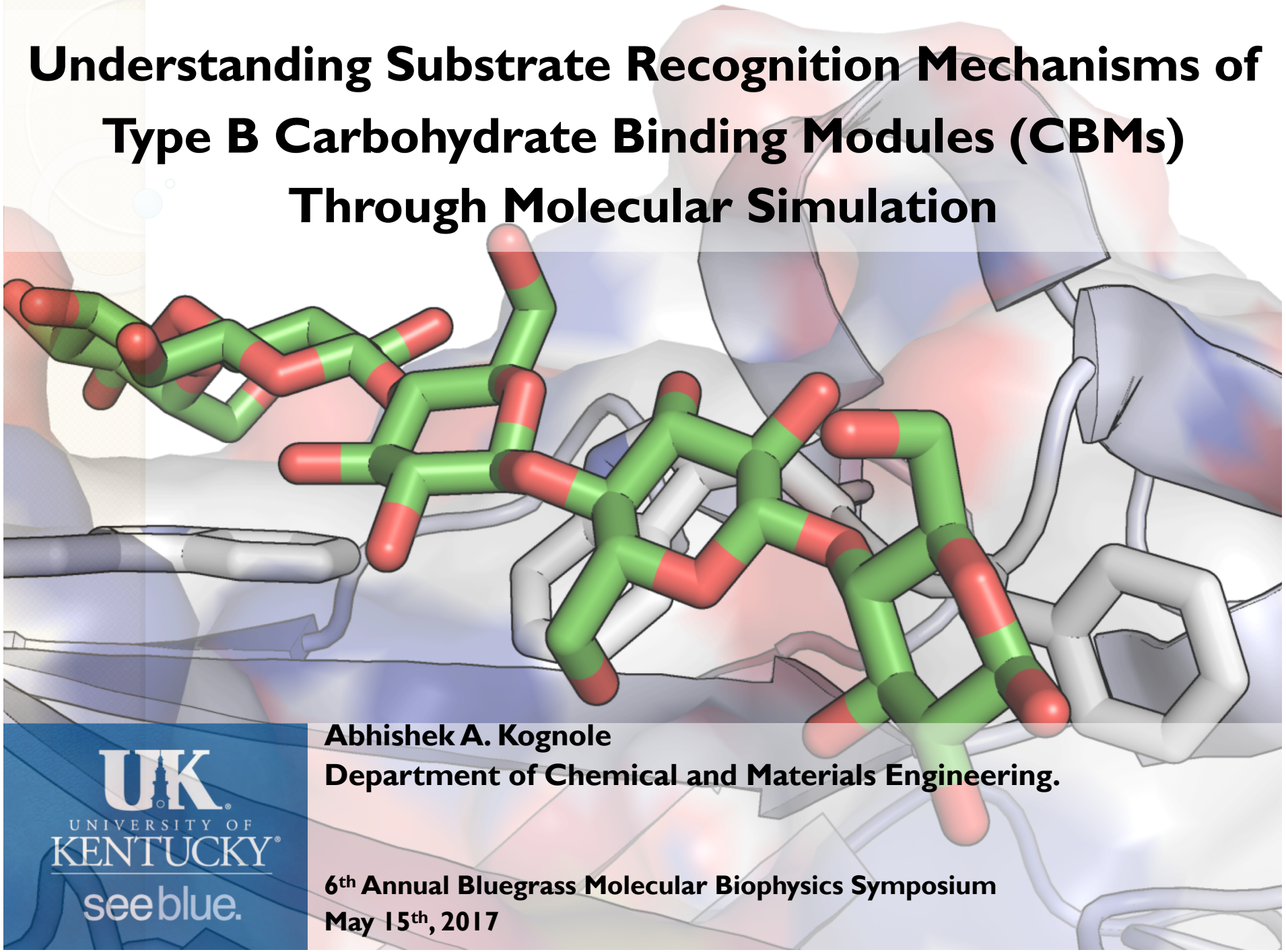
Understanding Substrate Recognition Mechanisms of Type B Carbohydrate Binding Modules (CBMs) Through Molecular Simulation

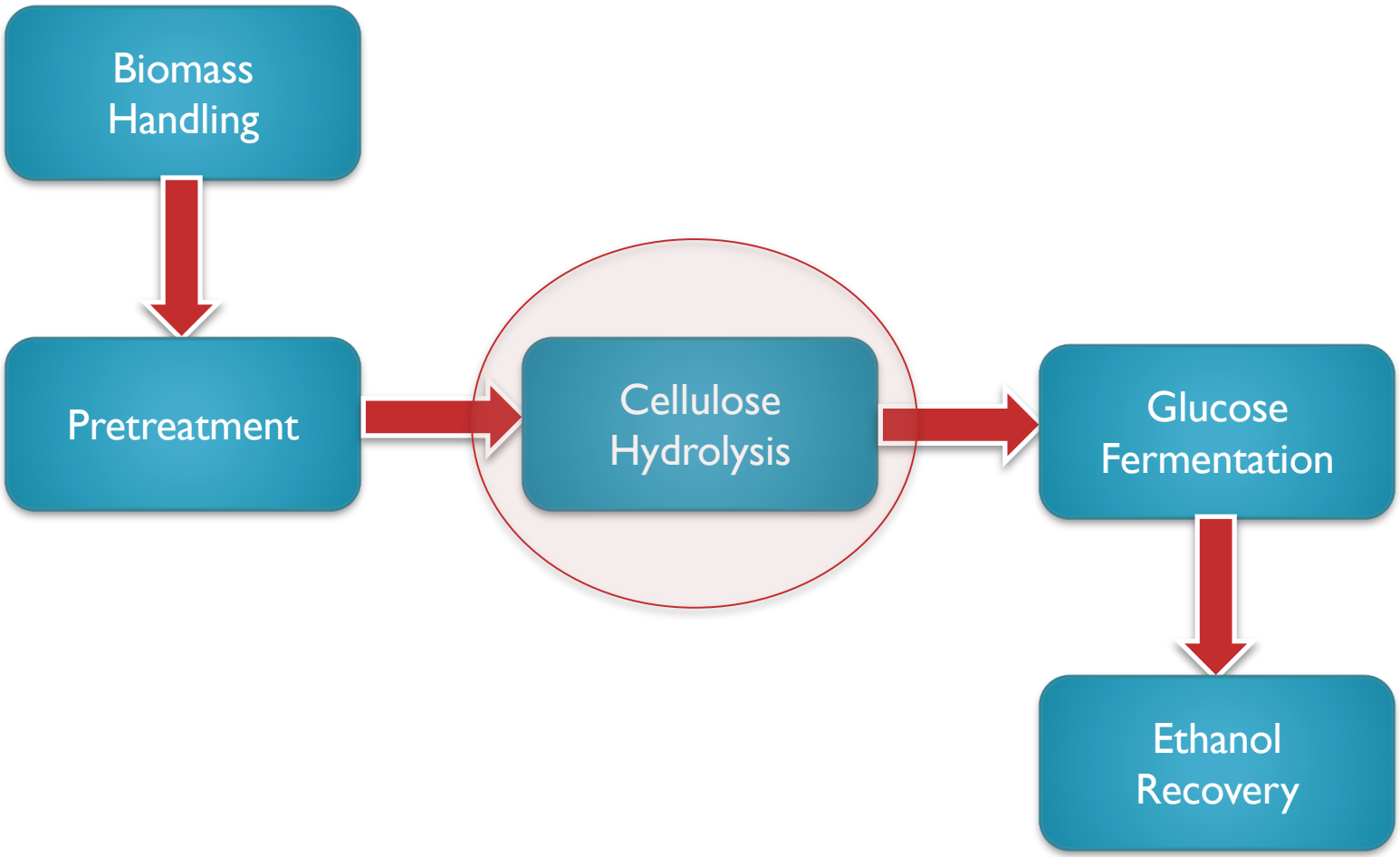


Abhishek A. Kognole
Department of Chemical and Materials Engineering.

UNIVERSITY OF KENTUCKY
seeblue.

6th Annual Bluegrass Molecular Biophysics Symposium
May 15th, 2017

Introduction: Biochemical biomass conversion



Background

Sustainable Energy

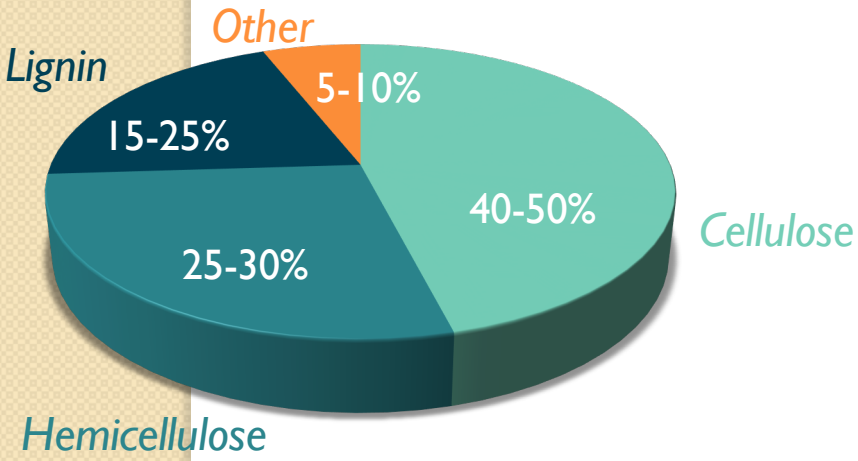
↓

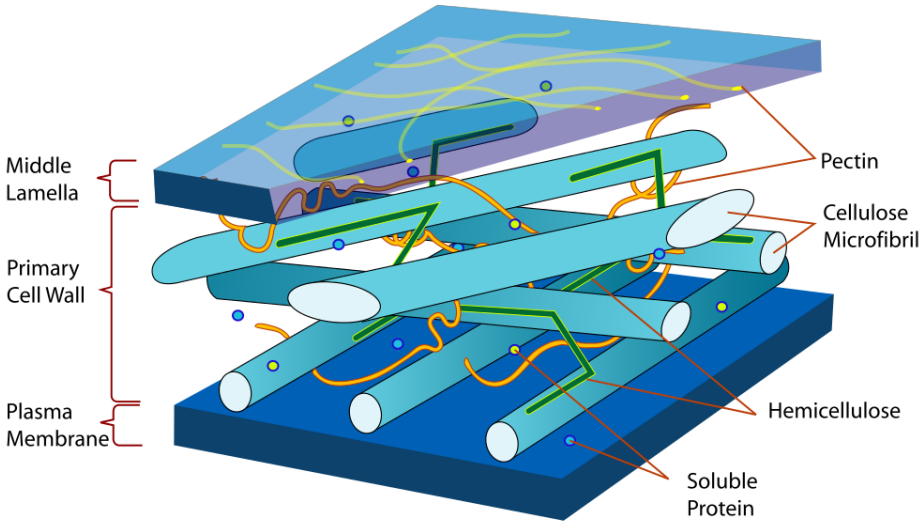
Biofuels

↓

Biomass - Plant Cell Wall

Image: Singh et al., *Renewable and Sustainable Energy Reviews*, 2014.





