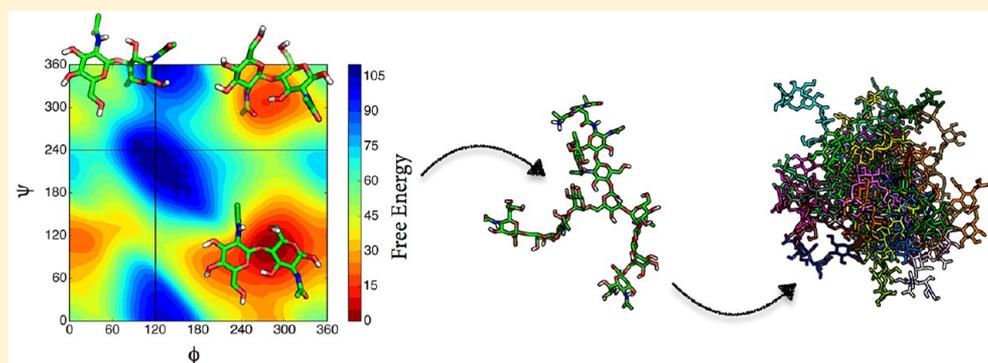


# Modeling of Oligosaccharides within Glycoproteins from Free-Energy Landscapes

Aysegül Turupcu and Chris Oostenbrink\*<sup>1</sup>

Institute of Molecular Modeling and Simulation, University of Natural Resources and Life Sciences, Muthgasse 18, 1190 Vienna, Austria

## Supporting Information



**ABSTRACT:** In spite of the abundance of glycoproteins in biological processes, relatively little three-dimensional structural data is available for glycan structures. Here, we study the structure and flexibility of the vast majority of mammalian oligosaccharides appearing in N- and O-glycosylated proteins using a bottom up approach. We report the conformational free-energy landscapes of all relevant glycosidic linkages as obtained from local elevation simulations and subsequent umbrella sampling. To the best of our knowledge, this represents the first complete conformational library for the construction of N- and O-glycan structures. Next, we systematically study the effect of neighboring residues, by extensively simulating all relevant trisaccharides and one tetrasaccharide. This allows for an unprecedented comparison of disaccharide linkages in large oligosaccharides. With a small number of exceptions, the conformational preferences in the larger structures are very similar as in the disaccharides. This, finally, allows us to suggest several efficient approaches to construct complete N- and O-glycans on glycoproteins, as exemplified on two relevant examples.

## INTRODUCTION

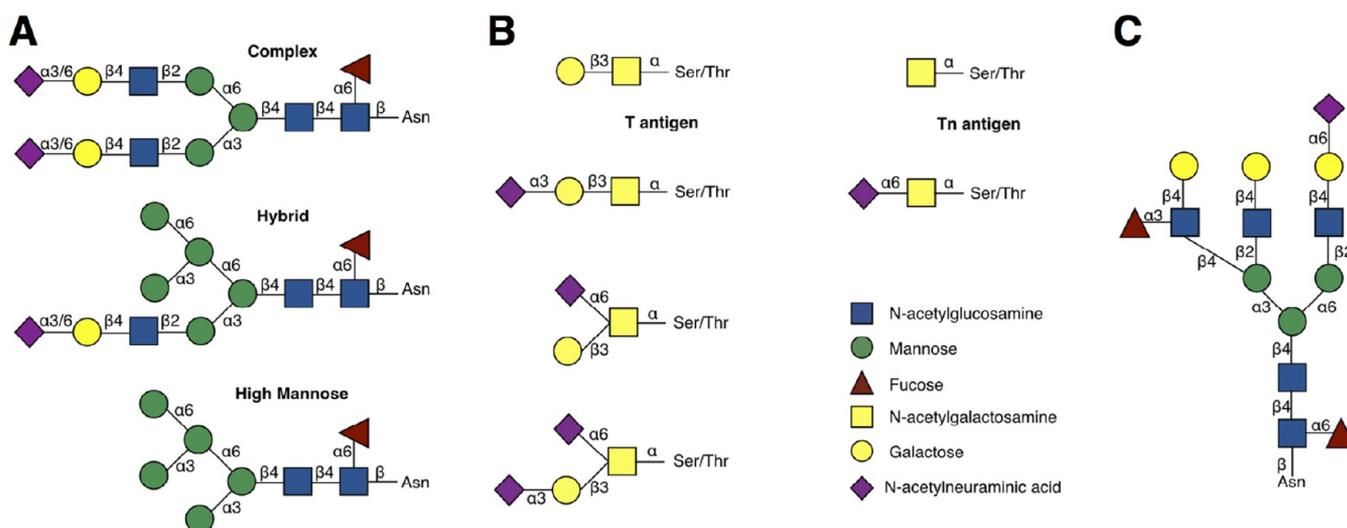
Most of the processes in living cells take place via some form of carbohydrate interaction.<sup>1</sup> Therefore, understanding the effect of glycosylation, which occurs on more than 50% of the eukaryotic proteins, is crucial to reveal biological mysteries. Glycans, the carbohydrate moieties of glycoproteins, serve as biological markers at various stages of cellular differentiation and proliferation and also play significant roles in cell–cell recognition, cell adhesion, and signal transduction. This broad range of functionality is attributed to their structural diversity in terms of size, sequence, branching, and linkage types that can be formed from a variety of possible monomeric units. However, there is a lack of reliable information on the structure of glycoproteins for several reasons. Most of the experimentally solved biomolecular structures have undergone extensive manipulation of oligosaccharides before X-ray crystallography or NMR spectroscopy due to their inherent flexibility and high degree of coordination with water. The available X-ray structures either have different types of glycoforms than the physiologically relevant forms or show only short sequences of glycan units. Also, structures from X-

ray crystallography offer the crystalline form of glycoproteins, lacking a more diverse solution representation. NMR can offer structures in solution; however, these represent averages of simultaneously occurring conformers, limiting the amount of structural information. Furthermore, glycan structures in the databases often contain wrong or missing information about stereochemistry and nomenclature.<sup>2</sup> Due to its detailed spatial and temporal resolution, molecular dynamics simulation emerges as a powerful tool for the modeling of glycoproteins. Some progress has been made in search of the conformational preferences of glycan units with different carbohydrate force-fields; including AMBER,<sup>3</sup> GLYCAM,<sup>4</sup> CHARMM,<sup>5</sup> and GROMOS.<sup>6,7</sup> In the current work, we will focus on the conformational preferences of carbohydrate structures that are relevant for glycoproteins.

It has become increasingly accepted that the tertiary structure of a protein around a glycosylation site plays a more crucial role than the primary structure in determining the occurrence and

Received: June 10, 2017

Published: August 17, 2017



**Figure 1.** (A) Major classes of N-glycan. (B) Some core types of O-glycan. (C) Typical complex type of N-glycan in mature glycoproteins.

**Table 1. Overview of the Studied Dimers, along with the Relative Free Energies of the Most Relevant Conformations<sup>a</sup>**

	$G_A$ [kJ/mol]	$G_B$ [kJ/mol]	$G_C$ [kJ/mol]	$G_D$ [kJ/mol]
$\alpha$ -D-Neup5Ac-(2 $\rightarrow$ 6)- $\beta$ -D-Galp	0.0	4.4	3.2	
$\alpha$ -D-Neup5Ac-(2 $\rightarrow$ 6)- $\beta$ -D-GalpNAc	0.0	14.9	4.0	
$\alpha$ -D-Neup5Ac-(2 $\rightarrow$ 3)- $\beta$ -D-Galp	14.6	0.0	36.2	8.6
$\alpha$ -D-Neup5Ac-(2 $\rightarrow$ 8)- $\alpha$ -D-Neup5Ac	0.0	24.0	9.6	
$\beta$ -D-Galp-(1 $\rightarrow$ 4)- $\beta$ -D-GlcpNAc	18.9	42.2	0.0	12.6
$\beta$ -D-GlcpNAc-(1 $\rightarrow$ 2)- $\alpha$ -D-Manp	61.7	21.9	39.2	0.0
$\alpha$ -D-Manp-(1 $\rightarrow$ 6)- $\beta$ -D-Manp	0.0	39.0		
$\alpha$ -D-Manp-(1 $\rightarrow$ 3)- $\beta$ -D-Manp	23.9	0.0	70.4	42.7
$\beta$ -D-Manp-(1 $\rightarrow$ 4)- $\beta$ -D-GlcpNAc	13.4	39.8	0.0	12.2
$\beta$ -D-GlcpNAc-(1 $\rightarrow$ 4)- $\beta$ -D-GlcpNAc	21.3	42.0	0.0	14.0
$\alpha$ -L-Fucp-(1 $\rightarrow$ 6)- $\beta$ -D-GlcpNAc	45.3	0.0		
$\alpha$ -L-Fucp-(1 $\rightarrow$ 3)- $\beta$ -D-GlcpNAc	72.8	39.3	12.9	0.0
$\alpha$ -D-Manp-(1 $\rightarrow$ 2)- $\alpha$ -D-Manp	31.4	0.0	81.6	44.4
$\beta$ -D-Galp-(1 $\rightarrow$ 3)- $\alpha$ -D-GalpNAc	43.5	14.4	25.6	0.0
$\beta$ -D-GlcpNAc $\rightarrow$ ASN	16.6	0.0		
$\alpha$ -D-GalpNAc $\rightarrow$ SER	0.0	44.2		
$\alpha$ -D-GalpNAc $\rightarrow$ THR	0.0	19.5	45.4	86.0

<sup>a</sup>Conformations are labeled with letters A to D as indicated in Figure 3. Conformational free energies are computed by integrating the free energy landscapes over the appropriate regions. The most favored conformation is set to 0.0 kJ mol<sup>-1</sup>.

level of glycosylation.<sup>8</sup> Even though glycosylation is a cotranslational process, processing of glycan units and their variation also depends on the accessibility of the glycosylation site and the glycans to the related processing enzymes.<sup>9</sup> Therefore, modeling of the glycoproteins not only contributes a three-dimensional visualization of the glycosylated proteins but also offers explanations for variations in the glycan units as characterized by experimental techniques.<sup>10</sup>

