

# Coarse-Grained Molecular Model for the Glycosylphosphatidylinositol Anchor with and without Protein

Pallavi Banerjee, Reinhard Lipowsky, and Mark Santer\*

Cite This: *J. Chem. Theory Comput.* 2020, 16, 3889–3903

Read Online

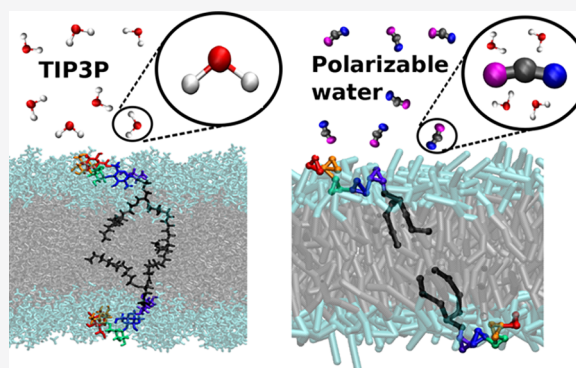
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

**ABSTRACT:** Glycosylphosphatidylinositol (GPI) anchors are a unique class of complex glycolipids that anchor a great variety of proteins to the extracellular leaflet of plasma membranes of eukaryotic cells. These anchors can exist either with or without an attached protein called GPI-anchored protein (GPI-AP) both *in vitro* and *in vivo*. Although GPIs are known to participate in a broad range of cellular functions, it is to a large extent unknown how these are related to GPI structure and composition. Their conformational flexibility and microheterogeneity make it difficult to study them experimentally. Simplified atomistic models are amenable to all-atom computer simulations in small lipid bilayer patches but not suitable for studying their partitioning and trafficking in complex and heterogeneous membranes. Here, we present a coarse-grained model of the GPI anchor constructed with a modified version of the MARTINI force field that is suited for modeling carbohydrates, proteins, and lipids in an aqueous environment using MARTINI's polarizable water. The nonbonded interactions for sugars were reparametrized by calculating their partitioning free energies between polar and apolar phases. In addition, sugar–sugar interactions were optimized by adjusting the second virial coefficients of osmotic pressures for solutions of glucose, sucrose, and trehalose to match with experimental data. With respect to the conformational dynamics of GPI-anchored green fluorescent protein, the accessible time scales are now at least an order of magnitude larger than for the all-atom system. This is particularly important for fine-tuning the mutual interactions of lipids, carbohydrates, and amino acids when comparing to experimental results. We discuss the prospective use of the coarse-grained GPI model for studying protein-sorting and trafficking in membrane models.



## INTRODUCTION

The plasma membrane of eukaryotic cells contains a large variety of functionally active proteins, such as transmembrane proteins acting as ion channels or RAS proteins which have a simple fatty acid tail tethering them to the plasma membrane. The so-called glycosylphosphatidylinositols (GPIs) provide a particularly intriguing anchoring mechanism. They are covalently added to the C-terminus of proteins through post-translational modification in the endoplasmic reticulum. The structure of GPI consists of a highly conserved pseudopenta-saccharide glycan core **Man- $\alpha$ (1 $\rightarrow$ 2)-Man- $\alpha$ (1 $\rightarrow$ 6)-Man- $\alpha$ (1 $\rightarrow$ 4)-GlcN- $\alpha$ (1 $\rightarrow$ 6)-myo-inositol** that is further connected to a lipid tail which inserts into the plasma membrane. GPI-anchored proteins (GPI-APs) are involved in many cellular functions such as signal transduction,<sup>1,2</sup> adhesion,<sup>3</sup> and apical sorting.<sup>4,5</sup> GPIs are also found on the cell surfaces of protozoan parasites such as *Toxoplasma gondii*, *Trypanosoma brucei*, and *Plasmodium falciparum*,<sup>6</sup> either with or without an attached protein, as an end product of metabolic processes in the latter case. Figure 1 shows the chemical structure of a GPI with its attached protein. At the trailing mannose (Man3), a phosphoethanolamine bridge (EtNP) connects the protein to the GPI. In spite of the conserved core, GPIs are of

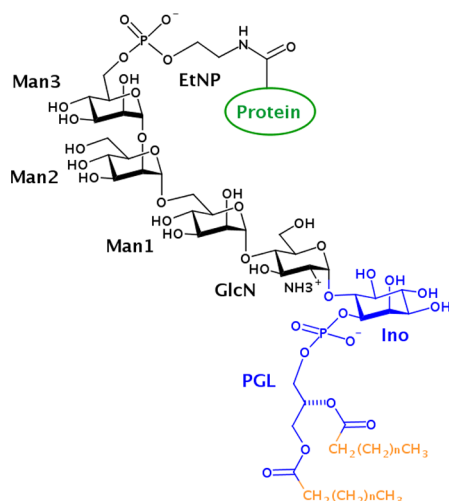
heterogeneous structure through various types of sugar side branches, the composition of which can vary even with the very same protein (microheterogeneity).

Ever since the discovery of GPIs, the question of the relationship between their exceptional structure and functions has been a matter of debate until today. One of the many controversial subjects is the conformation of GPIs and the orientation and placement of GPI-APs relative to the membrane they are embedded in. Conclusions vary with the type of experiment conducted. One scenario is that GPI-APs lie in close proximity to the membrane, almost flopping down on it.<sup>8,9</sup> Lehto and Sharom conducted a FRET-based study on lipid bilayer vesicles to conclude that the fluorescent tag on a GPI-anchored placental alkaline phosphatase (PLAP) is at most 10–14 Å away from the lipid–water interfacial region,

Received: January 18, 2020

Published: May 11, 2020





**Figure 1.** Chemical structure of a GPI anchor. The GPI core consists of Man3-Man2-Man1-GlcN-Ino. The core is connected to a phosphoglycerol (PGL) head which further connects to the lipid tail. A phosphoethanolamine linker (EtNP) attaches the protein to GPI. Ino+PGL are shown in blue to indicate the transition between the two force-field domains of GLYCAM06h (black) and Lipid14 (orange) that have been merged<sup>7</sup> to provide a molecular model of the full structure.

implying that the protein could be resting on the surface of the bilayer.<sup>8</sup> On the other hand, through diffusivity studies of synthetic GPI analogs in both supported bilayers and live cells, Paulick and co-workers concluded that the rapidly diffusing GPI analogs do not interact much with the membrane, thereby preventing interactions between the membrane and the attached protein.<sup>10,11</sup> On the other hand, a combined experimental and computational investigation of GPI-anchored Thy1 protein showed that GPI could substantially influence the conformation of the attached protein, suggesting considerable interactions between the two.<sup>9</sup> One may also wonder about how much this would impact another persisting controversy regarding the localization and partitioning of GPIs in functionally active, dense membrane microdomains, frequently referred to as lipid rafts. Some of the experiments conducted to address the partitioning behavior of GPIs in different lipid domains are quite contradictory.<sup>12–14</sup>

However, structural and dynamical details at the atomic level and high temporal resolution, for instance the femto-second time scale, are difficult to assess through experiments. Computer simulations of atomistic models provide powerful tools for filling these gaps, but only a few numerical studies have been conducted for GPIs so far, most of them advertising the idea that a GPI-AP may essentially be viewed as a rather rigid molecular arrangement rather than a vivid, dynamically changing object.<sup>15–17</sup> In our previous work, we devised an all-atom model of GPI using GLYCAM06h, Lipid14, and AMBER-ff14SB force fields to elucidate the conformational flexibility of the glycan core in solution<sup>18</sup> and to study a full GPI-AP embedded in lipid bilayer patches.<sup>7</sup> Through plain and biased MD simulations, the GPI core was revealed to behave effectively as a hinge, with two rather rigid disaccharide units connected via a flexible Man- $\alpha$ (1 $\rightarrow$ 6)-Man linkage. With a lipid tail attached and inserted into a bilayer, GPIs tend to assume a hooklike conformation with the glycan core partially immersed in the lipid headgroup region. In the simulations, all three species—lipids, proteins, and GPI (carbohydrates)—were

seen to mutually interact. In general, one may envisage several avenues to further develop the hybrid model of GPI-AP via a reasonable refinement of force-field parameters. We want to recall, however, that the situation of three disparate, mutually interacting biomolecular species is not covered by the usual process of force-field development to begin with. The effect of a reparametrization will, however, experimentally only be visible in an extended context such as the dynamic behavior of GPI-APs in heterogeneous membrane patches, and the lack of sufficient statistical sampling will inevitably impose a strict limit on how an all-atom model can be tested. The mapping of our atomistic GPI model to a numerically efficient coarse-grained representation is thus highly desirable.

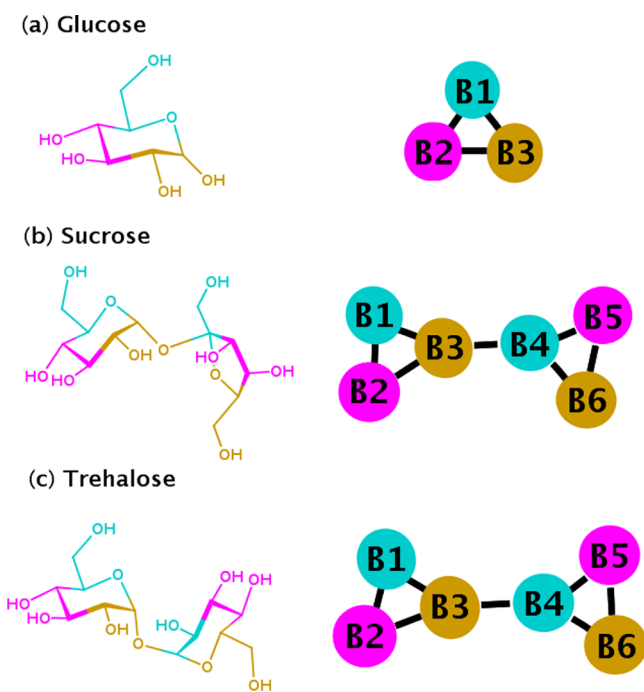
The MARTINI force field is a coarse-grained representation for biomolecular systems composed of lipids, proteins, glycolipids, and nucleotides, as well as, nanoparticles and a variety of polymers.<sup>19</sup> The MARTINI model is designed based on mapping 3 to 4 heavy atoms to one spherical superatom (bead). The interaction potentials between beads are inferred from the partitioning free energy of small coarse-grained molecules determined from their relative distributions in polar and apolar phases. MARTINI performs well in mimicking various types of lipids and replicating protein–lipid interactions as demonstrated for processes such as formation of pores and nanodisks, lipid-mediated protein clustering, and protein-mediated lipid flip-flop.<sup>20</sup>

In the present work, we devise a coarse-grained model of a full GPI and GPI-anchored green fluorescent protein (GFP) based on the MARTINI force field with polarizable water which has been proven to work consistently for modeling membranes in aqueous environment. After exhibiting our parametrization strategy and the definition of new parameters, we compare the behavior of the coarse-grained free-GPI and GPI-anchored-GFP with corresponding all-atom simulation results. We then discuss how the coarse-grained model may be used to study GPI-anchored proteins in membrane environments and how to deal with the situation that an optimally balanced parameter set for a GPI molecule is a priori unknown.

