Molecular Dynamics Investigations of Oligosaccharides Recognized by Family 16 and 22 Carbohydrate Binding Modules

Supporting Information

Ruihan Wang[1] and Dingguo Xu[1,2]*

[1]MOE Key Laboratory of Green Chemistry and Technology, College of Chemistry, Sichuan University, Chengdu, Sichuan, 610064 PR China

[2]Material Genome Engineering Research Center, Sichuan University, Chengdu, Sichuan, 610065, PR China

*To whom Correspondence should be addressed: dgxu@scu.edu.cn (D. Xu)

Computational Details for the Principal Component Analysis

We applied a Python package MDAnalysis[1] for the principal component analysis (PCA) [2]. PCA sorts a simulation into 3N directions of descending variance, with N being the number of atoms. These directions are called the principal components. The dimensions to be analyzed are reduced by only looking at a few projections of the first principal components. The PCA problem is solved by solving the eigenvalue problem of the covariance matrix, a $3N \times 3N$ matrix where the element (i,j) is the covariance between coordinates i and j. The principal components are the eigenvectors of this matrix. In addition, before PCA, align the whole trajectory to the structure of the specified frame to eliminate the effects of translation and rotation. The free energy surface graph describes the free energy of various conformations of macromolecules. This requires knowing the probability density $P(x)$ of various conformations, and then calculating the relative free energy $G(x)$:$G(x) = -kT \times \ln(P(x))$ by Boltzmann relation. The two coordinate axes of the free energy surface graph are taken as PC1 and PC2, respectively. The trajectory is projected onto these two coordinate axes according to the corresponding eigenvectors. The scatter distribution on the graph is transformed into probability distribution $P(x)$ by *ddtpd* program[3], and then $G(x)$ is obtained from the preceding formula. The free energy surface is obtained by making its contour map or color mapping.

Reference

1. R. J. Gowers, M. Linke, J. Barnoud, T. J. E. Reddy, M. N. Melo, S. L. Seyler, D. L. Dotson, J. Domanski, S. Buchoux, I. M. Kenney, and O. Beckstein. MDAnalysis: A Python package for the rapid analysis of molecular dynamics simulations. In S. Benthall and S. Rostrup, editors, Proceedings of the 15th Python in Science Conference, pages 98-105, Austin, TX, 2016. SciPy.

2. More detail in Berk Hess. Convergence of sampling in protein simulations. Phys. Rev. E 65

3. The author thanks Tian Lu at University of Science and Technology Beijing for providing *ddtpd* program (*ddtpd*--Converting dot distribution to probability distribution, version 1.3)