Background

Sustainable Energy

↓

Biofuels

↓

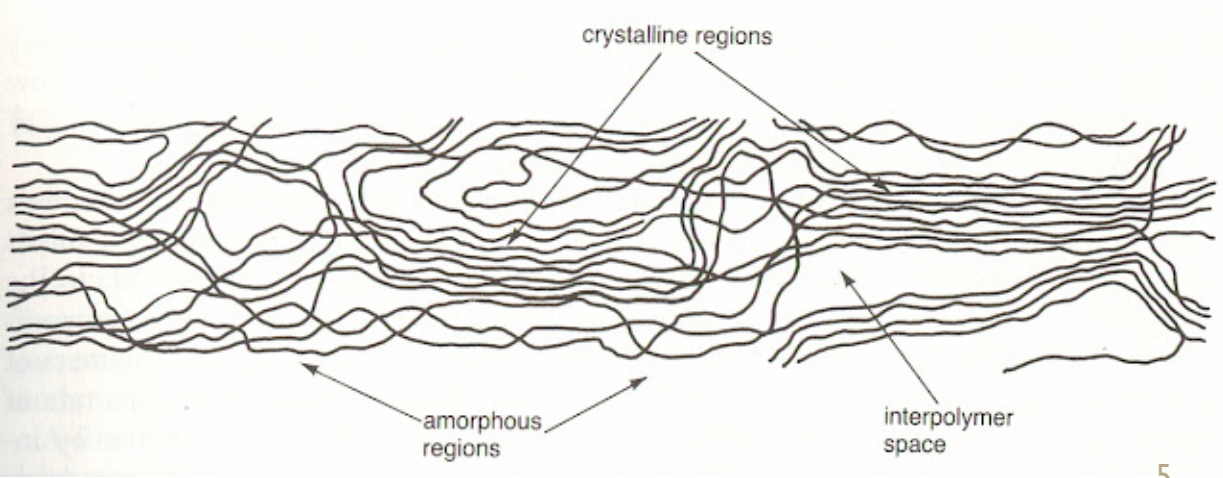
Biomass - Plant Cell Wall

↓

Pretreated Biomass

Cellulose morphologies after pretreatment:

- Cello-oligosaccharides
- Amorphous regions
  - Separate insoluble polysaccharides
  - Convoluted insoluble polysaccharides
  - Partially decrystallized polysaccharides
- Crystalline regions



Gohl and Vilensky. *Textile Sci.* 1989

Background

Sustainable Energy

↓

Biofuels

↓

Biomass - Plant Cell Wall

↓

Pretreated Biomass

↓

Enzymatic Hydrolysis

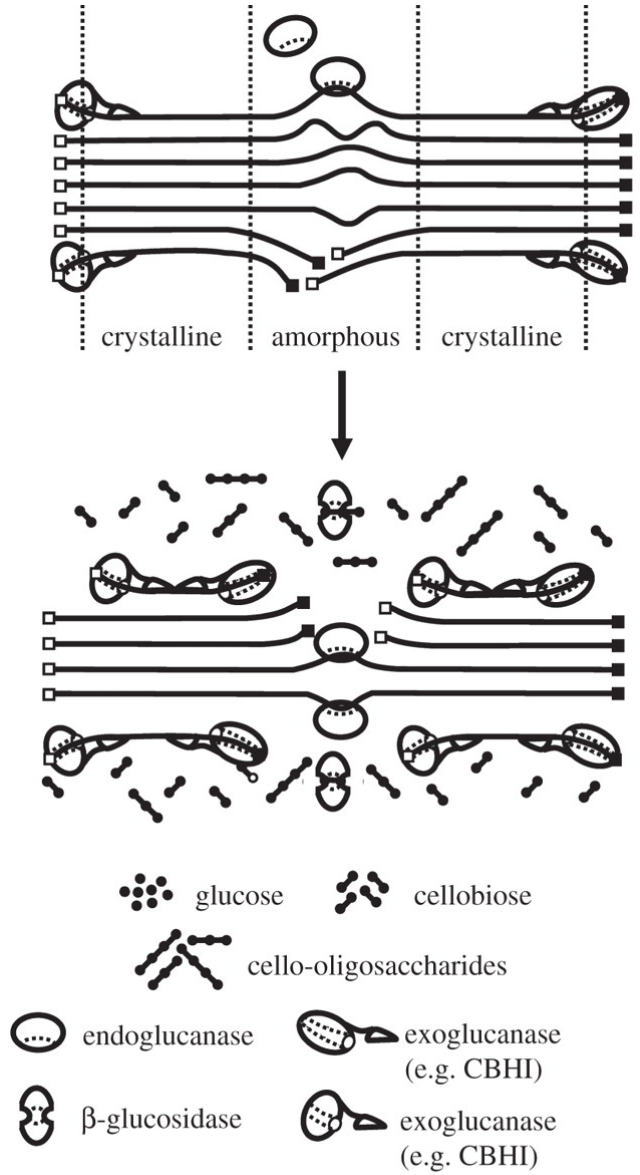
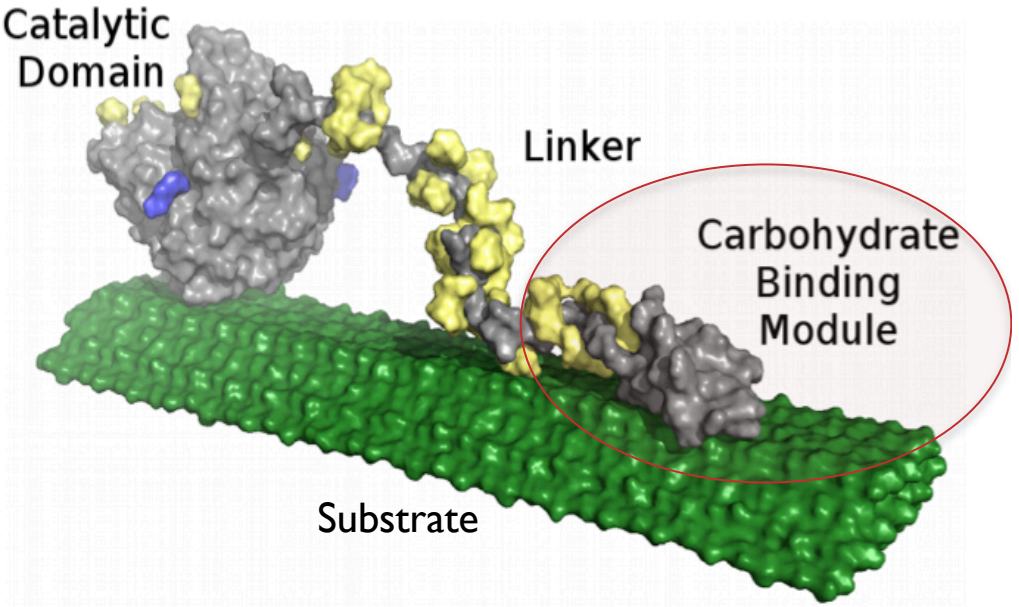
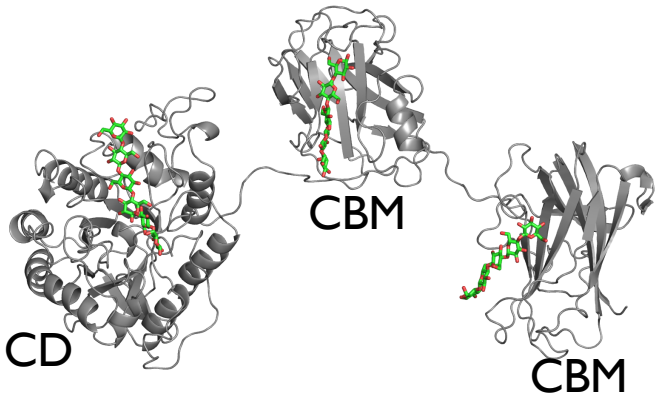


Image: Lynd et al., *Microbiology and Molecular Biology Reviews*, 2002.

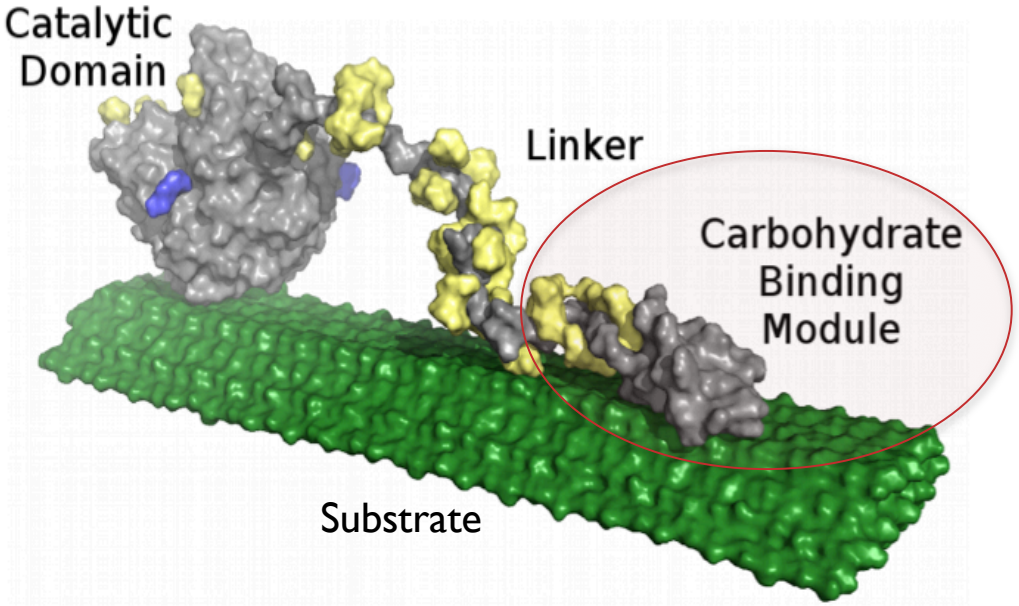
Carbohydrate Binding Modules (CBMs)

- Sustainable Energy
- Biofuels
- Biomass - Plant Cell Wall
- Pretreated Biomass
- Enzymatic Hydrolysis
- Multi-modular Glycoside Hydrolases
- Carbohydrate Binding Modules





Carbohydrate Binding Modules (CBMs)



Functions of CBM:

1. Maintain proximity to substrate
2. Target specific regions
3. Disrupt surface crystallinity

Other Applications:

