STRUCTURAL BIOLOGY

Volume 76 (2020)

Supporting information for article:

Structural basis of carbohydrate binding in domain C of a type I pullulanase from *Paenibacillus barengoltzii*

Ping Huang, Shiwang Wu, Shaoqing Yang, Qiaojuan Yan and Zhengqiang Jiang

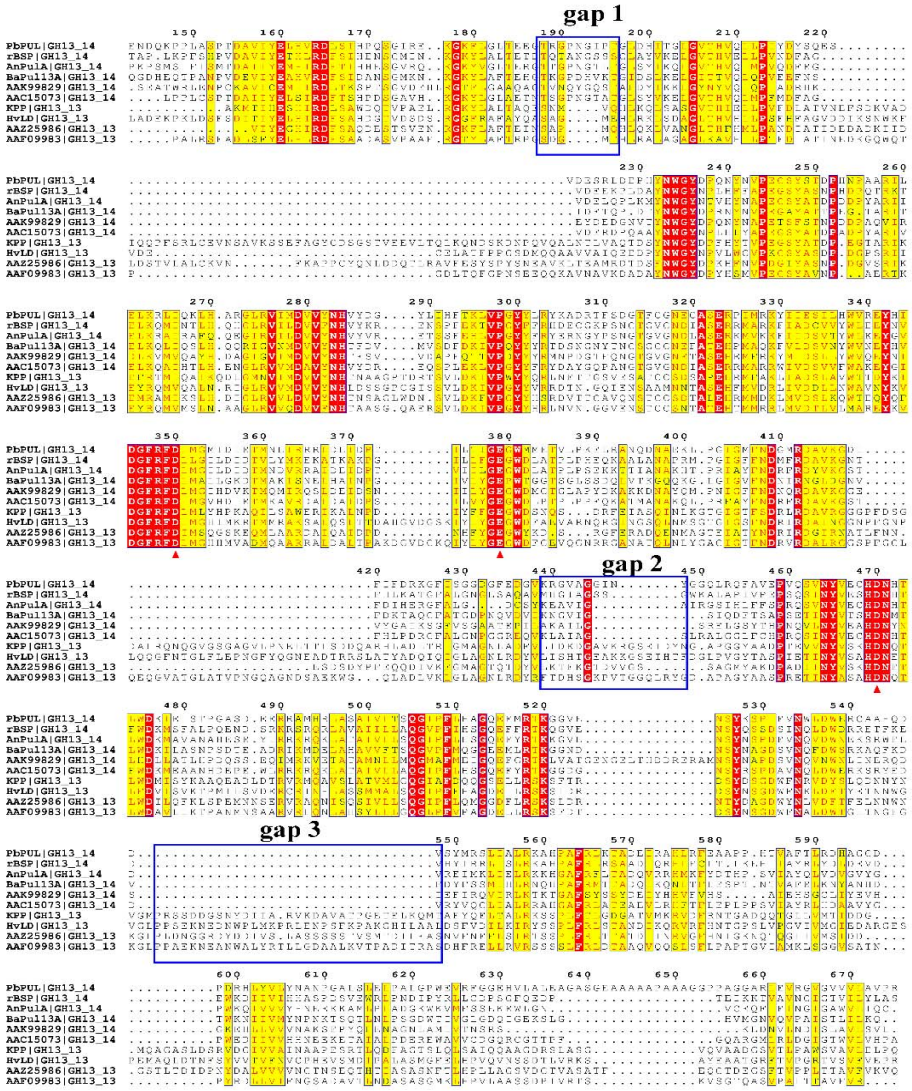

Figure S1 Multiple sequence alignment of structure-reported or characterized enzymes from GH13_13 and GH13_14. The sequences were labeled with abbreviations or GenBank accession numbers (obtained from the reference by Stam *et al.*, 2006). Identical residues are shown in white on red background. Three catalytic residues of *Pb*PulA (Asp350, Glu379 and Asp470) are marked by red triangle. The blue boxes highlight the gaps present in one subfamily but not in the other. The multiple sequence alignment confirmed that the corresponding subfamilies are closely related, but distinct, and that the main difference between GH13_13 and GH13_14 are three gaps (gap1 presents in GH13_13; gap2 and gap3 present in GH13_14) (Stam *et al.*, 2006). rBSP (*Bacillus subtilis* str. 168 pullulanase; PDB entry 2E8Z), *An*PulA (*Anoxybacillus* sp. LM18-11 pullulanase; PDB entry 3WDJ), *Ba*Pul13A (*Bacillus acidopullulyticus* pullulanase; 2WAN), KPP (*Klebsiella pneumoniae* pullulanase; PDB entry 2FHF), *Hv*LD (*Hordeum vulgare* limit dextrins; PDB entry 2Y4S), AAK99829 (*Streptococcus*