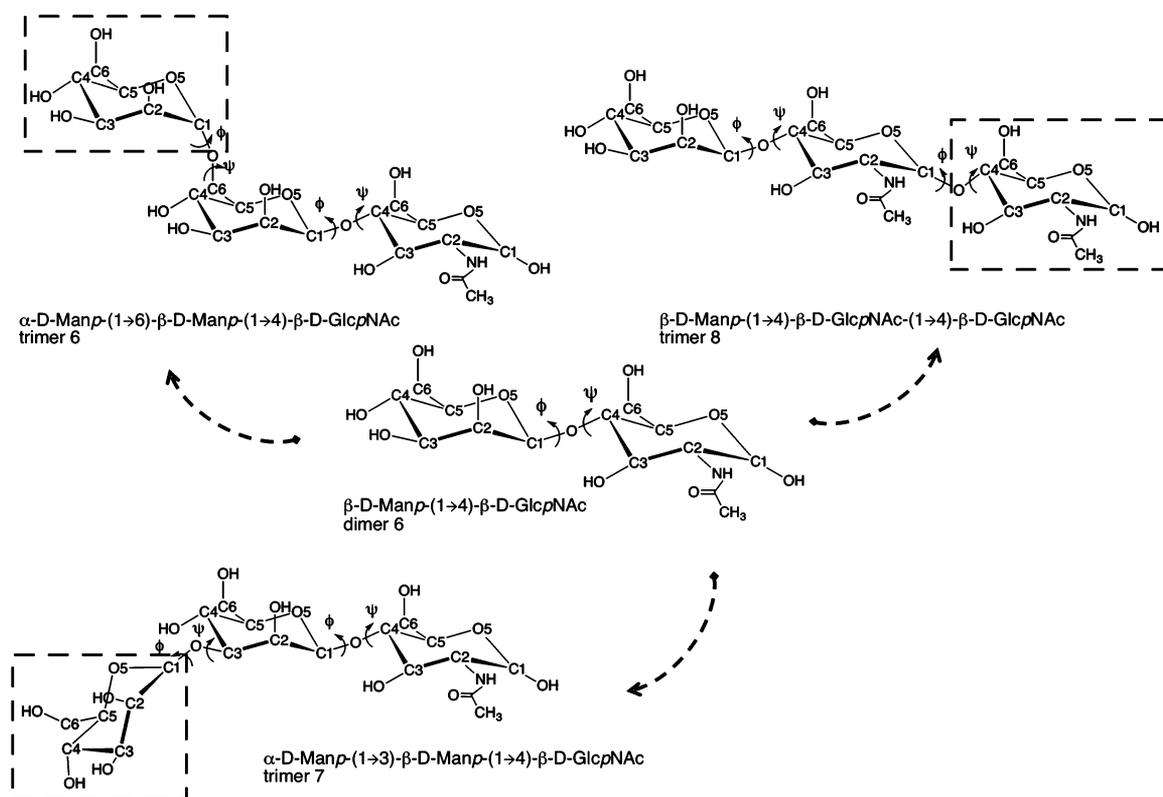
Depending on the variations of the 1,3- or 1,6-branches, N-linked oligosaccharides are classified into three groups: high mannose, hybrid, and complex types (see Figure 1). All types start with two N-acetylglucosamine ( $\beta$ -D-GlcpNAc) units that are linked to an asparagine side chain. The high mannose type subsequently consists of only  $\alpha$ -D-Manp residues. In the complex type there are different types of monosaccharides present (for example  $\beta$ -D-GlcpNAc,  $\beta$ -D-Galp,  $\alpha$ -D-Neup5Ac,  $\beta$ -D-GalpNAc, and  $\alpha$ -L-Fucp) occurring in multiple levels of

branching or antennas. In hybrid types, the 1,6-branch may be typically high mannose while the 1,3-branch is complex.

O-Glycosidic linkages are formed between different monosaccharides and amino acid residues. The most common glycoproteins carrying O-glycans are mucins and in some cases they carry only a GalNAc residue linked to serine or threonine residues, while in others the O-glycan may be a disaccharide ( $\alpha$ -D-Neup5Ac-(2  $\rightarrow$  6)- $\beta$ -D-GalpNAc;  $\beta$ -D-Galp-(1  $\rightarrow$  3)- $\alpha$ -D-GalpNAc) or an oligosaccharide consisting of  $\beta$ -D-Galp or  $\beta$ -D-GlcpNAc, elongated from the core units.

We will use a bottom up approach, investigating the conformational preferences of all relevant disaccharides, then moving on to trisaccharides and larger structures. In particular, we will focus on the conformational preferences of the glycosidic linkages and the influence of additional substituents on the central units.

We start by analyzing 17 disaccharide or amino acid-monomer units that are fragments of the vast majority of



**Figure 2.** Representation of the  $\beta$ -D-Manp-(1  $\rightarrow$  4)- $\beta$ -D-GlcpNAc dimer and demonstration of the nomenclature (center). All the disaccharide fragments from complex type of glycan are simulated with their connecting third monosaccharide; trimer6 represents an additional  $\alpha$ -D-Manp-(1  $\rightarrow$  6) link, trimer7 a  $\alpha$ -D-Manp-(1  $\rightarrow$  3) link, and trimer8 a  $\beta$ -D-GlcpNAc link. See Table 2 for the rest of the set.

mammalian glycan trees belonging to mature glycoproteins, involving N- and O-glycans. Furthermore, for the first time, we systematically study the effect of neighboring residues on the conformational preferences of the disaccharides by extensively sampling all relevant trisaccharides. The observed conformational preferences of the disaccharides can be subsequently used to build the common major classes of N- and O-glycans, including typical modifications (like core fucosylation). With the selected disaccharide units, typical N-glycans found in mature glycoproteins can be created. Additionally, for O-glycans, the Tn and T antigen, blood group related antigens Lewis x, and their Sialyl versions can be build. The full list of the studied fragments is presented in Table 1.

The most relevant conformations of the dimers are characterized through the glycosidic  $\phi$  and  $\psi$  angles. For a (1  $\rightarrow$  X) linkage, where X is 2, 3, 4, or 6 for a reducing disaccharide, the glycosidic dihedral angles,  $\phi$  and  $\psi$ , are defined as O5-C1-OX-CX' and C1-OX-CX'-C(X-1)', respectively. For  $\alpha$ -D-Neup5Ac-(2  $\rightarrow$  X) linkage,  $\phi$  and  $\psi$ , are defined as O6-C2-OX-CX' and C2-OX-CX'-C(X-1)', respectively. See Figure 2 for the numbering in the  $\beta$ -D-Manp-(1  $\rightarrow$  4)- $\beta$ -D-GlcpNAc dimer as an example. This figure also shows the subsequent steps of our analysis. To investigate the effect of neighboring residues on the conformational preferences of the dimer, all relevant trimers are studied as well, in the case of the dimer in Figure 2; three different dimers are studied.

Previously, similar simulations investigating the conformational preferences of disaccharides as well as N-glycans have been described. Fernandes et al.<sup>11</sup> simulated some of the related linkages using the GROMOS 43a1 force field with plain MD.

Initial structures were taken from a systematic search in vacuum and used in 100 ns explicit solvent simulation. Perić-Hassler et al.<sup>12</sup> investigated glucose-based disaccharides using the same enhanced sampling method as the current work. Yang et al.<sup>13</sup> applied a protocol based on replica exchange to sample the conformations of N-glycans. The same technique was used to extensively sample complex glycans<sup>14</sup> and high mannose oligosaccharides.<sup>15</sup> Even though these studies provide understanding about conformational preferences of several disaccharides, either the studied systems and techniques are different or the sampling of all relevant conformations is insufficient. To the best of our knowledge, the current work represents a first description of a consistent conformational library for the vast majority of relevant linkages in mammalian glycans.

The organization of this paper is as follows. In the methods section, the theory of the enhanced sampling method used is presented, followed by the computational details for its application. Then, the procedure for the analysis of the simulations is introduced. In the Results and Discussion section, a comparison of the conformational preferences as obtained by the simulations with the available experimental data and literature are reported. Furthermore, the effect of neighboring residues on these preferences are systematically studied. We subsequently propose protocols for the construction of the higher oligosaccharides and describe how our approach may be used in future work. Finally, the main findings are summarized in the Conclusion.

## METHODS

This work aims to suggest protocols to build carbohydrate moieties of glycoproteins in the absence of previous structural information on carbohydrates. As outlined above, we start with disaccharides as they carry all the rotational degrees of freedom of the glycosidic linkage. The conformational characteristics of a disaccharide can be represented by a plot of  $\phi$  against  $\psi$ , in analogy to the Ramachandran plot for amino acids. As the relative orientations of the successive monosaccharides differ in each linkage type, the  $\phi$ ,  $\psi$  map of a disaccharide is distinct. Therefore, an analysis of all possible conformations of a disaccharide is a key for creating higher oligosaccharides. Molecular dynamics (MD) simulations are capable of sampling broad ranges of conformations and are therefore the method of choice. However, it turns out that relevant, distinct conformations, represented by different regions on the  $\phi$ ,  $\psi$  map are often separated by high barriers, hampering the computational efficiency of the method. Similarly, some conformations may be thermodynamically accessible, but only observed very rarely in plain MD simulations, even if they are experimentally observed. Therefore, we apply an enhanced sampling method to ensure a near-to-complete sampling of the  $\phi$ ,  $\psi$  maps, local elevation umbrella sampling (LEUS).<sup>16,17</sup> Application of this method offers enhanced coverage of those regions in the  $\phi$ ,  $\psi$  maps which are sterically allowed. It relies on first building up a memory-based biasing potential in the desired subspace and a subsequent umbrella sampling phase to sample the conformational space using the “frozen” biased potentials. We used the adaptable nature of the LEUS method by creating disaccharide biasing potential libraries in the glycosidic torsional angle subspace ( $\phi$  and  $\psi$ ) and construct the larger carbohydrate moiety of glycoproteins using the relevant biasing potentials from the glycosidic linkages of all disaccharide fragments. A similar kind of approach has been proposed for proteins with peptide  $\phi$  and  $\psi$  dihedral angles.<sup>18,19</sup>

**MD Simulation Settings.** All MD simulations were performed using the GROMOS11 biomolecular simulation package (<http://www.gromos.net>)<sup>20</sup> with the 53A6glyc parameter set<sup>7,21</sup> of the GROMOS force field for carbohydrates. Minor modifications of this parameter set were applied to ensure compatibility with the protein force field and to ensure the stereochemistry of sialic acid. These modifications are listed in Table S2 of the Supporting Information. In particular, we have updated the N-Acetyl linkage in GlcNAc and Neu5Ac to the current protein force field and adjusted the GlcNAc-Asn linkage accordingly.

Initial structures of the disaccharide units of complex glycans were modeled in the molecular operating environment (MOE; Chemical Computing Group, Inc. 2014) from the crystal structure of IgG Fc carrying a sialylated complex type of glycan (PDB ID 4BYH<sup>22</sup>). The rest of the initial structures of units were modeled using the glycosidic dihedral angle preferences according to the Glycosciences portal.<sup>23</sup> Hydrogen atoms were added according to geometric criteria followed by a short energy minimization using the steepest-descent algorithm. The compounds were placed in a periodic cubic water box with simple point charge (SPC) water molecules<sup>24</sup> and initialized with a 1.4 nm minimum distance of the solute to the box walls. The system was further relaxed by a steepest descent minimization with position restraints on the solute atoms. Prior to the production simulations, the systems were equilibrated with initial random velocities generated from a

Maxwell–Boltzmann distribution at 60 K. The systems were heated up to 300 K in five discrete steps with a simulation length of 20 ps, while simultaneously, position restraints on the solute atoms were reduced from  $2.5 \times 10^4$  to  $0.0 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ .