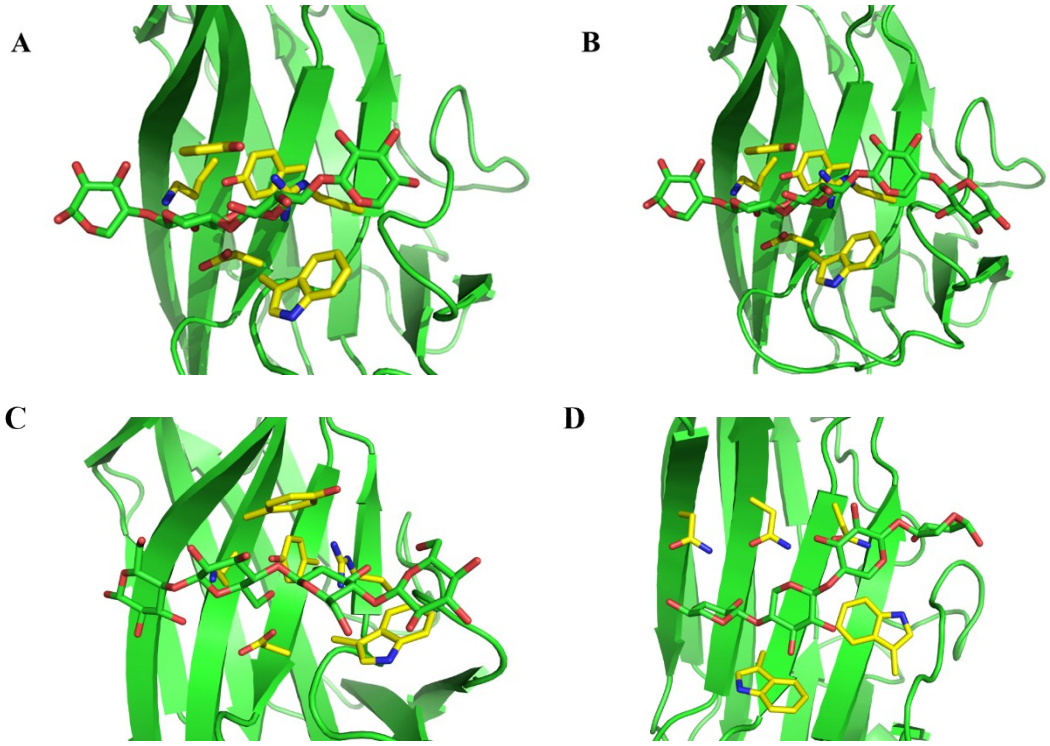

Figure S1. The initial structures of CBM22-2 in complex with (A) xylotetraose (B) xylopentaose (C) cellotetraose and the initial structures of CBM16-1 in complex with (D) xylotetraose for MD simulations after Autodock-vina simulation.

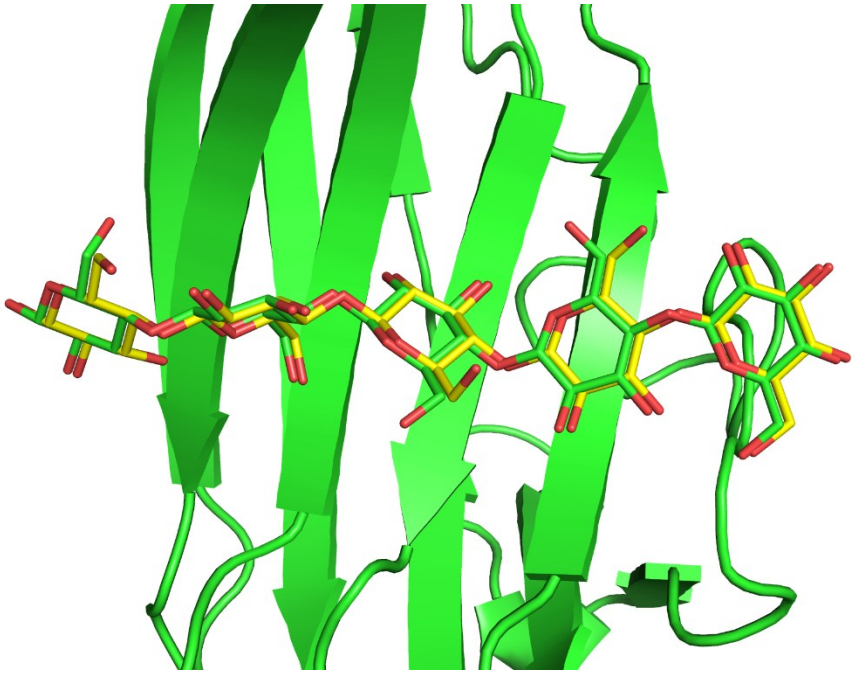

Figure S2. Overlap the crystal structure (carbon in green) of the CBM16-1 with CE5 and structure from the docking using AutoDock vina (carbon in yellow).

