H3	H3
0.10						H3	H3	H3									HP1	L2			H3	H3
0.13						H3	H3										HP1	HP1, L2				H3
0.17							H3										HP1	HP1	HP1			
0.20																	HP1	HP1				
0.23																	HP1	HP1				H3
0.27																						H3
0.30																					H3	H3
0.33																						
0.37																					H3	H3
0.40																					H3	H3
0.43																					H3	H3
0.47																HP1						H3
0.50															HP1	HP1	HP1					
0.53															HP1		HP1					
0.57															HP1							
0.60							H3															
0.63																						
0.67																				H3		
0.70																						
0.73							H3	H3														
0.77							H3	H3														
0.80								H3												H3		
0.83							H3	H3					HP1							H3		
0.87							H3	H3					HP1					L7				
0.90							H3	H3	L8				HP1					L7				
0.93	HP1							H3										L7				
0.97		HP1						H3	L8			HP1	HP1	HP1				L7				
1.00	HP1	HP1	HP1				L7					HP1	HP1	HP1				L7				

Fig. S4 The coupling strengths of the HP1 and L7 segments of Cel7B CD to the TGC and NNC parts of cellulose during the threading pathway of complexation. The profiles of individual protein residues with strong coupling strengths are also shown to illustrate the underlying details.

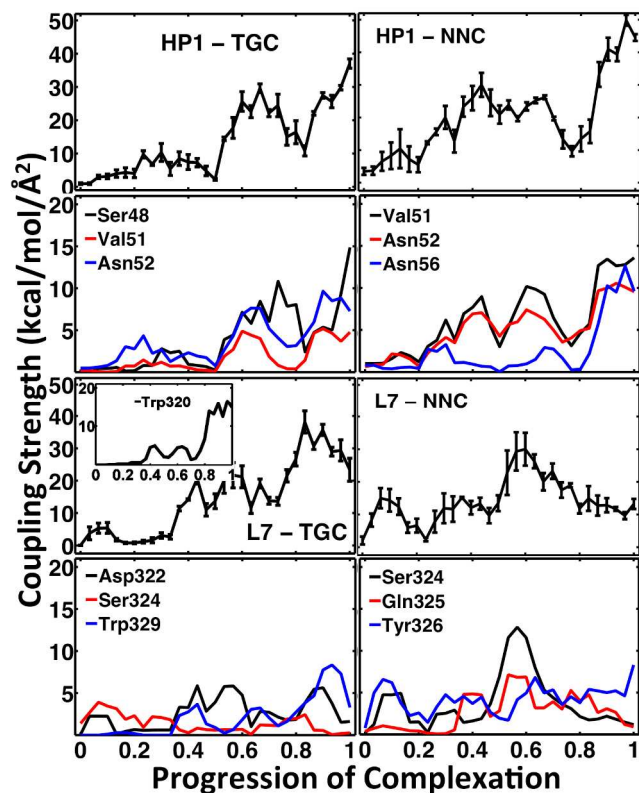


Fig. S5 The coupling strengths of Cel7B CD to different parts of cellulose in all the three models: (Top) coupling of Cel7B CD to TGC and NNC; (Middle) coupling of Cel7B CD to RNC; (Bottom) coupling of Cel7B CD to whole cellulose.