The simulations were performed at a constant temperature of 300 K and a constant pressure of 1 atm using a weak coupling scheme for both temperature and pressure with coupling times  $\tau_T = 0.1 \text{ ps}$  and  $\tau_P = 0.5 \text{ ps}$  and an isothermal compressibility of  $4.575 \times 10^{-4} \text{ kJ}^{-1} \text{ mol nm}^3$ . Newton's equations of motion were integrated using the leapfrog scheme with a time step of 2 fs. The SHAKE algorithm<sup>25</sup> was used to maintain the bond distances at their desired values. Long-range electrostatic interactions beyond a cutoff of 1.4 nm were truncated and approximated by a generalized reaction field with a dielectric permittivity of 61.<sup>26</sup> Nonbonded interactions up to a distance of 0.8 nm were computed at every time step using a pairlist that was updated every 5 steps. Interactions up to 1.4 nm were computed at pairlist updates and kept constant in between.

**Creating Biased Potential Libraries with LE and Sampling with US.** The local elevation (LE) biasing potential is defined as the sum of grid-based functions along the LE coordinates  $Q$ , that is (eq 1)

$$\mathcal{U}_{\text{LE}}(Q; \{n_k\}) = c \sum_{k=1, n_k \neq 0}^{N_g^{N_l}} n_k \prod_{i=1}^{N_l} g(\text{MI}[Q_i - Q_{k,i}^*]; \sigma) \quad (1)$$

In the original implementation  $g$  is defined as a one-dimensional truncated Gaussian function;<sup>17</sup> in this study, a polynomial type is used instead (eq 2), following its advantages discussed in ref 16.

$$g(\Delta Q; \sigma) = H\left(1 - \frac{|\Delta Q|}{\sigma}\right) \left[ 1 - 3\left(\frac{|\Delta Q|}{\sigma}\right)^2 + 2\left(\frac{|\Delta Q|}{\sigma}\right)^3 \right] \quad (2)$$

$N_g$  is the number of grid points along the  $N_l$  dimensional subspace and  $n_k$  is the number of visits to the corresponding grid cell. MI is the minimum image function ensuring the choice of  $Q_i$  when it is periodic.  $H$  is the Heaviside step function. The force constant,  $c$ , and width,  $\sigma$ , are the free parameters which need to be tuned for the building up procedure.

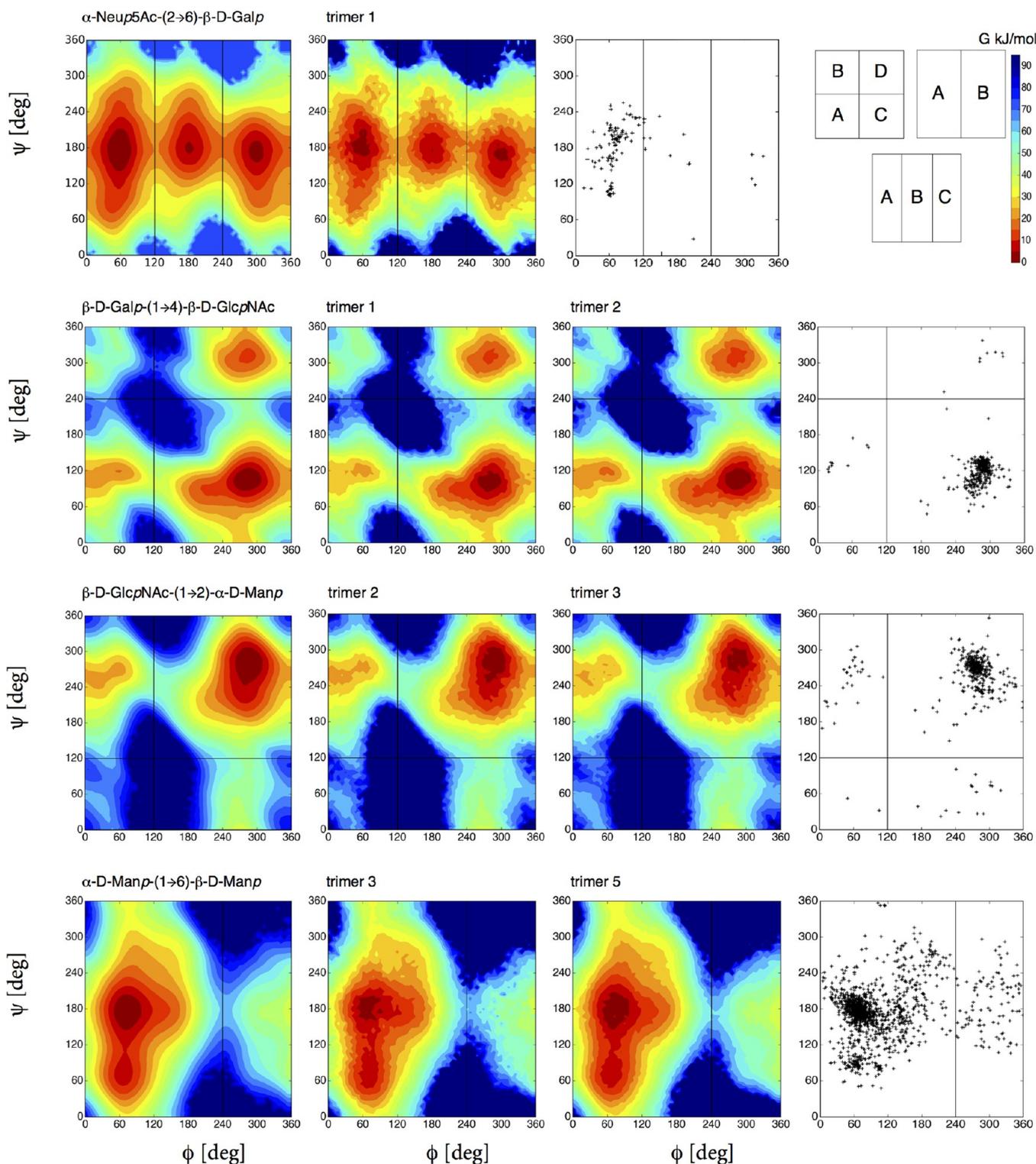
After the LE phase comes an equilibrium umbrella sampling (US) phase during which the biased Hamiltonian is time-independent. That is the values of  $n_k$  are no longer increased

$$\mathcal{H}_B(p; q) = \mathcal{H}(p; q) + \mathcal{U}_{\text{LEUS}}(Q) \quad (3)$$

The real (unbiased) probability to find the system at a value  $Q$  can be recovered from the LEUS (biased) simulations by reweighting:

$$P_{\text{UB}}(Q^\circ) = \frac{\langle \delta(Q - Q^\circ) \exp[\beta \mathcal{U}_{\text{LEUS}}(Q)] \rangle_B}{\langle \exp[\beta \mathcal{U}_{\text{LEUS}}(Q)] \rangle_B} \quad (4)$$

where  $\delta$  is the Dirac delta function, and  $\beta$  is  $1/k_B T$  with  $k_B$  as the Boltzmann constant and  $T$ , the temperature. The angular brackets indicate an ensemble average over the biased US simulation. The corresponding free-energy profile can be obtained as



**Figure 3.** Free-energy maps  $G(\phi, \psi)$  in the glycosidic dihedral angle subspace from unbiased LEUS simulations compared with available crystal and NMR data of the related linkages taken from [glycanstructure.org](http://glycanstructure.org).<sup>27</sup> The first column represents the disaccharide simulation, and the second and/or third column corresponds to the disaccharide within a trisaccharide simulation (for trimer codes see Table 2). The division into distinct conformational states (A, B, C, and D) is illustrated at right top panel with colorbar used for free-energy maps with 5 kJ mol<sup>-1</sup> spacing starting from global minimum. Regions that were never visited are shown in dark blue. For the rest of the free-energy maps, see Figures 4, 6, and 7 as well as Figure S1 in the Supporting Information.

$$G(Q) = -\beta^{-1} \ln P_{UB}(Q) \quad (5)$$

In this study, the LEUS subspace is defined by the glycosidic dihedral angles  $\phi$  and  $\psi$ . After testing a number of different

possible combinations, parameters for the LEUS method are chosen as  $N_g = 36$ ,  $\sigma = 360^\circ/N_g$ ,  $c = 0.005$  kJ mol<sup>-1</sup>,  $N_l = 2$  for the  $\phi$  and  $\psi$  angle of a glycosidic linkage. The LE phase is run for  $t_{LE} = 100$  ns and the US phase for  $t_{US} = 100$  ns. During the

Table 2. Relative Free Energy Values of Conformational States from Trisaccharide and Disaccharide LEUS Simulations after Unbiasing<sup>a</sup>

		$G_A$ [kJ/mol]	$G_B$ [kJ/mol]	$G_C$ [kJ/mol]	$G_D$ [kJ/mol]
$\alpha$ -D-Neup5Ac-(2 → 6)- $\beta$ -D-Galp-(1 → 4)- $\beta$ -D-GlcpNAc	trimer1	0.0	3.7	2.1	
$\alpha$ -D-Neup5Ac-(2 → 6)- $\beta$ -D-Galp	dimer1	0.0	4.4	3.2	
$\alpha$ -D-Neup5Ac-(2 → 6)- $\beta$ -D-Galp-(1 → 4)- $\beta$ -D-GlcpNAc	trimer1	20.6	41.0	0.0	11.7
$\beta$ -D-Galp-(1 → 4)- $\beta$ -D-GlcpNAc(1 → 2)- $\alpha$ -D-Manp	trimer2	17.7	42.0	0.0	11.6
$\beta$ -D-Galp-(1 → 4)- $\beta$ -D-GlcpNAc	dimer2	18.9	42.2	0.0	12.6
$\beta$ -D-Galp-(1 → 4)- $\beta$ -D-GlcpNAc-(1 → 2)- $\alpha$ -D-Manp	trimer2	60.1	20.9	38.9	0.0
$\beta$ -D-GlcpNAc(1 → 2)- $\alpha$ -D-Manp-(1 → 6)- $\beta$ -D-Manp	trimer3	58.7	20.7	36.7	0.0
$\beta$ -D-GlcpNAc-(1 → 2)- $\alpha$ -D-Manp-(1 → 3)- $\beta$ -D-Manp	trimer4	60.5	20.8	37.9	0.0
$\beta$ -D-GlcpNAc-(1 → 2)- $\alpha$ -D-Manp	dimer3	61.7	21.9	39.2	0.0
$\beta$ -D-GlcpNAc-(1 → 2)- $\alpha$ -D-Manp-(1 → 6)- $\beta$ -D-Manp	trimer3	0.0	38.8		
$\alpha$ -D-Manp-(1 → 3)-[ $\alpha$ -D-Manp-(1 → 6)]- $\beta$ -D-Manp	trimer5	0.0	37.7		
$\alpha$ -D-Manp-(1 → 6)- $\beta$ -D-Manp-(1 → 4)- $\beta$ -D-GlcpNAc	trimer6	0.0	42.5		
$\alpha$ -D-Manp-(1 → 6)- $\beta$ -D-Manp	dimer4	0.0	39.0		
$\beta$ -D-GlcpNAc-(1 → 2)- $\alpha$ -D-Manp-(1 → 3)- $\beta$ -D-Manp	trimer4	27.9	0.0	72.4	48.2
$\alpha$ -D-Manp-(1 → 3)-[ $\alpha$ -D-Manp-(1 → 6)]- $\beta$ -D-Manp	trimer5	24.0	0.0	71.2	43.5
$\alpha$ -D-Manp-(1 → 3)- $\beta$ -D-Manp-(1 → 4)- $\beta$ -D-GlcpNAc	trimer7	23.2	0.0	70.0	44.6
$\alpha$ -D-Manp-(1 → 3)- $\beta$ -D-Manp	dimer5	23.9	0.0	70.4	42.7
$\alpha$ -D-Manp-(1 → 3)- $\beta$ -D-Manp-(1 → 4)- $\beta$ -D-GlcpNAc	trimer7	13.2	41.1	0.0	10.2
$\alpha$ -D-Manp-(1 → 6)- $\beta$ -D-Manp-(1 → 4)- $\beta$ -D-GlcpNAc	trimer6	15.4	42.1	0.0	10.8
$\beta$ -D-Manp-(1 → 4)- $\beta$ -D-GlcpNAc-(1 → 4)- $\beta$ -D-GlcpNAc	trimer8	14.5	40.2	0.0	11.8
$\beta$ -D-Manp-(1 → 4)- $\beta$ -D-GlcpNAc-(1 → 4)- $\beta$ -D-GlcpNAc-ASN	tetramer	16.5	39.5	0.0	12.4
$\beta$ -D-Manp-(1 → 4)- $\beta$ -D-GlcpNAc	dimer6	13.4	39.8	0.0	12.2
$\beta$ -D-Manp-(1 → 4)- $\beta$ -D-GlcpNAc-(1 → 4)- $\beta$ -D-GlcpNAc	trimer8	21.8	39.6	0.0	13.4
$\beta$ -D-GlcpNAc-(1 → 4)-[ $\alpha$ -L-Fucp-(1 → 6)]- $\beta$ -D-GlcpNAc	trimer9	21.8	42.0	0.0	14.2
$\beta$ -D-GlcpNAc-(1 → 4)-[ $\alpha$ -L-Fucp-(1 → 3)]- $\beta$ -D-GlcpNAc	trimer10	27.1	38.9	0.0	18.8
$\beta$ -D-Manp-(1 → 4)- $\beta$ -D-GlcpNAc-(1 → 4)- $\beta$ -D-GlcpNAc-ASN	tetramer	23.2	44.8	0.0	15.0
$\beta$ -D-GlcpNAc-(1 → 4)- $\beta$ -D-GlcpNAc	dimer7	21.3	42.0	0.0	14.0
$\beta$ -D-GlcpNAc-(1 → 4)-[ $\alpha$ -L-Fucp-(1 → 6)]- $\beta$ -D-GlcpNAc	trimer9	45.4	0.0		
$\alpha$ -L-Fucp-(1 → 6)- $\beta$ -D-GlcpNAc	dimer8	45.3	0.0		
$\beta$ -D-GlcpNAc-(1 → 4)-[ $\alpha$ -L-Fucp-(1 → 3)]- $\beta$ -D-GlcpNAc	trimer10	69.8	38.4	20.1	0.0
$\alpha$ -L-Fucp-(1 → 3)- $\beta$ -D-GlcpNAc	dimer9	72.8	39.3	12.9	0.0

