Supporting Information

Investigating Conformational Dynamics of Lewis Y Oligosaccharide and Elucidating Blood Group Dependency of Cholera Using Molecular Dynamics

Rajarshi Roy[1], Biplab Ghosh[2], Parimal Kar [1, *]

[1]*Discipline of Biosciences and Biomedical Engineering, Indian Institute of Technology Indore, Khandwa Road, Indore 453552, India*

[2]*High Pressure and Synchrotron Radiation Physics Division, Bhabha Atomic Research Centre, Trombay, Mumbai 400085, India*

*Corresponding Author: Parimal Kar | Email: parimal@iiti.ac.in | Phone: +91 731 2438700 (Ext. 550)

Table S1: Distance between GlcNAc-GalNAc and GalNAc-Fuc3 in the crystal structure.

Distance (Å)	cCTB/A-Penta BGA	El-Tor/A-Penta BGA
GlcNAc-GalNAC	8.90	9.05
	9.22	9.25
	9.17	9.24
	9.37	-
	-	-
GalNAc-Fuc3	9.19	9.40
	9.52	9.37
	9.48	9.39
	9.55	-
	-	-

Table S2: Occupancy of hydrogen bonds between glycan and protein during MD simulation.

Acceptor	Donor	Distance (Å)	Angle (°)	Occupancy (%)
cCT/H-tetra BGA				
Fuc3@O4	H94@N-H	2.88	164.08	69.09
Fuc3@O5	T47@N-H	2.92	158.08	27.12
Fuc3@O2	Q3'@NE2-H	2.87	157.29	11.18
T47@O	Fuc3@O4-H	2.71	155.82	95.43
Q3'@OE1	Fuc3@O2-H	2.69	161.22	63.35
G45@O	GlcNAc@N2-H	2.86	154.27	58.46
ET CT/H-tetra BGA				
Fuc3@O4	H94@N-H	2.90	162.46	45.18
Fuc3@O5	I47@N-H	2.92	160.40	30.63
Fuc3@O2	Q3'@NE2-H	2.86	162.64	17.39
Gal@O2	H94@NE2-H	2.87	148.53	12.50
I47@O	Fuc3@O4-H	2.71	153.42	94.26
G45@O	GlcNAc@N2-H	2.84	156.80	59.38
Q3'@OE1	Fuc3@O2-H	2.70	161.30	52.50
Q3'@OE1	Fuc3@O3-H	2.75	159.24	20.96
Y18@OH	Gal@O4-H	2.81	154.65	17.34
cCT/A-penta BGA				
Fuc3@O4	H94@N-H	2.88	164.15	61.79
Fuc3@O5	T47@N-H	2.91	162.90	24.67
Fuc3@O2	Q3'@NE2-H	2.88	162.10	23.55
GlcNAc@O4	H18@NE2	2.87	152.90	18.25
GalNAc@O3	T19@N-H	2.88	155.39	12.50
T47@O	Fuc3@O4	2.69	156.66	96.31
Q3'@OE1	Fuc3@O2	2.70	161.70	44.12
N44@O1	OH@O1	2.72	157.23	37.89
G45@O	GlcNAc@N2	2.88	161.55	35.21
ET CT/A-penta BGA				
Fuc3@O4	H94@N	2.89	161.69	43.57
Gal@O4	H94@N	2.88	151.15	26.32
Fuc3@O5	I47@N	2.92	162.74	23.04
GalNAc@O3	N89@ND2	2.86	159.79	17.57
Fuc3@O2	Q3'@NE2	2.86	162.63	16.77
Fuc3@O5	Y18@OH	2.78	159.35	10.29
I47@O	Fuc3@O4-H	2.71	153.50	94.41
G45@O	GlcNAc@N2-H	2.86	160.50	57.60
Q3'@OE1	Fuc3@O2-H	2.70	161.32	57.18
N89@O	GalNac@O3-H	2.73	159.94	40.73
Q3'@OE1	Fuc3@O3-H	2.74	158.59	22.24

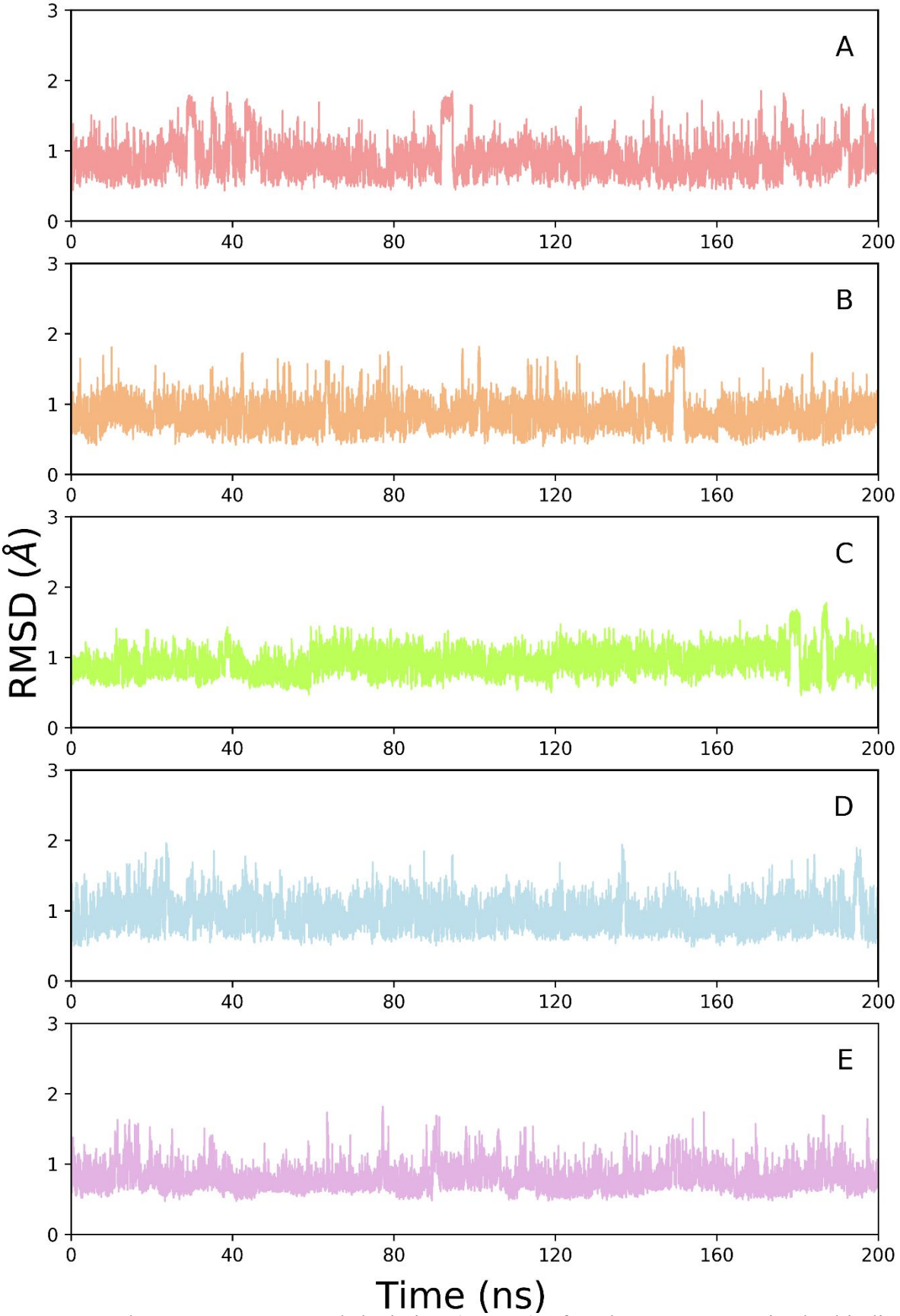

Figure S1: The root mean squared deviation (RMSD) of each H-tetra BGA in the binding site of cCT as a function of simulation time: (A) Chain A; (B) Chain B; (C) Chain C; (D) Chain D; and (E) Chain E.