## PARAMETRIZATION STRATEGY

**Mapping Scheme, Bonded Interactions, and Bead Types.** Due to microheterogeneity of naturally occurring GPIs and the inherent difficulties of synthesizing sufficient amounts of pure GPI species,<sup>21</sup> molecular-level studies of GPIs are difficult, and concise experimental data are lacking. To build a coarse-grained model of GPI, parametrization of simple sugars (mono- or disaccharides) was necessary. Parametrizing glucose (monosaccharide), sucrose, and trehalose (disaccharides) was sufficient to model the whole GPI glycan in a building-block manner as the mapping strategy was consistent across all these saccharides, entailing similar bead types. Moreover, properties such as partitioning free energies, which will be used in turn to derive nonbonded interactions, are well-known for these species. Note that at the coarse-grained level, there is no difference between the nonbonded parameters of different epimers of sugars such as glucose, mannose, and galactose. The differences are contained in the bonded parameters that are derived straight from the atomistic systems. The MARTINI coarse-grained force field is based on mapping 3 or 4 heavy atoms of the underlying atomistic system to one coarse-grained bead. To coarse-grain monosaccharides, we followed a similar mapping scheme as in the original work of the MARTINI team where the model for carbohydrates<sup>22</sup> was introduced. One

saccharide unit is composed of three coarse-grained beads, connected together like a triangle. Unlike in ref 22 where polysaccharides were mapped linearly, we adopted a triangular mapping protocol (see Figure 2). The glycosidic linkages were

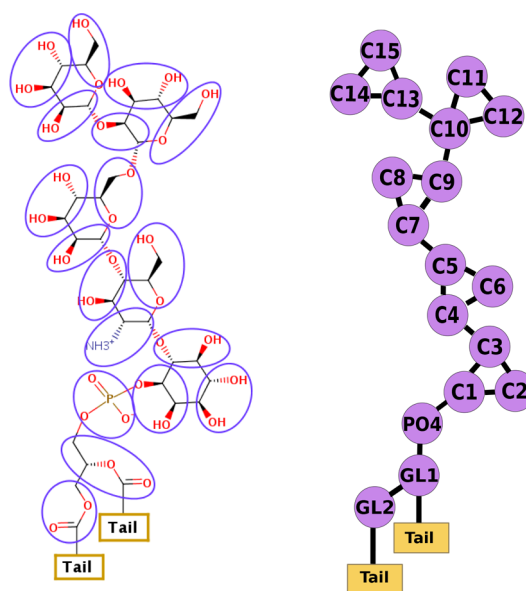


**Figure 2.** Mapping scheme for coarse-graining sugars: (a) glucose, (b) sucrose, and (c) trehalose. The colors of the coarse-grained beads encode the mapped groups of the atomistic molecules.

represented by just one bond in the coarse-grained landscape. There appears to be no general advantage of preferring the linear mapping over the triangular in the MARTINI scheme. The choice is usually made according to numerical stability of the simulation. In the present study, the triangular mapping scheme with a time step of 5 fs worked consistently for all the simulations. All systems in our study were parametrized against MARTINI's polarizable water as the aqueous medium. The polarizable water model implements the dielectric screening of bulk water through the orientational polarizability induced by its three-bead water model.<sup>23</sup> This water model is known to give more realistic and closer to atomistic results for processes involving membranes, such as pore formation,<sup>24</sup> phase transition,<sup>25</sup> and adsorption of charged peptides on membranes.<sup>26</sup>

**Bonded Interactions.** Bonded potentials for the simple sugars were obtained from 200 ns all-atom trajectories of one sugar molecule in water. GPIs were mapped from the atomistic structure using the same triangular mapping scheme as for the simple sugars (see Figure 3). Potentials for bonds, angles, and dihedrals were derived from a 1  $\mu$ s long all-atom trajectory of one GPI glycan in water. In this way, bonded parameters as a function of just the intramolecular interactions and the effect of the solvent were captured. The potentials were obtained from the all-atom trajectories through simple Boltzmann inversion. Bonds between coarse-grained beads were imposed by harmonic potentials

$$V_b(r) = \frac{1}{2}K_b(r - r_0)^2 \quad (1)$$



**Figure 3.** Mapping of GPI anchor from atomistic to coarse-grained representation.

where  $K_b$  is the spring force constant, and  $r_0$  is the equilibrium bond length. Similarly, an angle connecting three consecutively placed beads is defined by a cosine-harmonic potential

$$V_a = \frac{1}{2}K_a[\cos(\theta) - \cos(\theta_0)]^2 \quad (2)$$

where  $K_a$  and  $\theta_0$  are the force constant and equilibrium angle, respectively. Lennard-Jones interactions between beads connected by bonds and angles were excluded from the nonbonded force calculation. This exclusion was necessary in order to incorporate all the crucial bonded potentials while avoiding numerical instabilities. The same strategy was employed by Gu et al.<sup>27</sup> to model the glycolipids GM1 and GM3. Torsions were incorporated through a proper dihedral potential with multiplicity ( $m$ ) = 1, unless otherwise specified

$$V_d = K_d[1 + \cos(m\phi - \phi_0)] \quad (3)$$

where  $K_d$  is the force constant, and  $\phi_0$  is the equilibrium dihedral angle. Improper torsions were included wherever explicitly mentioned, the potential energy of which is described by a harmonic function, with  $K_i$  as the harmonic force constant and  $\xi_0$  as the equilibrium dihedral angle

$$V_i = \frac{1}{2}K_i(\xi - \xi_0)^2 \quad (4)$$

Equilibrium values of the potentials for all bonds, angles and dihedrals were picked from target distributions at the atomistic level. The bonded parameters of the coarse-grained sugars and GPIs are listed in Table 2.

**Partitioning Free Energy.** Nonbonded or Lennard-Jones parameters of the coarse-grained molecules are contained in the assigned bead types. The bead types of the simple sugars—glucose, sucrose, trehalose—were assigned by considering the octanol–water partition coefficient ( $\log P_{OW}$ ) obtained from free energy calculations. Free energies of solvation of the sugars in (polarizable) water and water-saturated octanol were calculated separately to obtain  $P_{OW}$ . The amount of water in water-saturated octanol was 25 mol %. Only one sugar molecule was coupled/decoupled with the solvent. Solvation

free energy ( $\Delta G$ ), i.e., the free energy difference ( $\Delta F$ ) of the solute in vacuum ( $F_X$ ) and in the condensed phase ( $F_Y$ ), was calculated using thermodynamic integration according to

$$\Delta G = \Delta F_{YX} = F_Y - F_X = \int_{\lambda_X}^{\lambda_Y} d\lambda \left\langle \frac{\delta U(\lambda)}{\delta \lambda} \right\rangle_{\lambda} \quad (5)$$

The coupling parameter  $\lambda$  defines the strength of the potential energy  $U$  between the solute and the solvent.  $\lambda$  lies in the range between 0 (no interaction) and 1 (full interaction between the two). A soft core approach was used to couple nonbonded interactions in order to remove singularities from the potential energy calculation.<sup>28</sup> Bonded interactions were linearly interpolated.  $\delta U/\delta \lambda$  was calculated at 25 regularly spaced  $\lambda$  intervals between 0 and 1. The simulation time at each such window was 30 ns. The free energy curve was then integrated by the trapezoidal rule to obtain the final value of  $\Delta G$ . Block averaging was done at every  $\lambda$  value to calculate the statistical error in free energy. Partition coefficients were obtained from the difference in the two solvation energies, given by

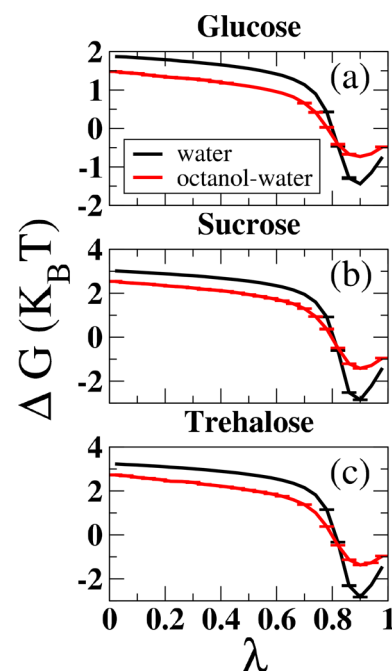
$$\Delta \Delta G_{OW} = \Delta G_O - \Delta G_W = -2.3RT \log P_{OW} \quad (6)$$

Here, the subscript *O* refers to water-saturated octanol, and the subscript *W* refers to water. The obtained free energy values are listed in Table 1, and the solvation free energy profiles from which these values were derived are shown in Figure 4. The calculated partition coefficients compare well with experimental values.

**Table 1. Octanol–Water Partitioning Coefficients ( $\log P_{OW}$ ) of Glucose, Sucrose, and Trehalose Compared to Experimental Values**

	$\Delta \Delta G$ ( $K_B T$ )	$\log P_{OW}$ (calc)	$\log P_{OW}$ (exp) <sup>29</sup>
glucose	6.81	-2.95	-2.8
sucrose	7.28	-3.16	-3.3
trehalose	9.37	-4.06	-3.78

**Bead Types.** To arrive at the final bead types comprising the simple sugars, an iterative process of trial-and-error was carried out to arrive at their respective experimental octanol–water partitioning coefficients. The bead types examined here were taken from the database of MARTINI's polarizable force field and assigned through the parametrization procedure described in the section above. The distribution of bead types within the same sugar ring was determined based on the polarities of the beads relative to each other. For example, the two GP3 beads (B2, B3) in glucose have two free OH groups making them more polar than the GP2 bead (B1) that contains one free OH and one ether oxygen (see Figure 2 and Table 2). The bead types of GPI were assigned on the basis of the newly devised bead types of simple sugars, the chemical nature of the bead, and the interaction matrix of MARTINI. The glycan was constructed in a modular fashion from the models of mono- and disaccharides. Charged beads were used to represent the groups containing  $PO_4^-$  and  $NH_3^+$ . The bead types together with the bonded parameters making up the simple sugars and GPI are listed in Table 2. Alessandri et al. pointed out that short bonds in MARTINI could give rise to discrepancies in the hydrophilic/hydrophobic interactions of the molecule.<sup>30</sup> To take this possibility into account, we have used small (S) beads wherever short bonds (<0.3 nm) had to be included to facilitate finer mapping (see Table 2).



**Figure 4.** Free energy profiles  $\Delta G$  as a function of the coupling parameter  $\lambda$  for (a) glucose, (b) sucrose, and (c) trehalose obtained from the thermodynamic integration of one sugar molecule in water (black) and in water-saturated octanol (red) separately.

**Parametrizing EtNP Linker.** To study the behavior of GPI-anchored GFP placed in lipid bilayers, a crucial step was to model the linker connecting protein and GPI. In all the GPI-APs discovered so far, this bridging linker is the same—phosphoethanolamine (EtNP). The EtNP linker was individually coarse-grained in an aqueous environment of polarizable water. Coarse-grained bonded parameters of the EtNP linker were derived from 200 ns all-atom simulations of the molecule shown in Figure 5a. The simulations were conducted in an aqueous medium of TIP3P water. For the GPI-anchored GFP, the EtNP linker is the bridge between the protein GFP and GPI. Therefore, in order to maintain the same connectivities, the linker was connected to amino acid residues: Threonine-Isoleucine-Glycine-Terminal Cap (THR-ILE-GLY-T), in the same order as in GFP, as shown in Figure 5. The terminal cap (T) is an acetyl group that was added to end the amino-acid chain. At the other end, the linker was connected to the last two mannose residues (Man3-Man2) of GPI. The EtNP linker was represented by two beads: a neutral L1 bead to substitute for ethanolamine and a negatively charged L2 bead to represent the phosphate group. The bead definitions and bonded parameters of the entire molecule in Figure 5b are listed in Table S1 of the SI. Coarse-grained simulations of the molecule in Figure 5b were also conducted for 200 ns to compare with the all-atom system. The derived bonded parameters involving beads L1 and L2 were plugged into the coarse-grained model of GPI-anchored GFP.