<sup>a</sup>Conformational free energies are computed by integrating the free energy landscapes over the appropriate regions (indicated in Figure 3).

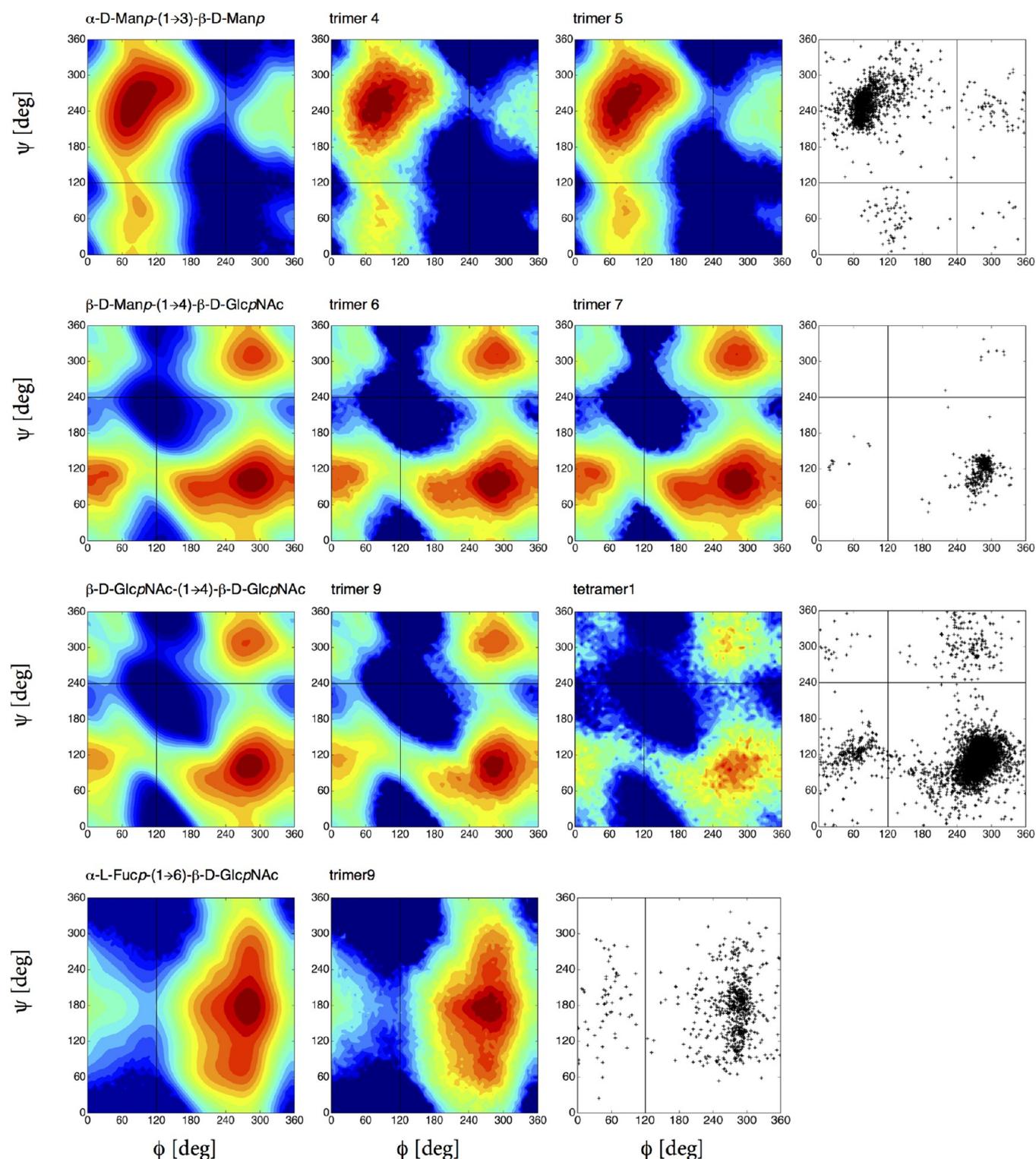
US phase, the trajectory is saved for every 0.1 ps resulting in 10<sup>6</sup> frames for 100 ns to meet the statistical efficiency as discussed in ref 12.

For each of the 17 disaccharides, plain MD simulations were carried out for 50 ns after equilibration. Glycosidic linkage libraries were created by LE and free-energy maps were constructed after US. To assess the effect of neighboring linkages on consecutive units, we used disaccharide frozen potentials and applied umbrella sampling on trisaccharides. Thereby, proving our concept that disaccharide motion libraries based on glycosidic linkages can be taken as a model for oligosaccharides. For each of the 10 trisaccharides simulated, the initial structures were modeled based on the free-energy maps of the disaccharides by setting their relevant glycosidic linkage angles to the global minimum value in the disaccharide. After equilibration, US sampling was applied on two glycosidic linkages with four  $\phi$  and  $\psi$  dihedral angles of the trimer unit by

using the respective frozen LE potentials evaluated from the corresponding disaccharides.

**Analysis.** Free-energy maps  $G(\phi, \psi)$  were created from the LEUS simulations after reweighting of the biased energy following eqs 4 and 5. The global minimum of each map was set to  $G = 0$  kJ mol<sup>-1</sup> and the value for grid points that were never visited set to  $G_{\max}$ .

The free-energy maps were divided into distinct conformational regions (A, B, C, up to D) with different  $\phi, \psi$  cutoff values for different disaccharides as illustrated in Figure 3. This scheme has previously been used in refs 6 and 12 for glucose-based disaccharides. After the unbiasing procedure, the free energies of these states ( $G_A, G_B, G_C$ , and  $G_D$ ) are calculated by integration over regions enclosed by the cutoff values. Subsequently, the values were rescaled by setting the state having the minimum free energy to 0.