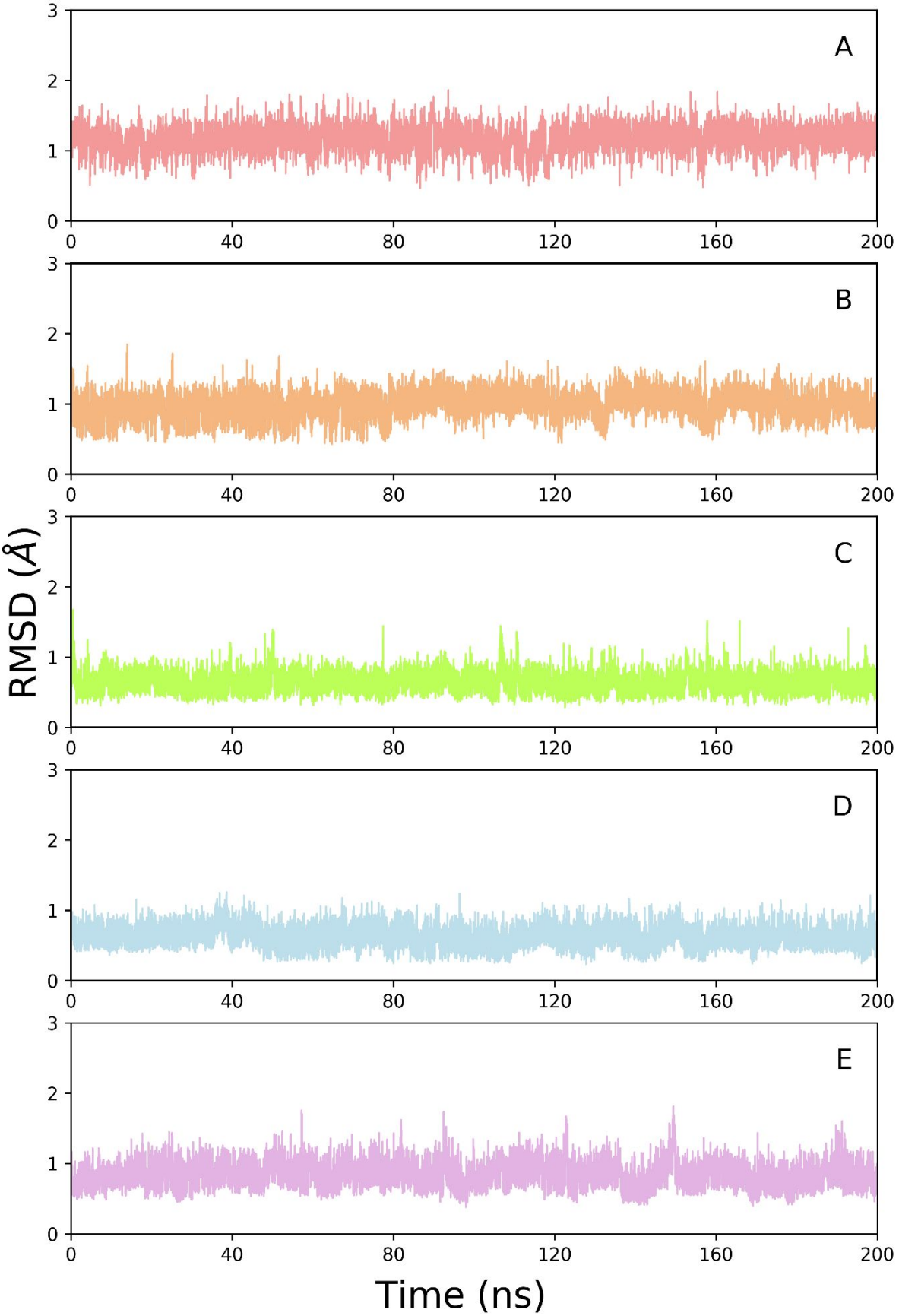

Figure S2: The root mean squared deviation (RMSD) of each H-tetra BGA in the binding site of ET CT as a function of simulation time. (A) Chain A; (B) Chain B; (C) Chain C; (D) Chain D; and (E) Chain E.

