ACS OMEGA

http://pubs.acs.org/journal/acsodf

Article

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

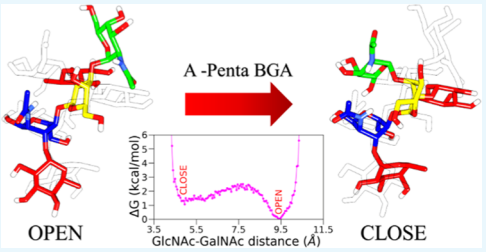
Rajarshi Roy, Biplab Ghosh, and Parimal Kar*

Cite This: *ACS Omega* 2020, 5, 3932−3942

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Cholera is caused by *Vibrio cholerae* and is an example of a blood-group-dependent disease. Recent studies suggest that the receptor-binding B subunit of the cholera toxin (CT) binds histo−blood group antigens at a secondary binding site. Herein, we studied the conformational dynamics of Lewis Y (Le^Y) oligosaccharides, H-tetrasaccharides and A-pentasaccharides, in aqueous solution by conducting accelerated molecular dynamics (aMD) simulations. The flexible nature of both oligosaccharides was displayed in aMD simulations. Furthermore, aMD simulations revealed that for both oligosaccharides in the free form, $^4C_1$ and $^1C_4$ puckers were sampled for all but GalNAc monosaccharides, while either the $^4C_1$ (GlcNAc, Gal, GalNAc) or $^1C_4$ (Fuc2, Fuc3) pucker was sampled in the CT-bound forms. In aMD, the complete transition from the $^4C_1$ to $^1C_4$ pucker was sampled for GlcNAc and Gal in both oligosaccharides. Further, we have observed a transition from the open to closed conformer in the case of A-pentasaccharide, while H-tetrasaccharide remains in the open conformation throughout the simulation. Both oligosaccharides adopted an open conformation in the CT binding site. Moreover, we have investigated the molecular basis of recognition of Le^Y oligosaccharides by the B subunit of the cholera toxin of classical and El Tor biotypes using the molecular mechanics generalized Born surface area (MM/GBSA) scheme. The O blood group determinant, H-tetrasaccharide, exhibits a stronger affinity to both biotypes compared to the A blood group determinant, A-pentasaccharide, which agrees with the experimental data. The difference in binding free energy between O and A blood group determinants mainly arises due to the increased entropic cost and desolvation energy in the case of A-pentasaccharide compared to that of H-tetrasaccharide. Our study also reveals that the terminal Fuc3 contributes most to the binding free energy compared to other carbohydrate residues as it forms multiple hydrogen bonds with CT. Overall, our study might help in designing glycomimetic drugs targeting the cholera toxin.



INTRODUCTION

There is a strict variation in the blood group type and frequencies among the human population throughout the world, which is a result of the genetic pressure giving protection from different infectious diseases.[1] Among all of these blood-group-dependent diseases, cholera is one of the deadliest and epidemic diseases in human history and most prevalent in the African subcontinent. More than 1.2 million cases were reported in 2017 all over the world, with a fatality rate of 0.5. Cholera is caused by the pathogen *Vibrio cholerae*.[2] The relationship between blood groups and cholera was first shown by Barua and Paguio.[3] People with the O blood group tend to acquire a higher risk in the development of severe cholera symptoms.[4−7] Further, evolutionary history also suggests an association between the blood group O and cholera. The Ganges delta is a region where cholera is prevalent from ancient times. As a result of the genetic pressure, the O blood group is decidedly less prominent in the Ganges delta region.[8] All of these studies indicate an active role of the blood groups, especially the O blood group, in the infection and development of cholera. Despite several experimental analyses, a detailed understanding of the molecular mechanism underlying this cause is still missing, and this needs to be addressed.[9]

Throughout the world, cholera infection can be divided into two serogroups, O1 and O139. Further, O1 can be classified into classical and El Tor biotypes.[10,11] Both biotypes produce cholera toxins (cCT and ET CT) differing in two residues of the B subunit, H18Y and T47I (classical and El Tor). The interaction of the blood group with El Tor was observed to be stronger compared to that with the classical strain.[12]

Blood group antigens (BGAs) are cell surface carbohydrates that are conjugated with glycolipids (e.g., GM1) or membrane glycoproteins expressed mainly in red blood cells and epithelial cells. The Lewis blood group antigens are mostly fucosylated glycans and possess a typical basic structure composed of

Received: October 12, 2019
Accepted: February 11, 2020
Published: February 19, 2020

https://dx.doi.org/10.1021/acsomega.9b03398
ACS Omega 2020, 5, 3932−3942

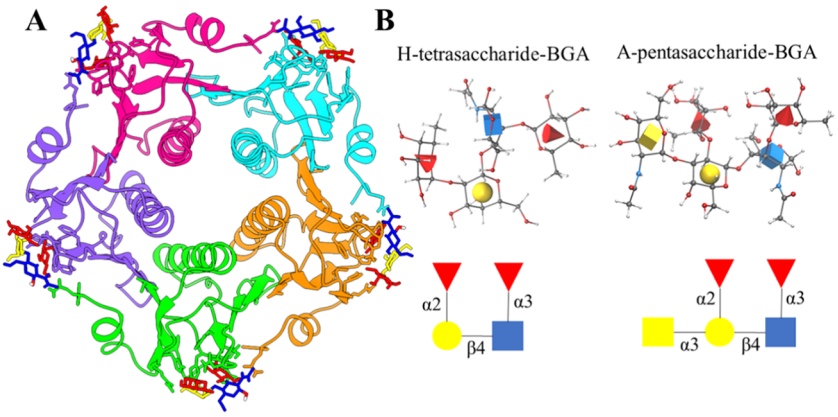


Figure 1. (A) Crystal structure of the pentameric cholera toxin B subunit in complex with glycans. Each protein chain is shown in a different color, and the glycan is represented in the ball-and-stick model. (B) Structure and nomenclature of Lewis Y (Le^Y) blood group determinants: H-tetrasaccharide and A-pentasaccharide BGAs. The structure and nomenclature of Le^Y blood group determinants are displayed in three-dimensional (3D) symbol nomenclature for glycans (SNFG) icon mode. In this representation, *N*-acetylglucosamine is shown as a blue cube, L-fucose is shown as a red cone, D-galactose is shown as a yellow sphere, and *N*-acetylgalactosamine is shown as a yellow cube. All images were generated using Visual Molecular Dynamics (VMD) with the help of the 3D-SNFG plugin available at www.glycam.org/3d-snfg. A detailed atomic structure is also shown in the ball-and-stick model, where the standard color scheme is used to depict different types of atoms.

galactose (Gal), *N*-acetylglucosamine (GlcNAc), and fucose (Fuc) residues.[13] There are four Lewis antigens, namely, Lewis A (Le^A), Lewis B (Le^B), Lewis X (Le^X), and Lewis Y (Le^Y). Further, Le^A or Le^X can often be found sialylated to yield sialyl Lewis A (sLe^A) or sialyl Lewis X (sLe^X), respectively. For Le^X, the sequence of the core trisaccharide is found to be Galβ(1−4)(Fucα(1−3))GlcNAcβ.[14] Le^Y is obtained when a second Fucα(1−2) is attached to the terminal Gal residue of Le^X. All of these determinants are characteristics of blood group O. An enzymatic addition of *N*-acetylgalactosamine (GalNAc) or galactose (Gal) yields A or B blood group, respectively. In this study, we consider only Le^Y oligosaccharides, as shown in Figure 1B. Recently, Heggelund et al.[15] reported the crystal structure of cCT and ET CT with Le^Y blood group determinants, H-tetra and A-penta BGAs, which opens the possibility of systematic dynamical studies of interactions between blood groups and cholera toxins. Moreover, it was shown that H-tetrasaccharide binds to CT with a much higher affinity in comparison to A-pentasaccharide.

The interaction between carbohydrates and the cholera toxin is highly specific and depends on the conformational variation. The cholera toxin binds to not only GM1 but also various analogues of GM1.[16] A previous study on the interaction of the cholera toxin with several monosialoganglio-sides describes the possible binding modes of carbohydrates to the toxin and elucidates the role of water-mediated hydrogen bonds that stabilize the interaction.[17] Several studies involving different analogues of sialic acid against the cholera toxin showed that it is successfully recognized and interacts with the analogues.[18,19] Cervin et al. show that CT can induce toxicity in the absence of GM1 and support a role for host glycoproteins in CT intoxication. It is worth noting here that these glycoprotein receptors are fucosylated since human intestinal epithelial cells are heavily fucosylated.[20] Fazil and co-workers have shown by utilizing molecular dynamics simulations that the binding efficiency of the carbohydrate units of GM1 receptors varies against different genotypes of CT.[21]

In this study, we focus on investigating the conformational dynamics of Lewis Y oligosaccharides (i.e., H-tetra and A-pentasaccharides) and elucidating the biophysical basis of their recognition by the cholera toxin of two biotypes. Previously, various computational methods have been employed to study the conformational dynamics of glycans.[22−29] In general, Lewis sugars behave as a rigid structure in solution stabilized by the stacking interaction between fucose and galactose rings.[30] In this study, we have utilized the accelerated molecular dynamics (aMD)[31] technique for investigating the conformational dynamics of free Le^Y oligosaccharides. We compare the conformational preferences in solution and bound to CT by constructing the underlying free energy surfaces corresponding to the glycosidic linkages. It is worth noting here that the conformational dynamics of H-tetra and A-pentasaccharides have not been studied previously. Finally, we calculate the binding free energy of each glycan complexed with both biotypes of CT by employing the popular MM-GBSA[32] scheme, which is an average estimation of molecular mechanics and solvation free energy. Since the binding event occurs in aqueous solution,[33] MD simulations were conducted in an explicit water environment.

