



# CCRC Glyfinder-aided Studies of Cellulose

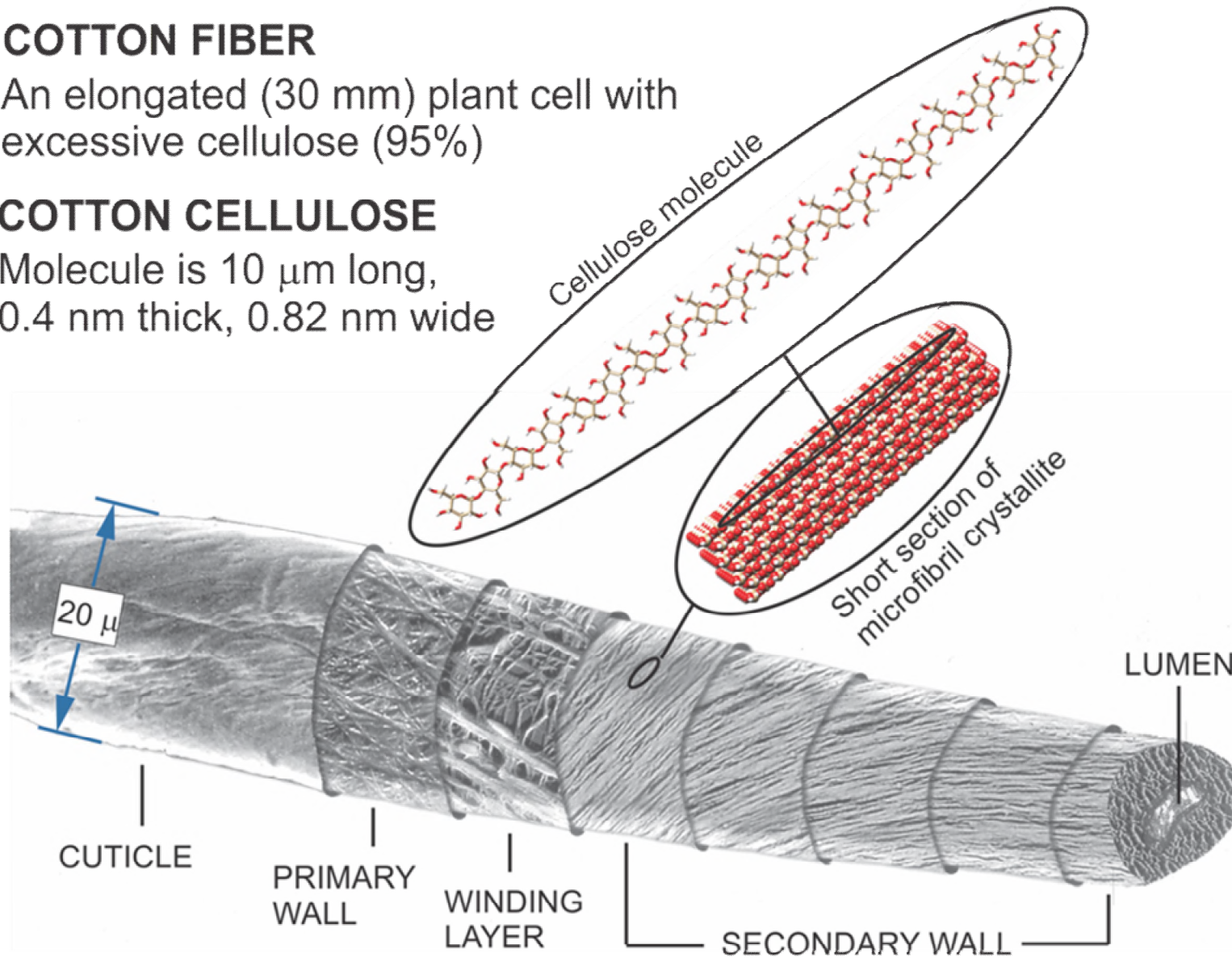
Alfred D. French, New Orleans, Louisiana, USDA Retired Collaborator  
Editor-in-Chief, Cellulose  
With J. Vincent Edwards and Nicolette Prevost, USDA

## COTTON FIBER

An elongated (30  $\mu\text{m}$ ) plant cell with excessive cellulose (95%)

## COTTON CELLULOSE

Molecule is 10  $\mu\text{m}$  long, 0.4 nm thick, 0.82 nm wide



*Adapted from figure by Wilton Goynes by A. D. French*

Cotton fibers are complete cells. They are said to be composed of crystalline and amorphous regions. Many reactions occur in amorphous regions, yet they are not well understood.