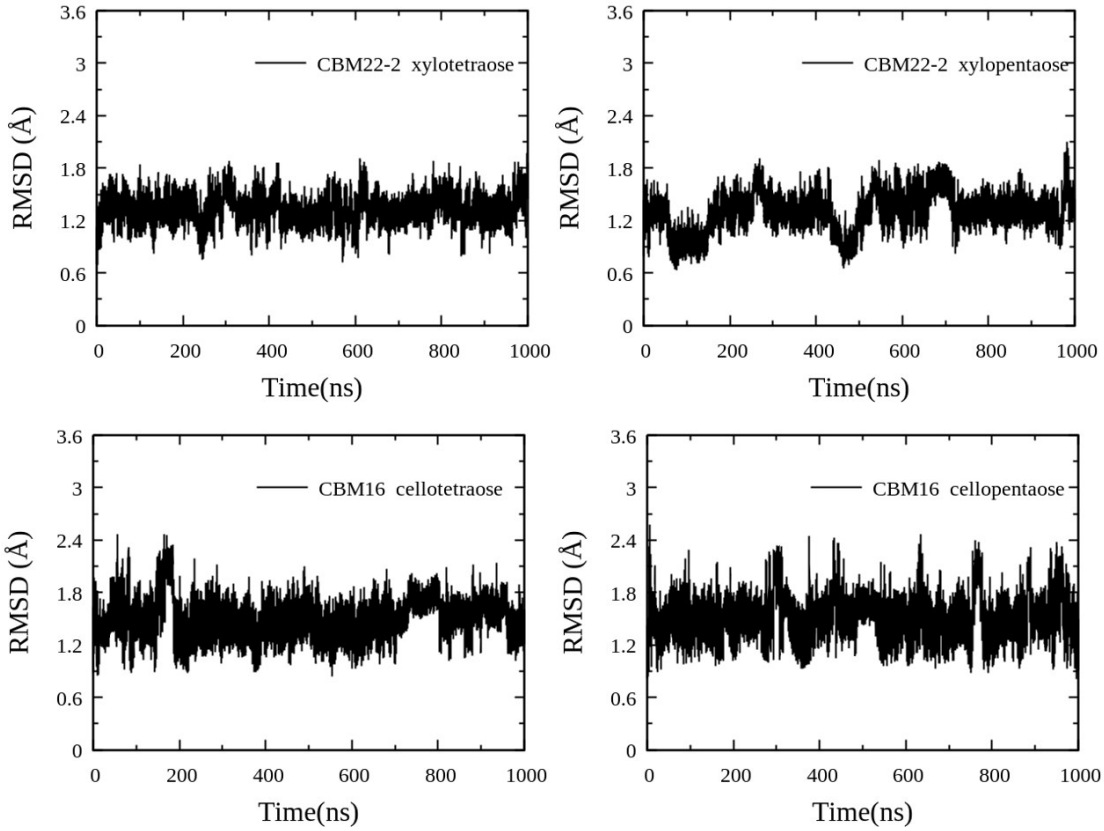

Figure S3. Root mean square deviations (RMSDs) for the alpha carbon atoms for four models.

Table S1. Average torsion angles between putative molecular planes of sugar ring and side chains of aromatic amino acids for complexes of CBM16-1/CE5 and CBM22-2/XL5.

CBMs		Pyranose ring	Torsion Angles
Family	Residues		
16-1	W23	S4	25.8±8.2°
16-1	W128	S3	156.1±9.8°
22-2	W53	S3	23.9±8.5°
22-2	Y101	S2	158.9±7.3°

Table S2 Binding free energy for CE4/CBM16-1 and decomposition to electrostatic interaction, van der Walls interaction, solvation free energies and entropy. Energies are in kcal/mol. Numbers in parentheses are standard deviations.

	complex	receptor	ligand	Delta
E_{vdW}	-1172.58(13.83)	-1122.06(13.74)	-13.94(1.62)	-36.59(2.98)
E_{ele}	-	-		
	11154.05(85.42)	10719.97(84.61)	-393.95(8.51)	-40.13(7.15)
E_{GB}	-2265.55(73.63)	-2246.81(72.11)	-69.08(3.54)	50.35(5.23)
E_{surf}	45.85(1.32)	44.67(1.22)	5.93(0.05)	-4.76(0.32)
G_{gas}	-	-		
	12326.63(85.63)	11842.03(84.80)	-407.88(8.27)	-76.72(6.96)
G_{solv}	-2219.69(73.00)	-2202.14(71.43)	-63.15(3.52)	45.59(5.23)
$E_{gas} + G_{sol}$	-	-		
	14546.32(44.84)	14044.17(44.30)	-471.03(7.83)	-31.12(3.90)
$-TS_{total}$	-1622.68(6.73)	-1562.16(5.57)	-84.75(0.39)	24.24(2.94)
ΔG_{bind}^{cal}				-6.88
ΔG_{bind}^{exp}				-6.60

* Experimental binding free energy is obtained from Ref 18.