pneumoniae R6 thermostable pullulanase), AAC15073 (*Thermus* sp. IM6501 pullulanase), AAZ25986 (*Colwellia psychrerythraea* 34H putative pullulanase) and AAF09983 (*Deinococcus radiodurans* R1 α-dextran endo-1,6-α-glucosidase)were aligned by ClustalX2 (Larkin *et al.*, 2007) and the figure was produced in ESPript 3 (Robert *et al.*, 2014)

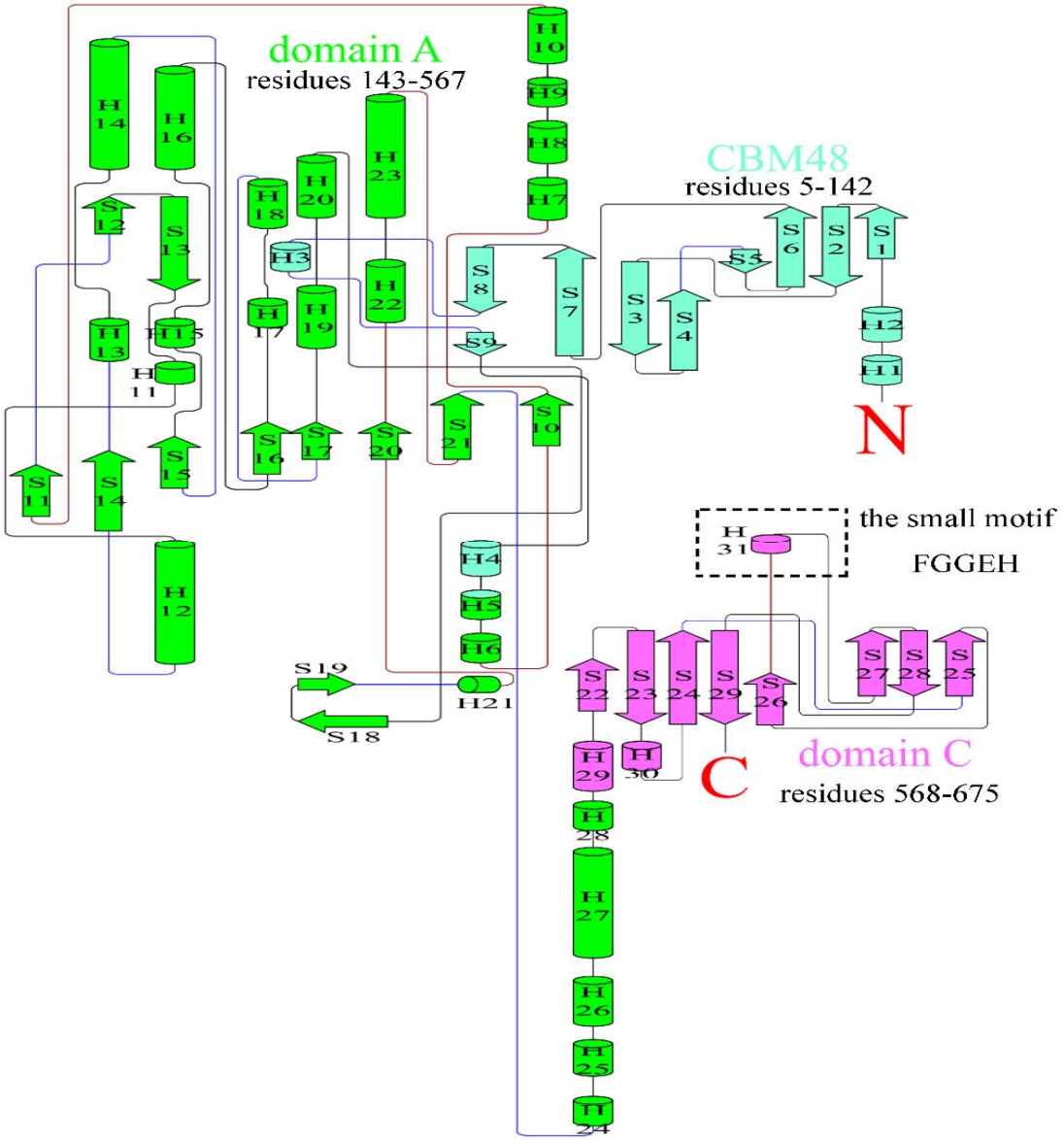

Figure S2 Topology diagram of *Pb*PulA. Helices (α-helix and 3_{10}-helix) are shown as cylinders labeled H; β-strands are shown as arrows labeled S. CBM48, residues 5-142, aquamarine; domain A, residues 143-567, green; domain C, residues 568-675, lightmagenta. S10, S11, S14, S15, S16, S17, S20, and S21 correspond to β1, β2, β3, β4, β5, β6, β7, and β8 of $(\beta/\alpha)_8$-like barrel, respectively, and H10, H12, H14, H16, H18, H19, H22, and H27 correspond to α1, α2, α3, α4, α5, α6, α7, and α8 of $(\beta/\alpha)_8$-like barrel, respectively. The small motif (FGGEH), which may allow substrate binding in domain C, is surrounded by the dotted box.