**Coarse Graining GFP.** GFP was modeled based on the ELNEDYN<sup>31</sup> framework of MARTINI. ELNEDYN, or the elastic network approach, is built on the philosophy of combining a structure-based coarse-grained model with a thermodynamics-based coarse-grained force field to model a protein. Secondary and tertiary structures of proteins are stabilized to a large extent by h-bonds, but this vital information is lost in the coarse representation. Therefore, to

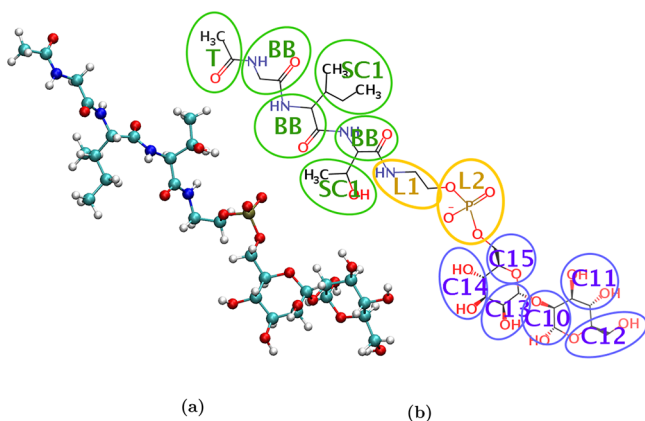
Table 2. Bead Definitions and Bonded Parameters for the Carbohydrates Incorporated in Our Study: Glucose, Sucrose, Trehalose, and GPI<sup>a</sup>

molecule	bead name	bead type	bonds	$r_0$ (nm)	$K_b$ (kJ/mol)	angle	$\theta_0$ (deg)	$K_a$ (kJ/mol)	dihedral	$\phi_0$ (deg)	$K_d$ (kJ/mol)
glucose	B1	GP2	B1–B2	0.328	35000						
	B2	GP3	B1–B3	0.375	35000						
	B3	GP3	B2–B3	0.311	50000						
sucrose	B1	GP2	B1–B2	0.325	30000	B1–B3–B4	85	10	B1–B3–B4–B5	108	14
	B2	GP3	B2–B3	0.311	35000	B2–B3–B4	143	160	B1–B3–B4–B6	166	15
	B3	GP2	B1–B3	0.379	35000	B3–B4–B5	93	165	B2–B3–B4–B5	143	8
	B4	GSN0	B3–B4	0.335	5000	B3–B4–B6	80	280			
	B5	GP3	B4–B5	0.327	10000						
	B6	GP2	B5–B6	0.302	10000						
			B4–B6	0.406	10000						
trehalose	B1	GP2	B1–B2	0.329	20000	B1–B3–B4	77	150	B1–B3–B4–B5	2.9	50
	B2	GP3	B2–B3	0.311	35000	B2–B3–B4	107	300	B1–B3–B4–B6	–54	28
	B3	GP2	B1–B3	0.379	35000	B3–B4–B5	96	300	B2–B3–B4–B5	44	50
	B4	GSP1	B3–B4	0.376	30000	B3–B4–B6	69	250			
	B5	GP3	B4–B5	0.299	50000						
	B6	GP2	B5–B6	0.329	25000						
			B4–B6	0.399	30000						
GPI	C1	GP2	C1–C2	0.325	40000	C1–C2–C3	55	600	C1–PO4–GL1–GL2	39.3	2.5
	C2	GP3	C1–C3	0.307	35000	C1–C3–C2	60.5	600	C2–C1–PO4–GL1 ( $m = 2$ )	23	5
	C3	GP2	C2–C3	0.34	40000	C1–C3–C4	88	200	C3–C1–PO4–GL1	15.4	3
	PO4	GQa	C3–C4	0.37	20000	C3–C1–PO4	112	70	C3–C4–C5–C7	–32.3	6.2
	C4	GSQd	C4–C5	0.30	40000	C2–C1–PO4	144	450	C1–C3–C4–C5	–5.7	20
	C5	GP2	C4–C6	0.40	35000	C2–C3–C4	142	400	C1–C3–C4–C6	–54.4	25
	C6	GP2	C5–C6	0.32	20000	C3–C4–C5	90	500	PO4–C1–C3–C4	–163.3	80
	C7	GSP1	C5–C7	0.35	20000	C3–C4–C6	63	550	C4–C5–C7–C8	12.6	10
	C8	GP3	C7–C8	0.28	35000	C4–C5–C6	80	400	C4–C5–C7–C9	–44	7.8
	C9	GNa	C7–C9	0.34	20000	C4–C6–C5	48	500	C5–C7–C9–C10	46	25
	C10	GSN0	C8–C9	0.32	20000	C4–C5–C7	172	500	C7–C9–C10–C11	114	4.7
	C11	GP3	C9–C10	0.40	15000	C5–C7–C8	114	350	C7–C9–C10–C12	57.7	6
	C12	GP2	C10–C11	0.28	40000	C5–C7–C9	90	250	C7–C9–C10–C13	–91.14	4
	C13	GSP1	C10–C12	0.35	30000	C6–C5–C7	109	300	C9–C10–C13–C14	–80	14
	C14	GP3	C11–C12	0.33	30000	C7–C8–C9	69	300	C9–C10–C13–C15	–132	15
C15	GP2	C10–C13	0.36	20000	C7–C9–C8	50	400				
L1	GNa	C13–C14	0.28	40000	C7–C9–C10	126	50				
L2	GNa	C13–C15	0.35	30000	C8–C9–C10	118	80				
			C14–C15	0.33	30000	C9–C10–C11	100	120			
			C1–PO4	0.30	3000	C9–C10–C12	82	90			
			PO4–GL1	0.40	5000	C9–C10–C13	140	30			
			GL1–GL2	0.34	3000	C10–C11–C12	69	400			
					C10–C12–C11	49	500				
					C10–C13–C14	94	300				
					C10–C13–C14	67	300				
					C11–C10–C13	120	100				
					C12–C10–C13	128	120				
					C13–C14–C15	69	200				
					C13–C15–C14	48	150				
					C1–PO4–GL1	112	20				
					PO4–GL1–GL2	96	50				

<sup>a</sup>MARTINI bead types are prefixed with ‘G’ to indicate the redefined nonbonded parameters.

replicate the secondary, tertiary, and quaternary structures more realistically, an elastic network was imposed on the protein through the ELNEDYN approach. Mapping of amino acids and assignment of bead types is done according to the same protocol as in ref 31, where the center of the backbone bead is located on the  $C_\alpha$  atom of the respective all-atom amino acid. When the distance between the nearest-neighbor beads was less than the imposed cutoff  $R_C$ , a harmonic spring

potential of force constant  $K_S$  was turned on between the two. The equilibrium lengths of these artificial bonds were set to the distances obtained from an equilibrated structure of atomistic GFP in water, and the values of  $R_C$  and  $K_S$  were kept uniform across all such pairs of beads. Nonbonded potentials among the backbone beads connected through a spring force are excluded from the calculation of the system potential. Bonded parameters (bonds, angles) were derived straight from the



**Figure 5.** (a) All-atom representation of EtNP linker. (b) Mapping of the all-atom model in (a) to a coarse-grained parametrization consisting of beads, with the green beads representing the amino acid residues in the following order: THR-ILE-GLY-T, starting from the linkage at L1. BB beads are backbone beads, and SC are side chain beads. The yellow beads make up the EtNP linker, and the blue beads represent GPI's last two mannose residues. Beads are shown with their bead names.

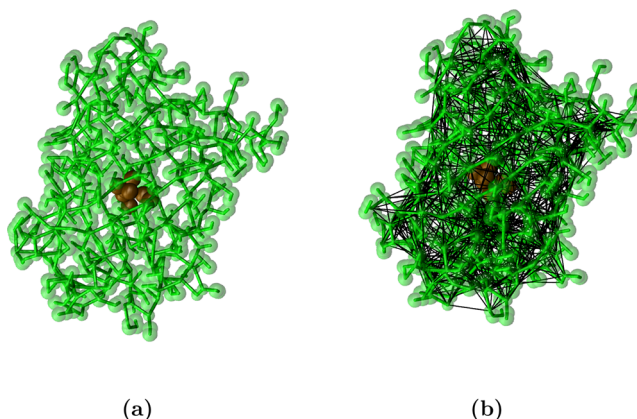
corresponding atomistic simulations of GFP in water. Along with the protein, a coarse-grained representation of the chromophore situated inside the barrel of GFP was also modeled from the all-atom system. We observed that the presence of the chromophore affected the size of GFP and hence was important to model the protein more realistically. Details of the chromophore model are provided in the SI with the mapping scheme illustrated in Figure S3, and the corresponding bonded parameters are listed in Table S2.

As per the work of Periole and co-workers,<sup>31</sup> the optimal values of the elastic scaffold parameters could range from 0.8 to 1.0 nm for  $R_C$  and from 500 to 1000 kJ/mol for  $K_S$ . We observed that for our system of GFP in polarizable water, the combination of  $R_C = 1.0$  nm and  $K_S = 500$  kJ/mol replicates the atomistic system sufficiently well. Mapping of atoms to coarse beads was conducted on an equilibrated structure of GFP from the atomistic simulations. Note that the crystal structure of protein should not be directly mapped to coarse-grained representation, as the protein changes in size upon solvation and equilibration. As the elastic network ensures that the structure and size of the protein are maintained throughout the simulation, the atomistic system to be mapped should be chosen carefully. Figure 6 shows the coarse-grained representation of the protein with and without the elastic network.

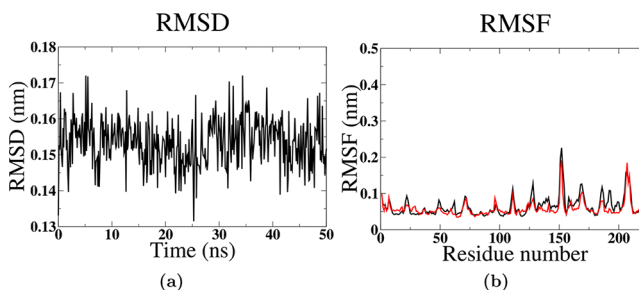
To compare with the crystal structure, we calculate the root-mean-square deviation (RMSD) of GFP. RMSD is a metric used to quantify the degree of similarity between two corresponding, superimposed structures. It is calculated by the following relation

$$\text{RMSD}(t) = \sqrt{\frac{1}{M} \sum_{i=1}^N m_i |r_i(t) - r_i^{\text{ref}}|^2} \quad (7)$$

where  $M = \sum_i m_i$ , the sum of masses of all atoms,  $r_i(t)$  is the position of atom  $i$  at time  $t$  of the simulation trajectory, and  $r_i^{\text{ref}}$  is the position of atom  $i$  in the reference structure. For the calculation of RMSD, only the backbone beads are taken into account. As shown in Figure 7a, RMSD stays well within the resolution of determination of crystal structure, i.e., 0.19 nm,<sup>32</sup> throughout the trajectory, suggesting that the protein is



**Figure 6.** Coarse-grained representation of GFP (a) without and (b) with elastic bonds. The black mesh in (b) depicts the elastic network imposed on the backbone beads of GFP. The chromophore is shown as brown beads in the center of the barrel.



**Figure 7.** (a) Root-mean-square deviation (RMSD) of coarse-grained GFP compared to the crystal structure. (b) Comparison of root-mean-square fluctuation (RMSF) of the all-atom (black) and coarse-grained (red) GFPs in water.