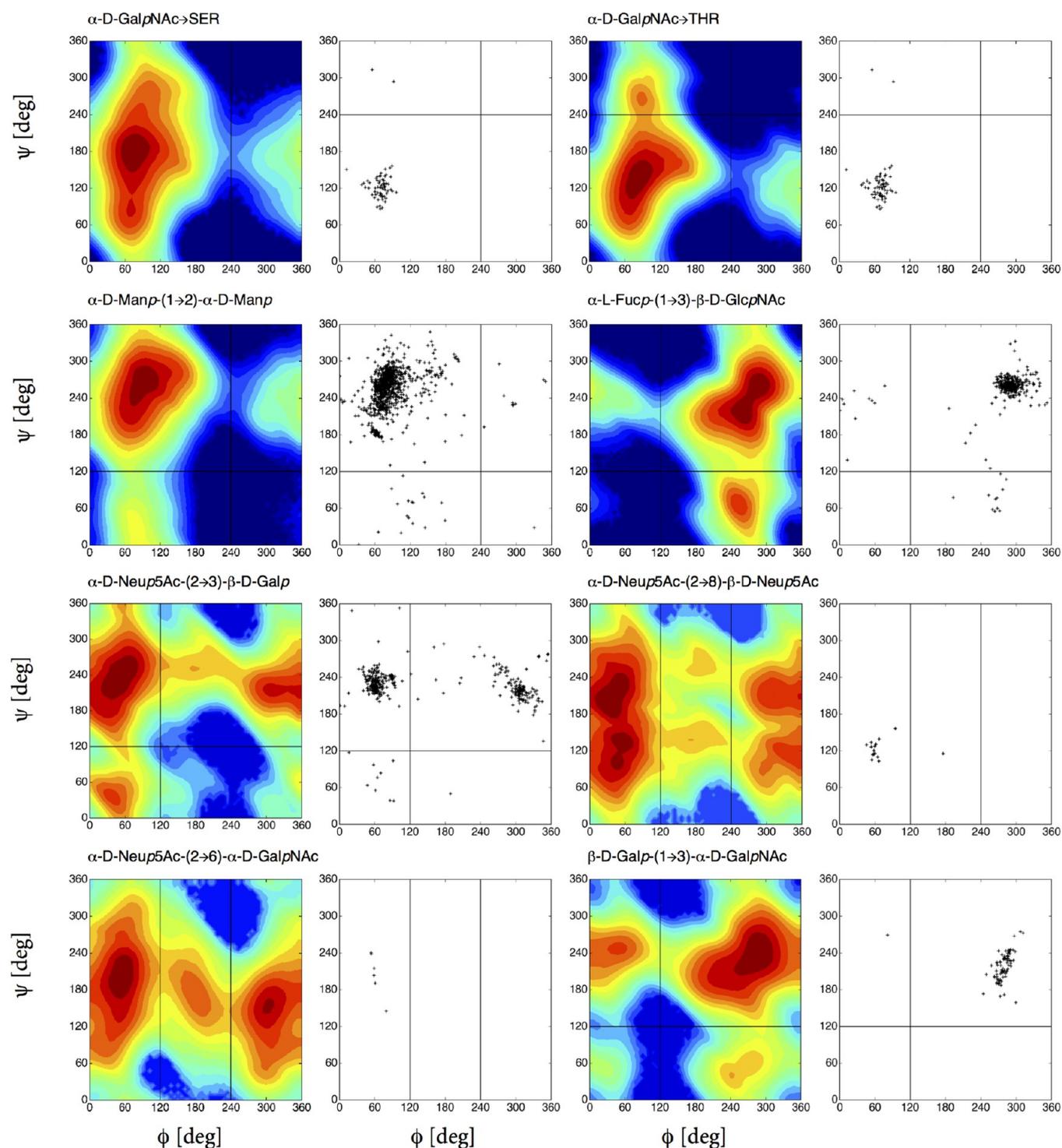


**Figure 4.** Free-energy maps  $G(\phi, \psi)$  in the glycosidic dihedral angle subspace from unbiased LEUS simulations compared with available crystal and NMR data of the related linkages taken from [glycanstructure.org](http://glycanstructure.org) (last column).<sup>27</sup> The first column represents disaccharide simulation, and the second and/or third column corresponds to the disaccharide within a trisaccharide simulation and one for tetramer (for trimer/tetramer codes, see Table 2). Color coding as in Figure 3.

## RESULTS AND DISCUSSION

The conformational landscapes of 17 disaccharides (see Table 1), 10 trisaccharides, and 1 tetramer were studied (see Table 2). The analyzed disaccharides constitute the most typical mammalian N- and O-glycans (see Figure 1). Glycoproteins

carrying those units have crucial biological functions, often performed through their glycans. However, how different types of glycans change the function of the glycoproteins often remains unsolved because of incomplete information on their structure. To shed light on the structure–function relation we explored how addition of a subsequent monosaccharide affects



**Figure 5.** Free-energy maps  $G(\phi, \psi)$  of disaccharide fragments of O-glycan cores and sialic acid (Neup5Ac) linkage variations in the glycosidic dihedral angle subspace from LEUS simulations compared with available crystal and NMR data of the related linkages taken from [glycanstructure.org](http://glycanstructure.org).<sup>27</sup>

the conformational preferences. Our approach is to examine this effect through comparing the free-energy landscape of trisaccharides with their relevant disaccharide units.

**Comparison of Disaccharide Free-Energy Maps with Experimental Data.** The glycosidic linkage conformations of 17 disaccharides and 10 trisaccharides evaluated from LEUS simulations are compared with available experimental data in [Figures 3–7](#). In [Figures 3 and 4](#), the first column represents the

free-energy map of the disaccharide from LEUS simulation, while the second and third column corresponds to the conformation of the same linkage as observed in the trisaccharide LEUS simulations. In the last column, available crystal and NMR structure data of the corresponding linkage gathered from [glycanstructure.org](http://glycanstructure.org) are plotted.<sup>27</sup> In the free-energy maps, conformational landscapes are divided into states (A, B, C, up to D) which is illustrated in the right top panel of

**Figure 3.** Contour maps were drawn with 5 kJ mol<sup>-1</sup> spacing starting from the global minimum which is set to 0 kJ mol<sup>-1</sup>. The regions that were never visited are shown in dark blue in all free-energy maps. In **Figure S5**, the remaining disaccharide free-energy maps are reported and compared to  $\phi$ ,  $\psi$  maps from experimental data. Comparison of the free-energy maps shows that the simulations of disaccharides are in good agreement with the available crystal and NMR structures. Not only the most populated regions but also the shape of the conformational maps and the density of observations was captured when there were enough available structures. Besides the most stable states, there are other thermally accessible states present. Some of the disaccharide free-energy landscapes which are previously reported from plain MD simulations are also in agreement considering the most visited energy basin.<sup>11</sup> We applied two-dimensional local elevation along  $\phi$  and  $\psi$  dihedral angle; however, in the 1  $\rightarrow$  6, 2  $\rightarrow$  6, and 2  $\rightarrow$  8-linked disaccharides, sampling along the extra dihedral angles ( $\omega$  for 1  $\rightarrow$  6, 2  $\rightarrow$  6;  $\chi_1$  and  $\chi_2$  for 2  $\rightarrow$  8) was free (see **Figures S1 and S2**) as was also seen in the work of Perić-Hassler et al.<sup>12</sup> For the  $\beta$ -N-glycosidic linkage, a one-dimensional local elevation potential was built along O5–C1–N $\delta$ –C $\gamma$  torsional angle.

**Analysis of Disaccharides.** The free energies of the individual conformational states is given in **Table 1**, and the respective minimum energy conformations in terms of the glycosidic angles is presented in **Table S1**. We consider states within 25 kJ mol<sup>-1</sup> ( $\sim 10k_B T$ ) to be thermally accessible. Except from disaccharides containing sialic acid,  $\alpha$ -Neup5Ac, for all the  $\alpha$ -linked disaccharides and O-linkages with a nonreducing residue involved in  $\alpha$  linkage ( $\alpha$ -D-Manp-(1  $\rightarrow$  6)- $\beta$ -D-Manp,  $\alpha$ -D-Manp-(1  $\rightarrow$  3)- $\beta$ -D-Manp,  $\alpha$ -D-Manp-(1  $\rightarrow$  2)- $\alpha$ -D-Manp, and linkages  $\alpha$ -D-GalpNac  $\rightarrow$  SER and  $\alpha$ -D-GalpNac  $\rightarrow$  THR) we observe only one significant minimum for the  $\phi$  dihedral angle, in the [60°; 100°] range. While for  $\beta$ -linked (also for  $\alpha$ -L) disaccharides and N-linkage with the nonreducing residue involved in  $\beta$  linkage ( $\beta$ -D-Galp-(1  $\rightarrow$  4)- $\beta$ -D-GlcpNac,  $\beta$ -D-GlcpNac-(1  $\rightarrow$  4)- $\beta$ -D-GlcpNac,  $\beta$ -D-GlcpNac-(1  $\rightarrow$  4)- $\beta$ -D-Manp,  $\beta$ -D-Manp-(1  $\rightarrow$  4)- $\beta$ -D-Manp,  $\beta$ -D-Galp-(1  $\rightarrow$  3)- $\beta$ -D-GalpNac,  $\alpha$ -L-Fucp-(1  $\rightarrow$  3)- $\beta$ -D-GlcpNac,  $\alpha$ -L-Fucp-(1  $\rightarrow$  6)- $\beta$ -D-GlcpNac, and  $\beta$ -D-GlcpNac-ASN) the  $\phi$  dihedral angle is observed in the [0°; 100°] and [270°; 300°] ranges, with the latter one being the most populated. These observations of the  $\phi$  dihedral angle are in agreement with conformations predicted by the exoanomeric effect. The minima of the  $\phi$  dihedral angle in  $g^+$  and  $g^-$  conformations are preferred, whereby the latter one is disfavored for the nonreducing residue involved in  $\alpha$ -linkage due to steric considerations.<sup>8</sup>

When we compare X-(1  $\rightarrow$  4)- $\beta$ -D-GlcpNac where X is  $\beta$ -D-GlcpNac,  $\beta$ -D-Manp, and  $\beta$ -D-Galp, we observe that for all disaccharides conformation C is preferred, with conformation D representing the second energetically favored state at 12–14 kJ mol<sup>-1</sup> above the minimum. However, when  $\beta$ -D-Manp is the nonreducing residue (X), the difference between state D (second energetically favored) and state A (third energetically favored) is only 1.2 kJ mol<sup>-1</sup>, significantly smaller than the rest, for which state A is around 7 kJ mol<sup>-1</sup> higher.

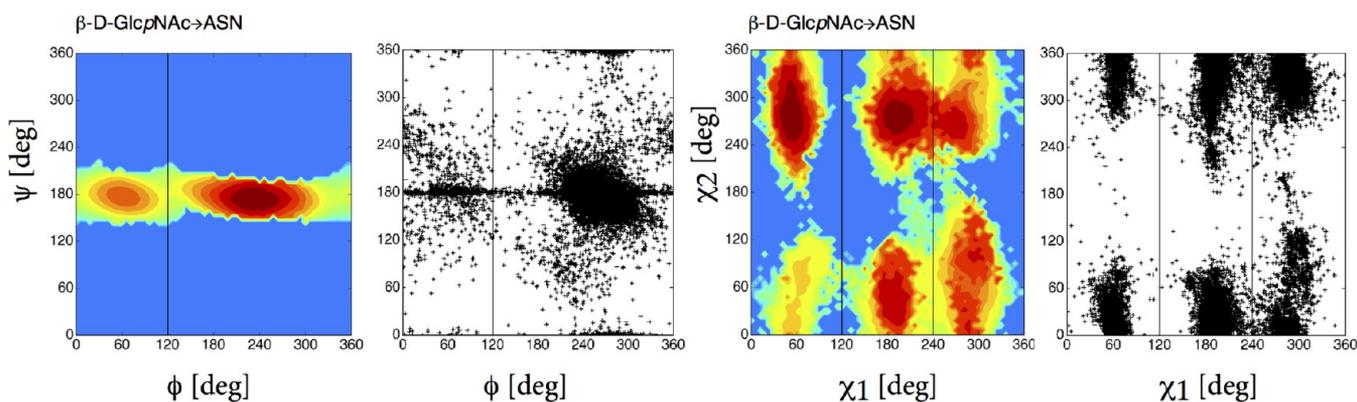
Turning our attention to the  $\psi$  dihedral angle, we observe for disaccharides with a R-configuration at CX' (1  $\rightarrow$  4 linkages having D-GlcpNac at the reducing end) preferred values in the [60°; 120°] range and a secondary visited state with a  $\psi$  dihedral angle of about 300°. The ones with a S-configuration at CX' (1  $\rightarrow$  3 or 2  $\rightarrow$  3 linkages having D-Manp, D-Galp, D-GalpNac, and D-GlcpNac at the reducing end; 1  $\rightarrow$  2 linkage

having D-Manp at the reducing end), however, prefer conformations in the [220°; 300°] range. For (1  $\rightarrow$  6)-linked disaccharides, preferred values of the  $\psi$  dihedral angle lie within the [160°; 180°] range. These results are also compatible with the previous studies from the glucose-based disaccharide simulations<sup>12</sup> and also with MM3 maps.<sup>28</sup>

Disaccharides containing neuraminic (sialic) acid at the nonreducing end ( $\alpha$ -D-Neup5Ac-(2  $\rightarrow$  6)- $\beta$ -D-Galp and its (2  $\rightarrow$  6)- $\alpha$ -D-Galp isomer) are commonly found in the carbohydrate moiety of glycoproteins and glycolipids, typically found at outermost ends of glycan units. It is a key component of receptors in molecular recognition for many viruses and bacterial toxins<sup>29</sup> and the receptor specificity of influenza viruses is defined in terms of their recognition of the type of sialic acid glycosidic linkage.<sup>30</sup> Ligabue-Braun et al. evaluated the conformational ring puckering properties of sialic acid along with 53A6glyc compatible parameters using metadynamics indicating <sup>2</sup>C<sub>5</sub> as preferred and confirmed this behavior by 1000 ns of unbiased MD simulations.<sup>31</sup> Free-energy maps of the disaccharides show three distinct allowed regions for the  $\phi$  dihedral angle (60°, 180°, and 300°) unlike other  $\alpha$  linked disaccharides. The related states (A, B, C) have very close free-energy values of 0.0, 4.4, and 3.2 kJ mol<sup>-1</sup>, respectively. The additional degree of freedom,  $\omega$ , in these linkages is almost freely rotating as can be seen in **Figure S1**.