- Bioprocessing
- Cell immobilization
- Protein Engineering

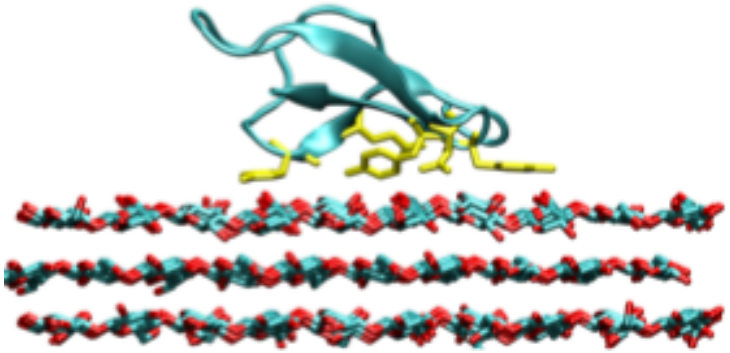
Boraston et al., *Biochem. J.*, 2004

Shoseyov et. al., Microbio. & Mol. Bio. Reviews, 2006

Different Substrates – Different Types

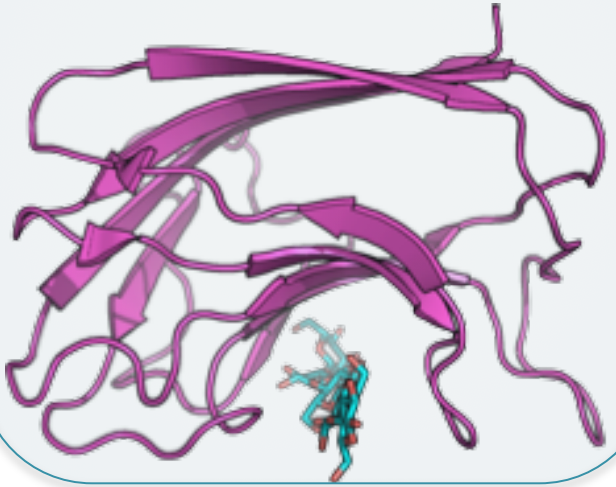
Type A

(crystalline polysaccharides)



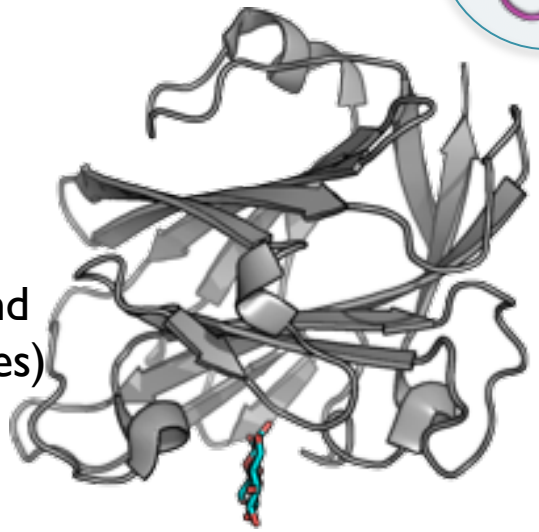
Type B

(single glycan chains)



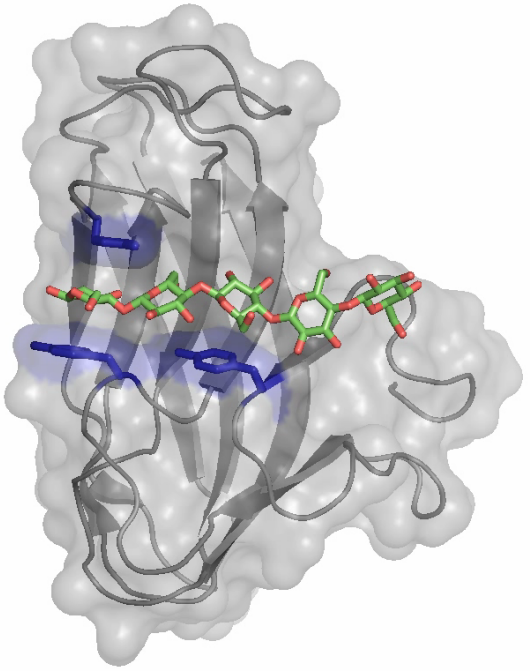
Type C

(glycan chain termini and mono-, di-, tri-saccharides)



Boraston et al., *Biochem. J.*, 2004
Gilbert et al., *Curr. Opin. Struct. Biol.*, 2013

Type B Carbohydrate Binding Modules



- Common characteristics:
- Binding site in the form of groove or cleft
- β- sandwich protein fold
- Target on single glycan chains
  - Soluble oligosaccharides
    - Cellotetraose, cellopentaose, cellohexaose etc.
  - Non-crystalline/amorphous polysaccharides
    - Individual insoluble chains (longer than oligomers)
    - Convoluted insoluble chains
    - Partially decrystallized chains

Nomenclature of CBMs

Abbreviation of CBM from certain family is CBM#, where # is its family number

Types

Based on functional activity

3 types as A, B and C

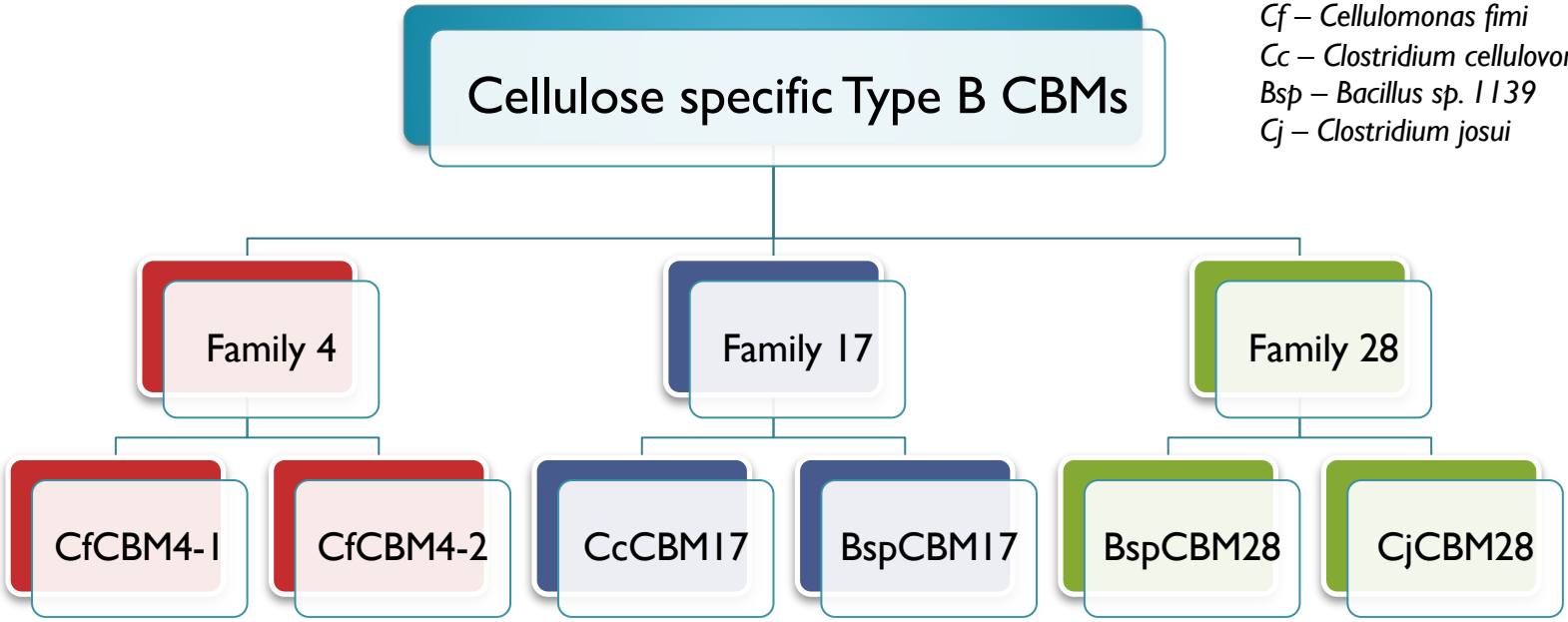
Boraston et al., *Biochem. J.*, 2004

Families

Based on peptide sequence homology

Currently 81 families in database available

http://www.cazy.org/



Cf – Cellulomonas fimi
Cc – Clostridium cellulovorans
Bsp – Bacillus sp. 1139
Cj – Clostridium josui

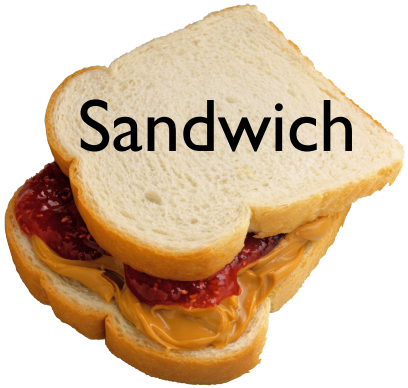
Three parts of the story…

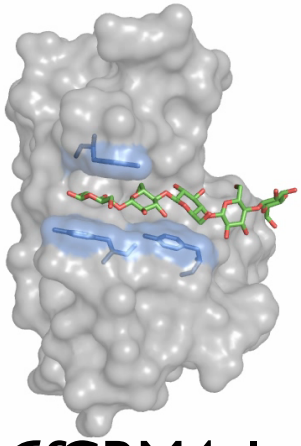
Role of binding site architecture in oligomeric substrate recognition

Bi-directional ligand binding in Type B CBMs

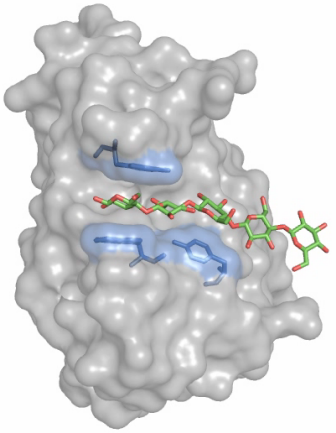
Non-crystalline substrate recognition with high- and low-affinity binding sites