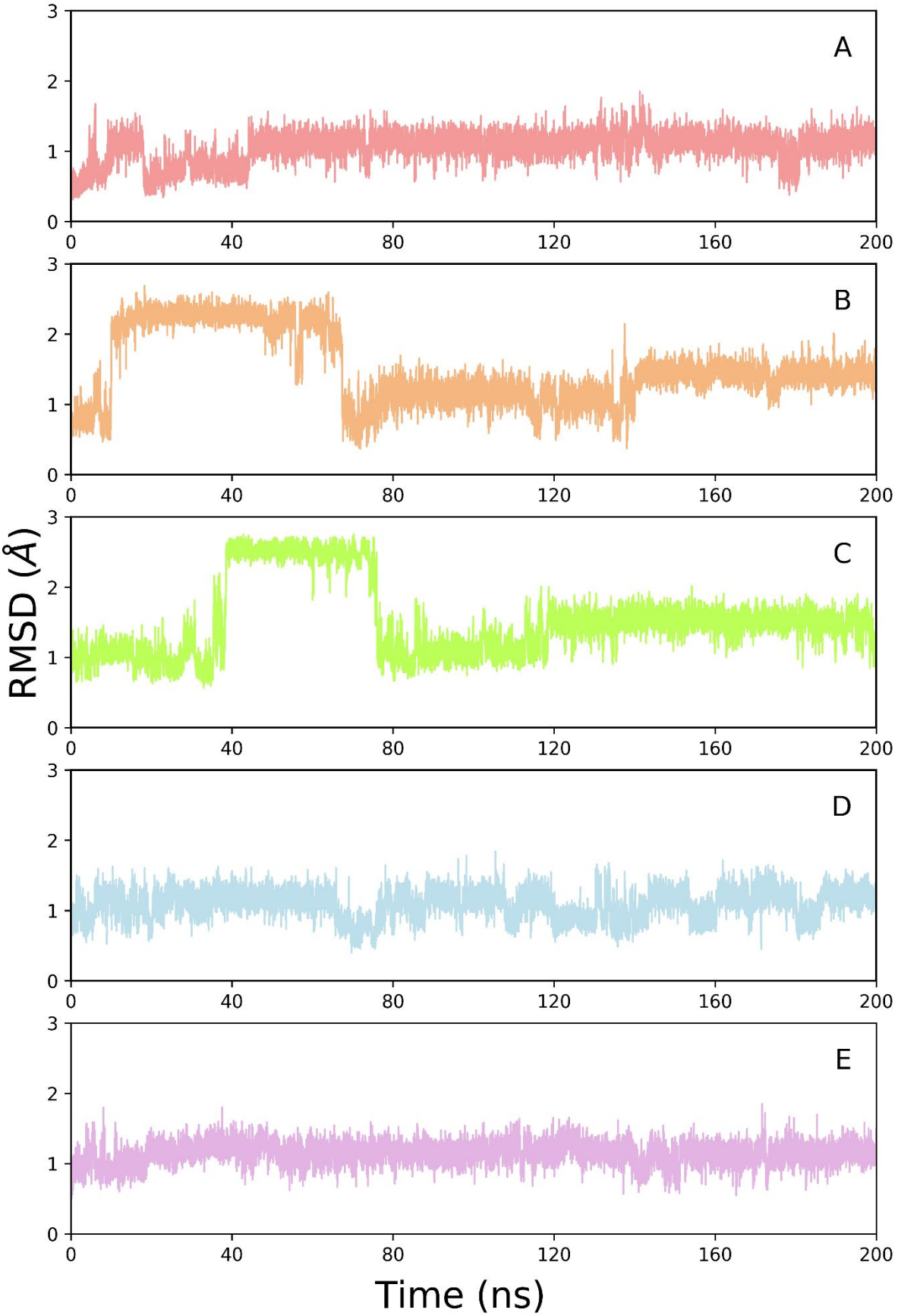

Figure S3: The root mean squared deviation (RMSD) of each A-penta BGA in the binding site of cCT as a function of simulation time. (A) Chain A; (B) Chain B; (C) Chain C; (D) Chain D; and (E) Chain E.

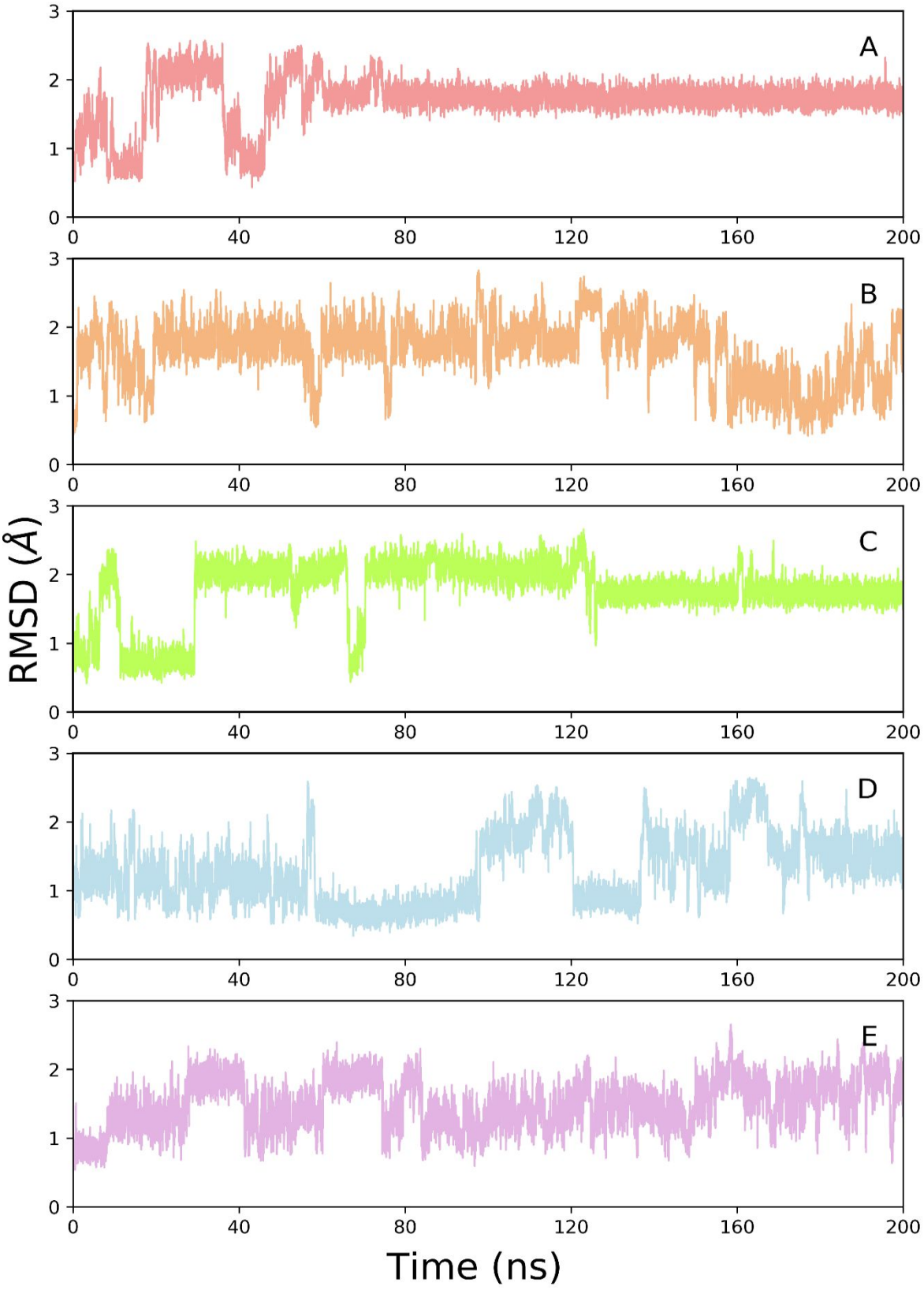

Figure S4: The root mean squared deviation (RMSD) of each A-penta BGA in the binding sites of ET CT as a function of simulation time. (A) Chain A; (B) Chain B; (C) Chain C; (D) Chain D; and (E) Chain E.