RESULTS AND DISCUSSION

Structural Stability and Flexibility of the Complexes.

First, we have computed the root-mean-squared deviations (RMSDs) of backbone atoms compared to the initial structure for all four protein−glycan systems, and the time evolution of RMSD is shown in Figure 2A. In the case of all four complexes, the RMSD values quickly reach the plateau and remain stable throughout simulations, suggesting the convergence of simulations. The average RMSD is estimated to be 1.2 Å for the classical strain bound to H-tetra and A-pentasaccharides. Similarly, average RMSDs of 1.3 and 1.4 Å are obtained when the El Tor (ET) strain is complexed with the H-tetra and A-pentasaccharides, respectively. The main reason for the structural stability of the pentameric cholera toxin could be a conserved intermonomer hydrogen bond.

Next, we investigated the flexibility of each residue by computing root-mean-squared fluctuations (RMSFs) of C_α atoms, which are shown in Figure 2B. A similar variation is observed for all four protein−glycan complexes. Further, it is revealed in Figure 2B that the region between residues 51 and 58 is the most flexible region. Another highly flexible region is

https://dx.doi.org/10.1021/acsomega.9b03398
ACS Omega 2020, 5, 3932−3942

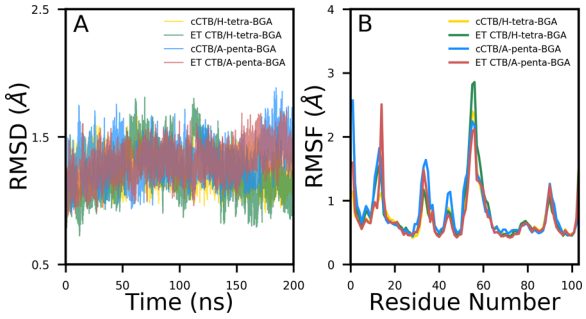


Figure 2. (A) Time evolution of the root-mean-square deviations (RMSDs) of backbone atoms with respect to the initial structure and (B) root-mean-square fluctuations (RMSFs) of C_α atoms for all four complexes.

located between residues 8 and 12. Both the areas correspond to a dynamic loop and are solvent-exposed.

Conformational Dynamics of Glycans Bound to CT. To investigate the conformational features of glycans, we have calculated the RMSD of each glycan bound to each monomer of the pentameric cholera toxin of both biotypes, which is shown in the Supporting Information (see Figures S1−S4). For both biotypes, the RMSD of each of the five H-tetrasaccharides is very stable (see Figures S1 and S2). In contrast, the RMSD of A-pentasaccharide is found to be varying and fluctuating (see Figures S3 and S4), suggesting a more flexible nature of A-pentasaccharide compared to that of H-tetrasaccharide in the binding site. Such a flexible nature was observed for the Le^X trisaccharide bound to RSL.[28] In the case of H-tetrasaccharide, an average RMSD of ∼0.9 Å is obtained when it is bound to either cCTB or ET CTB (see Table 1). On the other hand, average RMSDs of 1.3 and 1.6 Å are obtained for A-pentasaccharide bound to cCTB and ET CTB, respectively (see Table 1). Furthermore, it is evident from Figures S3 and S4 that the flexibility of A-pentasaccharide is biotype-specific as it is found to be more flexible when it is bound to ET CTB than to cCTB. An addition of GalNAc monosaccharide to A-pentasaccharide causes a higher degree of flexibility compared to that for H-tetrasaccharide.

Further, we analyzed the flexibility of the glycosidic linkages for both glycans bound to both biotypes by constructing two-dimensional (2D) free energy surfaces of φ/ψ glycosidic torsions (see Figures S5−S6, Supporting Information). Three glycosidic linkages of H-tetra are Galβ(1−4)GlcNAc, Fucα(1−2)Gal, and Fucα(1−3)GlcNAc, while four glycosidic linkages of A-penta are Galβ(1−4)GlcNAc, Fucα(1−2)Gal, Fucα(1−3)GlcNAc, and GalNAc(1−3)Gal. For constructing the φ/ψ map, we generated a 1 μs trajectory by extracting the trajectory of each BGA at the five binding sites and combining such five trajectories (5 × 200 ns). For both glycans bound to both biotypes, the free energy surfaces are similar, and we obtained a narrow and single free energy minimum for all of the glycosidic linkages. In the case of H-tetra, the principal minima corresponding to Galβ(1−4)GlcNAc, Fucα(1−2)Gal, and Fucα(1−3)GlcNAc linkages almost coincide with the corresponding linkages of A-penta. However, the φ/ψ distribution of the Fucα(1−2)Gal linkage shows a slightly broader free energy minimum in H-tetra compared to that in A-pentasaccharide. Moreover, the free energy surface of the GalNAcα(1−3)Gal linkage in A-pentasaccharide is also found to be a little broad. This observation suggests that the movements of the glycans get restricted while they are bound to CT.

Conformational Dynamics of Free Oligosaccharides in Aqueous Solution. The flexible nature of the glycan is revealed in 1 μs long aMD simulations. The time evolutions of RMSD of nonhydrogen atoms for both BGAs are displayed in Figure S7A,B. The average RMSD for both BGAs obtained from the aMD simulation is reported in Table 1. It can be noted from Table 1 that the average RMSD value of the free oligosaccharide is much higher compared to that of the oligosaccharide at the binding sites and is estimated as 3.3 and 3.9 Å for H-tetra and A-penta BGA, respectively (see Table 1). We have also calculated and plotted the probability distribution of RMSD for both glycans in Figure S7C,D. It displays a single peak at ∼3.3 Å for H-tetra BGA, while three peaks at 3.3, 3.7, and 4.7 Å are obtained for A-penta BGA. Furthermore, the conformational dynamics of both BGAs are illustrated by the free energy landscape (FEL) of RMSD and the radius of gyration (R_g) and are shown in Figure S8A,B. It is evident from Figure S8 that in H-tetra BGA, a single broad conformational space is explored with a global free energy minimum at RMSD/R_g ∼3.3/∼4.6 Å, while the FEL of A-penta BGA is characterized by three dispersed basins (RMSD/R_g: ∼3.3/∼5.3, ∼3.7/∼5.2, ∼4.7/∼4.9 Å). The global free energy minimum in A-penta BGA is located at ∼4.7, ∼4.9 Å. In contrast, in the case of bound H-tetra BGA, the free energy minimum is located at ∼0.8, ∼5.1 Å, as shown in Figure S9. On the other hand, in the case of A-penta BGA, the global free energy minimum is located at ∼1.0, ∼5.2 and ∼1.8, ∼5.1 Å for cCTB and ET CTB, respectively. Overall, this suggests that the conformational ensemble of A-penta BGA in solution is more diverse compared to that of H-tetra.

Next, we constructed the free energy surfaces of the glycosidic linkages for both H-tetra and A-penta BGAs, which are shown in Figure 3. Overall, Figure 3 indicates a broad sampling around the principal minimum for all of the

Table 1. Average RMSDs Obtained from Simulations of Free Glycans as Well as Complex Simulations[a]

		complex				free	
		cCT/H-tetra	ET CT/H-tetra	cCT/A-penta	ET CTB/A-penta	H-tetra aMD	A-penta aMD
chain	A	0.9 (0.2)	1.2 (0.2)	1.1 (0.2)	1.7 (0.3)	3.3 (0.7)	3.9 (0.7)
	B	0.9 (0.2)	1.0 (0.2)	1.5 (0.5)	1.7 (0.4)		
	C	0.9 (0.2)	0.6 (0.1)	1.5 (0.5)	1.8 (0.5)		
	D	0.9 (0.2)	0.6 (0.2)	1.1 (0.2)	1.3 (0.5)		
	E	0.8 (0.2)	0.8 (0.2)	1.2 (0.2)	1.5 (0.3)		
global		0.9 (0.2)	0.9 (0.2)	1.3 (0.4)	1.6 (0.4)		

[a]All values are in Å. Standard deviations are provided in parentheses.