Role of binding site architecture in oligomeric recognition

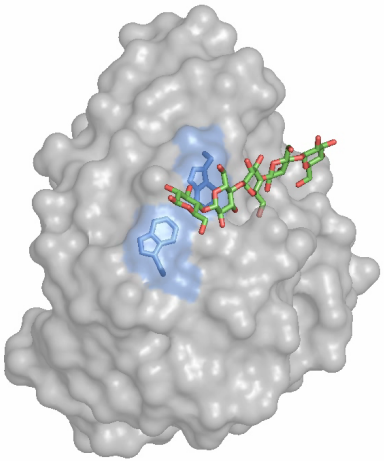




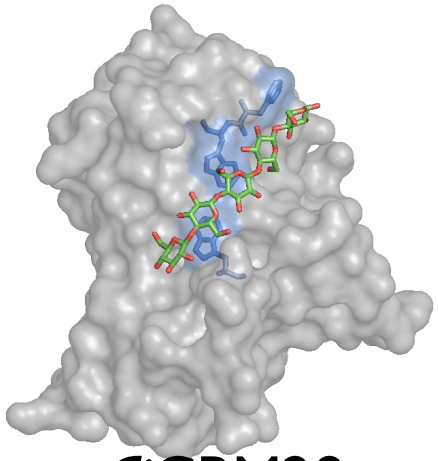
*Cf*CBM4-1



*Cf*CBM4-2



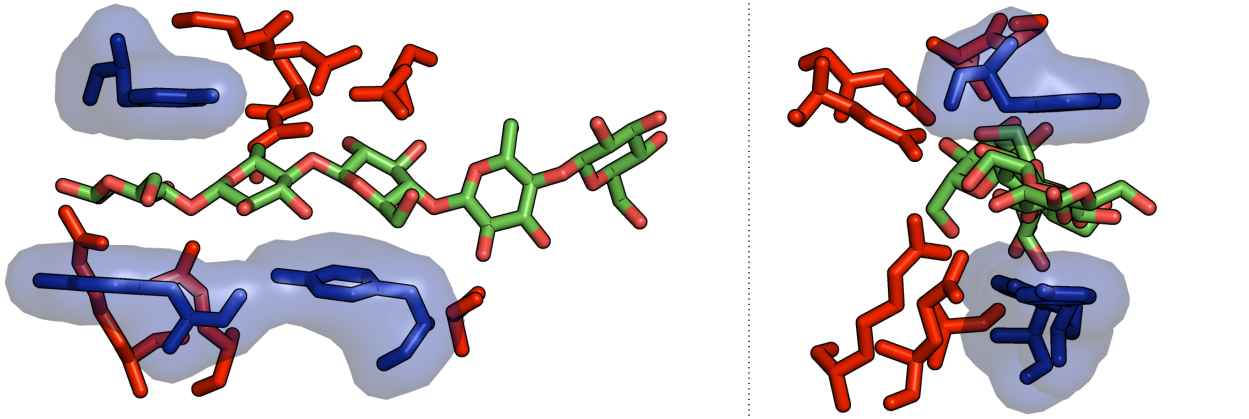
*Cc*CBM17



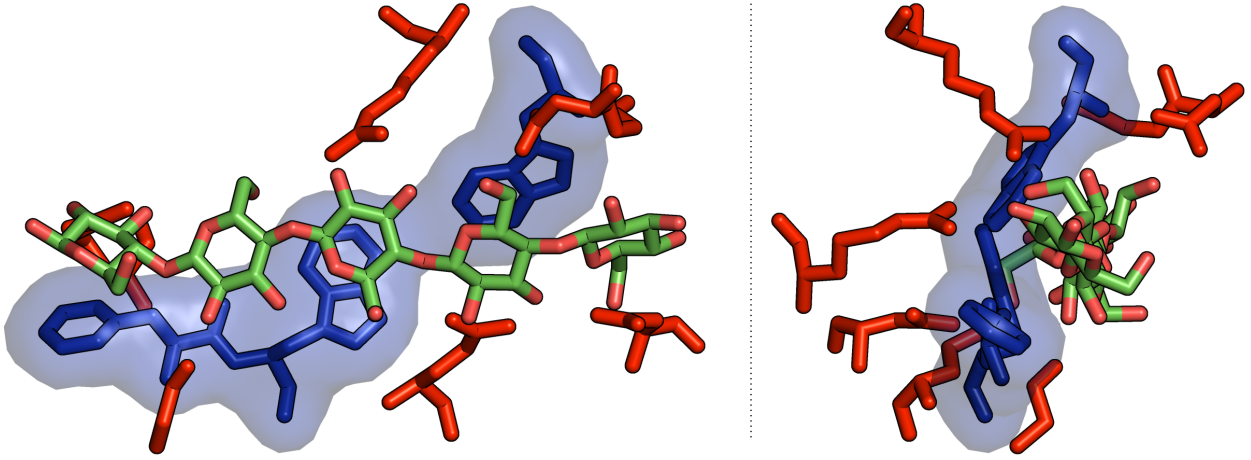
*Cj*CBM28

Differences in binding site topology and hydrogen bonding patterns

Sandwich Platform



Twisted Platform

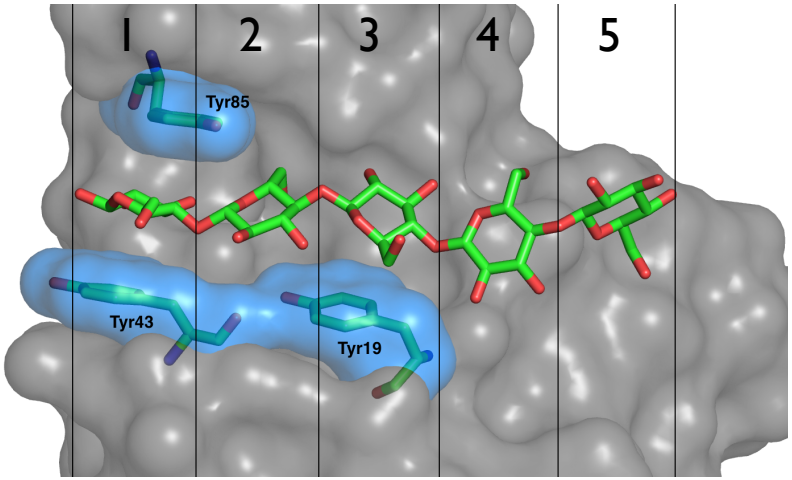


Ligand Binding Affinity

	ΔG (kcal/mol) of Cellopentaose binding to	**Experimental (ITC)**	**Computational (FEP/HREMD)**
Sandwich Platform	***Cf*CBM4-1**	- 5.24 ± 0.9 [1]	- 4.5 ± 1.3 [5]
	***Cf*CBM4-2**	- 5.80 ± 0.005 [2]	- 5.4 ± 1.3
Twisted Platform	***Cc*CBM17**	- 5.8 ± 0.025 [3]	- 6.9 ± 0.9
	***Cj*CBM28**	- 7.7 ± 0.6 [4]	- 6.3 ± 0.7

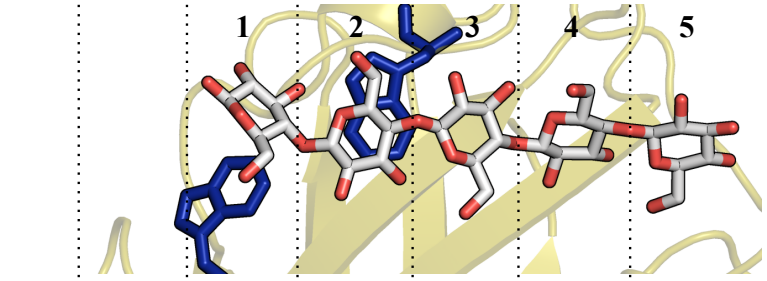
1. Tomme P, Creagh AL, Kilburn DG, and Haynes CA (1996) *Biochemistry*, **35**, 13885-13894.
2. Brun E, Johnson PE, Creagh AL, Tomme P, Webster P, Haynes CA, McIntosh LP (2000) *Biochemistry*, **39**(10), 2445-2458.
3. Notenboom V, Boraston AB, Chiu P, Freelove ACJ, Kilburn DG, Rose DR (2001) *J. Mol. Biol.*, **314,** 797-806.
4. Araki Y, Karita S, Tanaka A, Kondo M, and Goto M (2009) *Biosci. Biotechnol. Biochem.*, **73(5),** 1028-1032.
5. Kognole and Payne (2015) *Glycobiology*, **25**(10), 1100.

CfCBM4-1



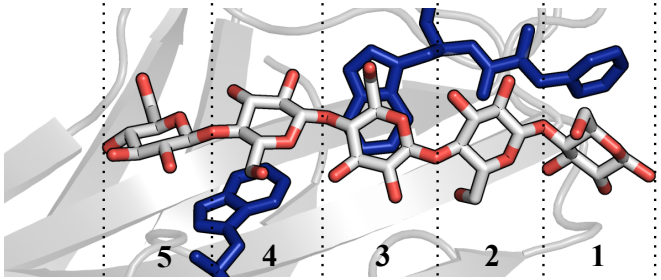
Highest affinity oligomer
- Cellopentaose

CcCBM17



Highest affinity oligomer
- Cellohexaose

CjCBM28



Extra sites available for Celloheptaose binding

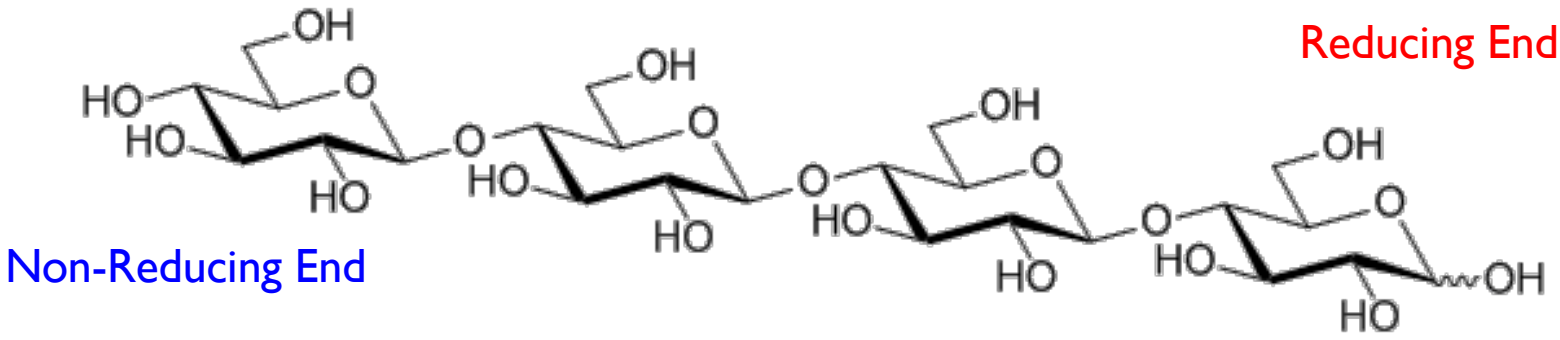
Conclusions – Sandwich vs Twisted

Open topology of twisted platform necessitates tighter binding of cello-oligomer as compared to closed sandwich platform.

Higher number of and well distributed hydrogen bonding partners along the twisted platform contribute significantly to favorable free energy change.

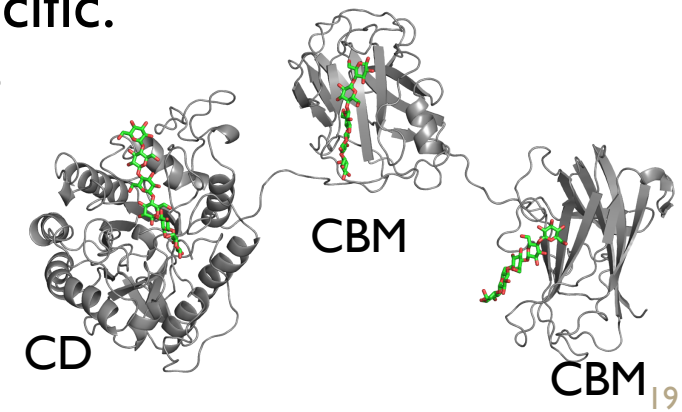
The twisted binding site may extend further to accommodate longer cello-oligomers and, ultimately, insoluble polysaccharides.