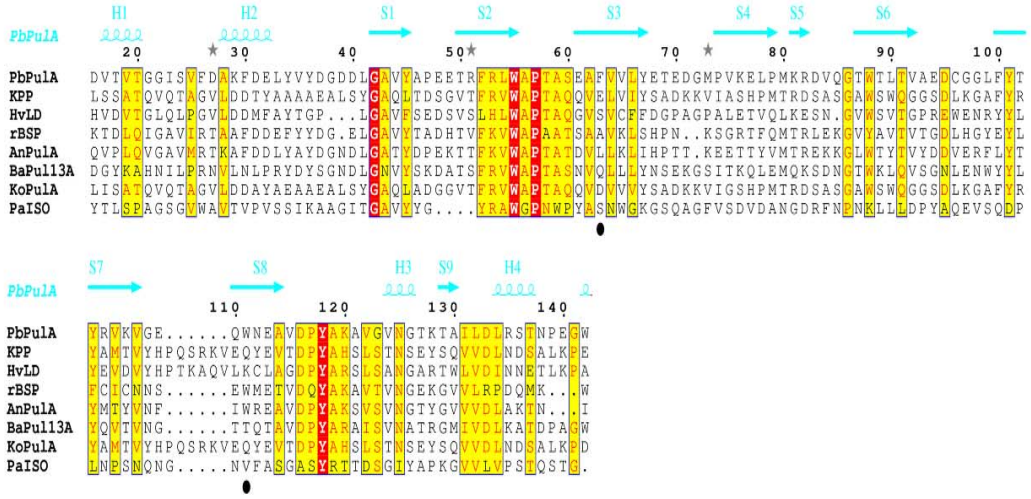

Figure S3 Sequence alignment of CBM48 domains from GH family 13 enzymes. Secondary structures (β strands, arrows; α-helices and 3_{10}-helices, helices) of CBM 48 domain from *Pb*PulA are presented above the alignment and color as aquamarine. Identical residues are shown in white on red background. Gaps inserted in the sequence alignment are indicated by dots. The two key aromatic residues (Phe63 and Trp111) responsible for β-CD binding are marked as black circle. KPP (*Klebsiella pneumoniae* pullulanase; PDB entry 2FHF), *Hv*LD (*Hordeum vulgare* limit dextrins; 2Y4S), rBSP (*Bacillus subtilis* str. 168 pullulanase; 2E8Z), *An*PulA (*Anoxybacillus* sp. LM18-11 pullulanase; 3WDJ), *Ba*Pul13A (*Bacillus acidopullulyticus* pullulanase; 2WAN), *Ko*PulA (*Klebsiella oxytoca* pullulanase; 2YOC), and *Pa*ISO (*Pseudomonas amyloderamosa* isoamylase; 1BF2) were aligned by ClustalX2 and the figure was produced in ESPript 3.