https://dx.doi.org/10.1021/acsomega.9b03398
ACS Omega 2020, 5, 3932−3942

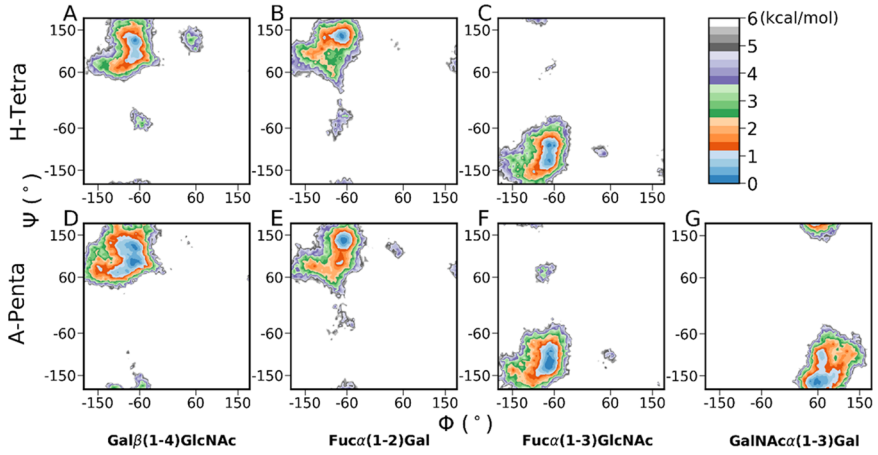


Figure 3. Free energy surfaces (kcal/mol) of the φ/ψ glycosidic angles of (A−C) H-tetra and (D−G) A-penta BGAs obtained from accelerated MD simulations. Top panel (H-tetra): (A) Galβ(1−4)GlcNAc, (B) Fucα(1−2)Gal, and (C) Fucα(1−3)GlcNAc. Bottom panel (A-penta): (D) Galβ(1−4)GlcNAc, (E) Fucα(1−2)Gal,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

glycosidic linkages. On comparing Figure 3 with Figure S5 (Supporting Information), we find out that the locations of principal minima for all glycosidic linkages are almost similar in the free and CT-bound states. However, the regions of minima of both glycans are much broader in solution compared to those in the CT-bound form. It is worth noting here that in the case of A-pentasaccharide, the core linkage [Galβ(1−4)-GlcNAc] displayed two distinct minima closely connected by a very shallow free energy barrier (Figure 3D), while a single minimum was obtained for H-tetrasaccharide (Figure 3A).

Investigating Closed and Open Conformations of Free A-Penta BGA in Solution. To further characterize the structural features, the distances between GlcNAc-GalNAc (R_1) and GalNAc-Fuc3 (R_2) for each CTB-bound A-penta BGA are calculated from the PDB (Protein Data Bank) structure for both biotypes and are listed in Table S1 (Supporting Information). It can be seen in Table S1 that R_1 varies between 8.9 and 9.4 Å while R_2 varies between 9.2 and 9.6 Å for both biotypes. This suggests that A-penta adopts an extended conformation in the binding site. To explore conformational preferences of unbound A-penta in aqueous solution, we constructed a two-dimensional potential of mean force (2D PMF) from the aMD simulation with R_1 and R_2 as reaction coordinates (see Figure 4). It is characterized by two distinct free energy wells separated by a relatively high free energy barrier. R_1 varies between 5 and 9 Å, suggesting that both open and closed conformations are explored by free A-penta in solution, which is in contrast to what has been observed in the CTB-bound X-ray structure. Further, Figure 4 suggests that the open conformation, characterized by R_1/R_2: ∼9.4/11.1 Å, is more probable compared to the closed structure (R_1/R_2: ∼5.3/9.3 Å). The three-dimensional structures of both conformations are also shown in Figure 4. The closed structure is characterized by a hydrogen bond between O5@GlcNAc and O6@GalNAc, which is absent in the open structure. The extended or open conformation is linear, and all of the terminal fucose residues are also perpendicular to the main branch. These two different forms of A-penta oligosaccharide can bind to different targets, as is the case with Lewis X.[34,35] The same 2D PMF was generated for A-penta BGA from the complex simulations and is shown in Figure S10 (Supporting Information). Here, for both biotypes, only the open conformation, characterized by R_1/R_2:

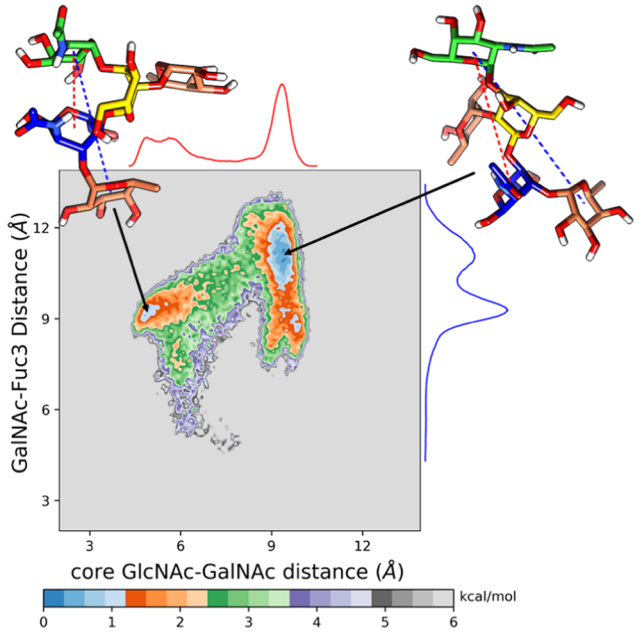


Figure 4. Two-dimensional potential of mean force for A-penta BGA in solution obtained from the aMD simulations. The distances between GlcNAc and GalNac and GalNac and terminal fucose residue (Fuc3) were considered as reaction coordinates.

∼10/9 Å, is explored. Overall, our simulations suggest that A-penta can adopt both open and closed conformations in solution, while only the open conformation is preferred when it is bound to pentameric CTB.

Puckering Conformations of the Pyranose Ring. Next, we characterized the shape variation of the individual sugar ring pucker for both oligosaccharides in solution and compared it with the protein-bound state. The IUPAC convention was used to calculate the canonical puckering conformations, and the Mercator representations of the Cremer−Pople sphere[36] depicting pyranose shapes for aMD simulations of both BGAs in solution are shown in Figure 5. Both chair conformations ($^4C_1$ and $^1C_4$) were sampled for all (GlcNAc, Gal, Fuc2, and Fuc3) but GalNAc monosaccharides, as is evident from Figure 5. On the other hand, either the $^4C_1$ (GlcNAc, Gal, GalNAc) or $^1C_4$ (Fuc2, Fuc3) pucker conformation was sampled in all

https://dx.doi.org/10.1021/acsomega.9b03398
ACS Omega 2020, 5, 3932−3942

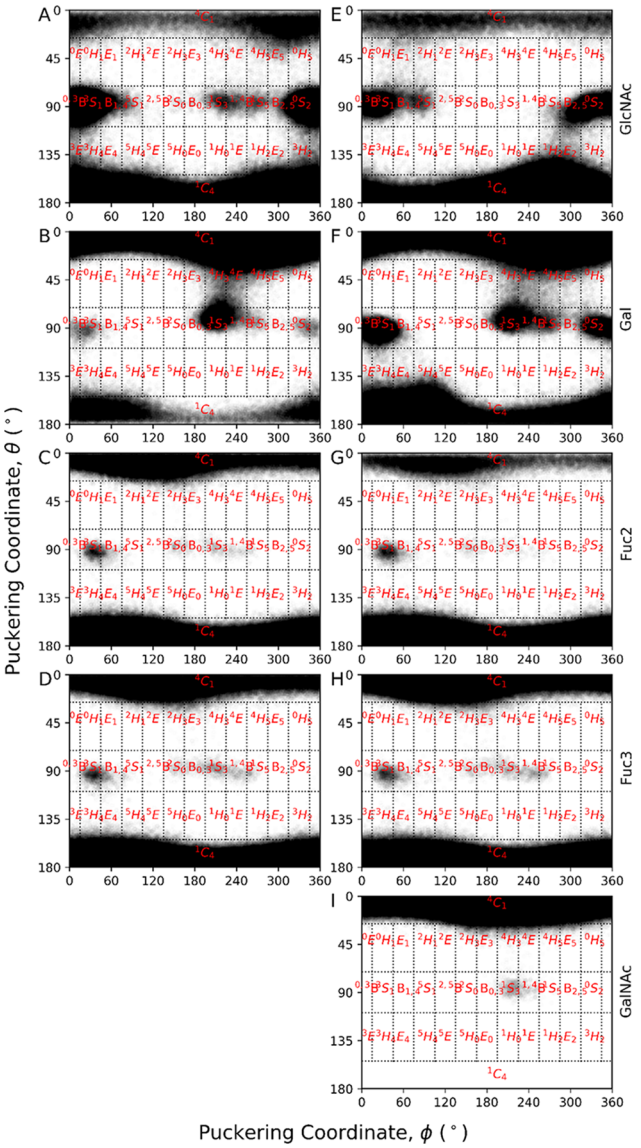


Figure 5. Mercator representation of the Cremer−Pople Sphere for all of the monosaccharides in solution derived from the accelerated MD simulations. Left-hand images (A)−(D) correspond to the GlcNAc, Gal, Fuc2, and Fuc3, respectively, from H-tetra BGA, whereas right-hand images (E)−(I) correspond to GlcNAc, Gal, Fuc2, Fuc3, and GalNAc, respectively, from A-penta BGA.

monosaccharides for both BGAs bound to CT (Figure S11, Supporting Information).