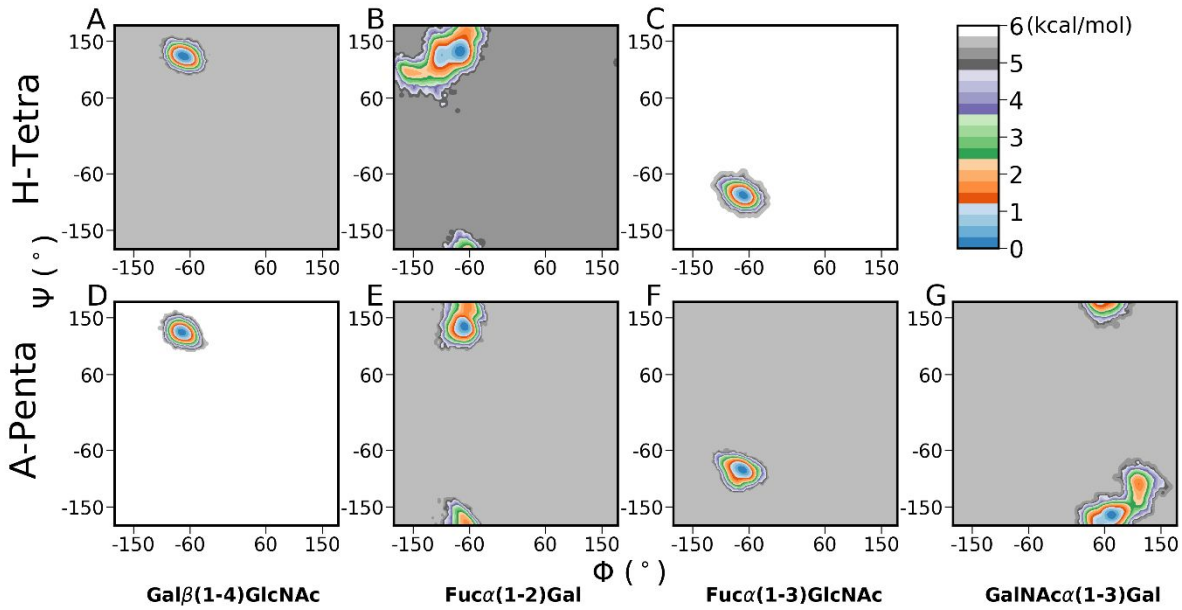

Figure S5: Free energy surfaces (kcal/mol) of the φ/ψ glycosidic torsions of cCT/H-tetra (A, B, and C) and cCT/A-penta (D, E, F, G) complexes obtained from conventional MD simulations. Top panel (H-tetra): A) Galβ(1-4)GlcNAc, B) Fucα(1-2)Gal, C) Fucα(1-3)GlcNAc. Bottom panel (A-penta): (D) Galβ(1-4)GlcNAc, (E) Fucα(1-2)Gal, (F) Fucα(1-3)GlcNAc, (G) GalNAcα(1-3)Gal.

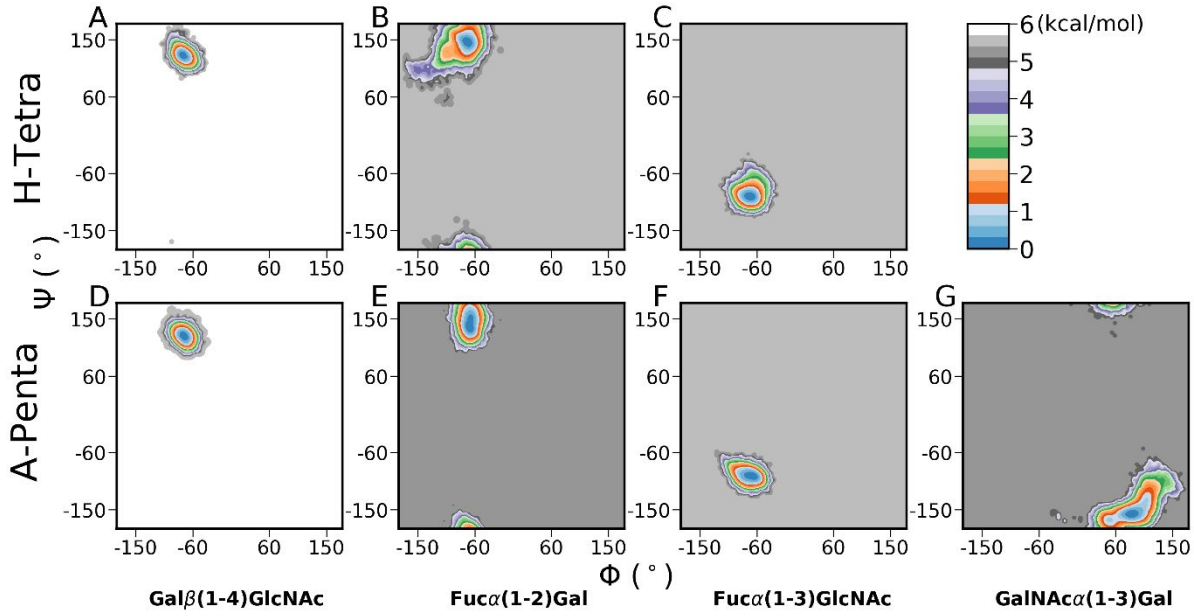

Figure S6: Free energy surfaces (kcal/mol) of the *φ*/*ψ* glycosidic torsions of ET CT/H-tetra (A, B, and C) and ET CT/A-penta (D, E, F, G) complexes obtained from conventional MD simulations. Top panel (H-tetra): A) Gal*β*(1-4)GlcNAc, B) Fuc*α*(1-2)Gal, C) Fuc*α*(1-3)GlcNAc. Bottom panel (A-penta): (D) Gal*β*(1-4)GlcNAc, (E) Fuc*α*(1-2)Gal, (F) Fuc*α*(1-3)GlcNAc, (G) GalNAc*α*(1-3)Gal.

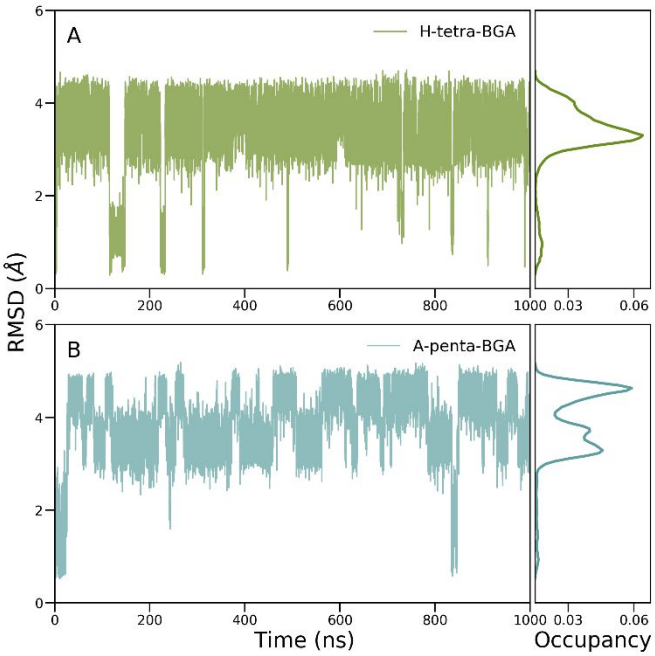

Figure S7: Time evolution of RMSD of H-tetra (A) and A-penta (B) BGA obtained from aMD simulations. Probability distributions of RMSD for H-tetra and A-penta BGAs obtained from aMD simulations are also shown in (A) and (B), respectively.