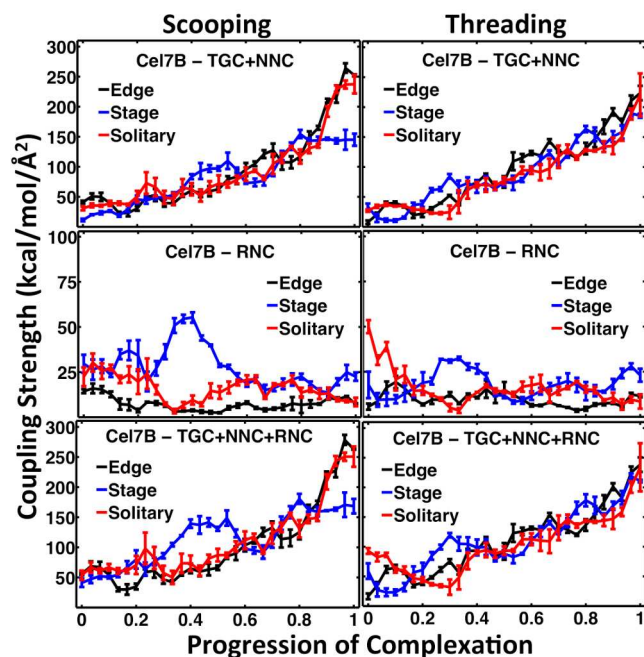


Fig. S6 Percentage of residue at specific positions in the globally aligned 340 homologues of the Cel7B CD collected from the NCBI database. Residue name is represented by 1-Letter; residue missed in homology is represented by dash symbol.

	HP1			L2			L7				L8	
	S48	V51	N52	Y94	S99	Y102	N323	S324	Q325	Y326	N334	S340
A	0.29	1.18	0.00	0.00	2.35	2.35	4.71	19.71	34.12	1.18	25.88	9.12
C	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.47	0.00	0.00
D	0.88	25.88	1.77	0.29	1.18	0.59	8.53	2.94	4.41	4.41	8.53	0.88
E	0.29	1.47	0.59	0.29	0.59	1.47	5.29	1.18	0.29	1.18	9.71	1.47
F	0.00	0.00	0.88	2.65	0.29	0.00	0.29	0.29	0.29	6.77	0.00	0.00
G	3.53	0.29	0.88	0.00	29.12	29.12	0.29	9.41	10.59	12.06	4.71	5.59
H	0.00	0.00	0.00	0.29	6.47	0.59	42.65	4.71	0.00	2.35	0.29	1.18
I	0.00	0.00	0.00	5.00	0.29	1.18	0.00	0.00	0.00	0.00	0.29	5.59
K	1.18	0.29	0.00	1.47	0.00	4.71	2.35	0.29	0.00	1.18	4.12	2.35
L	0.00	0.29	0.00	6.77	0.00	0.29	0.29	0.29	0.88	0.00	1.77	0.29
M	0.29	0.00	0.00	0.59	0.00	0.00	0.00	0.59	0.00	0.00	0.00	0.00
N	59.12	2.35	4.71	0.00	5.88	4.12	9.12	3.82	1.18	48.53	15.29	2.06
P	0.29	0.00	1.18	0.00	0.88	0.00	2.94	0.29	0.29	0.29	2.65	4.41
Q	0.00	0.29	0.00	0.00	0.00	14.12	1.77	1.18	5.88	2.94	2.94	2.65
R	0.00	0.00	0.29	0.88	0.00	0.59	0.29	0.29	0.59	0.00	0.29	1.47
S	13.24	3.53	5.29	0.00	19.12	22.65	0.59	12.65	13.53	6.47	10.00	17.35
T	0.29	38.24	0.29	0.29	2.06	2.06	0.29	5.59	1.77	0.00	2.06	38.24
V	0.00	6.47	0.29	64.41	0.00	0.59	0.88	1.18	13.53	0.00	6.47	2.35
W	0.00	0.00	5.00	0.88	1.47	0.00	0.00	0.00	0.00	1.18	0.00	0.29
Y	0.29	0.00	0.00	11.47	3.82	7.65	13.53	18.53	0.00	6.47	0.88	0.00
-	19.41	19.71	78.82	4.71	26.47	7.94	6.18	17.06	12.65	4.41	4.12	4.71

Fig. S7 Alignment of the sequences of the residue blocks: position 50 to 47 in segments HP1 and position 323 to 327 in L7 of EG I (Cel7B: 1EG1) and CBH I (Cel7A: 1CEL) from *T. reesei*.

	Residues 50~57 in HP1	Residues 323~326 in L7
Cel7B	TVNGGVNT	NSQY
Cel7A	YD-GTWSS	YYAN

SUPPORTING MOVIE LEGENDS

Movie S1. The movie of scooping pathway viewed from side view.

Movie S2. The movie of scooping pathway viewed from top view.

Movie S3. The movie of threading pathway viewed from side view.

Movie S4. The movie of threading pathway viewed from top view.