Next, we computed the one-dimensional (1D) puckering free energy profiles for all monosaccharides in both BGAs, which are shown in Figure 6. It is observed from Figure 6A,B that only for GlcNAc and Gal monosaccharides in both H-tetra and A-penta BGAs, the complete $^4C_1$ to $^1C_4$ pucker transition was sampled in solution. A similar trend was obtained by Alibay and co-workers for the Lewis X trisaccharide.[14] Figure 6C,D further reveals that for both BGAs bound to CTB, only the $^4C_1$ pucker was sampled for GlcNAc, GalNAc, and Gal monosaccharides, while the $^1C_4$ pucker conformation was populated for both fucoses (Fuc2, Fuc3). On computing the puckering angle ⊖ as a function of simulation time from aMD simulations of free BGAs in solution, we noted that for GlcNAc in both BGAs, the $^1C_4$ pucker was sampled more frequently than $^4C_1$ with significant time in $^{0,3}B$, $^3S_1$, and $^0S_2$ (Figure S12, Supporting Information). Further, Figure 6A,B suggests that the boat/skewed-boat conformations in GlcNAc are energetically more favorable by ∼1.5 kcal/mol for H-tetra than for A-penta. In contrast, for Gal in H-tetra BGA, the $^4C_1$ pucker was more populated compared to $^1C_4$, while both pucker conformations ($^4C_1$,$^1C_4$) were almost equally sampled in A-penta BGA (Figure S12, Supporting Information). This is because the $^4C_1$ pucker in Gal is energetically more favorable by ∼2 kcal/mol compared to the $^1C_4$ shape in H-tetra BGA (see Figure 6A,B). Furthermore, in the case of Gal, $^{0,3}B$, $^1S_3$, $^{1,4}B$, and $^0S_2$ puckers were also sampled more significantly in A-penta BGA than in H-tetra BGA, as these conformations have lower free energies in A-penta than in H-tetra BGA (see Figure 6A,B). Although both the chair conformations ($^4C_1$, $^1C_4$) were sampled for Fuc2 and Fuc3 monosaccharides in solution with a short time in the $^3S_1$ shape, the $^1C_4$ pucker conformation was found to be energetically more favorable than $^4C_1$ for both fucose monosaccharides, as is evident from Figure 6A,B. In the case of GalNAc, mostly, the $^4C_1$ pucker conformation was sampled with the only optimization passage to the $^1S_3$ and $^{1,4}B$ puckers when A-penta BGA is either in solution or at the binding sites (Figure 6B,D). Finally, Figure 6C,D displays that the boat/skewed-boat puckers are free energetically unfavorable for Gal, Fuc2, and Fuc3 monosaccharides when BGAs are bound to CTB.

Recognition of H-Tetra and A-Penta by the Cholera Toxin. *Binding Energetics.* To elucidate the mechanism of recognition of H-tetra and A-penta BGAs by the cholera toxin, we have calculated the total binding free energy and decomposed the same into various components including electrostatics, van der Waals, polar and nonpolar solvation free energies, and entropy (see Table 2). The total binding free energy and its different components for all four complexes are listed in Table 2. Overall, it is evident from Table 2 that the complex formation is mainly favored by the intermolecular electrostatic (ΔG_{elec}) and van der Waals (ΔG_{vdW}) interactions. The nonpolar component of the solvation free energy also favors the complexation. On the contrary, polar solvation free energy (ΔG_{polar}) and configurational entropy ($-T\Delta S$) oppose the binding of BGAs to CT.

As shown in Table 2, the calculated binding free energies (ΔG_{bind}) are −4.8, −2.7, −2.6, and −2.8 kcal/mol for cCTB/H-tetra, ET/H-tetra, cCTB/A-penta, and cCTB/A-penta, respectively. This suggests that H-tetra (i.e., the O blood group determinant) binds more strongly to both classical and El Tor (ET) biotypes compared to A-penta (the A blood group determinant). This observation is in agreement with the experimental study.[15] It should be noted that ΔG_{elec} varies between −39.1 and −47.6 kcal/mol, while ΔG_{vdW} varies between −32.9 and −42.8 kcal/mol for all four complexes. This means that ΔG_{elec} contributes more favorably to the complexation compared to ΔG_{vdW}. This is in contrast to what has been observed for HIV-1 protease-inhibitor[37−40] or with-no-lysine kinase-inhibitor complexes.[41,42] It is also worth mentioning here that for both oligosaccharides, ΔG_{elec} and ΔG_{vdW} are more favorable for the classical strain than for El Tor.

In the case of cCTB/A-penta, both ΔG_{elec} (−47.6 kcal/mol) and ΔG_{vdW} (−42.8 kcal/mol) are more favorable for the binding compared to the binding of H-tetra to cCTB (ΔG_{elec} = −41.1 kcal/mol, ΔG_{vdW} = −35.4 kcal/mol). However, in the case of cCTB/A-penta, both ΔG_{polar} (62.1 kcal/mol) and $-T\Delta S$ (31.2 kcal/mol) oppose more to the binding compared

https://dx.doi.org/10.1021/acsomega.9b03398
ACS Omega 2020, 5, 3932−3942

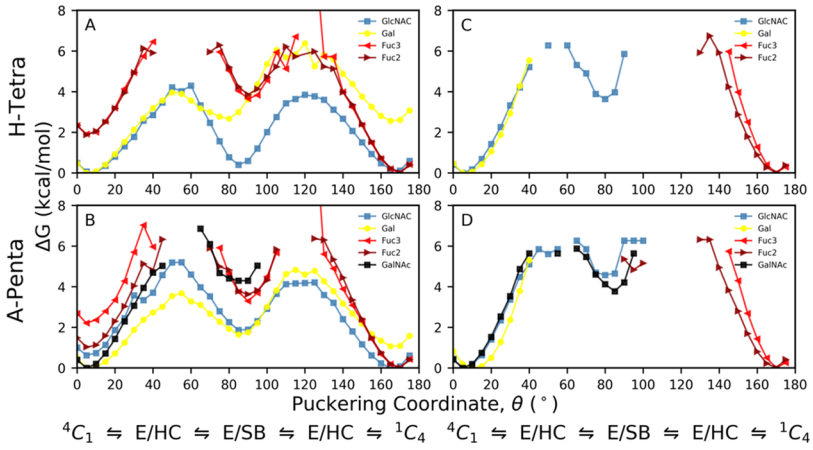


Figure 6. Cremer−Pople 1D ⊖ puckering profiles of H-tetra and B-penta BGAs (A, B) in solution and (C, D) in complex states. (A) and (B) were generated from aMD simulations, while (C) and (D) were obtained from cMD simulations of the complex structures.

Table 2. Binding Free Energies of All Four Complexes along with the Different Energy Components from the MM/GBSA Scheme[a]

energy term	cCT/H-tetra	ET CT/H-tetra	cCT/A-penta	ET CT/A-penta
ΔG_{vdW}	−35.4 (0.1)	−32.9 (0.1)	−42.8 (0.1)	−40.24 (0.1)
ΔG_{elec}	−41.1 (0.1)	−39.1 (0.1)	−47.6 (0.1)	−45.6 (0.1)
ΔG_{polar}	52.8 (0.1)	48.2 (0.1)	62.1 (0.1)	58.4 (0.1)
$\Delta G_{nonpolar}$	−4.4 (0.0)	−4.2 (0.0)	−5.5 (0.01)	−5.1 (0.0)
$-T\Delta S$	23.3 (1.6)	25.3 (1.7)	31.2 (2.1)	29.9 (1.8)
ΔG_{bind}	−4.8 (1.6)	−2.7 (1.7)	−2.6 (2.1)	−2.8 (1.8)
ΔG_{exp}	−4.1	−3.9	−3.7	−2.1

[a]All values are given in kcal/mol.

to cCTB/H-tetra (ΔG_{polar} = 52.8 kcal/mol, $-T\Delta S$ = 23.3 kcal/mol). Overall, A-penta binds less strongly to cCTB compared to H-tetra due to the increased polar solvation free energy and configurational entropy. However, in the case of the El Tor strain, a similar binding free energy (ΔG_{bind} = ∼−2.7 kcal/mol) is obtained for both H-tetra and A-penta.