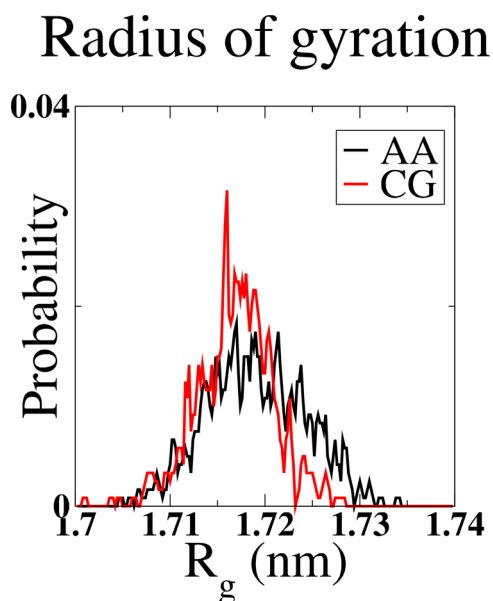
structurally stable. The flexibility of each residue of a protein can be measured by root-mean-square fluctuation (RMSF). RMSF is useful for characterizing local changes along the protein chain. It is calculated for the  $C_\alpha$  atoms in the all-atom case and backbone beads in the coarse-grained case. The RMSF for residue  $i$  is

$$\text{RMSF}_i = \sqrt{\langle (r_i - \langle r_i \rangle)^2 \rangle} \quad (8)$$

where  $r_i$  is the position of atom  $i$  in the residue after superposition with the reference structure, and  $\langle r_i \rangle$  is the average position of atom  $i$ . Figure 7b shows the comparison of root-mean-square fluctuation of each residue of the protein between the all-atom and coarse-grained systems. The local fluctuations/dynamics of the all-atom and coarse-grained GFPs turn out to be quite similar. We also compare the global structure of the protein in the two resolutions by calculating the radius of gyration of the backbone beads in Figure 8 and Table 3. Both RMSF and  $R_g$  plots show good overlap between the two resolutions, further validating the coarse-grained force field.

**Solute–Solute Adapted Nonbonded Interactions.** Nonbonded interactions between neutral beads in MARTINI are described by a Lennard-Jones 12-6 potential energy function

$$V_{\text{Lennard-Jones}}(r) = 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r} \right)^{12} - \left( \frac{\sigma_{ij}}{r} \right)^6 \right] \quad (9)$$



**Figure 8.** Radius of gyration  $R_g$  for GFP as obtained with the atomistic (AA) (black) and coarse-grained (CG) (red) models.

**Table 3.** Average Values of Radius of Gyration  $R_g$  for Atomistic and Coarse-Grained GFP

	atomistic	coarse-grained
$R_g$ (nm)	$1.725 \pm 0.005$	$1.717 \pm 0.004$

where  $r$  is the distance between two particles  $i$  and  $j$ ,  $\sigma_{ij}$  is the distance between them at which potential energy is zero, and  $\epsilon_{ij}$  is the well depth which is a measure of the strength of their interaction. Interaction between charged beads is represented both by aforementioned Lennard-Jones potential and a Coulombic potential energy function to describe the electrostatics

$$V_{\text{Coulomb}}(r) = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_{\text{rel}}r} \quad (10)$$

where  $q$  is the charge on the particle,  $\epsilon_0$  is the dielectric permittivity of vacuum, and  $\epsilon_{\text{rel}}$  is the relative dielectric permittivity of the medium. Charged nonbonded interactions are determined by the charge on the beads, and uncharged nonbonded Lennard-Jones interactions are dictated by the bead types, the parameters of which have been fit to reproduce partition coefficients of small organic molecules in polar-apolar solvent phases.<sup>33</sup> In accordance with the MARTINI parametrization, we did not alter the sugar–lipid interaction parameters because these interactions are taken care of through the octanol/water partitioning coefficients. A couple of studies have reported that MARTINI sugar–lipid parameters obtained through this parametrization scheme are well-characterized. Lopez et al. demonstrated the cryo- and anhydro-protective effect of MARTINI sugars on lipid bilayers.<sup>22</sup> In another study, MARTINI nonreducing disaccharides were shown to disrupt phase segregation in mixed membranes, whereas monosaccharides and reducing disaccharides had no such effect, as was also observed in experiments.<sup>34</sup>

The strategy of using octanol–water partitioning free energies to define nonbonded interactions naturally addresses carbohydrate–lipid or amino acid–lipid interactions, but it is

quite plausible that it cannot cover all conceivable situations met in biochemical modeling. Solute–solute interactions with sugars<sup>35</sup> and proteins<sup>36,37</sup> have previously been reported to turn out overestimated, leading to unnatural aggregation. The degree of aggregation, or stickiness, increases with the increase in length/size of the solute, as observed by Schmalhorst and co-workers.<sup>35</sup>

The MARTINI force field has already been extended to carbohydrates including simple sugars<sup>22</sup> and glycolipids,<sup>38</sup> however, their self-interactions are overestimated leading to unnatural aggregation both in solution and in membranes. Gu et al. proposed to use the small (S) beads of MARTINI which reduced the clustering propensity of glycolipids GM1 and GM3 when placed in membranes to better reproduce the clustering observed in the atomistic system.<sup>27</sup> Here, note that badly parametrized intermolecular vdW interactions are a general problem in force-field development, whether coarse-grained or atomistic.<sup>39,40</sup> Therefore, a coarse-grained model parametrized on the basis of atomistic cluster sizes cannot be trusted.

To fix this imbalance in interactions, a few strategies have been proposed based on the incorporation of solution observables in the parametrization process such as Kirkwood/Buff integrals,<sup>41–43</sup> osmotic pressure,<sup>44,45</sup> and osmotic coefficient.<sup>46</sup> Yet another way of optimizing potentials in MD simulations is by calculating the second virial coefficient of osmotic pressure  $B_{22}$ , a quantity that describes the deviation of a solution from ideality. It is related to the osmotic pressure  $\pi$  in the following way

$$\pi(T, c) = RT(c + B_{22}c^2 + B_{23}c^3 + \dots) \quad (11)$$

where  $c$  is the solution concentration,  $T$  is the temperature,  $R$  is the gas constant, and  $B_{ij}$  are coefficients of the virial expansion of osmotic pressure. The nonbonded forces between aggregating solutes can be scaled down by scaling down the pairwise amplitudes  $\epsilon_{ij}$ s of the Lennard-Jones potentials (eq 9) to match the experimental  $B_{22}$  values. This method has been applied on MARTINI for proteins by Elcock et al.<sup>36</sup> and for polysaccharides by Schmalhorst et al.,<sup>35</sup> in the environment of antifreeze water of MARTINI. We followed the same protocol to optimize the nonbonded interactions of simple sugars and GPIs in polarizable water as polarizability of the aqueous medium is essential to our study.

Based on the assumption that the total solute potential energy can be approximated as the sum of pairwise solute–solute interactions, McMillan and Mayer<sup>46</sup> derived a relation for  $B_{22}$  from the potential of mean force ( $w(r)$ ) between two particles separated by distance  $r$

$$B_{22} = -2\pi N_A \int_0^\infty \left[ \exp\left(-\frac{w(r)}{RT}\right) - 1 \right] r^2 dr \quad (12)$$

with  $N_A$  being Avogadro's constant. At thermodynamic equilibrium,  $w(r)$  can be approximately related to the radial distribution function (RDF)  $g(r)$  in the following way

$$g(r) = \exp\left(-\frac{w(r)}{RT}\right) \quad (13)$$

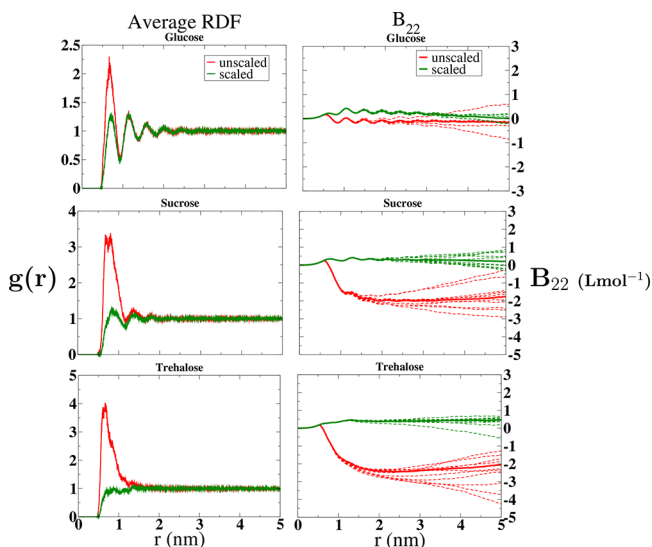
In order to calculate  $B_{22}$  from simulations, the integral in eq 12 needs to be finite

$$B_{22}(r') = -2\pi N_A \int_0^{r'} [g(r) - 1] r^2 dr \quad (14)$$

The value of  $r'$  should be high enough where the solute–solute interactions vanish and  $B_{22}(r') \rightarrow B_{22}(\infty)$ . In our systems, we found that a value of  $r' = 5$  nm worked consistently for all three sugar systems. For a two-component system, subscript 1 in  $B_{ij}$  stands for solvent, subscript 2 stands for solute. Thereby,  $B_{22}$  denotes solute–solute interactions. Positive values of  $B_{22}$  indicate net repulsion, and negative values indicate attraction between solute molecules. Its magnitude denotes the extent of aggregation. Experimentally,  $B_{22}$  can be obtained from static light scattering, and in an MD simulation it is derived from cumulative solute–solute RDF. Aqueous solutions of 100 mM sugar solutions were prepared and simulated for 1  $\mu$ s for monosaccharide (glucose) and 2  $\mu$ s for disaccharides (sucrose and trehalose). Cumulative RDFs were calculated for every 200 ns segment of the trajectories. Using eq 14,  $B_{22}$  was obtained by an integration over the solute–solute RDFs. Solute–solute interactions were varied by scaling down the  $\epsilon_{ij}$  of all the sugar–sugar pairwise nonbonded potentials of MARTINI, using a simple relation

$$\epsilon_{ij,new} = 2 + \gamma(\epsilon_{ij,old} - 2) \quad (15)$$

with  $\gamma$  as the scaling factor. This ansatz was also used by Schmalhorst et al. The constant, 2 kJ/mol, is the lowest value of  $\epsilon_{ij}$  in the MARTINI database. After systematically testing different scaling factors, we arrived at  $\gamma = 0.85$  that worked consistently for all the sugars in achieving more realistic osmotic pressure coefficients and eliminating aggregation in sugars. As can be seen in Figure 9, unscaled/original MARTINI resulted in  $B_{22}$  values in the attractive regime, whereas the experimentally obtained values suggest somewhat repulsive interactions. The  $B_{22}$  profiles obtained after the scale-down resulted in positive values with the averages close to those from experiments (see Table 4).



**Figure 9.** Sugar–sugar radial distribution functions (RDFs)  $g(r)$  as a function of distance  $r$  averaged over all 200 ns segments and corresponding  $B_{22}$  vs  $r$  profiles of all 200 ns segments put together for solutions of glucose, sucrose, and trehalose. In the  $B_{22}$  plots, the dotted lines come from the 200 ns intervals, and the solid line is the averaged profile over all the intervals. Profiles from unscaled  $\gamma = 1$  are shown in red, and profiles from scaled  $\gamma = 0.85$  are shown in green. The averaged constant value at the far end (at 5 nm) is the reported  $B_{22}$  value.

**Table 4.**  $B_{22}$  Values Collected at the Tail End of  $B_{22}$  vs  $r$  Profiles Calculated from Averaged RDFs

	$B_{22}(\text{L mol}^{-1})$		
	$\gamma = 1.0$	$\gamma = 0.85$	exp
glucose	−0.171	0.012	0.117 <sup>47</sup>
sucrose	−1.765	0.206	0.305 <sup>47</sup>
trehalose	−2.059	0.451	0.51 <sup>48</sup>

**Simulation Details.** All the Molecular Dynamics (MD) simulations in this work were performed with the simulation engine: GROMACS-2018.3.<sup>49</sup>

**All-Atom.** The all-atom models of the simple sugars considered in this study, glucose, sucrose, and trehalose, were built with the GLYCAM06h force field<sup>50</sup> with TIP3P water<sup>51</sup> in the background. Only one sugar solvated in water in cubic boxes was simulated for 200 ns each, so as to extract bonded information (bonds, angles, dihedrals) to build their coarse-grained representations.