Table S1 Primers used in site-directed mutagenesis

Primers	Primer sequence*（5’ to 3’）	Bases (bp)
D350A_sence	CTTTAATGGGAATGATCGACATCGAA	26
D350A_anti-sence	ATCATTCCCATTAAAGCAAAACGGAA	26
E379A_sence	CAGGCTGGATGATGGAGACAGTGCTC	26
E379A_anti-sence	TCCATCATCCAGCCTGCCCCGATTGT	26
D470A_sence	GCTAATCATACGTTATGGGACAAAATCG	28
D470A_anti-sence	CCCATAACGTATGATTAGCATGGCACT	27
W381A_sence	TCGGGGAAGGCGCGATGATGGAGAC	25
W381A_anti-sence	CGCGCCTTCCCCGATTGTCAGAATC	25
F63A_sence	ACCGCTTCCGAAGCCGCCGTCGTTCTG	27
F63A_anti-sence	GCGGCTTCGGAAGCGGTGGGGGCCCA	26
W111A_sence	AAGGTGGGGGAGCAGGCGAACGAGGC	26
W111A_anti-sence	GCCTGCTCCCCCACCTTCACACGGTACGTA	30
F627A_sence	CCGTGGGAGGTCCGGGCTGGCGGCGAG	27
F627A_anti-sence	GCCCGGACCTCCCACGGTCCGAGGGCG	27

*Mutations are indicated by boxes.

Table S2 Sugar puckering parameters and torsion angles

The table of Cremer-Pople parameters and torsion angles for G5, G6/α-CD, β-CD bound to the active site (a) and domain C (b). Cremer-Pople parameters were calculated using Cremer-Pople parameter calculator (http://enzyme13.bt.a.u-tokyo.ac.jp/CP/). The torsion angles were defined as O5-C1-O4'-C4' (φ) and C1-O4'-C4'-C5' (ψ) between Glc701 (include C1 atom) and Glc702 (include O4 atom; dash number) and calculated using the COOT program.

(a)

Bound-sugar	site	ϕ, θ(˚)	Q	conformation	φ, ψ (˚)
G5	**+2** (**GLC701**)	182.72, 4.21	0.58	$^4C_1$	
					101.89, -133.87
	+1 (**GLC702**)	303.26, 7.81	0.62	$^4C_1$	
					116.96, -122.74
	0 (**GLC703**)	51.97, 10.69	0.58	$^4C_1$	
	-1 (**GLC704**)	137.56, 22.27	0.60	$^4C_1$	
					82.78, -149.19
	-2 (**GLC705**)	33.03, 0.85	0.61	$^4C_1$	
					107.42, -126.76
	-3 (**GLC706**)	165.64, 4.27	0.58	$^4C_1$	
G6/α-CD	**+2** (**GLC701**)	196.69, 2.45	0.58	$^4C_1$	
					100.31, -134.50
	+1 (**GLC702**)	271.39, 5.79	0.61	$^4C_1$	
					120.42, -128.09
	0 (**GLC703**)	58.47, 12.20	0.59	$^4C_1$	
	-1 (**GLC704**)	129.22, 22.10	0.58	$^4C_1$	
					81.21, -149.87
	-2 (**GLC705**)	126.94, 1.68	0.60	$^4C_1$	
					108.69, -132.17
	-3 (**GLC706**)	132.54, 6.00	0.58	$^4C_1$	

	+2 (GLC711)	65.89, 4.67	0.58	$^4C_1$	
					102.89, -130.78
	+1 (GLC710)	45.53, 7.22	0.57	$^4C_1$	
					130.27, -112.16
	0 (GLC709)	93.11, 17.96	0.57	$^4C_1$	
					121.58, -108.57
	GLC708	86.43, 12.68	0.52	$^4C_1$	
β-CD					117.20, -122.70
	GLC714	45.43, 11.05	0.55	$^4C_1$	
					117.28, -114.67
	GLC713	14.39, 6.41	0.56	$^4C_1$	
					111.70, -107.23
	GLC712	21.49, 9.34	0.58	$^4C_1$	
					126.50, -122.19
	GLC711				

(b)