Quantification of per-Residue Contributions to ΔG_{bind}. To provide further insights into the mechanism of recognition of BGAs by CT, we calculated the contribution of each monosaccharide to ΔG_{bind} for all four complexes, which is listed in Table 3. For all four complexes, Fuc3 and GlcNAc contribute more favorably to the binding compared to other monosaccharides, as is evident from Figure 7. The contribution from GlcNAc ranges between −3.1 and −3.9 kcal/mol, while Fuc3 contributes between −8.6 and −8.9 kcal/mol toward the total binding free energy for all cases (see Table 2). The contribution from Fuc3 in H-tetra BGA accounts for ∼70% of the total binding energy for both strains, while it reduces to ∼55% in the case of A-penta complexed with either strain (see Figure 7). Further, Figure 7 reveals that GlcNAc is the next important monosaccharide for the binding for all four complexes, accounting for ∼25% of the binding energy. Interestingly, only in the case of A-penta BGA bound to the El Tor strain, the contribution of Gal (−2.6 kcal/mol) to ΔG_{bind} is significant compared to the other three cases (Figure 8). Gal in A-penta BGA contributes approximately 4 times more favorably to the binding with ET compared to the classical strain because of the ∼3-fold increase in favorable interatomic electrostatic interactions (T_{ele} = −5.5 kcal/mol) relative to classical (T_{ele} = −1.9 kcal/mol). A careful inspection of Table 3

Table 3. Decomposition of ΔG on a per-Monosaccharide Basis for All Four Complexes[a]

monosaccharide	T_{vdW}	T_{ele}	T_{GB}	T_{np}	T_{TOT}
cCTB/H-Tetra BGA					
reducing end OH	−0.2	−0.3	0.5	0.0	−0.1
GlcNAc	−4.3	−8.0	9.2	−0.6	−3.8
Gal	−2.3	−0.1	3.4	−0.3	0.7
Fuc3	−7.6	−12.3	12.5	−1.2	−8.7
Fuc2	−3.1	0.2	2.9	−0.6	−0.6
El Tor/H-Tetra BGA					
reducing end OH	−0.3	0.1	0.2	−0.1	−0.1
GlcNAc	−3.4	−6.9	7.7	−0.5	−3.1
Gal	−2.0	−1.7	3.9	−0.2	0.0
Fuc3	−7.3	−12.0	11.6	−1.2	−8.9
Fuc2	−3.5	1.0	2.8	−0.7	−0.4
cCTB/A-Penta BGA					
reducing end OH	−0.3	−1.6	2.2	−0.2	0.1
GlcNAc	−4.1	−6.6	7.2	−0.5	−3.9
Gal	−2.1	−1.9	3.5	−0.1	−0.6
GalNAc	−4.9	−1.5	5.2	−0.9	−2.1
Fuc3	−7.6	−12.9	13.0	−1.2	−8.6
Fuc2	−2.4	0.7	1.8	−0.5	−0.4
El Tor/A-Penta BGA					
reducing end OH	−0.6	−1.3	1.9	−0.2	−0.3
GlcNAc	−3.4	−6.6	7.2	−0.4	−3.1
Gal	−1.8	−5.5	4.8	−0.1	−2.6
GalNAc	−3.9	−0.3	4.7	−0.8	−0.2
Fuc3	−7.3	−11.2	10.8	−1.2	−8.8
Fuc2	−3.2	2.0	1.4	−0.6	−0.4

[a]Energies are shown as contributions from the van der Waals (vdW), electrostatic (ele), polar (GB), and nonpolar solvation energy (np) of the side chain (S), backbone (B), and their sum (TOT) for CT-BGA complexes. All values are given in kcal/mol.

reveals that the electrostatic (T_{ele}) contribution from GlcNAc and Fuc3 to the binding is more favorable compared to van der Waals (T_{vdW}) interactions and varies between −6.6 and −12.9 kcal/mol for all complexes. However, it gets nearly canceled due to unfavorable polar desolvation free energy (T_{GB}). Therefore, the key contributor to the binding from these two monosaccharides is $T_{vdW,}$ which varies between −3.4 and −7.6 kcal/mol.

Next, we calculated the contribution of each amino acid to the binding (Table 4) and plotted the CT/BGA interaction

https://dx.doi.org/10.1021/acsomega.9b03398
ACS Omega 2020, 5, 3932−3942

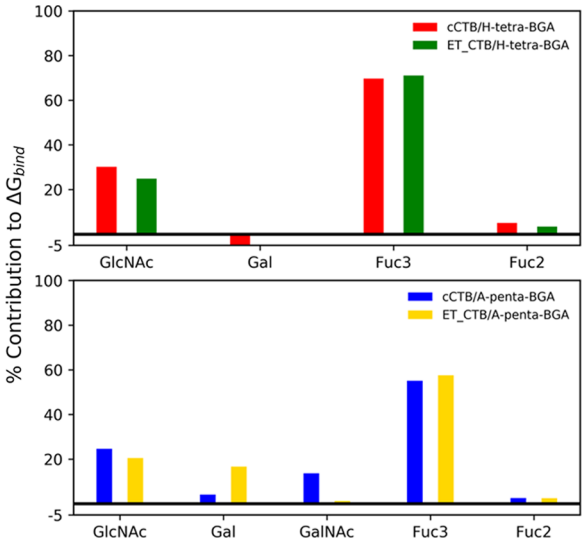


Figure 7. Percentage contribution by each monosaccharide in the total ΔG made by each BGA.

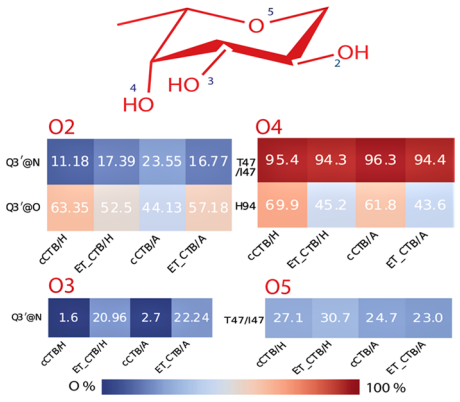


Figure 8. Hydrogen-bond profile of Fuc3.

energy spectra (Figure S13, Supporting Information). A similar spectrum is obtained for all four cases. We note that H18/Y18, A46 T47/I47, P93, H94, and Q3′ are important residues for the binding of BGA and contributed to the binding by ≥−1.5 kcal/mol (Table 4). In the case of the classical biotype, T47 contributed by ∼−3.1 kcal/mol, while the corresponding residue (I47) in El Tor contributed by ∼−4.0 kcal/mol. In T47 (i.e., classical), the backbone (T_B) contributed more favorably to the binding than the side chain (T_S), while T_S is more favorable than T_B in I47 (El Tor). Table 4 further reveals that in cases of A46, H94, and Q3′, T_S is more favorable than T_B for all four complexes. In contrast, T_B is more favorable than T_S in P93 for all complexes. Here, it is worth mentioning that T_{vdW} (−1.9 to −3.9 kcal/mol) is more favorable compared to T_{ele} (−0.2 to −3.5 kcal/mol) for both A46 and H94 and T_{ele} is overcompensated by T_{GB}.

Subsequently, we performed the hydrogen-bond (h-bond) analysis for all four complexes to complement the above results, and the critical h-bonds between BGA and CT with their occupancies are shown in Table S2 (Supporting Information). It is observed from Table S2 that H94, T47/I47, Q3′, and G45 form h-bonds with BGAs for all cases during MD simulations, explaining high intermolecular electrostatic contributions to the binding from these residues (Figure S14, Supporting Information). Further, we observe

Table 4. Decomposition of ΔG on a per-Residue Basis for All Four Complexes[a,b]

residue	T_{vdW}	T_{ele}	T_{GB}	T_{np}	T_S	T_B	T_{TOT}
			cCTB/H-Tetra BGA				
T47	−2.1	−3.4	2.5	−0.2	−0.5	−2.6	−3.1
H94	−2.5	−1.5	1.5	−0.3	−1.6	−1.2	−2.8
Q3′	−0.8	−5.5	4.5	−0.3	−2.3	0.1	−2.2
P93	−1.3	−1.5	0.7	−0.1	−0.7	−1.4	−2.1
A46	−2.1	−0.9	1.1	−0.2	−1.2	−0.9	−2.1
G45	−1.4	−2.5	2.3	−0.3	−0.4	−1.5	−1.9
			El Tor/H-Tetra BGA				
I47	−2.2	−3.7	2.2	−0.3	−1.9	−2.1	−4.0
H94	−2.9	−2.6	3.1	−0.4	−2.0	−0.8	−2.8
A46	−2.1	−0.5	0.7	−0.2	−1.3	−0.8	−2.1
Q3′	−0.4	−4.8	3.4	−0.2	−2.1	0.1	−2.0
P93	−1.0	−1.4	0.6	0.0	−0.6	−1.2	−1.8
			cCTB/A-Penta BGA				
H94	−3.2	−2.6	2.8	−0.5	−2.3	−1.1	−3.4
T47	−1.8	−3.7	2.5	−0.2	−0.6	−2.6	−3.2
Q3′	−0.8	−5.3	4.1	−0.3	−2.4	0.1	−2.3
G45	−1.5	−2.3	1.8	−0.2	−0.4	−1.8	−2.2
P93	−1.3	−1.3	0.7	0.0	−0.7	−1.3	−2.0
H18	−2.6	−2.5	3.5	−0.4	−1.9	−0.1	−2.0
A46	−1.9	−0.4	0.8	−0.2	−1.0	−0.7	−1.7
F48	−1.4	−0.7	0.7	−0.1	−1.1	−0.4	−1.5
			El Tor/A-Penta BGA				
I47	−2.1	−3.8	2.2	−0.3	−1.9	−2.1	−4.0
H94	−3.9	−3.5	4.4	−0.5	−2.7	−0.8	−3.5
Q3′	−0.4	−4.7	3.3	−0.2	−2.1	0.1	−2.0
A46	−2.2	−0.2	0.7	−0.2	−1.3	−0.5	−1.8
P93	−1.0	−1.2	0.6	0.0	−0.5	−1.2	−1.7

[a]Only residues with $|T_{TOT}| \geq 1.5$ kcal/mol are listed. [b]Energies are shown as contributions from the van der Waals (vdW), electrostatic (ele), polar (GB), and nonpolar solvation energy (np) of the side chain (S), backbone (B), and their sum (TOT) for CT-BGA complexes. All values are given in kcal/mol.

that mainly Fuc3 and GlcNAc monosaccharides participated in forming h-bonds with CT.