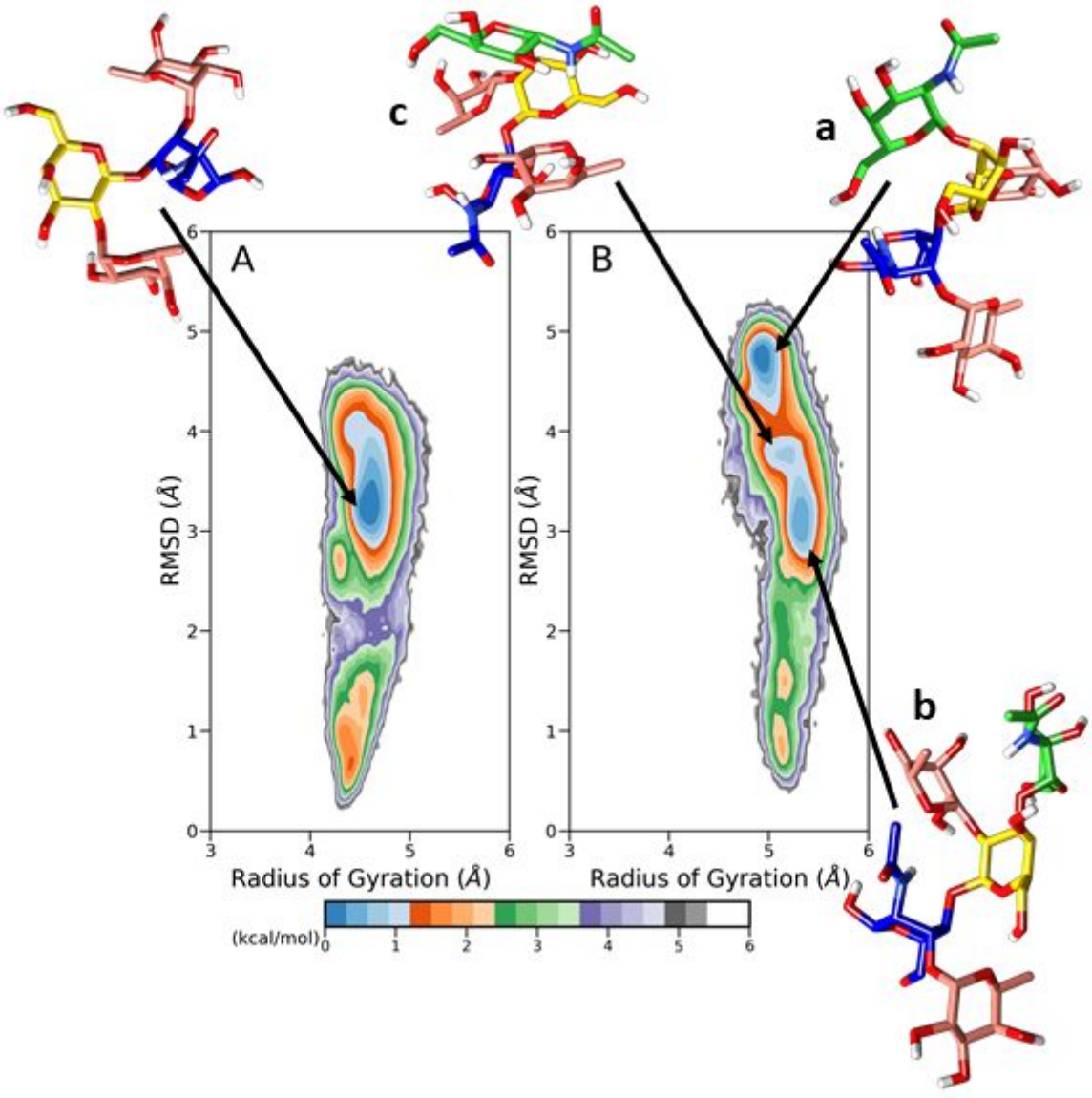

Figure S8: 2-dimensional free energy landscape of H-tetrasaccharide (A) and A-pentasaccharide (B) in solution obtained from aMD simulations. RMSD and Radius of gyration were taken as reaction coordinates. a) Structure corresponding to the principle minimum while (b) and (c) depict structures corresponding to secondary minima.

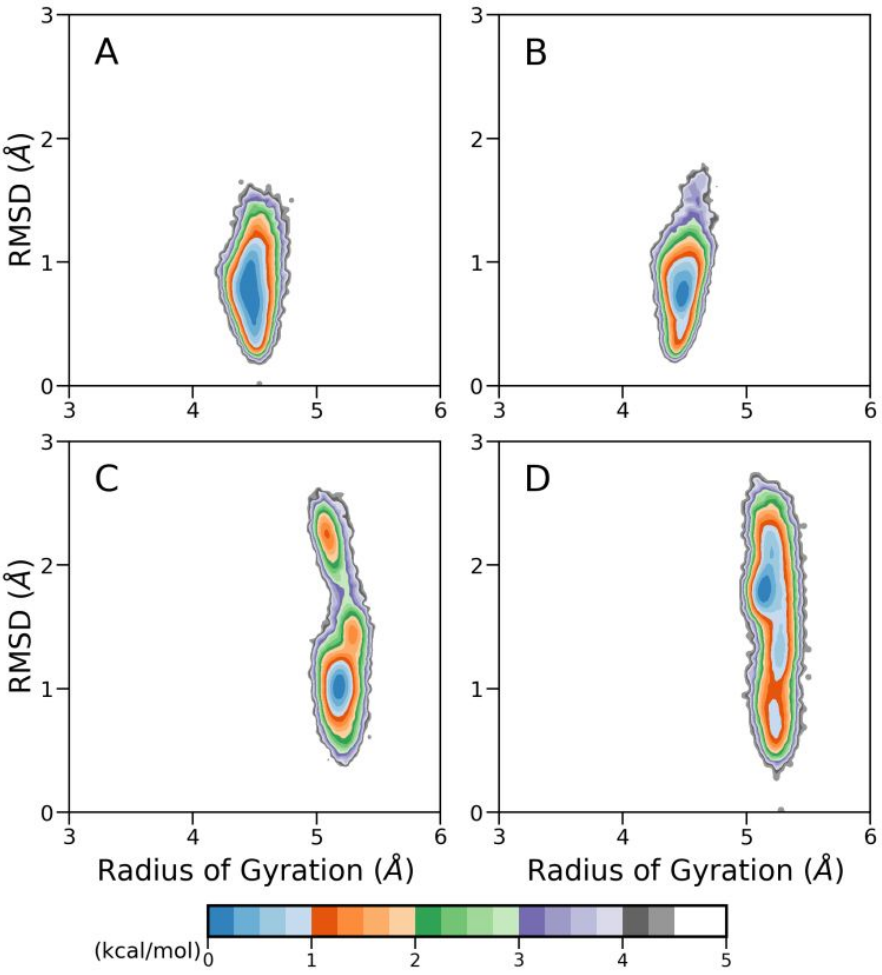

Figure S9: 2D PMF (potential of mean force) as a function of RMSD and radius of gyration: A) cCT/H-tetra, B) ET CT/H-tetra, C) cCT/A-penta and D) ET CT/A-penta complexes.