The all-atom model of the GPI anchor was constructed by merging two force-fields: GLYCAM06h to represent the glycan head and Lipid14<sup>52</sup> for the lipid tail. Figure 1 shows the transition between the two force-field domains. The inositol-together-with-phosphoglycerol (Ino+PGL) part of the molecule, shown in blue, is the hybrid, bridging moiety connecting the glycan head and the lipid tail. The atom types for this bridging residue were chosen through a careful mixing of the atoms from GLYCAM06h and Lipid14. Partial charges, angles, and torsions of this bridge were derived using quantum mechanical calculations, as described in our previous work.<sup>7</sup> We consider only pure DMPC lipid bilayer in this study, which was modeled with Lipid14. The lipid tail of the GPI is also a dimyristoyl. GFP was parametrized using AMBER's protein force field: ff14SB.<sup>53</sup> The aqueous phase was represented by TIP3P waters. The construction of the systems was achieved using the LEaP facility of AMBER. AMBER and GLYCAM topologies were converted to GROMACS format using a script that was originally written by Sorin and Pande<sup>54</sup> and was further modified by us to accommodate the specifics of GLYCAM06h.<sup>18</sup> One  $\mu$ s long simulations were conducted for free GPIs in water and in 8\*8 DMPC bilayers each, and 4 sets of 1  $\mu$ s long simulations amounting to a total of 4  $\mu$ s of simulation time were performed for GFP-GPIs embedded in larger 16\*16 DMPC bilayers. The detailed methodology of the all-atom model development has been described in our previous paper.<sup>7</sup>

**Coarse-Grained.** The coarse-grained GPI glycan was attached to a dimyristoyl lipid tail, the parameters of which were directly taken from the MARTINI lipid parameter set.<sup>55</sup> Bonded parameters to define the link between the phosphoinositol of GPI and the lipid tail were also taken from MARTINI's database. One GPI was inserted into each leaflet of an 8\*8 bilayer of pure, hydrated DMPC and simulated for 1  $\mu$ s. The system was assembled using the *insane* script of the Wassenaar group.<sup>56</sup> A single GFP-GPI was inserted into a 16\*16 pure, hydrated bilayer of DMPC to study its conformational behavior w.r.t. lipid bilayers. All the aforementioned coarse-grained systems were solvated in MARTINI's polarizable water. Nonbilayer systems were set up in cubic boxes with a minimum distance of 1.2 nm between the edges of the solute and the box. Bilayer systems were constructed in orthorhombic boxes. Counterions, represented as hydrated Na<sup>+</sup> beads, were added to the GFP-GPI-bilayer



Table 5. Technical Details of Simulation Settings for All the All-Atom and Coarse-Grained Systems Included in This Study

system	species	number	box size (nm)	time (ns)	system	species	number	box size (nm)	time (ns)	
<b>all-atom (AA) mapping (aqueous systems)</b>					<b>calculation of <math>B_{22}</math></b>					
glucose	glucose	1	4 × 4 × 4	200	trehalose	trehalose	1	5 × 5 × 5	200	
	water	876				water	543			
sucrose	sucrose	1	4 × 4 × 4	200	GPI	GPI	1	6 × 6 × 6	1000	
	water	2178				water	1188			
trehalose	trehalose	1	4 × 4 × 4	200	GFP	GFP	1	10 × 10 × 10	50	
	water	2170				water	4626			
GPI	GPI	1	5.3 × 5.3 × 5.3	1000	EtNP	EtNP	1	5 × 5 × 5	200	
	glycan					molecule				
GFP	water	4753			water	617				
	GFP	1	8.2 × 8.2 × 8.2	200	Na+	1				
	water	16608			<b>membrane systems</b>					
EtNP	Na+	7			glucose	glucose	420	28 × 28 × 28	1200	
	EtNP	1	5.2 × 5.2 × 5.2	200		water	59289			
	molecule				Na <sup>+</sup> Cl <sup>-</sup>	420				
water	water	4592			Ca <sup>2+</sup> Cl <sub>2</sub> <sup>-</sup>	42				
	Na+	1			sucrose	sucrose	420	28 × 28 × 28	2200	
	<b>membrane systems</b>					water	58738			
GPI in DMPC	GPI	2	8.4 × 8.4 × 15.4	1000	Na <sup>+</sup> Cl <sup>-</sup>	420				
	DMPC	126			Ca <sup>2+</sup> Cl <sub>2</sub> <sup>-</sup>	42				
	water	17095			trehalose	trehalose	420	28 × 28 × 28	2200	
GFP-GPI in DMPC	GFP-GPI	1	15 × 15 × 19	4 × 1000		water	58699			
	DMPC	511			Na <sup>+</sup> Cl <sup>-</sup>	420				
	water	81846			Ca <sup>2+</sup> Cl <sub>2</sub> <sup>-</sup>	42				
Na+	Na+	7			<b>membrane systems</b>					
	<b>coarse-grained (CG) mapping (aqueous systems)</b>					GPI in DMPC	GPI	2	6.5 × 6.5 × 16	1000
	glucose	glucose	1	5 × 5 × 5	200		DMPC	126		
water		338			water	4657				
sucrose	sucrose	1	5 × 5 × 5	200	GFP-GPI in DMPC	GFP-GPI	1	12.5 × 12.5 × 19.5	4000	
	water	545				DMPC	511			
	water				water	20581				
					Na+	7				

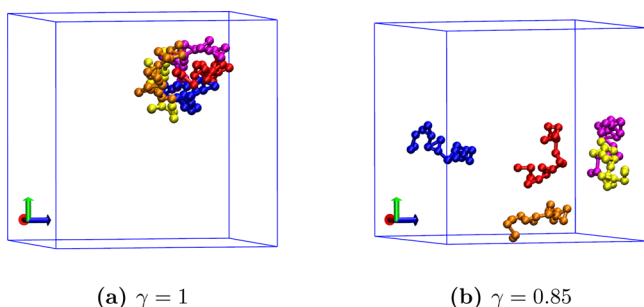
system to neutralize the net charge of  $-7$  on the protein. Energy minimization was performed for 10000 steps using the steepest descent algorithm, followed by an NPT equilibration for 1 ns. Postequilibration, the production run was carried out in an NPT ensemble. Protein-free GPIs in bilayers were simulated for 1  $\mu$ s. GFP-GPI-bilayer systems were simulated for 4  $\mu$ s. The first 10 ns of the production run of each system were excluded from analysis. The time step used for GPI simulations was 5 fs, which is relatively small compared to the typical range of time steps (10–40 fs) used in MARTINI models. Since GPI is structurally quite flexible, we avoided imposing constraints on the molecular conformation. The inclusion of rather tight bonds, some of them with force constants around 40000 kJ/mol, and the crucial glycosidic dihedrals made it necessary to limit the time step to 5 fs so as to avoid numerical instabilities. Besides, the choice of time step is in agreement with the study of MARTINI glycolipids where the small time step was required to avoid numerical instabilities arising from the tight force constants and a large number of angle and dihedral potentials used to maintain the complicated conformation of the atomistic glycolipids.<sup>38</sup> The cutoff (both vdW and Coulomb) for all the systems was 1.1 nm, imposed by the Verlet scheme.<sup>57</sup> The PME method<sup>58</sup> was employed for electrostatics, and the plain cutoff method was employed for vdW interactions. The vdW potential was shifted in energy to smoothly reduce it to zero at the cutoff. The relative dielectric constant was fixed at 2.5, the default value for polarizable water

in MARTINI. The leapfrog stochastic dynamics (sd) integrator<sup>59</sup> was used to integrate Newton's equations of motion. Temperature was controlled by the sd integrator with a time constant of 1 ps. For equilibration, the Berendsen barostat<sup>60</sup> was used to maintain the pressure at 1 bar, whereas for the production run the Parrinello–Rahman barostat<sup>61</sup> was employed. A time constant of 5 ps was used for the former, and a time constant of 12 ps was used for the latter. For all the cubic boxes, isotropic pressure coupling was applied, but for the bilayer systems semi-isotropic coupling was used, that is, isotropically only in  $x$  and  $y$  directions. Detailed information on the simulation settings can be found in Table 5.

## RESULTS AND DISCUSSION

### Scaled Solute–Solute Interactions: GPI and GFP-GPI.

The scaling factor,  $\gamma = 0.85$ , that was derived from simulations of sugar solutions was applied to nonbonded interactions between GPIs. To observe the aggregating tendencies of GPIs before and after scaling, 5 GPIs (GPI core + PGL) were solvated in water and simulated for 1  $\mu$ s. Figure 10 shows the snapshots taken at the end of the simulations with (a) unscaled and (b) scaled MARTINI parameters. With the original MARTINI parameters, all the GPIs ball up to form a globule which remains stable throughout the simulation. Upon scaling down the sugar–sugar interactions using the same scaling law (eq 15), we observed that GPIs freely float in water and

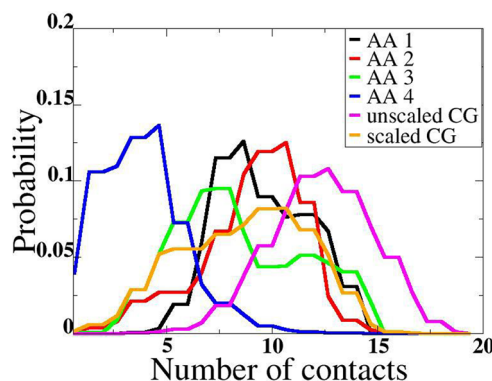


**Figure 10.** Snapshots taken at the end of 1  $\mu$ s long simulations of five GPI glycans in water modeled with (a) unscaled MARTINI at  $\gamma = 1$  and (b) scaled MARTINI at  $\gamma = 0.85$ . Each GPI molecule has a different color.

intermittently associate with each other. At no point do they aggregate into a solid, compact globule.

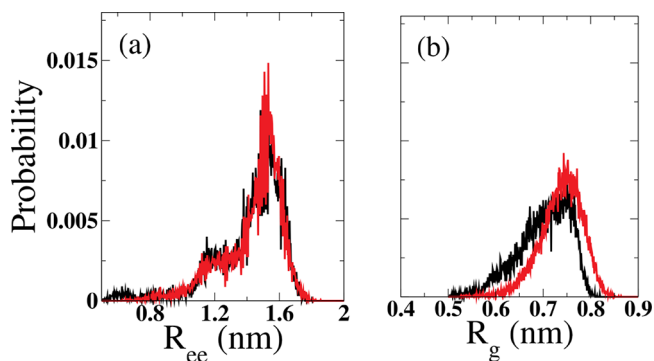
The combined model of GPI+EtNP+GFP was inserted into a pure 16\*16 lipid bilayer of DMPC to study the conformational behavior of GFP w.r.t. the bilayer. From a 4  $\mu$ s long simulation, it became apparent that the interactions between GFP and GPI were significantly stronger compared to the atomistic system. This is not surprising since the parametrization of nonbonded interactions in atomistic and coarse-grained systems follows different routes. We recall that the issue of overestimation of solute–solute interactions has been reported for both all-atom and coarse-grained systems (MARTINI in particular). In order to be consistent, we must, of course, make the coarse-grained model reflect the one at the atomistic level of a single molecular species (GFP-GPI) and weaken the sugar–protein interactions. Due to the lack of explicit experimental data on mixtures of sugars and amino acids, we tentatively use the scaling factor obtained for sugar–sugar interactions. Since the issue of aggregation has been reported both in proteins and sugars, it is not surprising that the interactions between proteins (GFP) and sugars (GPI) would also be similarly affected. To beat the excessive attractive force down, we applied the same scaling factor,  $\gamma = 0.85$ , to the Lennard-Jones potential between GFP and GPI beads. The scale-down presented results comparable with the all-atom system. The extent of interaction between molecules in close proximity can be quantified by the number of contacts formed between the two. We counted the number of contacts made by GFP as a whole with every atom of GPI. **Figure 11** shows how the unscaled and scaled coarse-grained versions compare with the all-atom system. Results of four different 1  $\mu$ s long atomistic trajectories are placed against those of 4  $\mu$ s long coarse-grained trajectories. A number of contacts made were counted within a shell of radius 0.6 nm. For a 1-to-1 comparison between the all-atom and coarse-grained resolutions, we mapped the atomistic GFP-GPI system to the coarse-grained form prior to calculating the frequency of contacts. The scaled coarse-grained force field (orange) covers the same range of contact frequencies as the all-atom system, whereas the unscaled coarse-grained force field lies far on the higher side, an uncharted regime (15–20) of the all-atom system. This shows that the interactions between GFP and GPI are overly strong in the regular MARTINI force field.