In Figure 8, we have shown h-bonds between Fuc3 and CT. The h-bond between T47/I47@O and Fuc3@O4 remains nearly conserved (∼95%) during the simulation for all four complexes (Figure 8). Furthermore, an h-bond with an occupancy of ∼23−30% was formed between Fuc3@O5 and T47/I47@N-H. Since h-bonds were formed with the backbone atoms, T_B was found to be more favorable in T47/I47 than T_S. The glutamine residue from the neighboring monomer (i.e., Q3′) was also found to be interacting with Fuc3 by forming h-bonds with O2 and O3 atoms of the fucose ring. We complemented the above result by calculating the pairwise correlation function or the radial distribution function, $g(r)$, of the water molecules with respect to the oxygen atoms of individual monosaccharide in the free (Figure S15A,B) and protein-bound (Figure S15C,D) states. In the case of free glycans, the intensity of the peak corresponding to Fuc2 and Fuc3 is found to be much higher compared to that of other monosaccharides (see Figure S15A,B). However, in the protein-bound form, the intensity of the peak of Fuc3 in both BGAs is greatly diminished (Figure S15C,D), while it remains unchanged for Fuc2. This is because Fuc3 was found to be involved in forming multiple h-bonds with the cholera toxin (Figure 8). Finally, for the whole glycan, the distribution

https://dx.doi.org/10.1021/acsomega.9b03398
ACS Omega 2020, 5, 3932−3942

of $g(r)$ is the same, irrespective of whether the glycan is in the protein-bound or free states.

Overall, Fuc3 is found to be critical in recognizing the cholera toxin by BGA as the maximum number of h-bonds is formed by Fuc3 with CT (Figure 8). This is in agreement with a previous experimental study, elucidating the role of fucosylated molecules in interaction with CT.[43]

CONCLUSIONS

First, in this study, we have elucidated the conformational dynamics of the Lewis Y blood group antigens, H-tetra, and A-penta oligosaccharides, in aqueous solution using 1 μs long accelerated molecular dynamics (aMD) simulations. We have also investigated the conformational preferences of BGAs bound to CT using cMD. Both oligosaccharides in aqueous solution were found to be flexible. A similar result was obtained for the core Le^X trisaccharide [Galβ(1−4)(Fucα(1−3))GlcNAcβ] in solution.[14,28] The aMD simulations suggest that the conformational ensemble of A-penta BGA is more diverse compared to that of H-tetra BGA. Three different conformers with varying populations were sampled for A-pentasaccharides, while a single conformer was populated in the case of H-tetrasaccharide. Moreover, the free energy surfaces of glycosidic linkages for both BGAs obtained from aMD simulations were found to be broader compared to those of the protein-bound form.

Interestingly, only aMD simulations revealed that A-pentasaccharide could adopt both open and closed forms in solution (Figure 5), while it takes an open conformation in the binding site. In contrast, H-tetrasaccharide remained in an open conformation in solution or bound form. Furthermore, aMD simulations revealed that for both oligosaccharides in the unbound state, $^4C_1$ and $^1C_4$ puckers were sampled for all (GlcNAc, Gal, Fuc2, and Fuc3) but GalNAc monosaccharides, while either the $^4C_1$ (GlcNAc, Gal, GalNAc) or $^1C_4$ (Fuc2, Fuc3) pucker conformation was sampled in cMD simulations of protein−glycan complexes. In the time scale of our cMD simulation, the other puckering states were not sampled due to the high energy barrier caused by the cyclization of the sugar ring.[44] The 1D free energy profile of the puckering angle (Θ) shows that only for the GlcNAc and Gal rings in H-tetra and A-pentasaccharides, the aMD simulations sampled the complete $^4C_1$ to $^1C_4$ pucker transition. In the case of Le^X trisaccharide, the complete transition was observed only for GlcNAc, while only the $^4C_1$ pucker was sampled for the Gal ring.[14]

Second, we have studied the mechanism of binding of H-tetra and A-penta BGAs with the cholera toxin using the MM/GBSA scheme. A relatively stronger binding affinity was obtained for the O blood group determinant, H-tetrasaccharide, complexed with both biotypes compared to the A blood group determinant, A-pentasaccharide. This agrees with the experimental data. In the case of the CT/H-tetrasaccharide complex, both polar solvation free energy and configurational entropy were less disfavored compared to those in CT/A-pentasaccharide, resulting in higher binding free energy. Per residue decomposition of the total binding free energy shows that the terminal Fuc3 monosaccharide contributes most to the binding free energy compared to other carbohydrate residues, and it forms multiple h-bonds with CT. Overall, our study provides a detailed overview of the conformational preferences of Le^Y BGAs in solution and the protein-bound form and molecular basis of recognition of BGAs by the cholera toxin, which may further help in designing a more potent and reliable cholera toxin inhibitor as well as a vaccine to fight this century-old deadly disease.

MATERIALS AND METHODS

MD Simulations of CT-BGA Complexes. We have simulated the pentameric CT-BGA complexes in aqueous solution using conventional molecular dynamics (cMD) simulations at 300 K. Initial coordinates of the four complexes were obtained from the Protein Data Bank (PDB IDs: 5ELB and 5ELC for H-tetra BGA with cCT and ET CT, respectively, and 5ELD and 5ELE for A-penta BGA with cCT and ET CT, respectively).[15] The missing part of the crystal structure was built using the University of California, San Francisco, Chimera software.[45] In cases where the glycan was missing in the binding site of a monomer, we placed it by geometrically translating the glycan from the neighboring monomer. For all complexes, the AMBER ff14SB force field[46] was used for modeling the protein, and glycam06j-1[47] was employed for describing the glycan. All of the systems were solvated using TIP3P[48] water molecules in a truncated octahedral box, having a 10 Å distance between the solute and the wall of the box. An appropriate number of sodium ions were added to neutralize the system. Bonds involving hydrogen atoms were constrained using the SHAKE algorithm,[49] and the long-range electrostatic interactions were estimated using the particle-mesh Ewald (PME)[50] method. The nonbonded cutoff was set to 10 Å, and the temperature was kept fixed at 300 K using the Langevin thermostat[51] with a collision frequency of 2 ps^{-1}. The energy minimization of the system was carried out in two stages: (i) 500 steps of steepest descent followed by another 500 steps of conjugate gradient with a harmonic restraint of 2 kcal/(mol Å^2) applied on the solute and (ii) without any restraint on the solute using the steepest descent followed by conjugate gradient schemes. After the minimization, the systems were heated stepwise from 0 to 300 K with a restraint force of 2 kcal/(mol Å^2) acting on the solute. A 50 ps simulation with the restraint of 2 kcal/(mol Å^2) on the solute was carried out at a pressure of 1 atm to equilibrate the density. After that, we conducted 1 ns simulation under the NPT ensemble without any restraint on the solute. Finally, the production simulations were carried for 200 ns in the NPT ensemble with a time step of 2 fs using the pmemd.cuda module of AMBER 16.[52−54] We recorded coordinates every 10 ps, resulting in 20 000 configurations.

aMD Simulations of Free Glycans. For better sampling, we also simulated BGAs in aqueous solution using the accelerated molecular dynamics (aMD) technique. We equilibrated the solute for 10 ns before applying the aMD boost to get the average potential energy and dihedral energy. After that, we used a dual boost potential using the following boost parameters. For H-tetrasaccharide, dihedral energies E_{dih} = 24.22 kcal/mol, α_{dih} = 3.5 kcal/mol; total potential energy E_{tot} = −13 222.0 kcal/mol; and α_{tot} = 807.1 kcal/mol were used. Similarly, in the case of A-pentasaccharide, dihedral energies E_{dih} = 16.6 kcal/mol, α_{dih} = 4.2 kcal/mol; total potential energy E_{tot} = −17 222.6 kcal/mol; and α_{tot} = 1047.0 kcal/mol were employed. Finally, aMD simulations were carried out for 1 μs under the NPT ensemble and the coordinates were stored every 10 ps.