Bi-directional binding in Type B CBMs

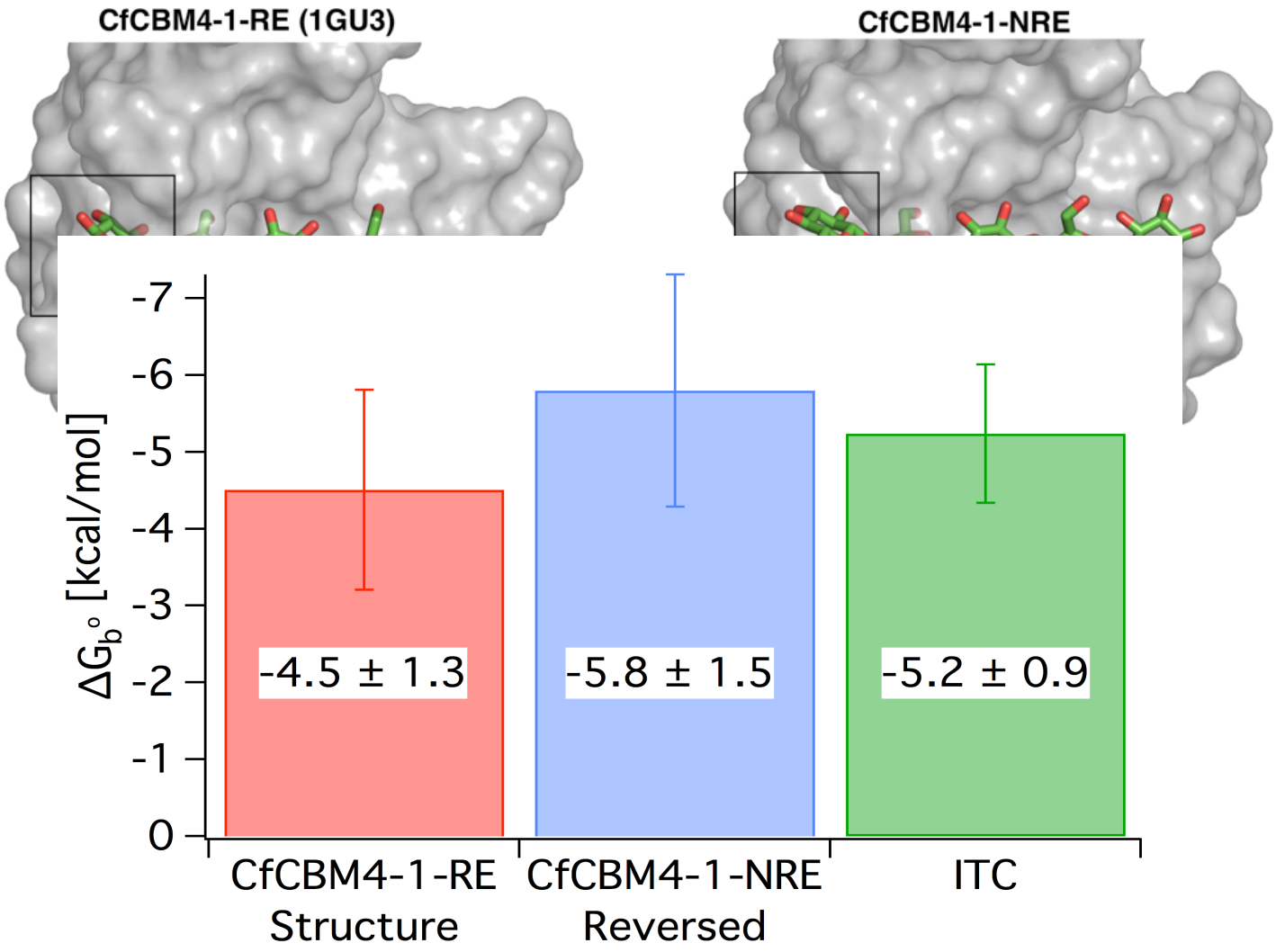


- Glucose is a reducing sugar. The polysaccharides of glucose have one reducing end and one non-reducing end.
- Catalytic domains of glycoside hydrolases are either reducing end specific or non-reducing end specific.
- What about the non-catalytic CBMs?

  Are they specific too?



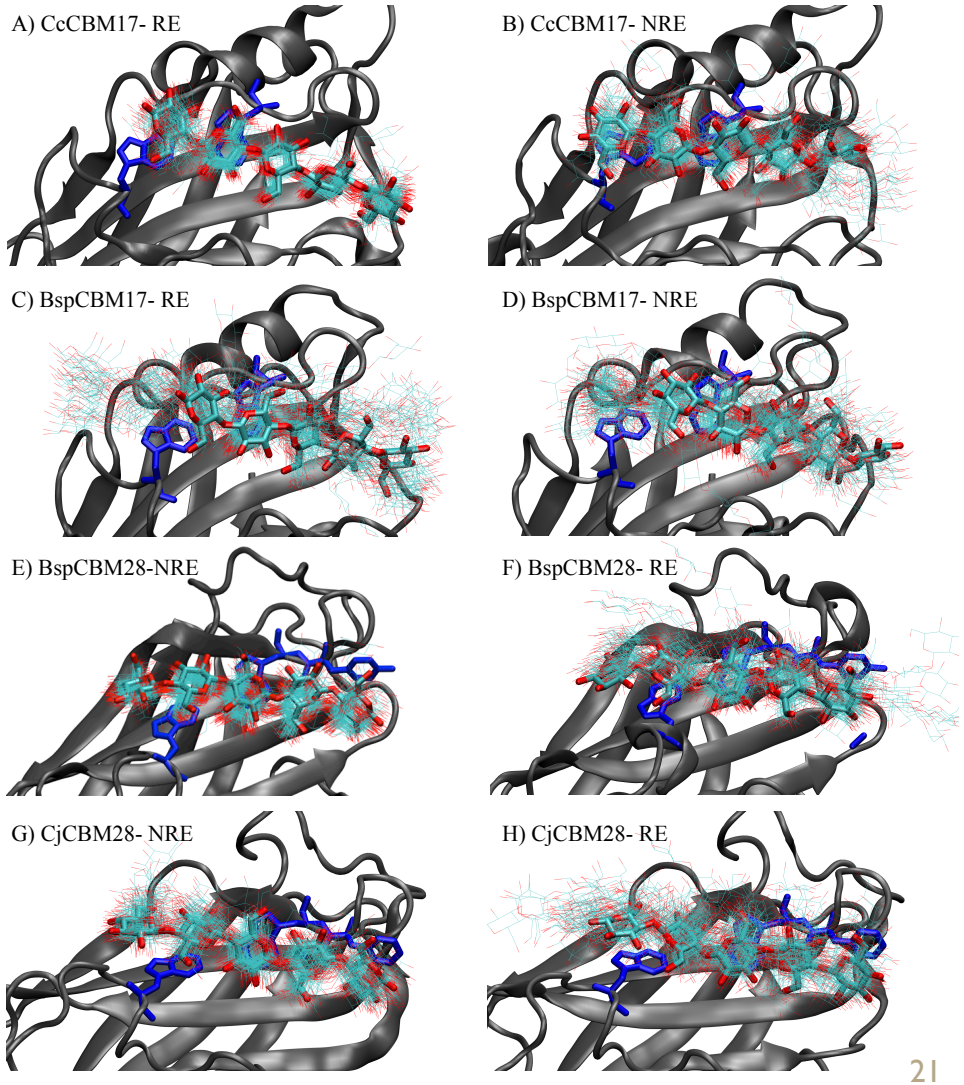
Bi-directional cello-oligomer binding in Family 4 CBMs



Kognole and Payne (2015) *Glycobiology* **25**, 10, 1100.

Bi-directional cello-oligomer binding extends to Family 17 and 28 CBMs

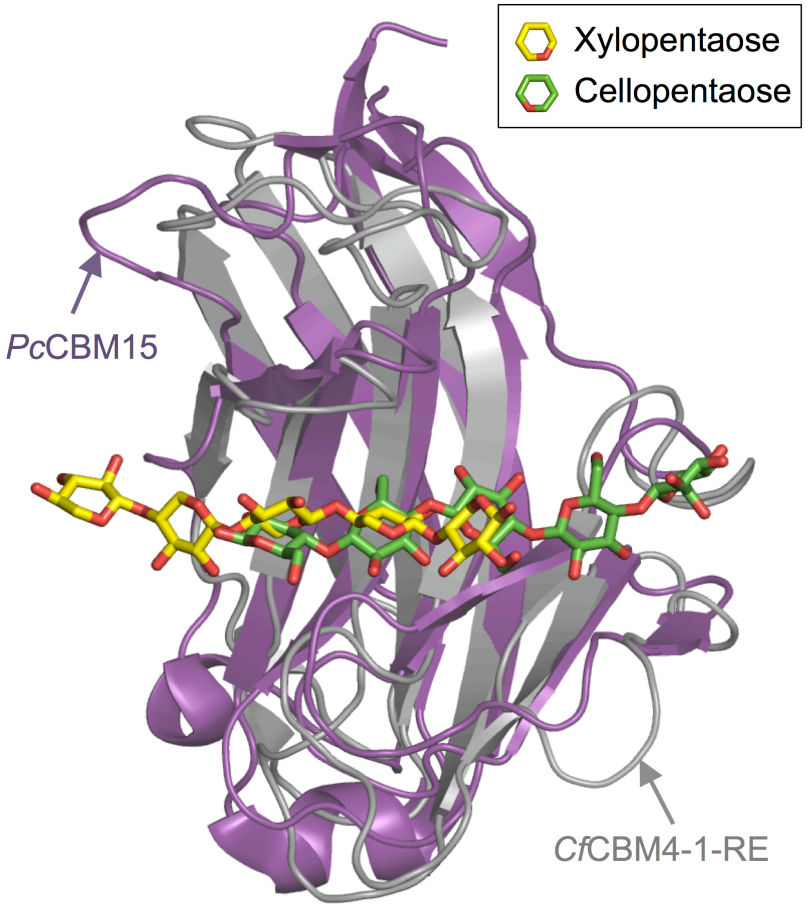
	*Cf*CBM4-1
	*Cf*CBM4-2
	*Cc*CBM17
	*Bsp*CBM17
	*Bsp*CBM28
	*Cj*CBM28



Note – 'β- sandwich' is a type of protein fold, not same as a binding site architecture.

General to β-sandwich CBMs?