The  $\alpha$ -D-Neup5Ac-(2  $\rightarrow$  8)- $\alpha$ -D-Neup5Ac disaccharide commonly occurs in gangliosides. They are mostly found at the end of the oligosaccharides of glycoproteins and they also form polysialic acid chains which have important function in cell adhesion and differentiation events. There are only few structures in the pdb carrying this linkage (see **Figure 5**); therefore, the conformational preferences of this linkage could not be compared with X-ray structural data. However, our results having the lowest energy for  $\phi$  in  $g^+$  and for  $\psi$  taking a broader range is compatible with existing NMR data.<sup>32,33</sup> Our simulations showed extensive sampling along the  $\chi_1$ , while  $\chi_2$  remains at a value of 120° throughout (see **Figure S2**). The predominant state for  $\chi_1$  was at  $g^+$  conformation which is in agreement with NMR studies.<sup>32,33</sup> Since the sampling is rather complex because of the additional two dihedral angles ( $\chi_1$ ,  $\chi_2$ ) in the glycosidic linkage, we also built a local elevation potential along the  $\phi$ ,  $\chi_2$  space, thereby including  $\chi_2$  in the LE biasing stage. Again the predominant  $\chi_2$  region was 120°; however, we observed a preference for  $\chi_2$  at 300° when  $\phi$  is in  $g^+$ . Forcing  $\chi_2$  angle to take 300° with a  $\phi$  value of 60° required 20 kJ mol<sup>-1</sup>. Our results suggest that sialic acid containing 2  $\rightarrow$  8 and 2  $\rightarrow$  6 linkages are more flexible and span a larger area than the 2  $\rightarrow$  3 linkage, because of the additional degrees of freedom.

The  $\alpha$ -D-Manp-(1  $\rightarrow$  2)- $\alpha$ -D-Manp linkage is highly abundant in high-mannosidic glycans. Säwén et al.<sup>34</sup> studied this disaccharide extensively with MD and NMR, derived its most probable conformation with glycosidic angles  $\phi = 80^\circ$  and  $\psi = 273^\circ$ . Our results for this disaccharide shows the most energetically favored state as B and minimum conformation at this state has  $\phi = 96^\circ$  and  $\psi = 276^\circ$  (see **Table S1**) showing strong agreement with their results.  $\alpha$ -D-GalpNac  $\rightarrow$  Thr and  $\alpha$ -D-GalpNac  $\rightarrow$  Ser have biologically important functions and also are of interest for therapeutics, especially in the development of vaccines for cancer treatment.<sup>35</sup> Also, this glycosylation has been known to force the peptide backbone to an extended conformation.<sup>35</sup> According to the work of Corzana et al., a change in the sequence of the glycopeptide from Thr to Ser can disturb the function of it. The only difference between



**Figure 6.** Free-energy map  $G(\phi, \psi)$  of ASN linkage in sampled  $\phi, \psi$  subspace. Comparison along glycosidic ( $\phi, \psi$ ) and torsional ( $\chi_1, \chi_2$ ) angles that are available in pdb structures.

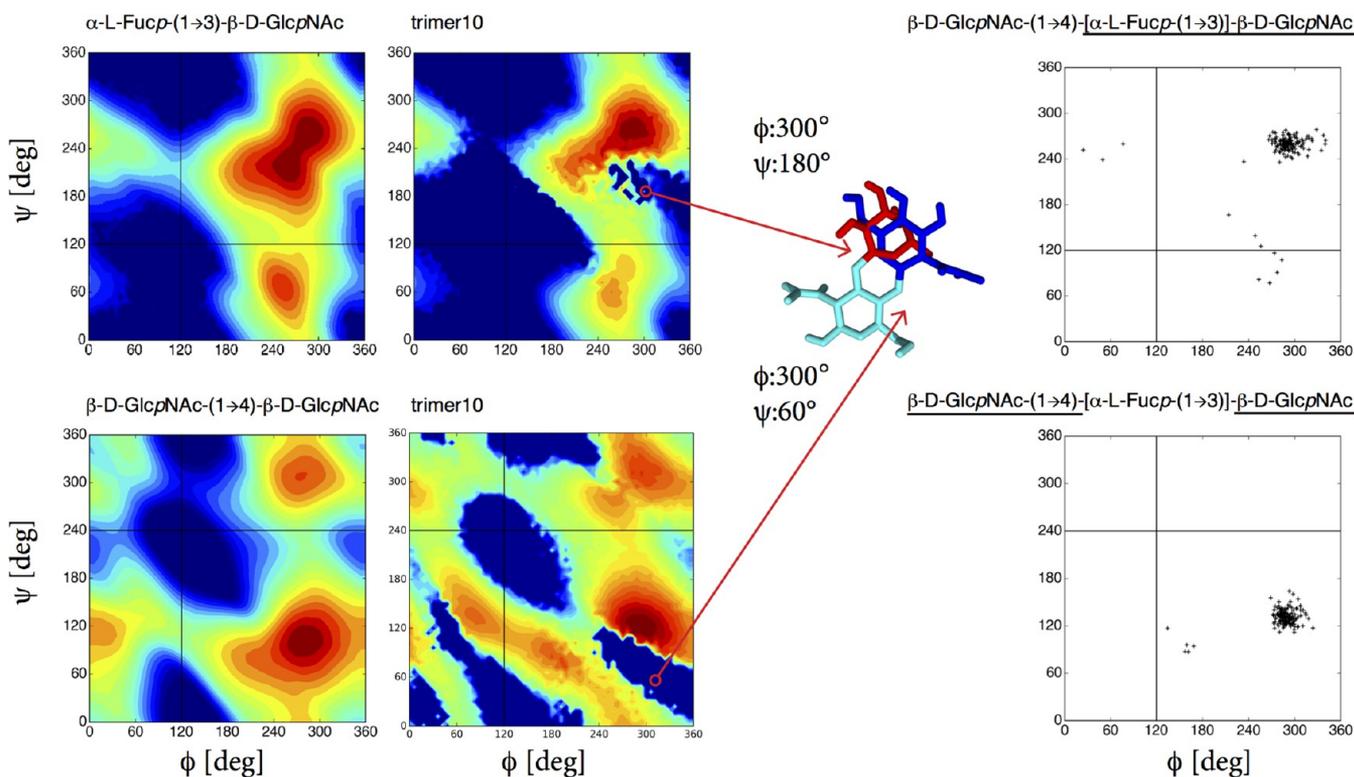
these two linkages is the additional methyl group in the Thr. One can see the impact of this methyl group on the change in  $\psi$  angle preference at the top of Figure 5. In the presence of a methyl group (Thr linkage) the free-energy landscape in the  $\psi$  dihedral angle splits into ranges  $[110^\circ; 170^\circ]$  and  $[250^\circ; 270^\circ]$  compared to one larger range  $[90^\circ; 200^\circ]$  observed in the Ser linkage. In the Ser linkage, the most populated state has a preference for the  $\psi$  angle around  $180^\circ$  resulting in a trans conformation while in Thr, it is at  $120^\circ$ . Also, N-linkage which occurs in most of the cases between Asn and N-acetylglucosamine (GlcNAc) has a global minimum at  $240^\circ$  and secondary preference of  $60^\circ$  along the  $O5-C1-N\delta_2-C\gamma$  and  $C1-N\delta_2-C\gamma-C\beta$  torsional angle. The  $C1-N\delta_2-C\gamma-C\beta$  torsional angle has a strong double bonded character (in analogy to a peptide bond). It remains around  $180^\circ$  and was not included in LE biasing. Coverage of preferences of the Asn side chain torsional angles  $\chi_1:N\delta_2-C\gamma-C\beta-C\alpha$  and  $\chi_2:C\gamma-C\beta-C\alpha-N$  was also captured as can be seen from Figure 6.

Although pyranose rings can adapt a number of ring conformations (chair, half-chair, envelope, boat, and skew-boat) for saturated rings, a chair is preferred since it relieves much of the torsional strain. Since there are high energy barriers between ring conformational states, they often adopt their minimum energy puckerings; for D-sugars, it is  ${}^4C_1$ , and for L-sugars,  ${}^1C_4$ . According to metadynamic studies of Pol-Fachin et al.<sup>7</sup> with the 53A6glyc force field, the free energies of transition from  ${}^4C_1$  to  ${}^1C_4$  is  $-77$ ,  $-54$ , and  $-59$  kJ/mol for  $\beta$ -D-Glcp,  $\beta$ -D-Galp, and  $\beta$ -D-Manp, respectively. During our LEUS simulations, each monosaccharide accordingly remained in their expected ring puckering conformation, and no transitions were observed toward minor states. All the simulations started from their expected ring puckering conformations which were  ${}^1C_4$  for L-Fucp,  ${}^2C_5$  for D-Neup5Ac and  ${}^4C_1$  for D-sugars. While the transition free energies<sup>7</sup> between differently puckered states is probably too high for this force field, restricting the simulations to the most relevant conformations avoids significant complications when analyzing the preferences along the glycosidic linkages.