Bound-sugar	site	ϕ, θ(˚)	Q	conformation	φ, ψ (˚)
	X1 (GLC707)	242.04, 0.38	0.57	$^4C_1$	
					106.42, -129.72
G5	**X2 (GLC708)**	163.48, 4.82	0.60	$^4C_1$	
					59.64, -164.83
	X3 (GLC709)	47.64, 14.80	0.53	$^4C_1$	
	X7 (GLC707)	81.74, 10.21	0.56	$^4C_1$	
					117.73, -112.82
	X1 (GLC708)	88.33, 4.17	0.54	$^4C_1$	
G6/α-CD					107.61, -127.02
	X2 (GLC709)	89.35, 6.30	0.61	$^4C_1$	
					111.66, -110.25
	X3 (GLC710)	53.76, 17.48	0.57	$^4C_1$	

					107.30, -109.17
	X4 (**GLC711**)	327.14, 3.51	0.54	$^4C_1$	
	X1 (**GLC716**)	20.49, 4.46	0.56	$^4C_1$	
					109.18, -131.47
	X2 (**GLC715**)	351.27, 1.01	0.59	$^4C_1$	
					118.70, -106.61
	X3 (**GLC721**)	78.73, 7.37	0.55	$^4C_1$	
					112.14, -109.61
	X4 (**GLC720**)	12.12, 9.00	0.55	$^4C_1$	
β-CD					115.34, -122.81
	X5 (**GLC719**)	43.64, 8.03	0.53	$^4C_1$	
					115.57, -121.96
	X6 (**GLC718**)	39.16, 7.16	0.57	$^4C_1$	
					114.68, -110.81
	X7 (**GLC717**)	5.27, 9.69	0.57	$^4C_1$	
					116.47, -111.47
	X1 (**GLC716**)				

Table S3 Sugar puckering parameters and torsion angles

The table of Cremer-Pople parameters and torsion angles for β-CD bound to CBM48. Cremer-Pople parameters were calculated using Cremer-Pople parameter calculator (http://enzyme13.bt.a.u-tokyo.ac.jp/CP/). The torsion angles were defined as O5-C1-O4'-C4' (φ) and C1-O4'-C4'-C5' (ψ) between Glc701 (include C1 atom) and Glc702 (include O4 atom; dash number) and calculated using the COOT program.

sugar	site	ϕ, θ(˚)	Q	conformation	φ, ψ (˚)
	GLC701	68.42, 4.41	0.53	$^4C_1$	
					110.99, -122.18
	GLC702	66.50, 11.40	0.53	$^4C_1$	
					117.11, -112.95
	GLC703	65.53, 6.17	0.56	$^4C_1$	
					114.23, -114.72
	GLC704	61.03, 11.49	0.53	$^4C_1$	
β-CD					109.28, -115.03
	GLC705	353.00, 5.62	0.56	$^4C_1$	
					113.15, -113.22
	GLC706	95.73, 11.96	0.56	$^4C_1$	
					118.99, -114.61
	GLC707	78.35, 8.39	0.54	$^4C_1$	
					116.25, -106.22
	GLC701				

Table S4 Interaction between bound β-CD and residues of CBM48 & symmetry molecules

Distances were calculated using Contact of the CCP4 program suite. The distances of hydrogen bond, C-C clash contacts and C-C contacts were in the ranges of 2.45~3.24 Å, 3.25~3.94 Å and 3.95~4.44 Å, respectively. The asterisk indicates a residue in the symmetric molecules.

Sugar atom	Protein atom & water	Hydrogen bond distance (Å)	C-C contacts distance (Å)	C-C clash contacts distance (Å)
GLC701				
C1	CZ Phe63*			3.77
C1	CE1 Phe 63*		4.38	
C1	CE2 Phe 63*		4.33	
C2	CZ2 Trp111*		4.40	
C2	CH2 Trp111*		4.32	
C3	CZ2 Trp111*		4.27	
C4	CZ2 Trp111*			3.61
C4	CE2 Trp111*		4.01	
C4	CH2 Trp111*		4.00	
C5	CZ2 Trp111*		4.30	
C5	CH2 Trp111*		4.25	
C6	CD Arg104*		4.34	
C6	CD2 Trp111*		4.12	
C6	CE2 Trp111*		4.24	
C6	CE3 Trp111*		4.04	
C6	CZ2 Trp111*		4.29	
C6	CZ3 Trp111*		4.07	
C6	CH2 Trp111*		4.19	
O2	Wat1336	2.83		
GLC702				
C6	CD Glu77*		4.33	

C6	CB Glu77*		4.30	
C6	CG Glu77*		4.35	
O2	Wat1059	2.75		
O3	Wat1048	2.85		
O3	Wat1333	3.01		