What shapes can cellulose molecules have?

Express shapes in  $\phi$ ,  $\psi$

## Point-and-click Search Interface

Select a configuration or Wild Card.



## Monosaccharide

Man	Gal	Glc	Ido	All	Alt	Gul	Tal
Xyl	Lyx	Rib	Ara	Fru	Psi	Sor	Tag
					Fuc	Rha	Qui
					GalNAc	GlcNAc	ManNAc
					GalA	GlcA	IdoA
				Neu5Ac	KDN	KDO	Neu5Gc

## Linkage

1-*	1-1	1-2	1-3	1-4	1-5	1-6	1-7
						1-8	1-9
2-*	2-1	2-2	2-3	2-4	2-5	2-6	2-7
						2-8	2-9

## Aglycon

-OH	-OME	-OtBu	-ASN	-THR	-SER
-----	------	-------	------	------	------

## Isomer

L
D

## Ring Type

f
p

## Configuration

α
β

## Wild Card

*
---

Undo

Clear

Done

Glyfinder makes it easy to find and download cellooligomers in the Protein Data Bank.

Refine By

## Resolution

Min

Max

## Mean B-Factor

Min

Max

Mean  
Oligosaccharide  
B-Factor

Min

Max

## Notes

- Error
- Warning
- Comment

Download Results

Download PDB List

Go Back

Glyfinder Home

1 to 25 of 269 for \*DGlcpb1-4DGlc

«

1

2

3

4

...

10

»

Results Per Page ▾

DGalpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-  
4DGlcpb1-4DGlcpb1-4DGlcpb1-  
4DGlcpb1-4DGlcpb1-4DGlcpb1-  
4DGlcpb1-4DGlcpb1-4DGlcpb1-  
4DGlcpb1-4DGlcpb1-4DGlcpb1-  
4DGlcpb1-OH

5E1Y



3D Viewer

More



2YJQ



DGlcpb1-4DGlcpb1-4OXZ1-

 DGlcpb1-4DGlcpb1-4OXZ1-

3D Viewer

More

Almost instantly returns  
269 oligomers that  
contain one or more  
cellobiose linkages, named by  
the complexing protein

**Glyfinder** Step 1: Set Glycan Query Step 2: Refine and Download

Refine By

Download Results Download PDB List Go Back

Glyfinder Home

1 to 25 of 269 for \*DGlcpb1-4DGlc

« 1 2 3 4 ... 10 »