Analysis. All of the trajectories obtained from the unbiased and biased simulations were analyzed using the Cpptraj[55] module of AMBER 16. Glycosidic torsion angles were defined

https://dx.doi.org/10.1021/acsomega.9b03398
ACS Omega 2020, 5, 3932−3942

as $\varphi = O_5-C_1-O-C_n$, $\psi = C_1-O-C_n-C_{(n-1)}$ for $\alpha/\beta(1-n)$ linkages. The relative free energy surfaces corresponding to the glycosidic angles were computed using the expression $\Delta G_x = -k_B T \ln(\rho_x)$, where k_B is the Boltzmann constant, T is the temperature, and ρ_x is the probability density of the geometric coordinate x. As the value of this logarithm function is negative or zero, the final free energy value remains zero or positive. The state having the highest probability will always be 1; therefore, the minimum always stays at zero. For estimating the puckering conformation of the pyranose rings, we have adopted the generalized pseudorotation coordinates coined by Cremer and Pople.[36] Further, we followed the IUPAC convention to show different puckerings, which lead to the partition of Θ and φ plane in 38 regions. The one-dimensional free energy surface was also generated using the already stated Boltzmann formula, taking Θ from the puckering parameter as a reaction coordinate. For aMD simulations, we have employed the Pyreweighting scripts developed by the McCammon group[56] for constructing the potential of mean force (PMF) surface from the biased potential. This script uses the 10th-order Maclaurin series expansion to reweigh the total boost energy from each frame.

For estimating the binding free energy of CT-BGA complexes, the popular molecular mechanics generalized Born surface area (MM/GBSA)[32] scheme was used. For protein−glycan systems, it has been shown in previous studies[57,58] that MM/GBSA outperforms the computationally expensive MM/PBSA.[59−61] The MM/GBSA scheme has been discussed elsewhere.[37−39,41,62,63] We have used the option (igb = 2)[64] within the MM-GBSA framework for all cases. We extracted the trajectory of a dimer pair with one oligosaccharide from the pentameric complex trajectory for estimating the binding free energy. Two adjacent monomers were considered because the ligand was found to form a stable hydrogen bond with the glutamine residue of the neighboring monomer (Q3′). We have considered 8000 frames for computing the binding free energy. The normal mode analysis (NMA) method was employed for calculating the configurational entropy, and 50 frames were used for the calculation.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.9b03398.

Hydrogen-bond occupancy, 2D free energy landscapes, ligand RMSD, time evolution of puckering conformation, Mercator representations of puckering conformations, puckering free energy profiles, per-residue decomposition of free energy, and pairwise correlation function (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Parimal Kar − *Discipline of Biosciences and Biomedical Engineering, Indian Institute of Technology Indore, Indore 453552, India;* orcid.org/0000-0001-8451-9739; Phone: +91 731 2438700; Email: parimal@iiti.ac.in

Authors

Rajarshi Roy − *Discipline of Biosciences and Biomedical Engineering, Indian Institute of Technology Indore, Indore 453552, India;* orcid.org/0000-0002-5788-9424

Biplab Ghosh − *High Pressure and Synchrotron Radiation Physics Division, Bhabha Atomic Research Centre, Mumbai 400085, India*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.9b03398

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Department of Biotechnology, Government of India (grant number BT/RLF/Re-entry/40/2014, DBT-Ramalingaswami Re-entry Fellowship), and Department of Science and Technology, Government of India (grant number ECR/2017/000010). R.J. thanks the Indian Institute of Technology Indore for financial assistance.

■ REFERENCES

(1) Anstee, D. J. The relationship between blood groups and disease. *Blood* **2010**, *115*, 4635−4643.

(2) Sack, D. A.; Sack, R. B.; Nair, G. B.; Siddique, A. K. Cholera. *Lancet* **2004**, *363*, 223−233.

(3) Barua, D.; Paguio, A. ABO blood groups and cholera. *Ann. Hum. Biol.* **1977**, *4*, 489−492.

(4) Swerdlow, D. L.; Mintz, E. D.; Rodriguez, M.; Tejada, E.; Ocampo, C.; Espejo, L.; Barrett, T. J.; Petzelt, J.; Bean, N. H.; Seminario, L. Severe life-threatening cholera associated with blood group 0 in peru: implications for the latin american epidemic. *J. Infect. Dis.* **1994**, *170*, 468−472.

(5) Tacket, C. O.; Losonsky, G.; Nataro, J. P.; Wasserman, S. S.; Cryz, S. J.; Edelman, R.; Levine, M. M. Extension of the volunteer challenge model to study South American cholera in a population of volunteers predominantly with blood group antigen O. *Trans. R. Soc. Trop. Med. Hyg.* **1995**, *89*, 75−77.

(6) Harris, J. B.; Khan, A. I.; LaRocque, R. C.; Dorer, D. J.; Chowdhury, F.; Faruque, A. S.; Sack, D. A.; Ryan, E. T.; Qadri, F.; Calderwood, S. B. Blood group, immunity, and risk of infection with *Vibrio cholerae* in an area of endemicity. *Infect. Immun.* **2005**, *73*, 7422−7427.

(7) Harris, J. B.; LaRocque, R. C.; Chowdhury, F.; Khan, A. I.; Logvinenko, T.; Faruque, A. S.; Ryan, E. T.; Qadri, F.; Calderwood, S. B. Susceptibility to *Vibrio cholerae* infection in a cohort of household contacts of patients with cholera in Bangladesh. *PLoS Neglected Trop. Dis.* **2008**, *2*, No. e221.

(8) Glass, R. I.; Holmgren, J.; Haley, C. E.; Khan, M.; Svennerholm, A.; Stoll, B. J.; Hossain, K. B.; Black, R. E.; Yunus, M.; Barua, D. Predisposition for cholera of individuals with O blood group possible evolutionary significance. *Am. J. Epidemiol.* **1985**, *121*, 791−796.

(9) Harris, J. B.; LaRocque, R. C. Cholera and ABO Blood Group: Understanding an Ancient Association. *Am. J. Trop. Med. Hyg.* **2016**, *95*, 263−264.

(10) Mekalanos, J. J.; Swartz, D. J.; Pearson, G. D.; Harford, N.; Groyne, F.; de Wilde, M. Cholera toxin genes: nucleotide sequence, deletion analysis and vaccine development. *Nature* **1983**, *306*, 551.

(11) Sánchez, J.; Holmgren, J. Cholera toxin structure, gene regulation and pathophysiological and immunological aspects. *Cell. Mol. Life Sci.* **2008**, *65*, 1347−1360.

(12) Clemens, J. D.; Sack, D. A.; Harris, J. R.; Chakraborty, J.; Khan, M. R.; Huda, S.; Ahmed, F.; Gomes, J.; Rao, M.; Svennerholm, A.-M. ABO blood groups and cholera: new observations on specificity of risk and modification of vaccine efficacy. *J. Infect. Dis.* **1989**, *159*, 770−773.

(13) Yamamoto, F.-i.; Clausen, H.; White, T.; Marken, J.; Hakomori, S.-i. Molecular genetic basis of the histo-blood group ABO system. *Nature* **1990**, *345*, 229.

(14) Alibay, I.; Burusco, K. K.; Bruce, N. J.; Bryce, R. A. Identification of rare Lewis oligosaccharide conformers in aqueous