- 29 of the 69 CBM families demonstrate the β-sandwich protein fold
- 10 of these 29 families have glycan bound structures available (34 structures in total)
  - 22 structures bind the ligand in the same direction as 1GU3
  - 12 structures bind the ligand in the opposite direction of 1GU3



Gray: *Cf*CBM4-1 (1GU3)
Purple: *Pseudomonas cellulosa* CBM15 (1GNY)

Kognole and Payne, *Glycobiology*, 2015

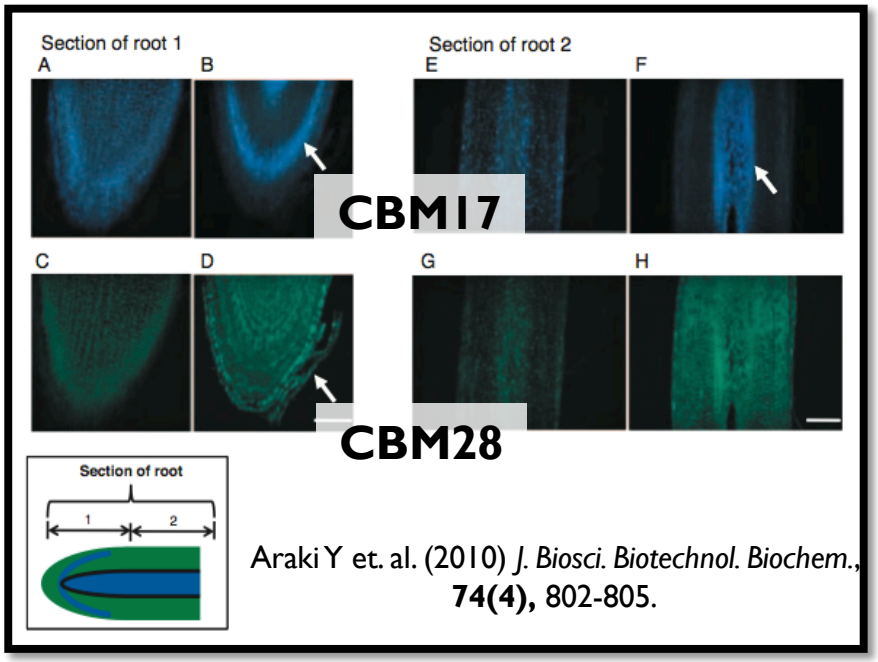
Conclusions – Bi-directionality

Cello-oligomers are recognized by family 4 CBMs in either orientation and there is no thermodynamic preference for reducing end.

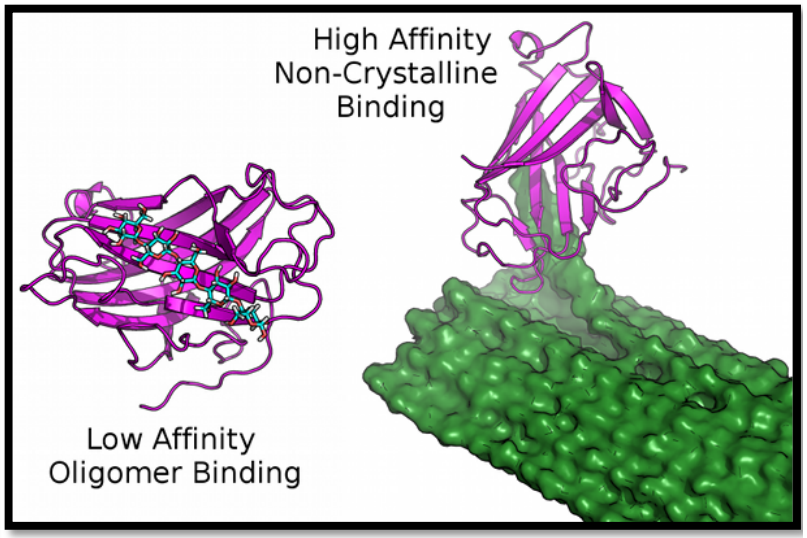
We confirm that the bi-directional binding of cello-oligomers extends to twisted platform of family 17 and family 28 CBMs.

Bi-directional binding phenomenon may not be limited to only cellulose specific Type B CBMs, potentially generalizes to all β-sandwich CBMs.

Non-crystalline substrate recognition



Araki Y et. al. (2010) *J. Biosci. Biotechnol. Biochem.*, **74(4),** 802-805.



Hypothesis

TABLE I

Adsorption parameters for the binding of CcCBM17 and BspCBM28 to Avicel™ and RC in 50 mM potassium phosphate, pH 7.0, at 25 °C

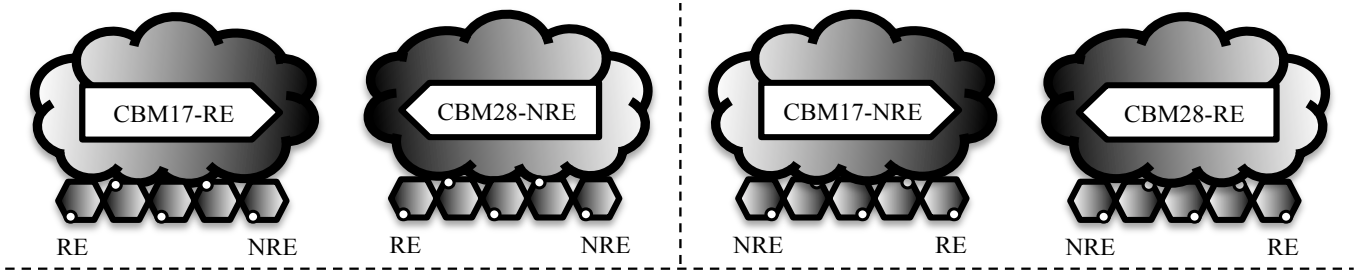
Errors represent the standard deviations of four binding experiments.

	Site 1			Site 2		
	K_{a1}	ΔG_1	$[N_1]_o$[a]	K_{a2}	ΔG_2	$[N_2]_o$[a]
	$\times 10^5$ M^{-1}	kcal/mol	μmol/g	$\times 10^5$ M^{-1}	kcal/mol	μmol/g
*Cc*CBM17						
Avicel	8.70 (±4.20)	−8.04 (±0.45)	0.26 (±0.06)	0.07 (±0.02)	−5.31 (±0.37)	5.01 (±0.88)
RC	11.30 (±1.40)	−8.41 (±0.32)	8.57 (±0.52)	0.18 (±0.05)	−5.88 (±0.36)	15.92 (±1.26)
*Bsp*CBM28						
Avicel	4.20 (±1.30)	−7.72 (±0.38)	0.08 (±0.02)	0.20 (±0.05)	−5.95 (±0.36)	0.79 (±0.05)
RC	9.90 (±2.30)	−8.28 (±0.35)	3.72 (±0.36)	0.21 (±0.07)	−5.93 (±0.38)	6.84 (±0.54)

[a] $[N_x]_o$, binding capacity, density of binding sites per gram of cellulose.

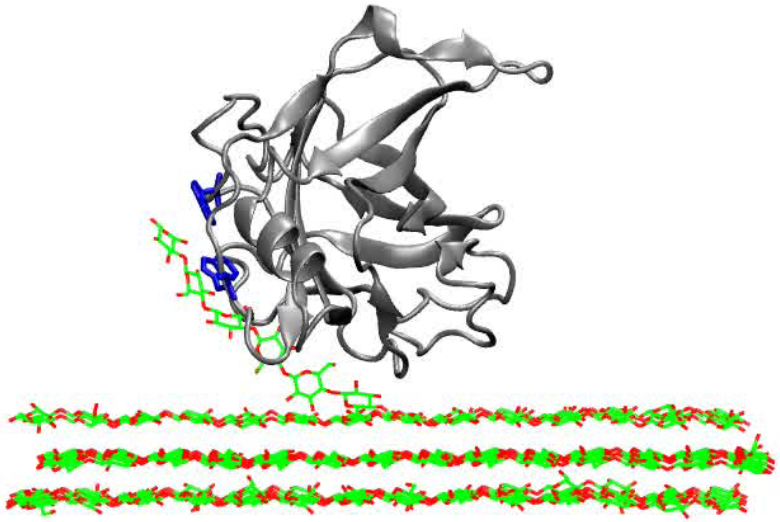
Boraston AB et. al. (2003) *J. Biol. Chem.*, **278(8),** 6120-6127.