**Comparison to All-Atom Simulations.** *GPI.* Having validated the modified MARTINI force field for simple sugars, the study was extended to model our system of interest, the GPI anchor, as outlined in the *Parametrization Strategy*. All the



**Figure 11.** Comparison of distributions of number of contacts made within a radius of 0.6 nm between GFP and GPI glycan between four different all-atom (AA) trajectories (black, red, green, blue) and coarse-grained (CG) trajectories (magenta for the unscaled and orange for the scaled force field). The plots show running averages over five neighboring data points to enhance legibility.

bonded parameters were derived from an all-atom system of 1 GPI core (without the lipid tail) in water. The comparison of the bonded potentials is shown in the SI in **Figures S4 and S5**. All the comparisons between the all-atom and coarse-grained systems were conducted between the mapped atomistic (in other words, pseudo-CG) and actual coarse-grained trajectories. To compare the global structures of the GPIs between atomistic and coarse-grained descriptions, we calculated their radius of gyration and end-to-end distance. Radius of gyration,  $R_g$ , gives an estimate of the size and conformation of a chainlike molecule, for, e.g., if the chain is coiled up or extended. End-to-end distance,  $R_{ee}$ , describes how much the polymer is stretched in structure. The comparison along with the values are shown in **Figure 12** and **Table 6**, respectively. As



**Figure 12.** Comparison of structural properties (a) end-to-end distance  $R_{ee}$  and (b) radius of gyration  $R_g$  between the all-atom (black) and coarse-grained (red) representations of a single GPI core in water.

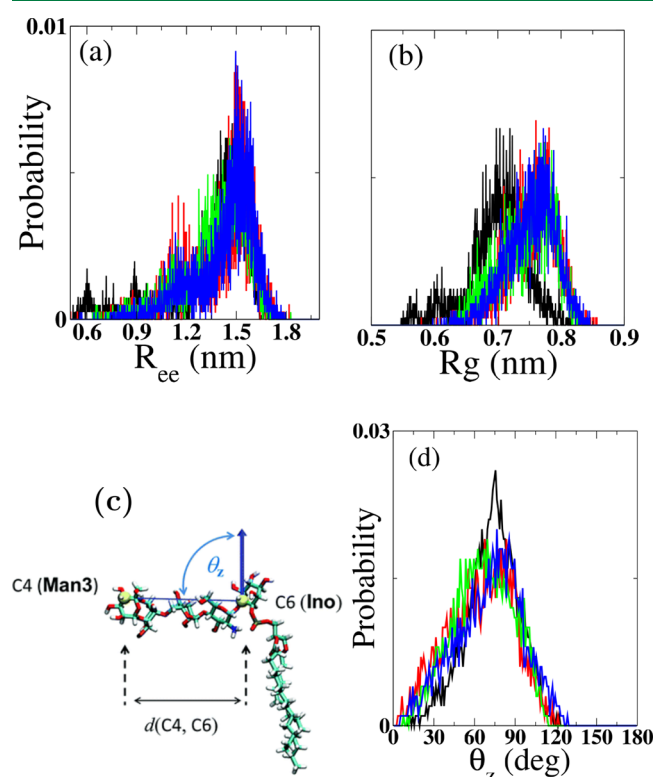
is evident from the overlapping plots and values, our coarse-grained GPI structurally represents its atomistic counterpart

**Table 6.** Average Values of End-to-End Distance  $R_{ee}$  and Radius of Gyration  $R_g$  between All-Atom (AA) and Coarse-Grained (CG) GPI Core in Water

	$R_{ee}$ (nm)	$R_g$ (nm)
AA	$1.41 \pm 0.22$	$0.70 \pm 0.05$
CG	$1.44 \pm 0.19$	$0.74 \pm 0.04$

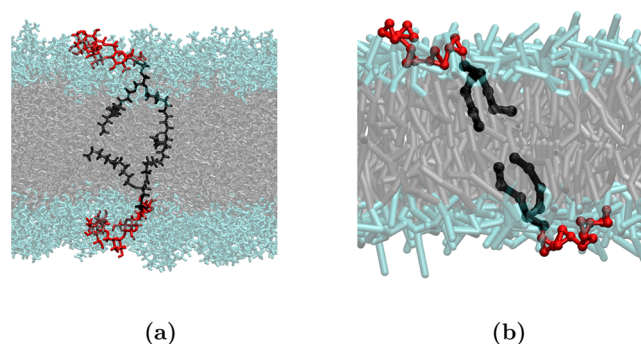
really well. The  $R_{ee}$  values match perfectly, whereas the coarse-grained  $R_g$  distribution (red) is slightly right-shifted, even though the modes are the same. This is because of the bigger sized coarse-grained particles that experience a basal LJ repulsion, which is absent in the pseudo-CG trajectory (black).

The GPI anchor was inserted into pure lipid bilayers of DMPC maintaining the same setup as in the corresponding atomistic system (see Table 5). Global structural properties, i.e., radius of gyration and end-to-end distance, were again compared between the all-atom and coarse-grained systems as shown in Figures 13a and 13b. Plots are shown for GPIs both



**Figure 13.** Comparison of structural properties (a) end-to-end distance,  $R_{ee}$ , and (b) radius of gyration,  $R_g$ , between all-atom and coarse-grained GPIs in a pure DMPC bilayer. Part (c) shows the description of tilt angle  $\theta_z$  of the GPI core, and its corresponding distribution profiles are displayed in (d). Profiles of all-atom GPI in the top leaflet are shown in black, in the bottom leaflet is shown in red, the coarse-grained GPI in the top leaflet is shown in green, and in the bottom leaflet it is shown in blue.

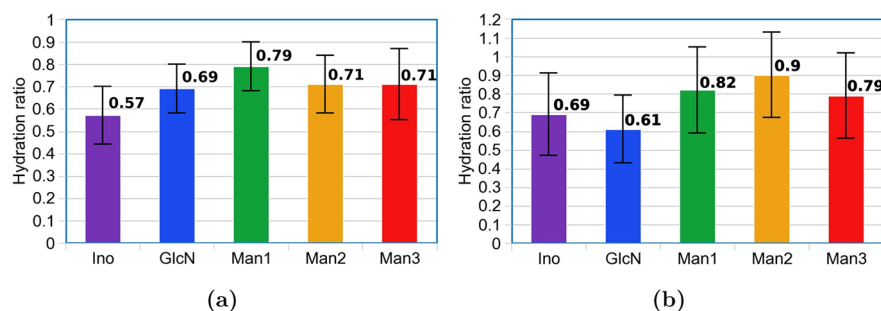
in upper and lower leaflets. To study the conformation adopted by the GPI with respect to the lipid bilayer, we calculate the angle of tilt formed by the GPI core with the bilayer normal. Figure 13c shows the definition of the tilt angle. In the atomistic system, it is the angle formed by the vector connecting the end points: C4 atom of Man3 and C6 atom of Ino, with the bilayer normal ( $z$  axis in this case). In the coarse system, this vector connects the beads containing the aforementioned atoms in the atomistic system. The distribution of the tilt angle of the coarse-grained GPI largely overlaps with that of the atomistic GPI (see Figure 13d). The peak value is  $\approx 80$  degrees, which implies that in both all-atom and coarse-grained representations GPIs flop down on the membrane, with the whole GPI core almost swimming in the headgroup region of the lipid bilayer (see Figure 14).



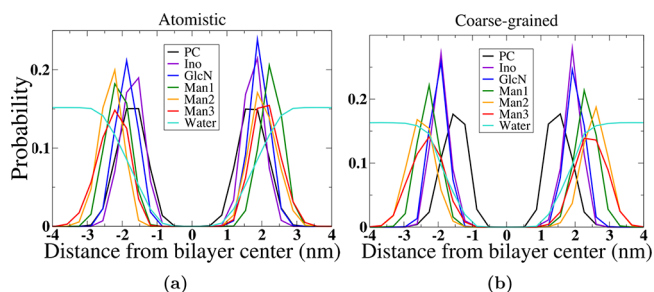
**Figure 14.** Snapshots at the end of 1  $\mu$ s long simulations of GPIs in DMPC bilayers for (a) the all-atom and (b) the coarse-grained model.

We characterized the embedding of the GPI within the lipid headgroup region by calculating the hydration number for each of the five sugar residues of the GPI, which is the number of water molecules lying within a radius of 5.5 Å from the atoms of the sugar residues. This distance criterion was applied only to the oxygen atoms of the waters in the all-atom system and to the central, neutral beads of the three-bead-waters in the coarse-grained system. Hydration numbers of each saccharide ring were compared between the purely aqueous system ( $N_{soln}$ ) where only 1 GPI is solvated in water and the bilayer system ( $N_{mem}$ ) with 1 GPI inserted into each leaflet. Figure 15 shows the hydration ratios  $\frac{N_{mem}}{N_{soln}}$  for the all-atom and coarse-grained GPIs. The relative hydration is lowest for the Ino (violet) and GlcN (blue) residues in both the all-atom and coarse-grained cases and highest for Man1 (green) in the all-atom and for Man2 (orange) in the coarse-grained systems. When comparing hydration ratios to the density profiles of each residue along the bilayer normal (see Figure 16), it is observed that either Man1 or Man2 can be the outermost residue or, in other words, the most solvent-exposed residue in the all-atom system. Ino and GlcN lie at about the same distance away from the bilayer center in the coarse-grained system, whereas a small difference can be seen in the all-atom system. In both the all-atom and coarse-grained systems, Man3 (red) lies closer to the bilayer head than either Man1 or Man2. The same is conveyed by the hydration ratio plots of Man3, indicating that GPIs flop down on the bilayer in both representations. Note that the embedding of GPI into the lipid head is more pronounced in the all-atom than the coarse-grained system. This difference arises from the differences in size of the all-atom and the coarse-grained particles they define. Atoms can percolate more easily into gaps between lipid heads than coarse-grained beads, thereby exposing them less to the solvent phase. Layering effects tend to occur for the coarse-grained system at a larger length scale compared to the all-atom system. Regardless, the overall qualitative picture of the conformation of GPI and its interaction with the membrane is retained in the coarse-grained representation.