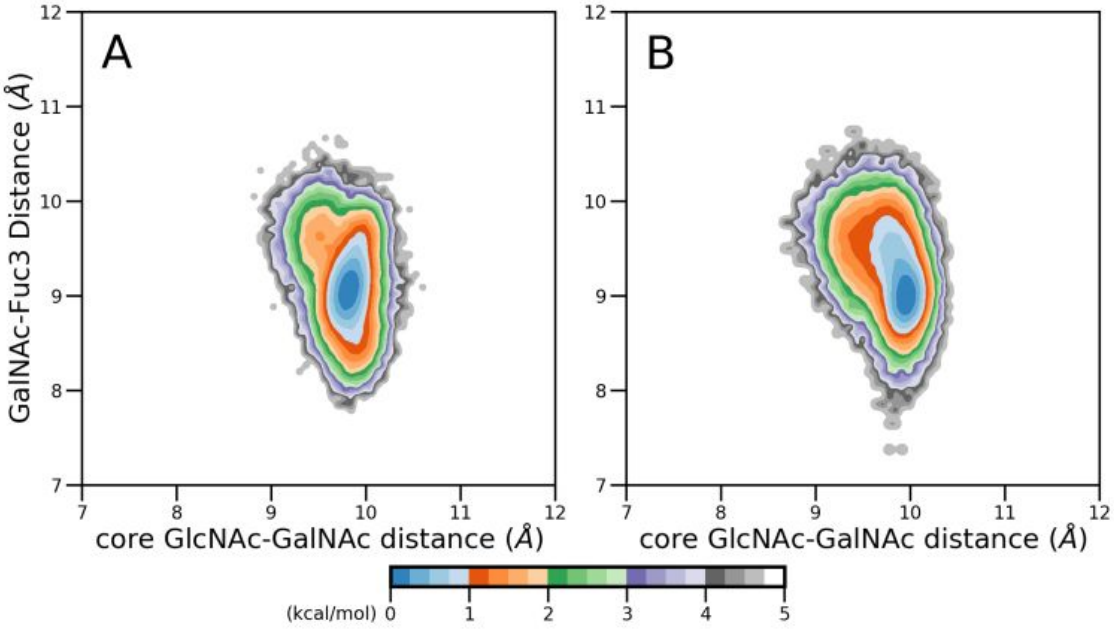

Figure S10: 2D PMF (potential of mean force) for A-Penta BGA in the binding site for classical (A) and El Tor (B) strains. The distances between GlcNAc and GalNac and GalNac and terminal Fuc3 residue were considered as reaction coordinates.

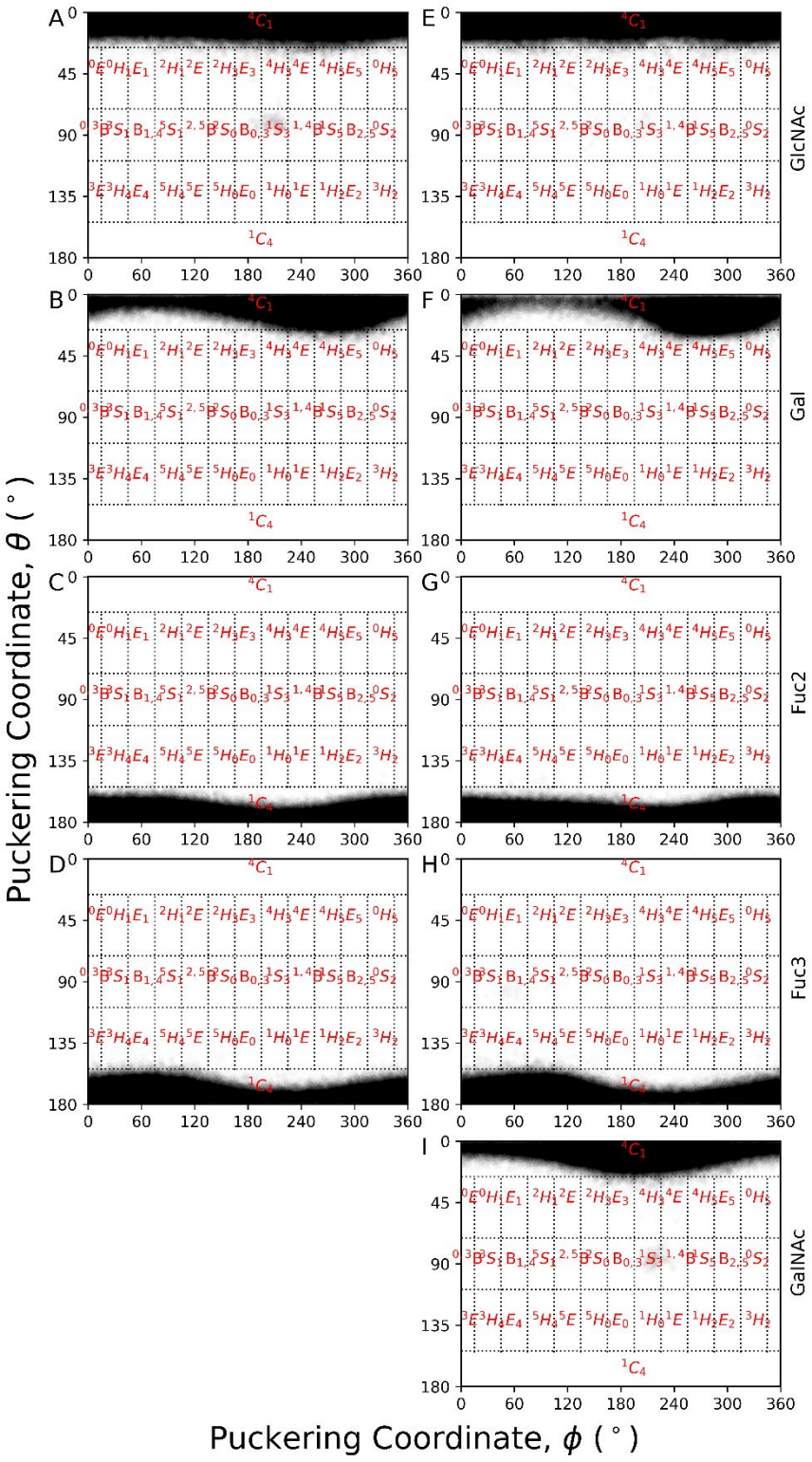

Figure S11: Mercator representation of the Cremer Pople Sphere for all the monosaccharide in solution derived from complex simulations. Left-hand images (A, B, C, D) correspond to the GlcNAc, Gal, Fuc2 and Fuc3 for cCT/H-tetra whereas right-hand images (E, F, G, H, I) correspond to GlcNAc, Gal, Fuc2, Fuc3 and GalNAc for cCT/A-penta.