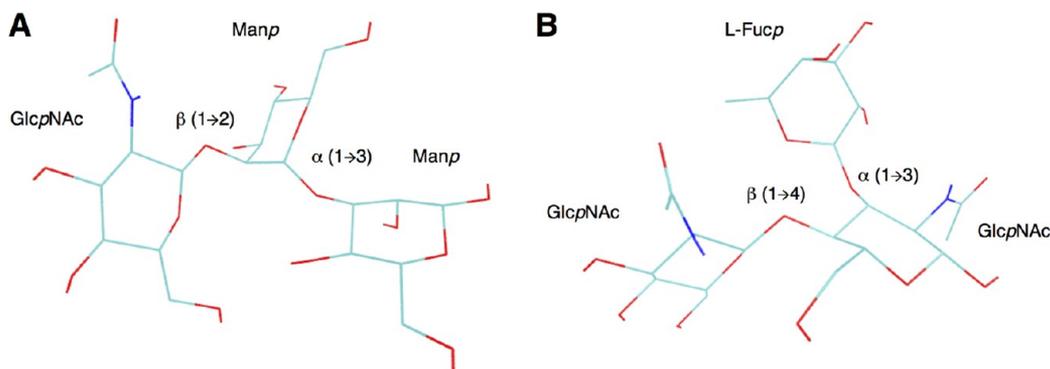
**Comparison of the Conformational Preferences of Trisaccharides.** The trisaccharides are compared to their respective dimer units in terms of their free-energy landscapes in Figures 3 and 4. While the unbiased free-energy landscapes are slightly more noisy in the trisaccharides the sampling using the disaccharide LE potentials is clearly sufficient to reconstruct the conformational preferences. The relative free-energies of the states are compared in Table 2. As can be seen from this

table, the most populated state in the trisaccharides always remains the same as the ones observed for the corresponding disaccharide. Not only the global minimum state, but also the preferred order of the states did not change. The preference for second, third, and fourth energetically favored states remained the same in the presence of an additional linkage. Overall, the relative free-energy values of the linkages as observed in the trisaccharides deviated by around  $2$  kJ mol<sup>-1</sup> up to a maximum of  $7.2$  kJ mol<sup>-1</sup> from their disaccharide simulations. This maximum deviation was seen in the trisaccharides where the two flanking monosaccharides were on adjacent carbon atoms of the central one (trimer4 and trimer10). Besides from these two cases, the maximum deviation in the thermally accessible states was around  $3$  kJ mol<sup>-1</sup>. Subsequently, we studied the occurrence of intramolecular hydrogen bonds in the di- and trisaccharides. The presence of a H-bond was determined using a geometric criterion, requiring a hydrogen-acceptor distance shorter than  $0.25$  nm and a donor hydrogen-acceptor angle larger than  $135^\circ$ . The occurrences as observed in the LEUS simulations were unbiased using umbrella sampling, to obtain average occurrences for the unbiased ensembles. In Table S3 in the Supporting Information, we report the occurrences of the most prominent (P1) and the second most relevant (P2) hydrogen bonds, using a cutoff of 2%.

Strong H-bonds are observed in the  $\beta(1 \rightarrow 4)$  linkage between the hydroxyl group vicinal to the linkage to the ring oxygen atom of the nonreducing residue ( $HO3'-O5$ ), with occurrences between 80 and 90%. The highest probability was observed in the  $\beta$ -D-GlcpNAc-(1  $\rightarrow$  4)- $\beta$ -D-GlcpNAc linkages, with 93%, which may be attributed to its rigidity. Strikingly, for the  $\beta$ -D-Galp-(1  $\rightarrow$  4)- $\beta$ -D-GlcpNAc linkage, this hydrogen bond was observed in the dimer for only 12%, which was increased to 70% in the trimer. Note that, according to Table 2, this was not resulting in a significant shift in the conformational preferences. For the  $\alpha(1 \rightarrow 3)$  linkage, the corresponding H-bonds were much weaker, with occurrences of 2–9% ( $HO4'-O5$  and  $HO2'-O5$ ). Similarly, the  $HO3'-O5$  hydrogen bond was populated for about 5% in the  $\beta(1 \rightarrow 2)$  linkage, for which an additional  $HO4'-O6$  H-bond was seen in the trimers. No H-bonds with a probability higher than 2% were observed for the (1  $\rightarrow$  6) linkages, as also reported by Perić-Hassler et al.<sup>12</sup> and Patel et al.<sup>36</sup> in their studies. Finally, sialic acid showed intramonomer H-bonding between its  $HO8$  hydroxyl group and either one of the carboxyl oxygens  $O1A/O1B$ . This H-bond also seems to be more pronounced in the trisaccharide than in the disaccharide.



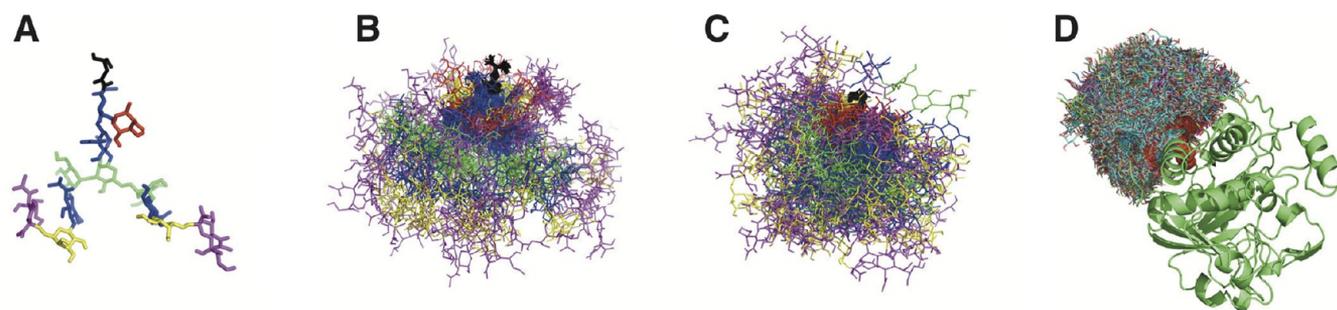
**Figure 7.** Comparison of free-energy maps  $G(\phi, \psi)$  of dimer9 and dimer7 to the values in trimer10 in disaccharide and trisaccharide simulations along with the comparison of available X-ray and NMR structures with exact matches of the  $\beta$ -D-GlcpNAc-(1  $\rightarrow$  4)-[ $\alpha$ -L-Fucp-(1  $\rightarrow$  3)]- $\beta$ -D-GlcpNAc trimer. A hypothetical molecule was constructed using nonobserved conformations of  $\phi, \psi$ , leading to an obvious steric clash. The color coding is as in Figure 3.



**Figure 8.** Representation of trimer4,  $\beta$ -D-GlcpNAc-(1  $\rightarrow$  2)- $\alpha$ -D-Manp-(1  $\rightarrow$  3)- $\beta$ -D-Manp (A) and trimer10,  $\beta$ -D-GlcpNAc-(1  $\rightarrow$  4)-[ $\alpha$ -L-Fucp-(1  $\rightarrow$  3)]- $\beta$ -D-GlcpNAc (B) where the glycosidic linkages are connected to carbon atoms next to each other in their global minimum conformations. In these two special cases, the free energy difference deviated most from their corresponding disaccharide values. For trimer4, maximum deviation compared to its dimers occurred in the  $\alpha$ -D-Manp-(1  $\rightarrow$  3)- $\beta$ -D-Manp linkage with 5.5 kJ/mol at state D. For trimer10, the maximum deviation compared to its dimers occurred in  $\alpha$ -L-Fucp-(1  $\rightarrow$  3)- $\beta$ -D-GlcpNAc linkage with 7.2 kJ/mol at state C.

From the similarity between conformational preferences of the linkages in disaccharides and trisaccharides, one can conclude that the effect of consecutive linkages on the previous linkage is limited and taking the disaccharides as a model for oligosaccharides is applicable. However, in the exceptional case  $\beta$ -D-GlcpNAc-(1  $\rightarrow$  4)-[ $\alpha$ -L-Fucp-(1  $\rightarrow$  3)]- $\beta$ -D-GlcpNAc there were regions that were not sampled in the trisaccharide simulations which are sampled in the disaccharide. We further analyzed those regions which were sterically not allowed in the trisaccharide case, as illustrated in Figure 7. This was expectable since the linkages are introduced right next to each other. However, when  $\alpha$ -L-Fucp was 1  $\rightarrow$  6 linked to the  $\beta$ -D-

GlcpNAc-(1  $\rightarrow$  4)- $\beta$ -D-GlcpNAc disaccharide (trimer9), all the states that were sampled in the disaccharide LEUS simulations are covered (see Figure 4). This comparison also confirms the additional conformational freedom in 1  $\rightarrow$  6 linkages. In the other trimer simulation where the linkage is connected to carbon atoms right next to each other,  $\beta$ -D-GlcpNAc-(1  $\rightarrow$  2)- $\alpha$ -D-Manp-(1  $\rightarrow$  3)- $\beta$ -D-Manp (trimer4) the conformational preferences compared to their disaccharide behavior were not effected as much as for trimer10. See Figure S3 for a detailed comparison of the conformational preferences of the dimer linkages in trimer4. The reduced effect for this trimer may be explained in terms of the stereochemical orientation in this



**Figure 9.** (A) Initial structure of a complex type of glycan created from related global minima of each glycosidic linkage after unbiasing LEUS simulations. (B) 50 frames aligned on ASN (shown in black) from plain MD and (C) from LEUS simulations. (D) 2100 Filtered structures of all the trajectory from LEUS which are fitted to protein environment (here it is ST6GAL-I, an important enzyme catalyzing the  $\alpha$ 2,6 sialylation on N-glycans, having a glycosylation site at ASN149, PDB ID 4JS1).

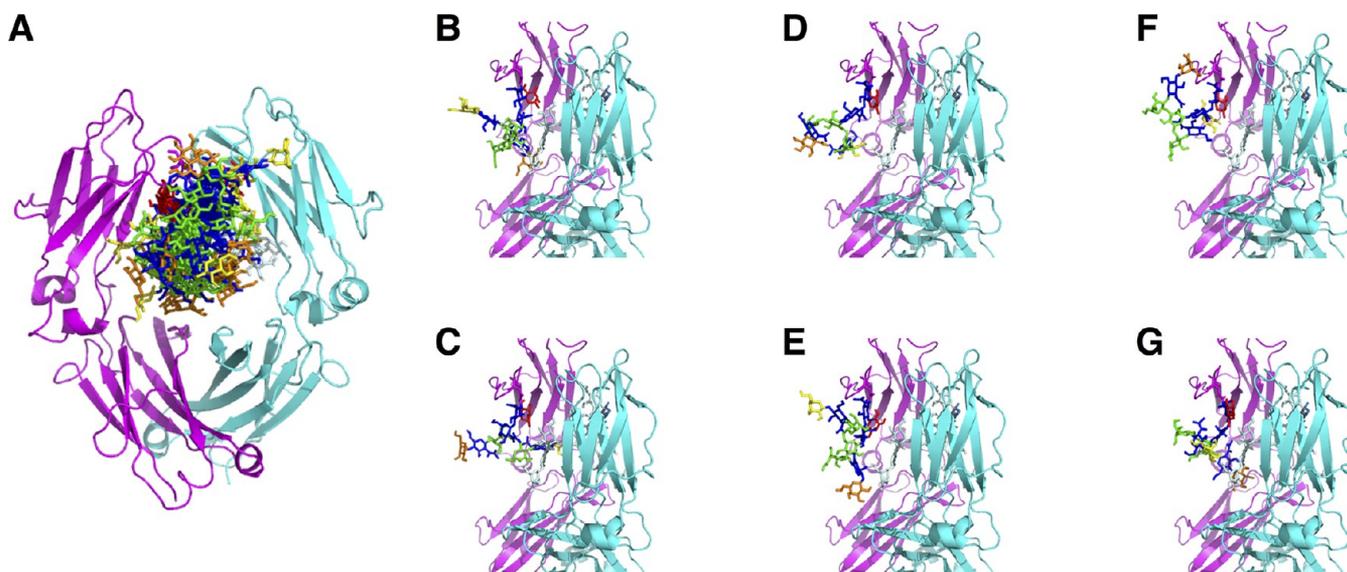
trimer as shown in Figure 8. Trimer10 was the case where the maximum deviation of 7.2 kJ/mol in free energy from its dimers were seen; at the  $\alpha$ -L-Fucp-(1  $\rightarrow$  3)- $\beta$ -D-GlcpNAc linkage, state C. This result emphasizes the role of the N-acetyl group in the monosaccharides, restricting the conformational space in the trisaccharides where the linkages are on the adjacent carbon atoms (see Figure 8). Besides from the presence of the extra N-acetyl group in trimer10, additional strain comes from the position of the linkages where the  $\beta$ (1  $\rightarrow$  4) linkage stays between the  $\alpha$ (1  $\rightarrow$  3) linkage at C3 and the methyl group at C5 (illustrated in Figure 8), thereby, affecting the free-energy landscape of  $\beta$ (1  $\rightarrow$  4) linkage as can be seen from Figure 7. An additional trimer unit  $\alpha$ -D-Manp-(1  $\rightarrow$  2)- $\alpha$ -D-Manp-(1  $\rightarrow$  2)- $\alpha$ -D-Manp commonly seen in high-mannose glycans was simulated to further study the effect of close-by linkages. The free-energy landscape of its dimer units showed similar conformational preferences compared to its dimer simulation (see Figure S3). The relative free energy values of both linkages showed  $\sim$ 3 kJ/mol difference in its conformational states compared to the disaccharide simulation (data not shown). For the relative free energy values of conformational states of  $\alpha$ -D-Manp-(1  $\rightarrow$  2)- $\alpha$ -D-Manp disaccharide refer to Table 1.