Results Per Page

**5E1Y** DGalpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-OH

3D Viewer More

**2YJQ** DGlcpb1-4DGlcpb1-4OXZ1-

DGlcpb1-4DGlcpb1-4OXZ1-

3D Viewer More

Resolution Min Max

Mean B-Factor Min Max

Mean Oligosaccharide B-Factor Min Max

Notes

- Error
- Warning
- Comment

5E1Y looks like a great find, with 16 useful linkages

**5E1Y** 4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-4DGlcpb1-OH

3D Viewer Less

**Bacterial Cellulose Synthase Bound To A Substrate Analogue**

J. T. Mcnamara, J. Zimmer

PMID: 26958837 | DOI: 10.1038/NATURE16966

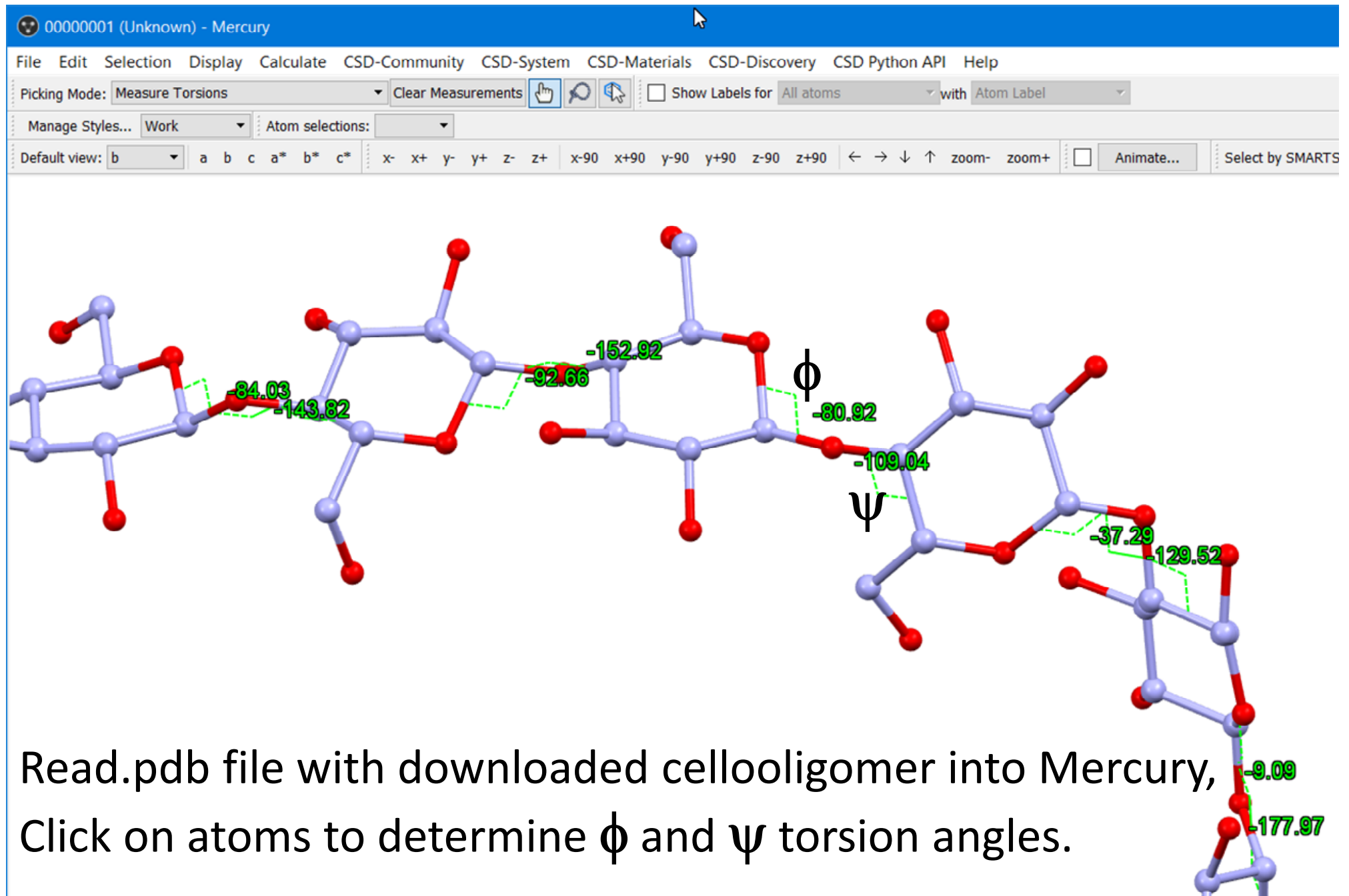
Resolution	2.95 Å
Mean B Factor	83.9
Oligosaccharide Mean B Factor	84.6