Sugar atom	Protein atom & water	Hydrogen bond distance (Å)	C-C contacts distance (Å)	C-C clash contacts distance (Å)
GLC703				
C1	CB Ala93		4.18	
C2	CB Ala93		4.21	
O2	O Thr91	2.60		
O3	O Thr91	3.18		
O3	Wat1092	2.93		
O3	Wat1234	3.00		
O5	Wat1329	3.19		
GLC704				
C1	CB Ala543*		4.43	
O2	Wat1210	3.02		
O2	Wat840	2.75		
O3	Wat819	2.69		
GLC705				
C2	CA Ala543*		4.29	
C3	CA Ala543*		4.26	
C4	CA Ala543*		4.21	
GLC706				
GLC707				
C1	CE2 Trp111*			3.77
C1	CD1 Trp111*		3.96	
C1	CD2 Trp111*		4.41	
C1	CZ2 Trp111*		4.15	
C2	CD1 Trp111*		4.35	

Table S5 The activity of wild-type and mutants of *Pb*PulA

PbPulA	Specific activity (U mg^{-1})	Relative activity (%)
Wild type	24.9±0.64	100
D350A	none	none
E379A	none	none
D470A	none	none
W381A	0.6±0.03	2.4
F63A	14.7±0.44	59
W111A	10.2±0.35	41
F627A	17.1±0.63	69

Table S6 Interaction between bound sugar moieties and residues of the active site

Distances were calculated using Contact of the CCP4 program suite. The distances of hydrogen bond, C-C clash contacts and C-C contacts were in the ranges of 2.45~3.24 Å, 3.25~3.94 Å and 3.95~4.44 Å, respectively.

Sugar atom	Protein atom & water	Hydrogen bond distance (Å)			C-C contacts	C-C clash contacts
		G5	**G6/α-CD**	**β-CD**		
Site +2		**GLC701**	**GLC701**	**GLC711**	**G5:** Trp381, Asp409, Phe420	**G5:** Trp381, Phe420
O2	OD2 Asp409	2.63	2.81	2.73		
O2	NH2 Arg412	2.95	2.91	2.93		
O3	NH1 Arg412	2.83	2.86	2.82	**G6/α-CD:** Trp381, Met383, Asp409, Phe420	**G6/α-CD:** Trp381, Phe420
O5	Wat877		2.99			
O5	Wat955	3.03				
O5	Wat1375			3.05		
O6	Wat873		2.65		**β-CD:** Trp381, Met383, Asp409 Ile419, Phe420	**β-CD:** Trp381, Phe420
O6	Wat877		2.91			
O6	Wat955	2.96				
Site +1		**GLC702**	**GLC702**	**GLC710**	**G5:** Leu351, Trp381, Met383, Asp470	
O2	ND2 Asn471	2.84		2.83		
O2	OD1 Asn471		2.83			
O3	OD1 Asn471	3.04		2.90		
O3	ND2 Asn471		2.98		**G6/α-CD:** Leu351, Trp381, Met383, Phe420, Asp470	**β-CD:**Phe420
O3	Wat807	3.16				
O6	Wat1040		2.90			
O6	Wat897	2.84				
O6	Wat983			2.70		
O6	Wat1132			3.07		

					β-CD: Phe315, Trp 381, Met383, Phe420	
Site 0		**GLC703**	**Glc703**	**GLC709**	**G5:** Phe315,	
O2	Wat908	2.73			Asn525	**M5:** Phe315
O2	Wat878		2.66			
O3	ND2 Asn525	3.23			**G6/α-CD:** Asn525	**G6/α-CD:**
O5	Wat1040		3.23			Phe315
O5	Wat975	3.21			**β-CD:** Phe315	