Table S3 Binding free energy for CE5/CBM16-1 and decomposition to electrostatic interaction, van der Walls interaction, solvation free energies and entropy. Energies are in kcal/mol. Numbers in parentheses are standard deviations.

	complex	receptor	ligand	Delta
E_{vdW}	-1164.18(14.87)	-1104.48(14.17)	-17.61(1.96)	-42.08(3.68)
E_{ele}	-	-		
	11283.28(78.61)	10742.43(77.60)	-491.83(9.35)	-49.01(8.23)
E_{GB}	-2253.60(70.09)	-2231.28(69.82)	-82.00(3.66)	59.67(6.26)
E_{surf}	47.81(1.21)	46.42(1.19)	7.14(0.06)	-5.75(0.32)
G_{gas}	-	-		
	12447.46(77.91)	11846.93(76.93)	-509.44(9.33)	-91.09(8.06)
G_{solv}	-2205.79(69.43)	-2184.85(69.15)	-74.86(3.65)	53.92(6.15)
$E_{gas} + G_{sol}$		-		
	-14653(36.27)	14031.78(34.50)	-584.30(8.54)	-37.17(4.20)
$-TS_{total}$	-1639.17(6.03)	-1565.63(5.89)	-102.05(0.56)	28.50(3.23)
ΔG_{bind}^{cal}				-8.67
ΔG_{bind}^{exp}				-7.66

* Experimental binding free energy is obtained from Ref 18.

Table S4 Binding free energy for CBM22-2/XL4 and decomposition to electrostatic interaction, van der Walls interaction, solvation free energies and entropy. Energies are in kcal/mol. Numbers in parentheses are standard deviations.

	complex	receptor	Ligand	delta
E_{vdW}	-1259.28(14.33)	-1218.19(12.98)	-9.26(1.15)	-31.84(4.22)
E_{ele}	- 10668.26(68.17)	- 10340.76(70.90)	-252.99(11.68)	-74.50(6.63)
E_{GB}	-1728.20(54.76)	-1755.96(57.09)	-49.85(2.64)	77.61(6.71)
E_{surf}	49.10(1.19)	48.60(1.05)	4.92(0.05)	-4.42(0.56)
G_{gas}	- 11927.55(69.71)	- 11558.95(71.33)	-262.26(11.68)	-106.33(8.37)
G_{solv}	-1679.10(54.70)	-1707.37(56.99)	-44.92(2.64)	73.19(8.21)
$E_{gas} + G_{sol}$	- 13606.65(45.32)	- 13266.32(42.29)	-307.18(11.10)	-33.14(3.34)
$-TS_{total}$	-1719.80(6.52)	-1673.98(5.75)	-71.02(0.25)	25.20(2.08)
ΔG_{bind}^{cal}				-7.94
ΔG_{bind}^{exp}				-7.09

* Experimental binding free energy is obtained from Ref 20.

Table S5 Binding free energy for CBM22-2/XL5 and decomposition to electrostatic interaction, van der Walls interaction, solvation free energies and entropy. Energies are in kcal/mol. Numbers in parentheses are standard deviations.

	complex	receptor	ligand	delta
E_{vdW}	-1259.43(16.95)	-1212.27(16.25)	-13.10(1.29)	-34.06(5.06)
E_{ele}	- 10663.55(72.32)	- 10298.83(63.53)	-280.74(20.72)	-83.97(16.85)
E_{GB}	-1724.83(46.14)	-1752.05(44.55)	-58.02(3.38)	85.23(13.58)
E_{surf}	50.54(1.29)	49.65(1.12)	5.77(0.09)	-4.88(0.76)
G_{gas}	- 11922.98(73.19)	- 11511.10(63.78)	-293.84(20.30)	- 118.03(18.39)
G_{solv}	-1674.29(45.62)	-1702.39(44.20)	-52.24(3.37)	80.34(12.96)
$E_{gas} + G_{sol}$	- 13597.28(45.86)	- 13213.49(37.97)	-346.08(19.74)	-37.69(4.79)
$-TS_{total}$	-1734.66(5.42)	-1679.17(6.14)	-84.86(0.69)	29.37(3.14)
ΔG_{bind}^{cal}				-8.32
ΔG_{bind}^{exp}				-7.23

* Experimental binding free energy is obtained from Ref 20.