Modeling of CBMs on non-crystalline cellulose

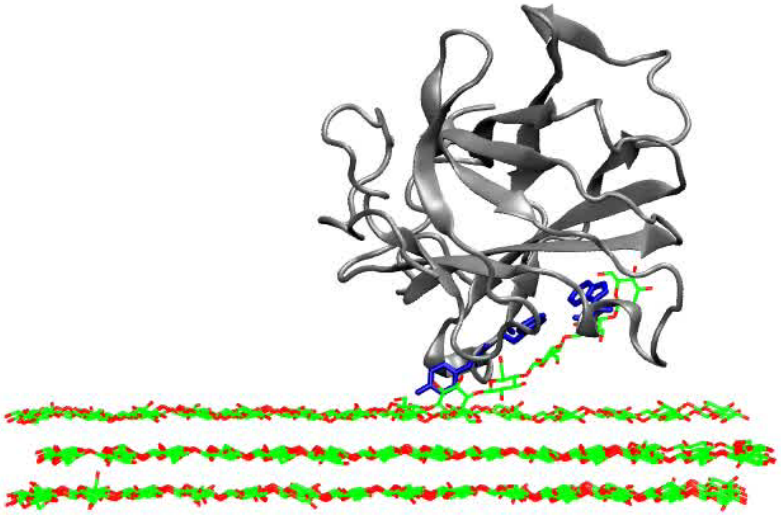


RE – Reducing End, NRE – Non-Reducing End, F – Forward, R – Reverse

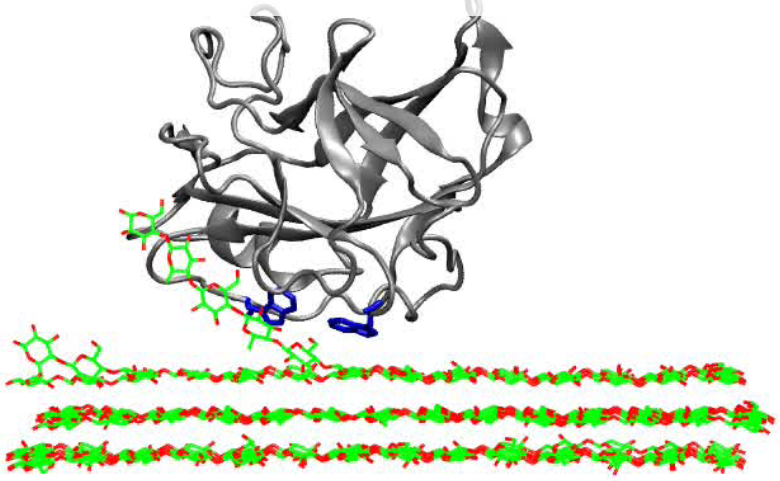
CcCBM17-Forward



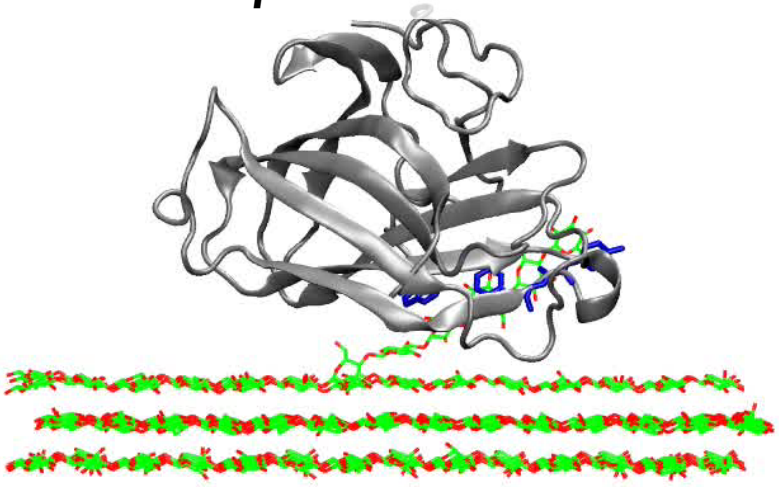
***Bsp*CBM28-Forward**



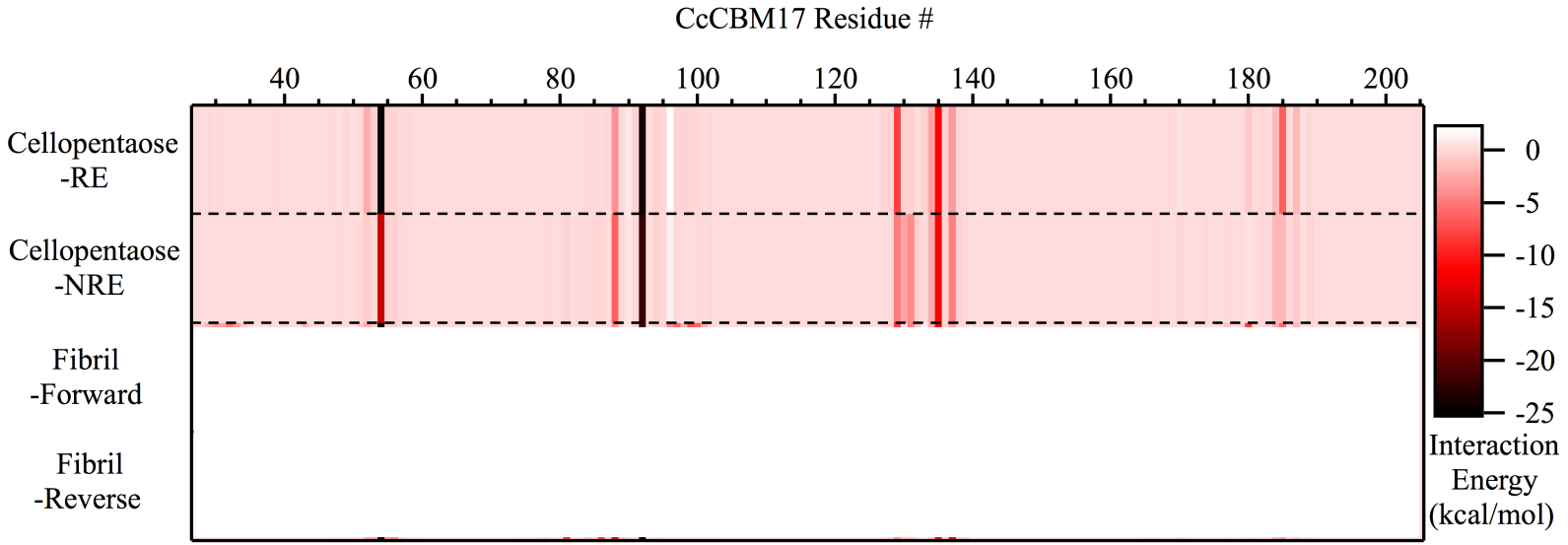
CcCBM17-Reverse

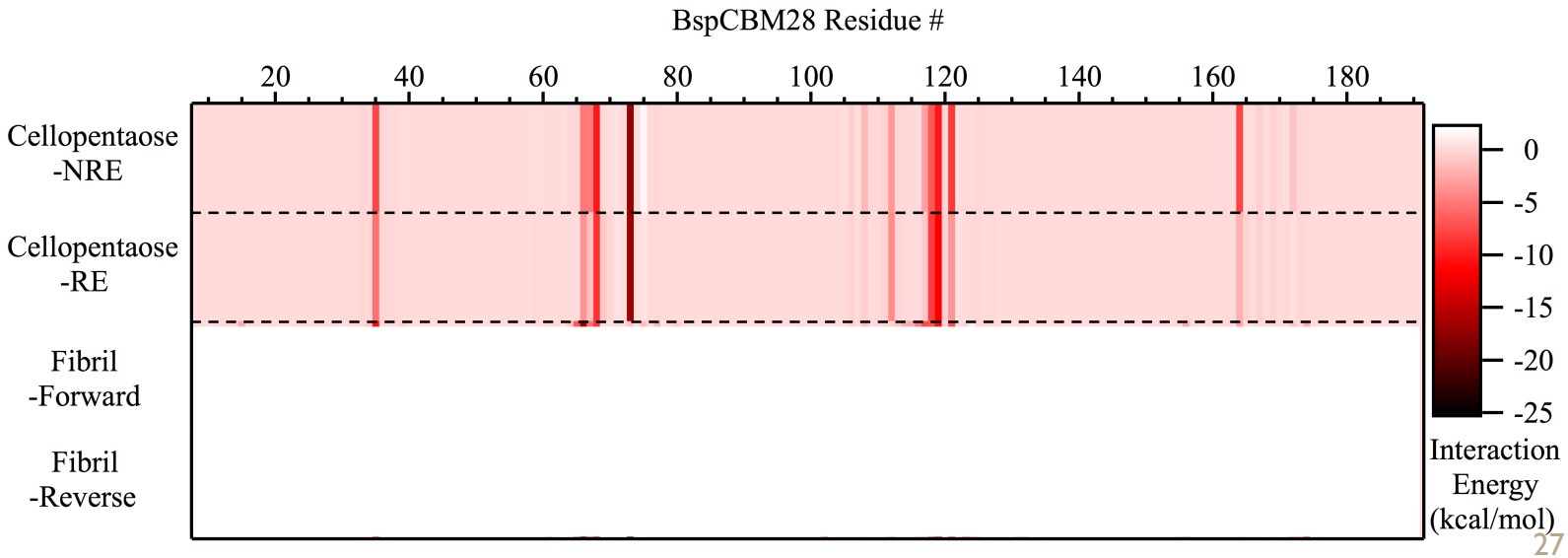


***Bsp*CBM28-Reverse**

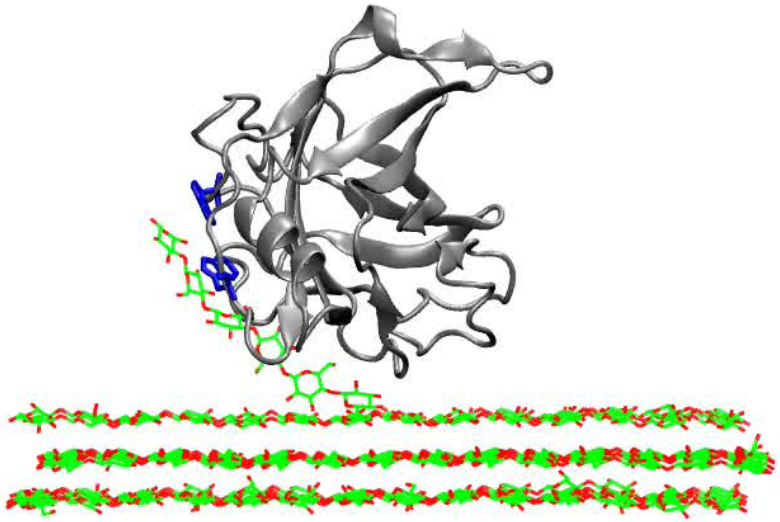


Total substrate interaction per CBM residue

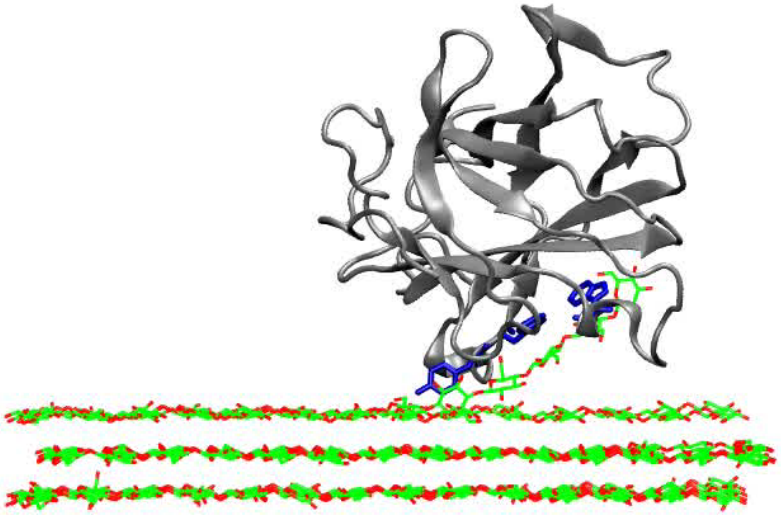




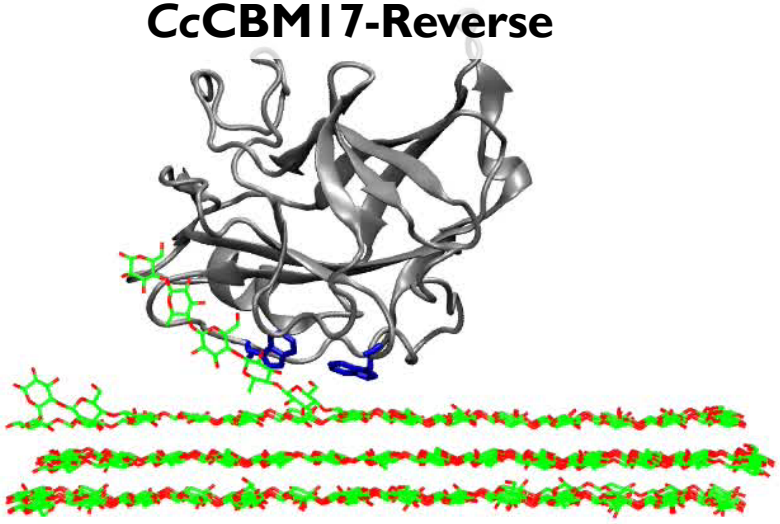
CcCBM17-Forward



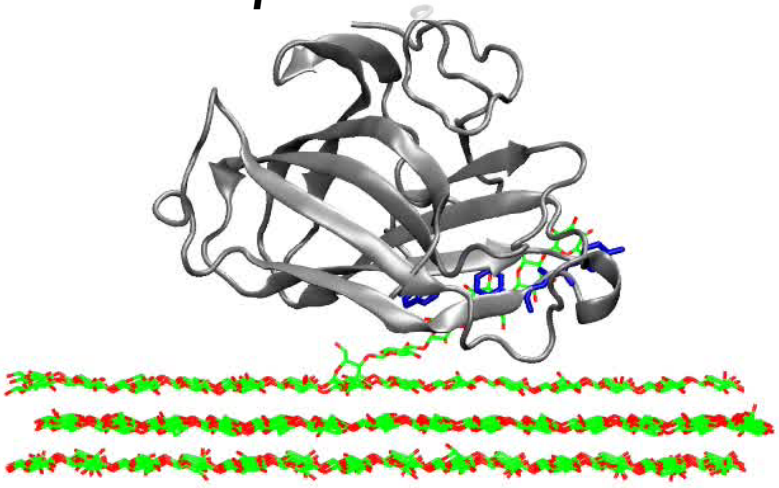
***Bsp*CBM28-Forward**



CcCBM17-Reverse



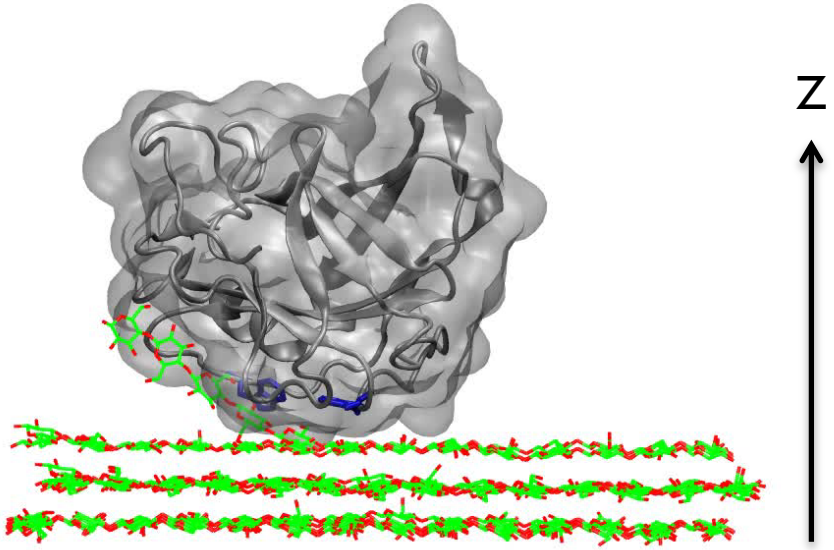
***Bsp*CBM28-Reverse**

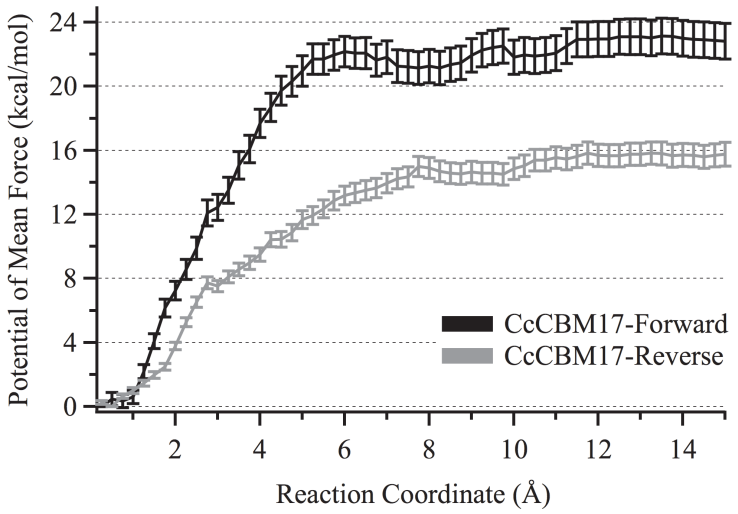


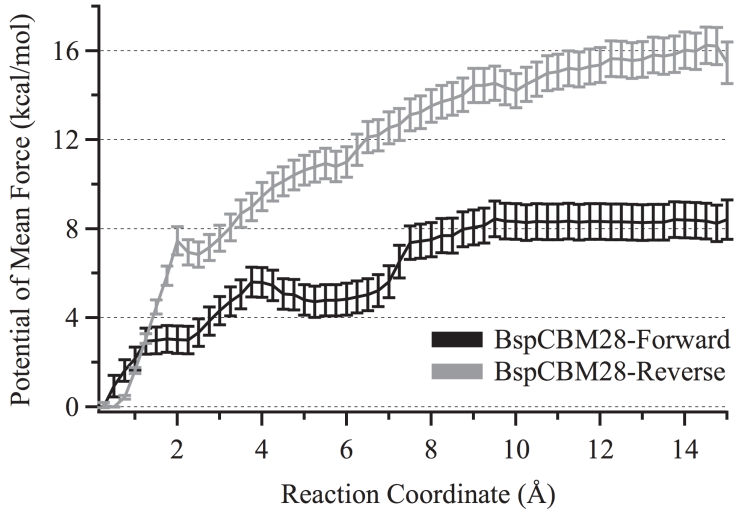
Umbrella sampling

Calculation of potential of mean force (PMF) to pull the CBM away from cellulose microfibril.

- Reaction coordinate is distance between projections of CBM and substrate on z-axis
- 30 windows of 0.5 Å each







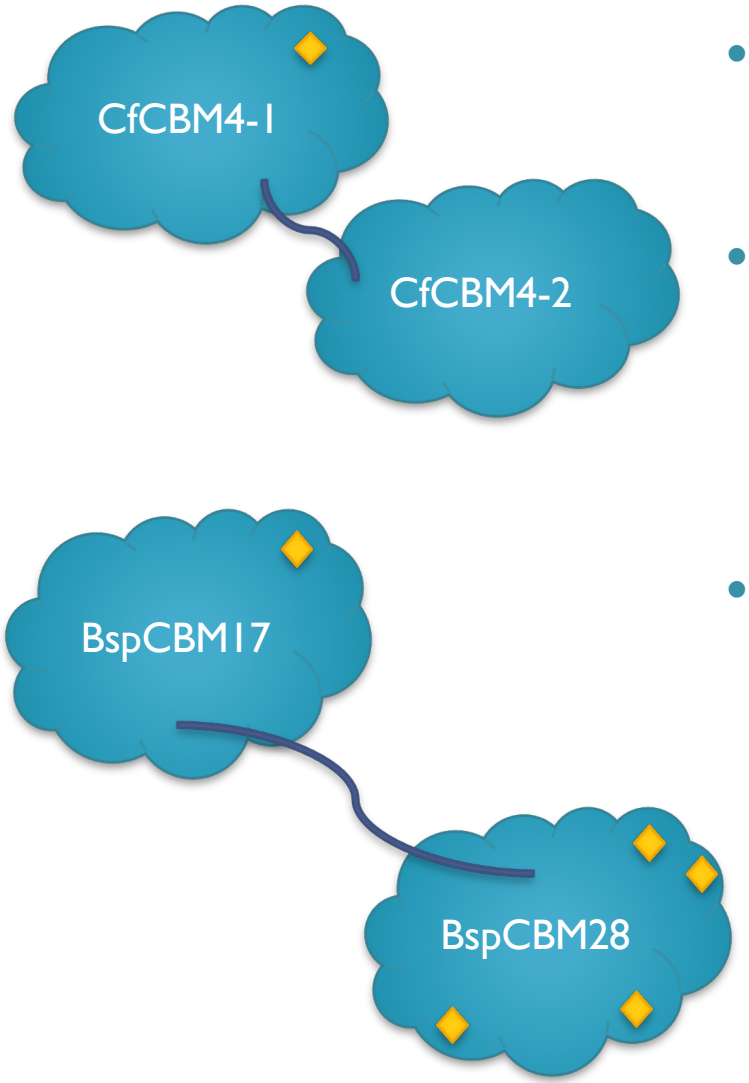
Conclusion – Non-crystalline binding

Non-crystalline recognition by Type B CBMs involves significant interactions of additional domains that are not involved in oligomeric recognition.

The high- and low-affinity binding sites for family 17 and 28 CBMs correspond to a range of non-crystalline substrate with increasing affinity as substrate approaches crystallinity.

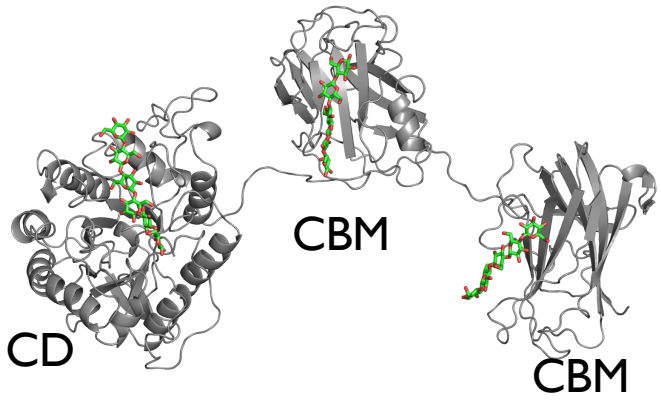