**GPI-Anchored GFP.** We recall that from the all-atom simulations from our previous work<sup>7</sup> we could convincingly infer the following properties of the mutual interaction of the three different molecular species: (i) the GPI core undergoes similar conformational changes as if free in solution; (ii) the GPI core lies in close contact with the lipid head groups for both the free GPI and with the GPI-AP; (iii) the GPI core makes contacts with the attached protein; and (iv) the EtNP-



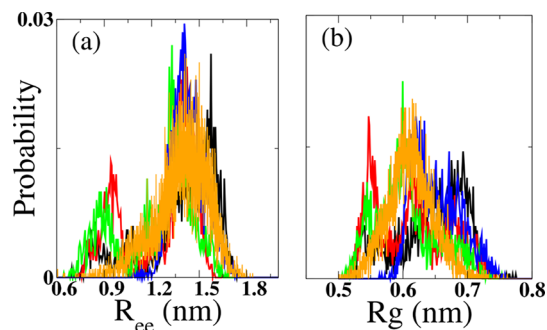
**Figure 15.** Hydration ratios for each carbohydrate residue of the GPI in the (a) all-atom and (b) coarse-grained system.



**Figure 16.** Comparison of density distributions of each residue of the GPI away from the bilayer center along the bilayer normal between the (a) all-atom and (b) coarse-grained systems.

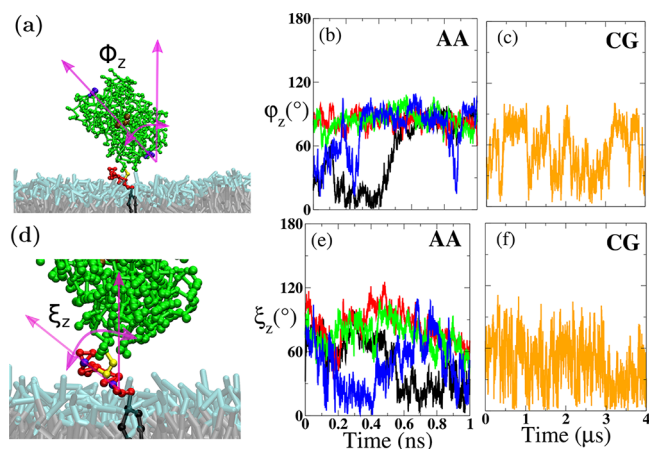
linker conveys extraordinary flexibility to the orientation of the protein w.r.t the bilayer.

We now verify the aforementioned properties with our coarse-grained model. In Figure 17, we compare the structural



**Figure 17.** Comparison of end-to-end ( $R_{ee}$ ) distance and radius of gyration ( $R_g$ ) of GFP-attached-GPI between four different  $1 \mu\text{s}$  long all-atom (black, red, green, blue) and a  $4 \mu\text{s}$  long coarse-grained (orange) trajectories.

properties, end-to-end distance,  $R_{ee}$ , and radius of gyration,  $R_g$ , of GPI when attached to GFP between the two resolutions. For both properties, the values from the coarse-grained system average around the peak values of the all-atom plots. The angle of tilt of both the GFP and GPI from the bilayer normal is a way of quantifying the extent of their communication with the lipid bilayers and of overall conformation in general. The definition of the tilt angle, along with the plots of comparison of the values, is illustrated in the schematic in Figure 18. The results from all four all-atom trajectories show that GFP eventually ends up reclining on the membrane, with its tilt angle saturating around  $70^\circ$ . The coarse-grained profile shows similar behavior of GFP until  $1 \mu\text{s}$ , beyond which the protein fluctuates greatly in its orientation. It is to be noted that the



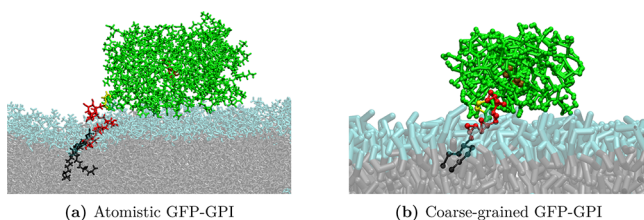
**Figure 18.** Comparison of tilt angle of (a) GFP and (d) GPI between four independent all-atom (black, red, green, blue) and coarse-grained (orange) systems. Parts (b) and (c) show tilt angles of GFP, and parts (e) and (f) show tilt angles of GPI. Tilt angle  $\phi_z$  of GFP is defined as the angle between the bilayer normal ( $z$  axis) and the vector connecting the purple residues (glutamine and histidine). (d) Tilt angle  $\xi_z$  of GPI is defined in the same way as in Figure 13c.

dynamics of a coarse-grained system is always faster than atomistic, about 4 times faster as has been reported for MARTINI. This is because of reduced degrees of freedom in the coarse-grained landscape that leads to loss of friction and hence faster dynamics. This implies that  $1 \mu\text{s}$  of coarse-grained simulation is equivalent to  $4 \mu\text{s}$  of all-atom simulation. Up until the same time frame as the atomistic simulations, coarse-grained GFP-GPI shows similar profiles of tilt angle. On running the simulation longer, it is revealed that, in fact, GFP does wobble a fair deal, instead of lying consistently flat on the membrane, a deceptive picture presented by the all-atom simulations as an offshoot of slow dynamics. The tilt angle of GPI also fluctuates between  $20$  and  $100^\circ$  in both the all-atom and coarse-grained systems. This shows that the GPI is equally flexible in structure in both the all-atom and coarse-grained systems.

Figure 19 shows snapshots of all-atom and coarse-grained simulations after  $700 \text{ ns}$  when GFP lies flat on the membrane.

## CONCLUSIONS

We developed a coarse-grained model of simple sugars—glucose, sucrose, and trehalose—, GPI and GPI-anchored GFP with a combined bottom-up and top-down approach to parametrize the bonded and nonbonded interactions, respectively. The model development is based on a modified version of the MARTINI force field that is suitable for



**Figure 19.** GFP-GPI inserted into DMPC lipid bilayers at atomistic and coarse-grained resolutions in (a) and (b), respectively.

modeling carbohydrates in the environment of polarizable water. The interaction potentials of lipid–lipid, sugar–lipid, and protein–lipid were retained from the MARTINI polarizable force field, but the potentials describing sugar–sugar and sugar–protein were altered by scaling down the amplitudes  $\epsilon_{ij}^s$  of the Lennard-Jones potentials to match the experimental and atomistic behavior. A scaling factor of  $\gamma = 0.85$  was sufficient to reproduce the experimental osmotic virial coefficients ( $B_{22}$ ) of simple sugars, which was extended to the bead types of GPI core. Using polarizable water was essential to the study because our objective was to study the conformational characteristics of GPI and GPI-AP inserted in lipid bilayers for which the interfacial interplay of interactions among lipid heads, carbohydrates, protein, and water needed to be well characterized. On comparing our model of GPI in polarizable water versus in standard MARTINI water, we observed that GPIs interact a great deal with the membrane in polarizable water, just as the atomistic case, whereas they barely interacted with the lipids in standard water causing the glycan, for the most part, to project out of the lipids like a brush. This shows that the water model has a strong effect on the GPI conformation.

GFP was individually coarse-grained in water with the ELNEDYN force field and was subsequently attached to GPI in a modular fashion with a EtNP linker, which also was separately coarse-grained from the atomistic system. GPI proves to be flexible both in the atomistic and coarse-grained landscapes, and the orientation of the attached protein (GFP) with respect to the lipid membrane fluctuates significantly. A plausible reason for this unsteady behavior could be the absence of specific adhesive interactions between GFP and the lipid bilayer. This phenomenon was observed in our control simulations where upon forcing GFP to lie in contact with the bilayer headgroups through a biased force for 300 ns and subsequently releasing the force, GFP moved away from the bilayer after about 500 ns. It has also been reported in experiments that GFP only negligibly binds to membranes.<sup>62</sup> The analysis of the similar number of contacts formed between GFP and GPI at the atomistic and coarse-grained resolutions suggests that they interact similarly in the two representations, providing further validation to our coarse-grained model. Our coarse-grained model of GPI along with its EtNP linker, both of which together form a conserved entity, can be combined with other GPI-anchored proteins like alkaline phosphatase, Thy1, MSP1 of *Plasmodium falciparum*, or even prion protein to address crucial questions concerning their general orientation, mechanisms of action, or pathogenesis.

The speed-up obtained from the coarse-graining was 16-fold in the GPI simulations and 10-fold in the GFP-GPI simulations. With this fast dynamics, we can further address challenging questions that entail larger systems and longer simulation runtime, like the role of GPIs in protein trafficking

which can be studied by observing their partitioning tendencies toward liquid-ordered or liquid-disordered regions of heterogeneous membranes consisting of a variety of lipids including gangliosides and cholesterol. The coarse-grained model of GPI presented herewith can be used in conjunction with the ever-expanding library of MARTINI lipid types to add another component toward building a complex plasma membrane.

## ■ ASSOCIATED CONTENT

### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.0c00056>.

Parameters of coarse-grained EtNP linker containing associated bead types and bonded potentials, Table S1; dihedral (Figure S1) and angle distributions (Figure S2) compared between atomistic and coarse-grained EtNP; mapping scheme of chromophore inside GFP, Figure S3; bead types and bonded parameters, Table S2; comparison of dihedral distributions for GPI in water between atomistic and coarse-grained systems (Figure S4) and of angle distributions (Figure S5); and procedure for backmapping coarse-grained system of GPI-anchored GFP, complemented by list of residue names of lipid DMPC (Table S3) and of GPI (Table S4) (PDF)

gpi.amber.map – mapping file for GPI; dmcp.amber.map – mapping file for DMPC; gfp-gpi-aa.top – atomistic topology of GPI-anchored GFP inserted in DMPC bilayer; and gfp-gpi-cg.gro – coarse-grained structure file of GPI-anchored GFP inserted in DMPC bilayer (ZIP)

## ■ AUTHOR INFORMATION

### Corresponding Author

**Mark Santer** – Max Planck Institute of Colloids and Interfaces, Potsdam 14476, Germany; [orcid.org/0000-0002-5545-5155](https://orcid.org/0000-0002-5545-5155); Email: [mark.santer@mpikg.mpg.de](mailto:mark.santer@mpikg.mpg.de)

### Authors

**Pallavi Banerjee** – Max Planck Institute of Colloids and Interfaces, Potsdam 14476, Germany; Institute of Biochemistry and Biology, University of Potsdam, Potsdam 14469, Germany; [orcid.org/0000-0002-1714-2780](https://orcid.org/0000-0002-1714-2780)

**Reinhard Lipowsky** – Max Planck Institute of Colloids and Interfaces, Potsdam 14476, Germany; Institute of Biochemistry and Biology, University of Potsdam, Potsdam 14469, Germany; [orcid.org/0000-0001-8417-8567](https://orcid.org/0000-0001-8417-8567)

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jctc.0c00056>

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was supported by the International Max Planck Research School (IMPRS) for Multiscale Biosystems. Open Access funding was provided by the Max Planck Society. We thank Marko Wehle for his contribution to the atomistic GPI model. We are grateful to Alex H. de Vries, Mateusz Sikora and Philipp S. Schmalhorst for the helpful discussions.