Sugar atom	Protein atom & water	Hydrogen bond distance (Å)			C-C contacts	C-C clash contacts
		G5	**G6/α-CD**	**β-CD**		
Site -1		**GLC704**	**GLC704**			
O1	OE2 Glu379	2.95	2.97		**G5:** Tyr236, His 284, Asp350, Leu351, Glu379, Asp470 **G6/α-CD:** Tyr236, His284, Asp350, Leu351, Glu379, His469, Asp470	**G5:** Asp350, Glu379, Asp470 **G6/α-CD:** Asp350, Glu379, Asp470
O2	NH2 Arg348	3.00	3.03			
O2	NE2 His469	2.96	2.88			
O2	OD2 Asp470	2.65	2.70			
O3	NE2 His469	3.00	2.87			
O3	OD1 Asp470	2.64	2.82			
O3	OD2 Asp470	3.20				
O5	OD1 Asp350	3.01	3.04			
O5	OD2 Asp350	3.24	3.23			
O6	OD2 Asp350	2.76	2.66			
Site -2		**GL705**	**GLC705**			
O2	OD1 Asn525	2.74	2.72		**G5:** Trp234, Tyr236, Asp237, Asn525, Tyr527	**G5:** Trp234, Tyr236, Tyr527 **G6/α-CD:** Tyr527
O3	ND2 Asn525	2.98	2.94			
O6	OD1 Asp237	2.74	2.77			
O6	Wat850	2.78				

O6	Wat888		2.97	**G6/α-CD:** Trp234,	
O6	Wat974		3.21	Tyr236, Asp237,	
O6	Wat893	3.03		Asn525, Tyr527	
Site -3		**GLC706**	**GLC706**	**G5:** Tyr527	**G5:** Tyr527
O4	O Phe315	3.22		**G6/α-CD:** Tyr527	**G6/α-CD:** Tyr527

Table S7 Interaction between bound sugar moieties and residues of the domain C

Distances were calculated using Contact of the CCP4 program suite. The distances of hydrogen bond, C-C clash contacts and C-C contacts were in the ranges of 2.45~3.24 Å, 3.25~3.94 Å and 3.95~4.44 Å, respectively.

Sugar atom	Protein atom & water	Hydrogen bond distance (Å)			C-C contacts	C-C clash contacts
		G5	**G6/α-CD**	**β-CD**		
X1		**GLC707**	**GLC708**	**GLC716**	**G5:** Tyr550,	
O2	OE2 Glu630	2.68	2.67	2.83	Ser553, Phe627,	**M5:** Phe627
O3	N Gly628	2.76	2.73	2.77	Glu630	
O3	OE1 Glu630	2.65	2.64	2.71		**G6/α-CD:**
O3	Wat1049	3.11			**G6/α-CD:** Tyr550,	Phe627
O4	Wat1049	2.85			Ser553, Phe627,	
O5	OG Ser553	3.22	3.19	3.15	Glu630,	**β-CD:**
O6	OG Ser553	2.83	2.75			Ser553,
O6	Wat941			2.92	**β-CD:** Ser553,	Phe627
O6	Wat1177			3.04	Phe627, Glu630	
X2		**GLC708**	**GLC709**	**GLC715**	**M5:** Arg495,	
O2	NH2 Arg495	2.83	2.94	3.09	Asp546, Ser549,	
O3	NH1 Arg495	2.87	2.85	3.02	Tyr550	
O5	OG Ser549	3.11	3.13	2.85		
O6	O Ser549	3.10	3.06		**G6/α-CD:**	
O6	OG Ser549	2.61	2.55	2.84	Arg495, Asp546, Ser549, Tyr550	
O6	OG Ser553	2.65	2.65	2.46	**β-CD:** Arg495, Asp546, Ser549, Tyr550	
X3		**GLC709**	**GLC710**	**GLC721**	**G6/α-CD:** Asp546	**M5:** Asp546

O3	OD1 Asp546	2.85	3.11			**G6/α-CD:**
O3	Wat918	2.79			**β-CD:** Asp546	Asp546
O6	Wat992		2.79			**β-CD:** Asp546
X4			**GLC711**	**GLC720**		
O2	Wat832			3.11		
X5				**GLC719**		
X6				**GLC718**		
X7			**GLC707**	**GLC717**		
O2	N Gly628		3.17	3.22	**G6/α-CD:**	
O2	N Gly629		2.72	2.99	Arg626, Phe627	
O3	Wat1020			2.90	**β-CD:** Arg626,	
O5	Wat1145			3.07	Phe627	
O5	Wat1177			3.17		

References

1. Stam, M. R., Danchin, E. G., Rancurel, C., Coutinho, P. M. & Henrissat, B. (2006). *Protein Eng. Des. Sel.* **19**, 555-562.

2. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). *Bioinformatics*, **23**, 2947-8.

3. Robert, X. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320-4.