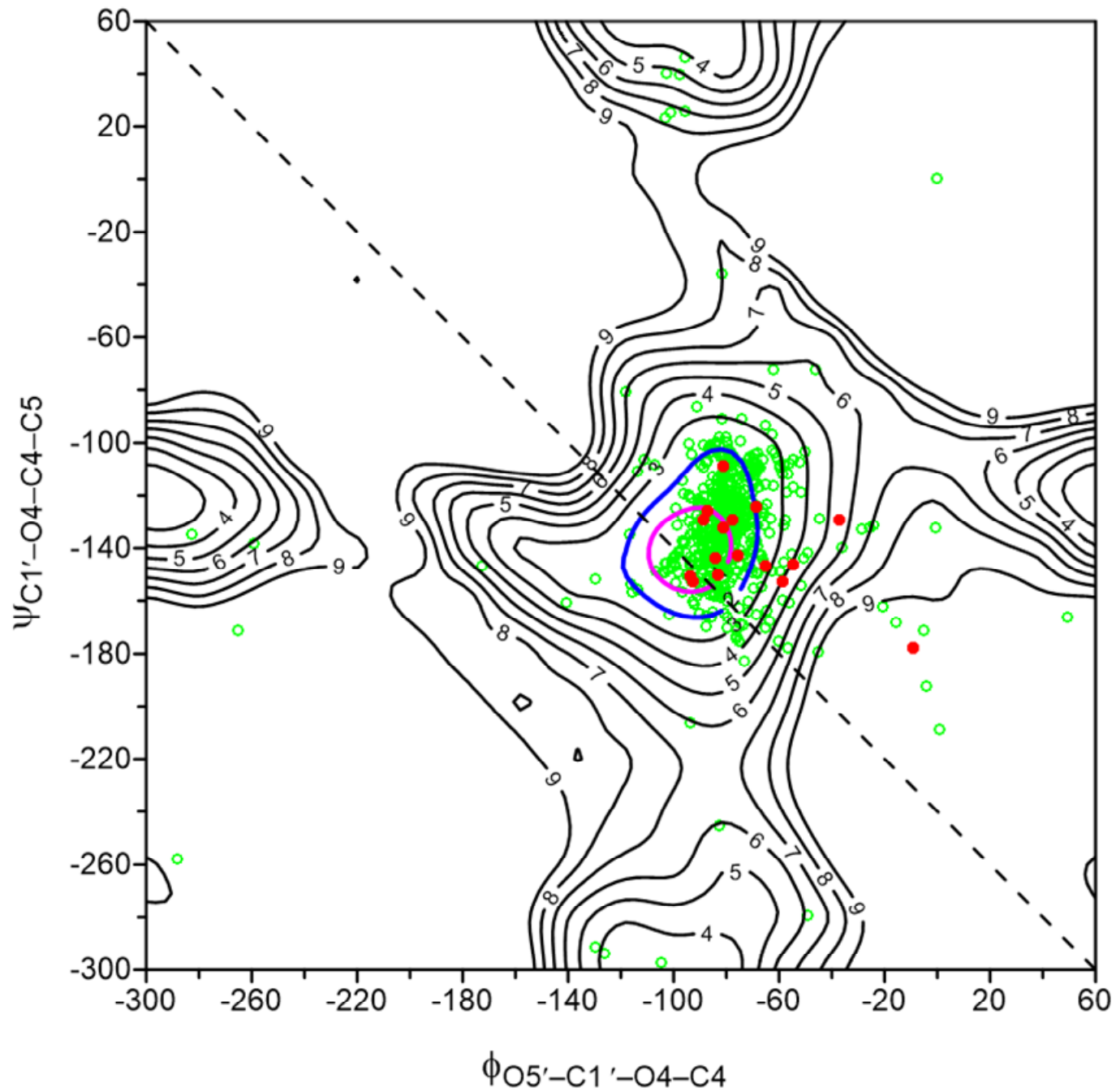
Notes

- Error
- Warning
- Comment

Download

Take note of reference, data quality 😞 . Then Download





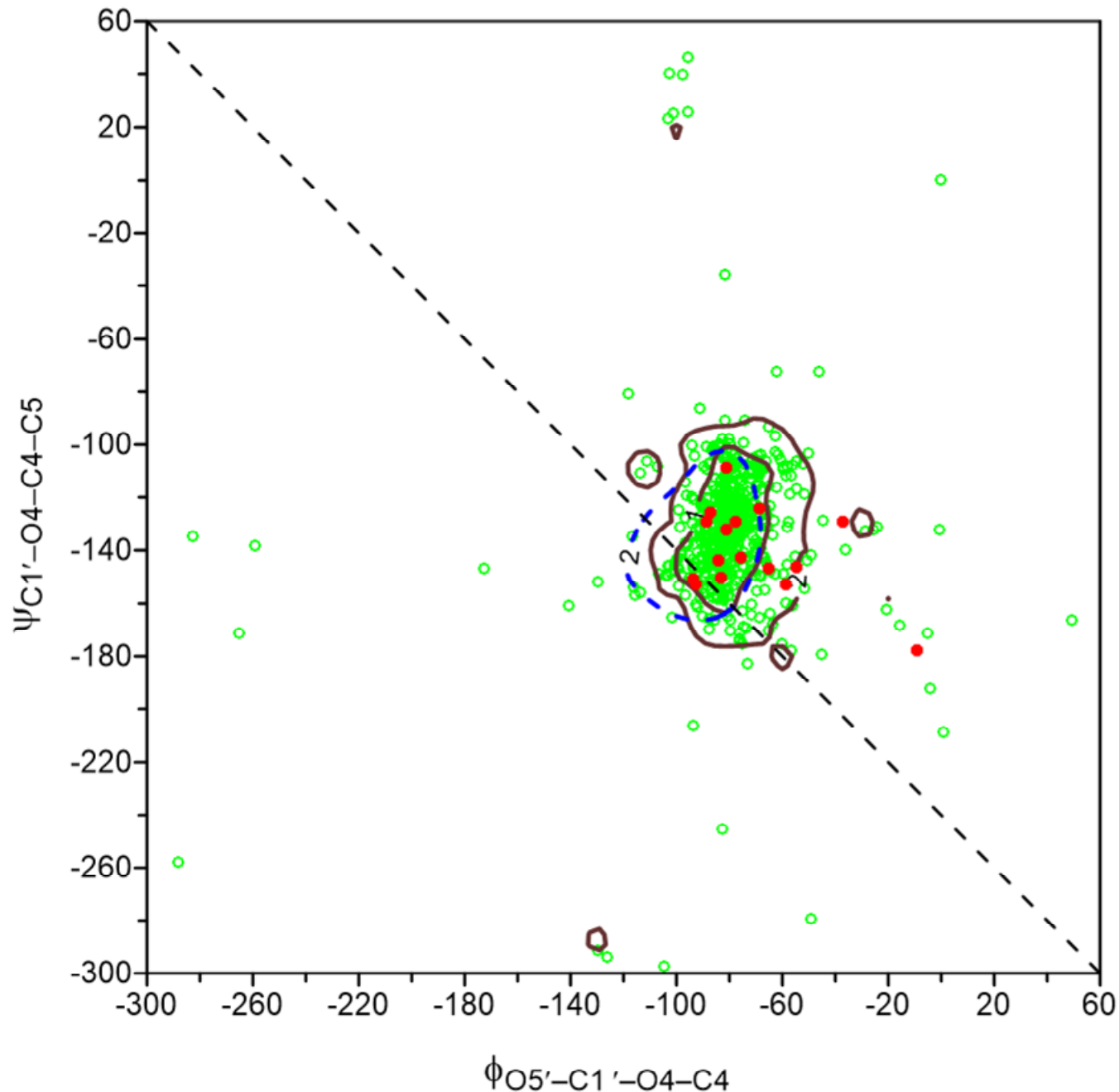
Ramachandran plot with Quantum Mechanics energies (SMD//B3LYP/6-31+G\*).

Energy contours are in increments of 1 kcal/mol.

Dashed line shows 2-fold screw axis shapes favored by crystalline cellulose.

Green dots show PDB structures located by Glyfinder. Red dots are for the just-found 5E1Y.