CBMs can have preferential affinity for certain substrate morphologies based on favorable binding orientations which, in turn, could result in uncompetitive binding.

Future work



Tandem CBMs

- Associate the individual the recognition mechanisms and binding affinities with evolution of tandem CBM systems.
- Investigate the dynamics of tandem systems to uncover the effects of protein-protein networking and linker chain lengths on additive/co-operative binding.
- Understand mechanism of feeding the substrate to catalytic domain (CD).



Summary

For faster hydrolysis, the oligomeric and non-crystalline cellulose which is a significant part of pretreated biomass can be targeted specifically by harnessing the abilities of Type B CBMs to do so.

Results from bi-directional binding studies suggest how Type B CBMs may have evolved certain mechanisms to increase the frequency of recognizing the substrate.

Characterization of protein-carbohydrate/protein-protein interactions through molecular simulations holds huge potential to uncover crucial insights that may have been left unexplored. They have significant implications in fields like structure based drug discovery as well.

Acknowledgements

Advisor - Dr. Christina M. Payne

Payne lab members –

Dr. Suvamay Jana

Yue Yu

Dr. Inacrist Geronimo

Ingvild Isaksen

Japheth Gado

István Tamás

Committee Members –

Dr. Tate T. H. Tsang

Dr. Brad Berron

Dr. Matthew Gentry

Funding Source

Division of Chemistry
Award # 1404849

Computational Resources

CCS Center for Computational Sciences

XSEDE Extreme Science and Engineering Discovery Environment

Thank You!

seeblue. in everything we do.