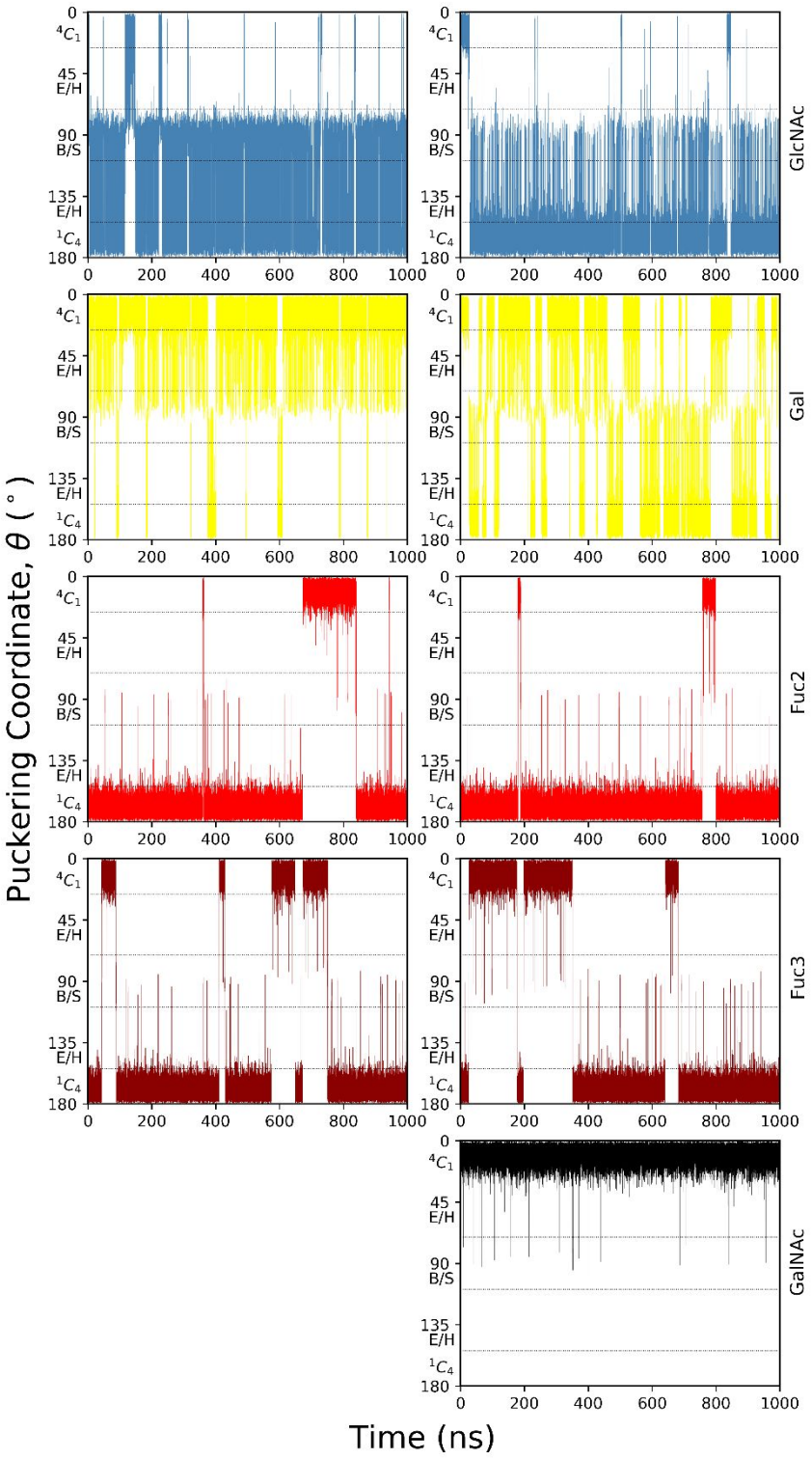

Figure S12: Time evolution of puckering coordinate for each monosaccharide in H-tetra and A-penta BGA obtained from aMD simulations.

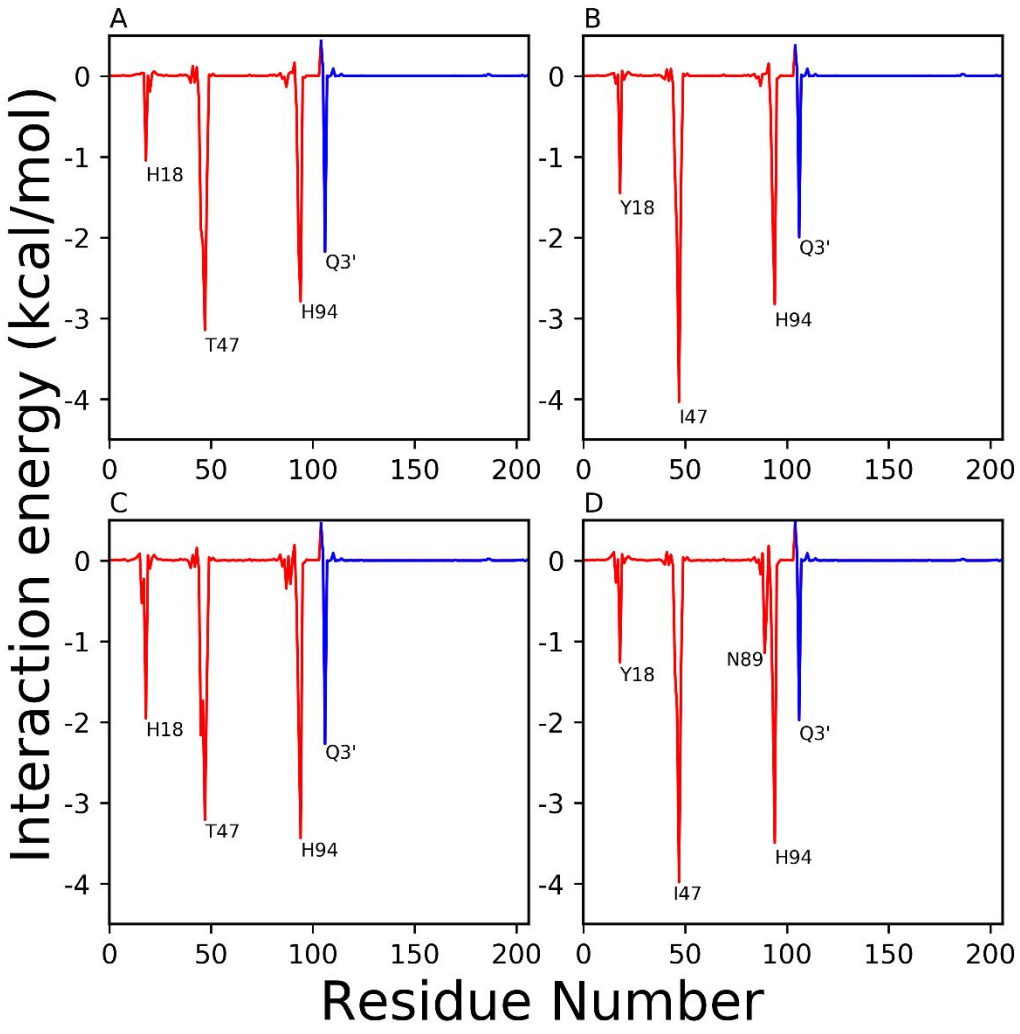

Figure S13: Decomposition of total binding free energy in each residue of all four complexes: (A) cCT/H-tetra; (B) ET CT/H-tetra; (C) cCT/A-penta; and (D) ET CT/A-penta. Two different colors denote two different monomers.