solution using enhanced sampling molecular dynamics. *J. Phys. Chem. B* **2018**, *122*, 2462−2474.
(15) Heggelund, J. E.; Burschowsky, D.; Bjørnestad, V. A.; Hodnik, V.; Anderluh, G.; Krengel, U. High-resolution crystal structures elucidate the molecular basis of cholera blood group dependence. *PLoS Pathog.* **2016**, *12*, No. e1005567.
(16) Lanne, B.; Schierbeck, B.; Ångström, J. Binding of Cholera Toxin B-Subunits to Derivatives of the Natural Ganglioside Receptor, GM 1. *J. Biochem.* **1999**, *126*, 226−234.
(17) Sharmila, D. J. S.; Veluraja, K. Monosialogangliosides and their interaction with cholera toxin—investigation by molecular modeling and molecular mechanics. *J. Biomol. Struct. Dyn.* **2004**, *21*, 591−613.
(18) Blessy, J. J.; Sharmila, D. J. S. Molecular simulation of N-acetylneuraminic acid analogs and molecular dynamics studies of cholera toxin-Neu5Gc complex. *J. Biomol. Struct. Dyn.* **2015**, *33*, 1126−1139.
(19) Sharmila, D. J. S.; Jino Blessy, J. Molecular dynamics of sialic acid analogues complex with cholera toxin and DFT optimization of ethylene glycol-mediated zinc nanocluster conjugation. *J. Biomol. Struct. Dyn.* **2017**, *35*, 182−206.
(20) Cervin, J.; Wands, A. M.; Casselbrant, A.; Wu, H.; Krishnamurthy, S.; Cvjetkovic, A.; Estelius, J.; Dedic, B.; Sethi, A.; Wallom, K.-L.; et al. GM1 ganglioside-independent intoxication by Cholera toxin. *PLoS Pathog.* **2018**, *14*, No. e1006862.
(21) Fazil, M. H. U. T.; Kumar, S.; Farmer, R.; Pandey, H. P.; Singh, D. V. Binding efficiencies of carbohydrate ligands with different genotypes of cholera toxin B: molecular modeling, dynamics and docking simulation studies. *J. Mol. Model.* **2012**, *18*, 1−10.
(22) Yang, M.; MacKerell, A. D., Jr. Conformational sampling of oligosaccharides using Hamiltonian replica exchange with two-dimensional dihedral biasing potentials and the weighted histogram analysis method (WHAM). *J. Chem. Theory Comput.* **2015**, *11*, 788−799.
(23) Islam, S. M.; Richards, M. R.; Taha, H. A.; Byrns, S. C.; Lowary, T. L.; Roy, P.-N. Conformational analysis of oligoarabinofuranosides: Overcoming torsional barriers with umbrella sampling. *J. Chem. Theory Comput.* **2011**, *7*, 2989−3000.
(24) Perić-Hassler, L.; Hansen, H. S.; Baron, R.; Hünenberger, P. H. Conformational properties of glucose-based disaccharides investigated using molecular dynamics simulations with local elevation umbrella sampling. *Carbohydr. Res.* **2010**, *345*, 1781−1801.
(25) Plazinski, W.; Drach, M. The influence of the hexopyranose ring geometry on the conformation of glycosidic linkages investigated using molecular dynamics simulations. *Carbohydr. Res.* **2015**, *415*, 17−27.
(26) Re, S.; Nishima, W.; Miyashita, N.; Sugita, Y. Conformational flexibility of N-glycans in solution studied by REMD simulations. *Biophys. Rev.* **2012**, *4*, 179−187.
(27) Nishima, W.; Miyashita, N.; Yamaguchi, Y.; Sugita, Y.; Re, S. Effect of bisecting GlcNAc and core fucosylation on conformational properties of biantennary complex-type N-glycans in solution. *J. Phys. Chem. B* **2012**, *116*, 8504−8512.
(28) Topin, J.; Lelimousin, M.; Arnaud, J.; Audfray, A.; Pérez, S.; Varrot, A.; Imberty, A. The hidden conformation of Lewis x, a human histo-blood group antigen, is a determinant for recognition by pathogen lectins. *ACS Chem. Biol.* **2016**, *11*, 2011−2020.
(29) Hansen, H. S.; Hünenberger, P. H. Using the local elevation method to construct optimized umbrella sampling potentials: calculation of the relative free energies and interconversion barriers of glucopyranose ring conformers in water. *J. Comput. Chem.* **2010**, *31*, 1−23.
(30) Lemieux, R.; Bock, K.; Delbaere, L.; Koto, S.; Rao, V. The conformations of oligosaccharides related to the ABH and Lewis human blood group determinants. *Can. J. Chem.* **1980**, *58*, 631−653.
(31) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *J. Chem. Phys.* **2004**, *120*, 11919−11929.
(32) Homeyer, N.; Gohlke, H. Free energy calculations by the molecular mechanics Poisson− Boltzmann surface area method. *Mol. Inf.* **2012**, *31*, 114−122.
(33) Wernick, N. L.; Chinnapen, D. J.-F.; Cho, J.; Lencer, W. I. Cholera Toxin: An Intracellular Journey into the Cytosol by Way of the Endoplasmic Reticulum. *Toxins* **2010**, *2*, 310−325.
(34) Topin, J.; Arnaud, J.; Sarkar, A.; Audfray, A.; Gillon, E.; Perez, S.; Jamet, H.; Varrot, A.; Imberty, A.; Thomas, A. Deciphering the glycan preference of bacterial lectins by glycan array and molecular docking with validation by microcalorimetry and crystallography. *PLoS One* **2013**, *8*, No. e71149.
(35) Houser, J.; Komarek, J.; Cioci, G.; Varrot, A.; Imberty, A.; Wimmerova, M. Structural insights into Aspergillus fumigatus lectin specificity: AFL binding sites are functionally non-equivalent. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2015**, *71*, 442−453.
(36) Cremer, D.; Pople, J. General definition of ring puckering coordinates. *J. Am. Chem. Soc.* **1975**, *97*, 1354−1358.
(37) Kar, P.; Knecht, V. Energetic basis for drug resistance of HIV-1 protease mutants against amprenavir. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 215−232.
(38) Kar, P.; Knecht, V. Energetics of mutation-induced changes in potency of lersivirine against HIV-1 reverse transcriptase. *J. Phys. Chem. B* **2012**, *116*, 6269−6278.
(39) Kar, P.; Knecht, V. Origin of decrease in potency of darunavir and two related antiviral inhibitors against HIV-2 compared to HIV-1 protease. *J. Phys. Chem. B* **2012**, *116*, 2605−2614.
(40) Sk, M. F.; Roy, R.; Kar, P. Exploring the potency of currently used drugs against HIV-1 protease of subtype D variant by using multiscale simulations. *J. Biomol. Struct. Dyn.* **2020**, 1−16.
(41) Jonniya, N. A.; Kar, P. Investigating specificity of the anti-hypertensive inhibitor WNK463 against With-No-Lysine kinase family isoforms via multiscale simulations. *J. Biomol. Struct. Dyn.* **2019**, 1−16.
(42) Jonniya, N. A.; Sk, M. F.; Kar, P. Investigating Phosphorylation-Induced Conformational Changes in WNK1 Kinase by Molecular Dynamics Simulations. *ACS Omega* **2019**, 17404.
(43) Wands, A. M.; Cervin, J.; Huang, H.; Zhang, Y.; Youn, G.; Brautigam, C. A.; Matson Dzebo, M.; Björklund, P.; Wallenius, V.; Bright, D. K.; et al. Fucosylated molecules competitively interfere with cholera toxin binding to host cells. *ACS Infect. Dis.* **2018**, *4*, 758−770.
(44) Wang, L.; Berne, B. Efficient sampling of puckering states of monosaccharides through replica exchange with solute tempering and bond softening. *J. Chem. Phys.* **2018**, *149*, No. 072306.
(45) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605−1612.
(46) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696−3713.
(47) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.* **2008**, *29*, 622−655.
(48) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(49) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(50) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log (N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(51) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press, 2017.
(52) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J.

https://dx.doi.org/10.1021/acsomega.9b03398
ACS Omega 2020, 5, 3932−3942

The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668−1688.
(53) Götz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized born. *J. Chem. Theory Comput.* **2012**, *8*, 1542−1555.
(54) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878−3888.
(55) Roe, D. R.; Cheatham, T. E., III PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **2013**, *9*, 3084−3095.
(56) Miao, Y.; Sinko, W.; Pierce, L.; Bucher, D.; Walker, R. C.; McCammon, J. A. Improved reweighting of accelerated molecular dynamics simulations for free energy calculation. *J. Chem. Theory Comput.* **2014**, *10*, 2677−2689.
(57) Sood, A.; Gerlits, O. O.; Ji, Y.; Bovin, N. V.; Coates, L.; Woods, R. J. Defining the Specificity of Carbohydrate−Protein Interactions by Quantifying Functional Group Contributions. *J. Chem. Inf. Model.* **2018**, *58*, 1889−1901.
(58) Mishra, S. K.; Koča, J. Assessing the Performance of MM/PBSA, MM/GBSA, and QM−MM/GBSA Approaches on Protein/Carbohydrate Complexes: Effect of Implicit Solvent Models, QM Methods, and Entropic Contributions. *J. Phys. Chem. B* **2018**, *122*, 8113−8121.
(59) Kar, P.; Wei, Y.; Hansmann, U. H.; Höfinger, S. Systematic study of the boundary composition in Poisson Boltzmann calculations. *J. Comput. Chem.* **2007**, *28*, 2538−2544.
(60) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889−897.
(61) Kar, P.; Seel, M.; Hansmann, U. H.; Höfinger, S. Dispersion Terms and Analysis of Size-and Charge Dependence in an Enhanced Poisson− Boltzmann Approach. *J. Phys. Chem. B* **2007**, *111*, 8910−8918.
(62) Kar, P.; Lipowsky, R.; Knecht, V. Importance of polar solvation for cross-reactivity of antibody and its variants with steroids. *J. Phys. Chem. B* **2011**, *115*, 7661−7669.
(63) Kar, P.; Knecht, V. Mutation-induced loop opening and energetics for binding of tamiflu to influenza N8 neuraminidase. *J. Phys. Chem. B* **2012**, *116*, 6137−6149.
(64) Onufriev, A.; Bashford, D.; Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins* **2004**, *55*, 383−394.

NOTE ADDED AFTER ASAP PUBLICATION

Figure 2 was corrected February 19, 2020.

https://dx.doi.org/10.1021/acsomega.9b03398
ACS Omega 2020, 5, 3932−3942