## ■ REFERENCES

- (1) Low, M. G.; Saltiel, A. R. Structural and functional roles of glycosyl-phosphatidylinositol in membranes. *Science* **1988**, *239*, 268–275.
- (2) Tachado, S. D.; Gerold, P.; Schwarz, R.; Novakovic, S.; McConville, M.; Schofield, L. Signal transduction in macrophages by glycosylphosphatidylinositols of Plasmodium, Trypanosoma, and Leishmania: activation of protein tyrosine kinases and protein kinase C by inositolglycan and diacylglycerol moieties. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 4022–4027.
- (3) Karageorgos, D. Neural GPI-anchored cell adhesion molecules. *Front. Biosci., Landmark Ed.* **2003**, *8*, s1304–20.
- (4) Brown, D. A.; Rose, J. K. Sorting of GPI-anchored proteins to glycolipid-enriched membrane subdomains during transport to the apical cell surface. *Cell* **1992**, *68*, 533–544.
- (5) Mayor, S.; Riezman, H. Sorting GPI-anchored proteins. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 110.
- (6) Taylor, D. R.; Hooper, N. M. *Post-translational modifications in health and disease*; Springer: New York, NY, USA, pp 39–55, DOI: 10.1007/978-1-4419-6382-6\_2.
- (7) Banerjee, P.; Wehle, M.; Lipowsky, R.; Santer, M. A molecular dynamics model for glycosylphosphatidylinositol anchors: flop down or lollipop? *Phys. Chem. Chem. Phys.* **2018**, *20*, 29314–29324.
- (8) Lehto, M. T.; Sharom, F. J. Proximity of the protein moiety of a GPI-anchored protein to the membrane surface: a FRET study. *Biochemistry* **2002**, *41*, 8368–8376.
- (9) Barboni, E.; Rivero, B. P.; George, A.; Martin, S.; Renoup, D.; Hounsell, E.; Barber, P.; Morris, R. The glycosylphosphatidylinositol anchor affects the conformation of Thy-1 protein. *J. Cell Sci.* **1995**, *108*, 487–497.
- (10) Paulick, M. G.; Wise, A. R.; Forstner, M. B.; Groves, J. T.; Bertozzi, C. R. Synthetic analogues of glycosylphosphatidylinositol-anchored proteins and their behavior in supported lipid bilayers. *J. Am. Chem. Soc.* **2007**, *129*, 11543–11550.
- (11) Paulick, M. G.; Forstner, M. B.; Groves, J. T.; Bertozzi, C. R. A chemical approach to unraveling the biological function of the glycosylphosphatidylinositol anchor. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 20332–20337.
- (12) Schroeder, R.; London, E.; Brown, D. Interactions between saturated acyl chains confer detergent resistance on lipids and glycosylphosphatidylinositol (GPI)-anchored proteins: GPI-anchored proteins in liposomes and cells show similar behavior. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 12130–12134.
- (13) Varma, R.; Mayor, S. GPI-anchored proteins are organized in submicron domains at the cell surface. *Nature* **1998**, *394*, 798.
- (14) Sevcsik, E.; Brameshuber, M.; Fölser, M.; Weghuber, J.; Honigsmann, A.; Schütz, G. J. GPI-anchored proteins do not reside in ordered domains in the live cell plasma membrane. *Nat. Commun.* **2015**, *6*, 6969.
- (15) Homans, S.; Edge, C.; Ferguson, M.; Dwek, R.; Rademacher, T. Solution structure of the glycosylphosphatidylinositol membrane anchor glycan of Trypanosoma brucei variant surface glycoprotein. *Biochemistry* **1989**, *28*, 2881–2887.
- (16) Chevalier, F.; Lopez-Prados, J.; Perez, S.; Martín-Lomas, M.; Nieto, P. M. Conformational study of GPI anchors: the common oligosaccharide GPI anchor backbone. *Eur. J. Org. Chem.* **2005**, *2005*, 3489–3498.
- (17) Zuegg, J.; Gready, J. E. Molecular dynamics simulation of human prion protein including both N-linked oligosaccharides and the GPI anchor. *Glycobiology* **2000**, *10*, 959–974.
- (18) Wehle, M.; Vilotijevic, I.; Lipowsky, R.; Seeberger, P. H.; Varon Silva, D.; Santer, M. Mechanical compressibility of the glycosylphosphatidylinositol (GPI) anchor backbone governed by independent glycosidic linkages. *J. Am. Chem. Soc.* **2012**, *134*, 18964–18972.
- (19) Periole, X.; Marrink, S.-J. *Biomolecular Simulations*; Humana Press: Totowa, NJ, USA, Vol. 924, pp 533–565, DOI: 10.1007/978-1-62703-017-5\_20.
- (20) Corradi, V.; Sejdiu, B. I.; Mesa-Galoso, H.; Abdizadeh, H.; Noskov, S. Y.; Marrink, S. J.; Tieleman, D. P. Emerging diversity in lipid–protein interactions. *Chem. Rev.* **2019**, *119*, 5775–5848.
- (21) Tsai, Y.-H.; Liu, X.; Seeberger, P. H. Chemical biology of glycosylphosphatidylinositol anchors. *Angew. Chem., Int. Ed.* **2012**, *51*, 11438–11456.
- (22) López, C. A.; Rzepiela, A. J.; De Vries, A. H.; Dijkhuizen, L.; Hunenberger, P. H.; Marrink, S. J. Martini coarse-grained force field: extension to carbohydrates. *J. Chem. Theory Comput.* **2009**, *5*, 3195–3210.
- (23) Yesylevskyy, S. O.; Schäfer, L. V.; Sengupta, D.; Marrink, S. J. Polarizable water model for the coarse-grained MARTINI force field. *PLoS Comput. Biol.* **2010**, *6*, No. e1000810.
- (24) Bennett, W. D.; Tieleman, D. P. Water defect and pore formation in atomistic and coarse-grained lipid membranes: pushing the limits of coarse graining. *J. Chem. Theory Comput.* **2011**, *7*, 2981–2988.
- (25) Kociurzynski, R.; Pannuzzo, M.; Bockmann, R. A. Phase transition of glycolipid membranes studied by coarse-grained simulations. *Langmuir* **2015**, *31*, 9379–9387.
- (26) Pluhackova, K.; Wassenaar, T. A.; Kirsch, S.; Bockmann, R. A. Spontaneous adsorption of coiled-coil model peptides K and E to a mixed lipid bilayer. *J. Phys. Chem. B* **2015**, *119*, 4396–4408.
- (27) Gu, R.-X.; Ingólfsson, H. I.; de Vries, A. H.; Marrink, S. J.; Tieleman, D. P. Ganglioside-lipid and ganglioside-protein interactions revealed by coarse-grained and atomistic molecular dynamics simulations. *J. Phys. Chem. B* **2017**, *121*, 3262–3275.
- (28) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; Van Gunsteren, W. F. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem. Phys. Lett.* **1994**, *222*, 529–539.
- (29) Mazzobre, M. F.; Román, M. V.; Mourelle, A. F.; Corti, H. R. Octanol–water partition coefficient of glucose, sucrose, and trehalose. *Carbohydr. Res.* **2005**, *340*, 1207–1211.
- (30) Alessandri, R.; Souza, P. C.; Thallmair, S.; Melo, M. N.; De Vries, A. H.; Marrink, S. J. Pitfalls of the Martini Model. *J. Chem. Theory Comput.* **2019**, *15*, 5448–5460.
- (31) Periole, X.; Cavalli, M.; Marrink, S.-J.; Ceruso, M. A. Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. *J. Chem. Theory Comput.* **2009**, *5*, 2531–2543.
- (32) Ormó, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J. Crystal structure of the Aequorea victoria green fluorescent protein. *Science* **1996**, *273*, 1392–1395.
- (33) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; De Vries, A. H. The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.
- (34) Moiset, G.; Lopez, C. A.; Bartelds, R.; Syga, L.; Rijpkema, E.; Cukkekane, A.; Baldus, M.; Poolman, B.; Marrink, S. J. Disaccharides impact the lateral organization of lipid membranes. *J. Am. Chem. Soc.* **2014**, *136*, 16167–16175.
- (35) Schmalhorst, P. S.; Deluweit, F.; Scherrers, R.; Heisenberg, C.-P.; Sikora, M. Overcoming the limitations of the MARTINI force field in simulations of polysaccharides. *J. Chem. Theory Comput.* **2017**, *13*, 5039–5053.
- (36) Stark, A. C.; Andrews, C. T.; Elcock, A. H. Toward optimized potential functions for protein–protein interactions in aqueous solutions: osmotic second virial coefficient calculations using the martini coarse-grained force field. *J. Chem. Theory Comput.* **2013**, *9*, 4176–4185.
- (37) Javanainen, M.; Martinez-Seara, H.; Vattulainen, I. Excessive aggregation of membrane proteins in the Martini model. *PLoS One* **2017**, *12*, No. e0187936.
- (38) López, C. A.; Sovova, Z.; van Eerden, F. J.; de Vries, A. H.; Marrink, S. J. Martini force field parameters for glycolipids. *J. Chem. Theory Comput.* **2013**, *9*, 1694–1708.
- (39) Petrov, D.; Zagrovic, B. Are current atomistic force fields accurate enough to study proteins in crowded environments? *PLoS Comput. Biol.* **2014**, *10*, No. e1003638.

- (40) Best, R. B.; Zheng, W.; Mittal, J. Balanced protein–water interactions improve properties of disordered proteins and non-specific protein association. *J. Chem. Theory Comput.* **2014**, *10*, 5113–5124.
- (41) Karunaweera, S.; Gee, M. B.; Weerasinghe, S.; Smith, P. E. Theory and simulation of multicomponent osmotic systems. *J. Chem. Theory Comput.* **2012**, *8*, 3493–3503.
- (42) Ploetz, E. A.; Benteñitis, N.; Smith, P. E. Developing force fields from the microscopic structure of solutions. *Fluid Phase Equilib.* **2010**, *290*, 43–47.
- (43) Ploetz, E. A.; Smith, P. E. A Kirkwood–Buff force field for the aromatic amino acids. *Phys. Chem. Chem. Phys.* **2011**, *13*, 18154–18167.
- (44) Luo, Y.; Roux, B. Simulation of osmotic pressure in concentrated aqueous salt solutions. *J. Phys. Chem. Lett.* **2010**, *1*, 183–189.
- (45) Sauter, J.; Grafmuller, A. Efficient Osmotic Pressure Calculations Using Coarse-Grained Molecular Simulations. *J. Chem. Theory Comput.* **2018**, *14*, 1171–1176.
- (46) Miller, M. S.; Lay, W. K.; Li, S.; Hacker, W. C.; An, J.; Ren, J.; Elcock, A. H. Reparametrization of protein force field nonbonded interactions guided by osmotic coefficient measurements from molecular dynamics simulations. *J. Chem. Theory Comput.* **2017**, *13*, 1812–1826.
- (47) Stigter, D. Interactions in aqueous solutions. II. Osmotic pressure and osmotic coefficient of sucrose and glucose solutions. *J. Phys. Chem.* **1960**, *64*, 118.
- (48) Davis, D. J.; Burlak, C.; Nicholas, P. Osmotic pressure of fungal compatible osmolytes. *Mycol. Res.* **2000**, *104*, 800–804.
- (49) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*, 19–25.
- (50) 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.
- (51) 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.
- (52) Dickson, C. J.; Madej, B. D.; Skjervik, Å. A.; Betz, R. M.; Teigen, K.; Gould, I. R.; Walker, R. C. Lipid14: the amber lipid force field. *J. Chem. Theory Comput.* **2014**, *10*, 865–879.
- (53) 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.
- (54) Sorin, E. J.; Pande, V. S. Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. *Biophys. J.* **2005**, *88*, 2472–2493.
- (55) Marrink, S. J.; De Vries, A. H.; Mark, A. E. Coarse grained model for semiquantitative lipid simulations. *J. Phys. Chem. B* **2004**, *108*, 750–760.
- (56) Wassenaar, T. A.; Ingólfsson, H. I.; Bockmann, R. A.; Tieleman, D. P.; Marrink, S. J. Computational lipidomics with insane: a versatile tool for generating custom membranes for molecular simulations. *J. Chem. Theory Comput.* **2015**, *11*, 2144–2155.
- (57) Verlet, L. Computer “experiments” on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.* **1967**, *159*, 98.
- (58) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- (59) Goga, N.; Rzepiela, A.; De Vries, A.; Marrink, S.; Berendsen, H. Efficient algorithms for Langevin and DPD dynamics. *J. Chem. Theory Comput.* **2012**, *8*, 3637–3649.
- (60) Berendsen, H. J.; Postma, J. v.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (61) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- (62) Kim, H.; Afsari, H. S.; Cho, W. High-throughput fluorescence assay for membrane-protein interaction. *J. Lipid Res.* **2013**, *54*, 3531–3538.