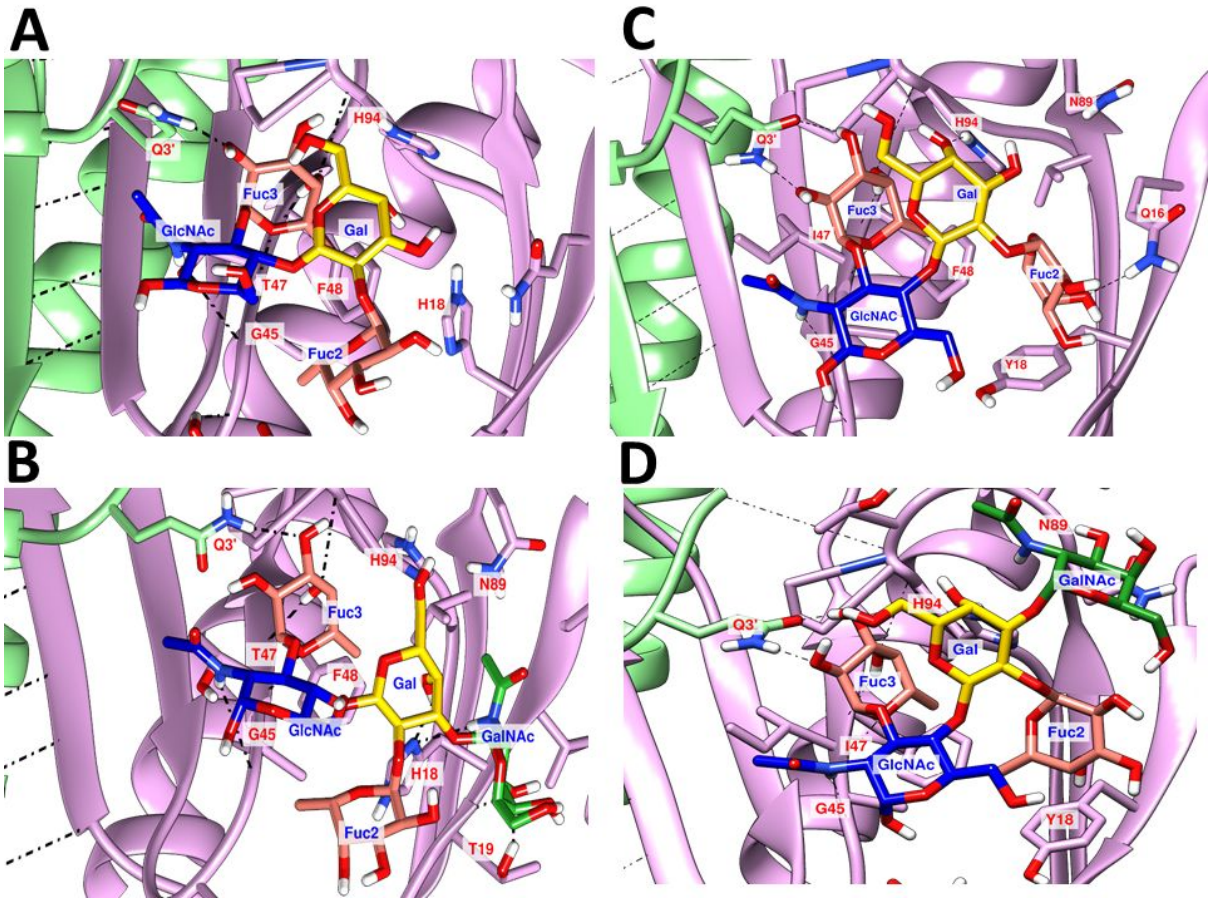

Figure S14: Carbohydrate-Protein interaction in Classical and El Tor strains with Lewis y oligosaccharides. A) cCTB with H-Tetra BGA, B) cCTB with A-Penta BGA, C) El Tor with H-Tetra BGA, D) El Tor with A-Penta BGA. Cholera toxin is shown in ribbon and different color depicts different chain. Hydrogen bonds are shown in the black dashed line. Glycans and interacting amino acids are shown in ball and stick model.

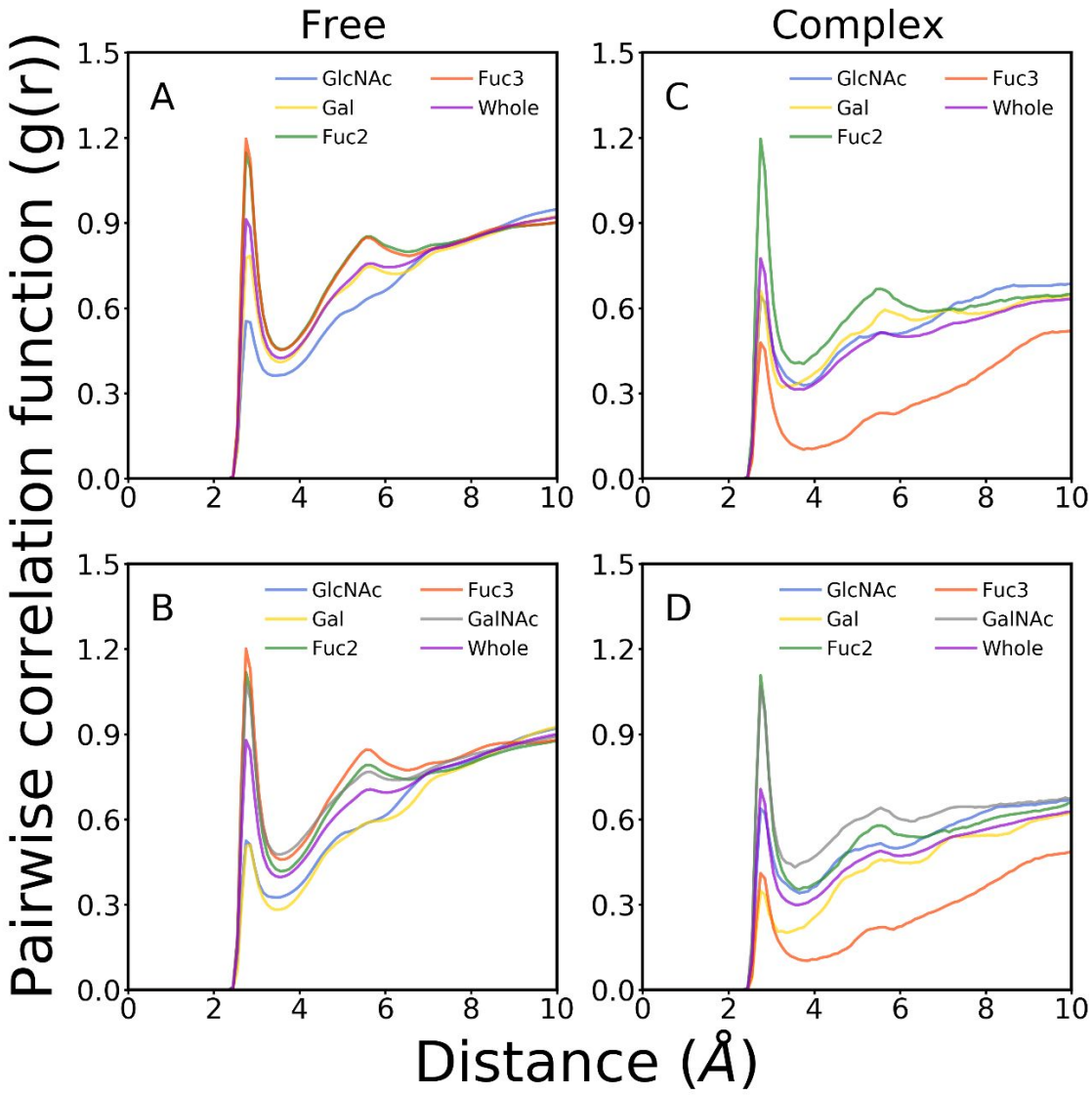

Figure S15: Pairwise correlation function, g(r) of water molecules in the simulation as a function of distance from the oxygen atoms of each monosaccharide residues as well as the whole (violet) line. A) & C) corresponds to H-tetra BGA in free and complex simulation respectively. B) & D) corresponds to A-penta BGA in free and complex simulation, respectively. Each of the monosaccharides is shown in a different color. GlcNAc: blue, Gal: yellow, Fuc2: green, Fuc3: orange, and GalNAc: grey.