Table S7 Binding free energy for CBM16-1/XL4 and decomposition to electrostatic interaction, van der Walls interaction, solvation free energies and entropy. Energies are in kcal/mol. Numbers in parentheses are standard deviations.

	complex	Receptor	Ligand	delta
E_{vdW}	-1155.23(14.99)	-1113.74(15.04)	-7.17(2.00)	-34.32(3.70)
E_{ele}	-	-		
	11072.42(82.57)	10777.59(80.98)	-269.72(7.22)	-25.11(9.25)
E_{GB}	-2201.26(71.16)	-2191.14(70.19)	-52.00(3.09)	41.88(6.40)
E_{surf}	45.52(1.21)	45.10(1.20)	4.90(0.04)	-4.48(0.47)
G_{gas}	-	-		
	12227.65(83.98)	11891.33(82.16)	-276.88(6.71)	-59.43(10.31)
G_{solv}	-2155.74(70.41)	-2146.04(69.53)	-47.09(3.07)	37.40(6.17)
$E_{gas} + G_{sol}$	-	-		
	14383.39(35.63)	14037.37(35.14)	-276.88(6.71)	-22.04(4.53)
$-TS_{total}$	-1608.45(5.70)	-1562.48(5.44)	-69.83(0.51)	23.84(3.70)
ΔG_{bind}^{cal}				1.80
ΔG_{bind}^{exp}				[a]NB

[a]NB indicates no binding detected.

Table S8 Binding free energy for CBM22-2 /CE4 and decomposition to electrostatic interaction, van der Walls interaction, solvation free energies and entropy. Energies are in kcal/mol. Numbers in parentheses are standard deviations.

	complex	Receptor	Ligand	delta
E_{vdW}	-1248.48(15.52)	-1208.85(15.45)	-14.03(1.64)	-25.60(3.75)
E_{ele}	-	-		
	10767.44(63.14)	10309.59(60.27)	-390.51(9.70)	-67.34(18.16)
E_{GB}	-1763.28(43.79)	-1773.57(42.68)	-60.22(4.06)	70.51(13.30)
E_{surf}	50.84(1.13)	48.87(1.05)	5.97(0.07)	-4.01(0.38)
G_{gas}	-	-		
	12015.93(63.71)	11518.45(60.43)	-404.54(9.70)	-92.94(18.09)
G_{solv}	-1712.44(43.44)	-1724.69(42.44)	-54.25(4.04)	66.50(13.09)
$E_{gas} + G_{sol}$	-	-		
	13728.37(38.48)	13243.14(36.96)	-458.79(8.14)	-26.44(7.48)
$-TS_{total}$	-1733.09(6.56)	-1671.16(6.54)	-85.13(0.61)	23.20(2.84)
ΔG_{bind}^{cal}				-3.24
ΔG_{bind}^{exp}				NE

[b]NE indicates no experimental value

Table S9. Sphere radius means the radius of the "solvent" sphere that will be used to compute the surface area accessible to solvent. 1.4 is a classical value employed as water radius. Threshold that determines if a residue is considered as buried or exposed. It corresponds to the ratio of the actual area computed and that of a reference. Here, the reference area is that of the residue in a peptide Ala-X-Ala in helical conformation. Classical threshold values are 0.15.

	Bck	S.Chn	Total	SC Ref.	Percent	SC Class
CBM16-1						
W23	29.39	52.72	82.11	267.74	19.69	E
W128	0.00	97.01	97.01	267.74	36.23	E
CBM22-2						
W53	0.61	138.41	139.02	267.74	51.70	E
Y103	22.95	82.32	105.27	232.25	35.44	E
Y136	0.00	2.35	2.35	232.25	1.01	B
CBM29-2						
W24	18.33	108.65	126.98	267.74	40.58	E
W26	10.36	85.76	96.11	267.74	32.03	E
Y46	0.00	119.11	119.11	232.25	51.28	E