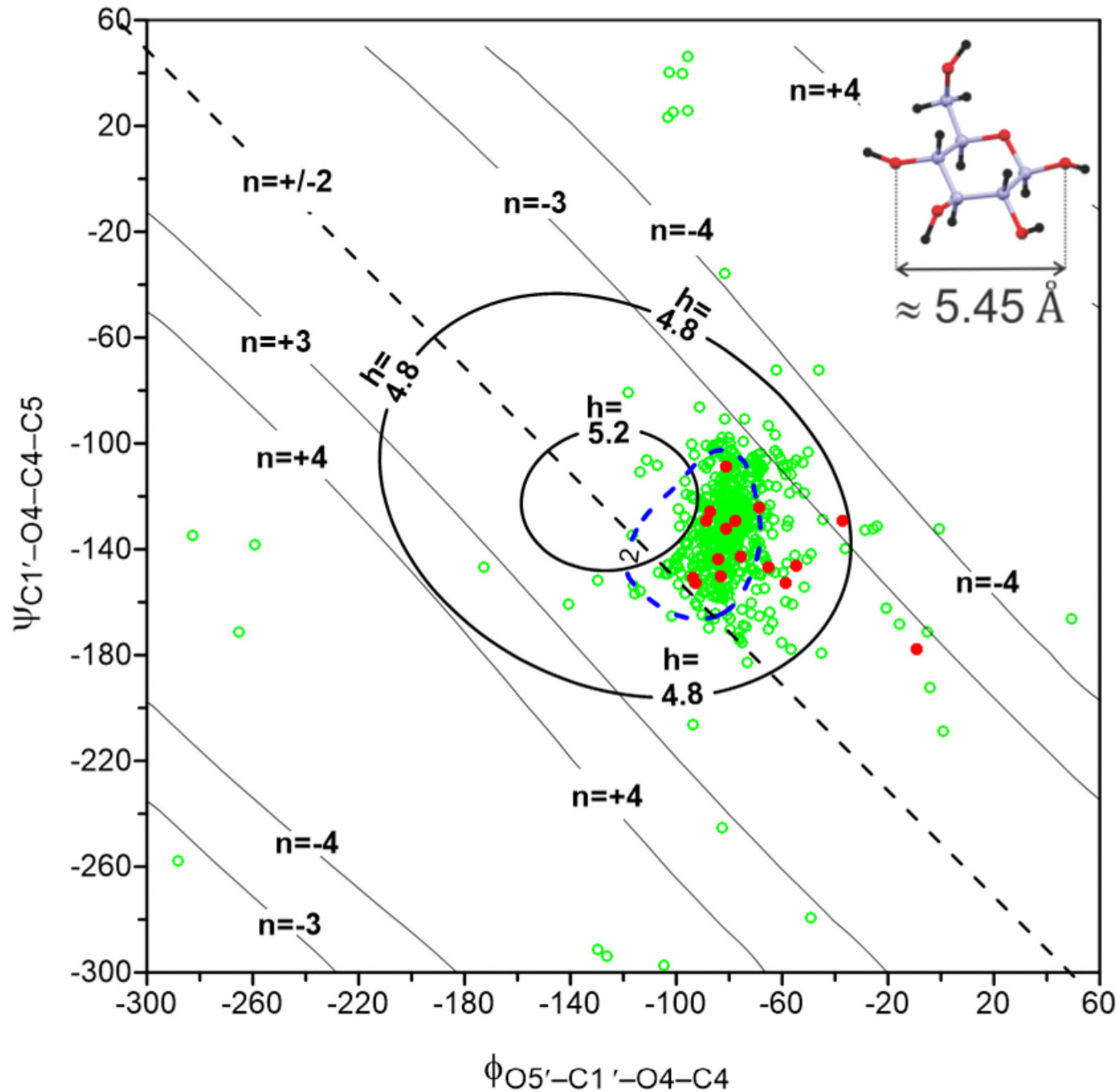
Only one CSD structure has an energy >2 kcal/mol.



An energy map can also be made from the frequency of the different values of  $\phi$  and  $\psi$ . “Only” 559 values were available, with a maximum of 61 in a  $10^\circ \times 10^\circ$  cell, limiting the extent of contouring to 2 kcal/mol when temperature = 298 K.

The 2-kcal/mol contour from the DFT map is shown as a dashed blue line, with the 5EIY conformations as red dots.





The  $\phi$  and  $\psi$  values can be converted to the helix parameters  $n$  and  $h$ .

$n$  is the number of monomer units per helix turn, and  $h$  is the advance along the helix axis per monomer.

Linkages in amorphous cellulose are likely to be those for 2- to left-handed 3-fold helices by both PDB analogy and DFT calculation.

# Conclusions

- Glyfinder is valuable for locating oligomers with polysaccharide linkages of interest to agriculture (we have also looked at starch, xylan, lactose, and  $\beta$ -(1 $\rightarrow$ 3)-glucan).
- Carbohydrate structures in the PDB roughly agree with our quantum mechanics studies, although structures from the Cambridge Crystal Structure Database agree much better.
- The high-energy conformations in the PDB deserve comment in the publications, although they are not discussed often.
- We look forward to the improvements mentioned by David Montgomery in Glyfinder to assist tabulating conformational attributes like  $\phi$  and  $\psi$ .