**Construction of Oligosaccharides.** While the analysis in the previous paragraph implies that for most linkages, the local effect of neighboring monosaccharides will be limited, longer range influences in larger structure are still to be expected.<sup>33,37,38</sup> We developed several approaches to construct glycoproteins without any prior structural knowledge of its carbohydrate moiety. First, the glycan unit of a desired glycoprotein may be built according to the minimum free-energy conformations of its individual glycosidic linkages. This offers an initial model that can be used as it is, or refined further (Figure 9A). Alternatively, optimal trees of a glycan can subsequently be constructed either by randomly picking states for each glycosidic linkage according to their probabilities (computed from the relative free energies) or by applying US for each linkage with a frozen disaccharide LE potential energy library. This generates a broad bundle of glycan structures, which can be filtered in a protein environment after alignment with the glycosylation site of the protein (either Asn or Ser/Thr). Longer range effects of the conformational preferences, most explicitly steric clashes, will be reflected in this bundle of feasible glycan structures. In a filtering step all structures from the bundle are discarded that have a nonbonded energy (intraglycan and glycan–protein) above a given cutoff value. The remaining glycoprotein structures are minimized in

vacuum and optimal trees may be selected according to the minimum energy values.

As an example we demonstrate this procedure by constructing a sialylated complex-type N-glycan tree on the human  $\beta$ -galactoside  $\alpha$ 2,6-sialyltransferase-I (ST6GAL-I, PDB code 4JS1) see Figure 9. Human ST6GAL-I has two glycosylation sites on ASN149 and ASN161 which were suggested to be critical for its folding<sup>39</sup> and have an impact on its enzymatic function.<sup>40</sup> Here the glycan tree is built only on ASN149 for representation. The initial structure of the complex type of a glycan tree was created from the global minimum of each glycosidic linkages (after unbiasing of disaccharide LEUS simulations). For the construction of the initial glycan structure, Fernandes et al. used the most populated conformations of the disaccharides obtained from 100 ns plain MD simulations. In their work, they are building the glycoprotein only with these units and subsequently apply plain MD to refine the carbohydrate environment. However, since plain MD simulations may get caught in one energy minimum, we rather apply the US stage on all linkages simultaneously to generate a diverse trajectory. Such a fragment-based approach has been described before.<sup>19</sup> Construction of the glycan tree through enhanced methods is the crucial step, since there are evidence of experimentally obtained N-glycans taking highly flexible dynamic conformations. A fold-back conformer of high-mannose-type oligosaccharides explored by combination of NMR and REMD is one of the examples.<sup>15</sup> Use of LEUS enables us to capture such conformers as well (see Figure S4). The difference between plain MD simulation and the LEUS simulation can be seen from Figure 9B and C where 50 frames of the trajectories from plain MD and LEUS were randomly selected and aligned on the Asn residue. In the plain MD simulation, the oligosaccharide was only fluctuating around its initial structure, while the LEUS trajectories carried more diversity. After LEUS simulation of the oligosaccharide, these diverse structures were fitted on a protein environment, by alignment of the Asn residue. Using an energy cutoff of  $+10^4$  kJ mol<sup>-1</sup>, these trajectories were filtered. Next the protein (PDB ID 4JS1,<sup>41</sup> ST6GAL-1) carrying different possible conformations of its oligosaccharide unit are minimized in vacuum. In Figure 9D, one can see 2100 minimized glycan structures of the protein with its glycan. For clarity, only a single protein structure is shown.

In most of the glycoproteins, glycans are exposed to the surface and they are relatively free to move, as seen in Figure 9D. However, there are therapeutically important glycoproteins where glycans are not on the surface but rather embedded



**Figure 10.** (A) Representation of a fitting of the galactosylated glycoform on IgG Fc protein carrying glycans on both chains at its core. Protein and glycans respectively are represented in cartoon and stick form, respectively. N-Acetylglucosamine, mannose, and fucose units are shown in blue, green, and red, respectively. Galactose in the 1,6-branch is shown in yellow and in the 1,3-branch in orange. (B–G) Related views with different levels of glycan exposure: (B) the 1,6-branch is exposed; (C) the 1,3-branch is exposed; (D, E) both ends of the glycan are exposed; (F, G) examples of conformations where the glycan interacts with the protein or with the other glycan tree.

within the glycoprotein environment. In a second application, such a complicated system was chosen; the IgG Fc fragment carrying two complex-type biantennary N-glycans attached to ASN297 of the dimeric Fc fragment, buried in the cavity between the two chains. The Fc glycans are heterogeneous leading to diverse glycoforms which correlates with certain diseases.<sup>42</sup> In this application, we have studied the IgG Fc fragment containing a fucosylated complex-type N-glycan, terminated with galactose, due to its high occurrence in healthy humans.<sup>43</sup> Since not only the two glycan units are interacting with each other but they are also embedded within the core of glycoprotein, the number of optimal trees that was obtained was significantly reduced compared to fitting of glycans on the exposed protein surface, but some diversity remained among the structures after filtering (see Figure 10A). In addition to the conformation observed in the crystal structure (PDB ID 4BYH<sup>22</sup>), we have found other energetically possible conformers for the glycan units that are more exposed to the surface and more accessible, as has also been suggested by a NMR study.<sup>44</sup> The conformations obtained after minimization of the fitted structures showed that glycan units can take fully exposed conformations, either having the 1,6-branch (Figure 10B) or the 1,3-branch exposed (Figure 10C), and a more dynamic glycan conformation with both ends exposed (Figure 10D and E). Furthermore, conformations of glycan units having contacts with the protein and with the other glycan chain were seen and illustrated in Figure 10F and G, respectively. From previous NMR analysis of the Fc glycan it was concluded that the 1,6-branch terminus of the glycan is anchored to the protein surface while the 1,3-branch is exposed and readily accessible for enzymatic addition of sialic acid.<sup>45,46</sup> However, in the recent NMR spin relaxation study of Barb and Prestegard,<sup>44</sup> it was proposed that not only the 1,3-branch but also the 1,6-branch is highly dynamic and can take exposed conformations as well.

Note that the presampled bundle of glycan trees for a given chemical composition be reused time and again for different protein systems. By generating a library of typically occurring

glycan tree bundles, the model building can be done rather efficiently. These approaches are readily applicable using the GROMOS package for molecular simulations. The analysis programs to filter unfavorable glycan trees from the larger conformational bundle will also be made available in the GROMOS++ library of analysis programs. Finally, all local elevation biasing potentials of the disaccharides in Table 1 are made available in the Supporting Information.

## CONCLUSION

The conformations of glycans in conjugation with glycoproteins form a challenge both for experimental and theoretical methods. Their complexity is the result of the variety of possible monomeric units which are linked in a branched way and have differently populated conformational states. There is a pronounced lack of spatial information about them causing an obstacle to grasp the full picture about their biological functions. Especially toward the end of the glycan chains, the amount of available structural data is scarce; see for instance the rightmost panels for the  $\alpha$ -D-Neup5NAc-(2  $\rightarrow$  6)- $\beta$ -D-Galp linkage in Figure 3 and for various O-glycan core units in Figure 5.

In this article, we studied the glycosidic linkages that occur in virtually all of the biologically relevant N- and O-linked glycan units. The individual free-energy landscapes were derived from extensive local elevation and umbrella sampling molecular dynamics simulations. For the first time, we reported a complete conformational library in the form of the free-energy landscapes of all disaccharide linkages needed to construct biologically relevant glycans. Since the glycan structures have a great flexibility, not only the lowest-energy conformation is relevant, but also energetically less favorable conformations can be expected to be accessible. For linkages for which structural data is available, our simulations are in excellent agreement with observed conformational preferences, while linkages for which experimental data is sparse, such as the  $\alpha$ -D-Neup5NAc-(2  $\rightarrow$  6)- $\alpha$ -D-Galp,  $\alpha$ -D-Neup5NAc-(2  $\rightarrow$  8)- $\alpha$ -D-Neup5NAc linkage

and various O-glycan linkages, we provide conformational preferences that complement the experimental data. Many of these linkages have not been studied thoroughly before, other than using vacuum conformational searches or very short MD simulations. All relevant initial structures, topologies, and local-elevation biasing potentials of the disaccharides are made available in the [Supporting Information](#).

In a next step, we systematically studied the effect of an additional linkage on the free-energy landscape of the disaccharide. All relevant trisaccharides and one tetrasaccharide were extensively simulated and the free-energy landscapes were compared. We show that the conformational preferences are largely independent of the consecutive linkages. However, we identified branching linkages, for which neighboring molecules do limit the accessible conformational space (see [Figure 7](#)).

The conformational library for the disaccharides, combined with the LEUS method, enables us to build models of higher oligosaccharides without extensive additional sampling of the conformational space. At the hand of two realistic examples, we demonstrate possible modeling methods for glycans that are conjugated to glycoproteins. Overall, our work offers important handles to realistically model the three-dimensional structures and flexibility of glycans in the absence of experimental data. As such, our work opens the way to more realistically study the effect of protein glycosylation on its structure, dynamics, and function.

## ■ ASSOCIATED CONTENT

### ■ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.jcim.7b00351](https://doi.org/10.1021/acs.jcim.7b00351).

Glycosidic linkage conformations at free-energy minima of the corresponding states (Table S1). Overview of changes in the force field that was used in the current work as compared to the 53A6glyc parameters (Table S2). Average occurrence of intramolecular hydrogen bonds (Table S3). Free-energy maps along the additional degrees of freedom in (1–6) and (2–8) linkages (Figures S1 and S2). Detailed comparison of dimer conformational preferences to those in trimer4 (Figure S3). Examples of fold-back conformations observed in the conformational bundle of a complex glycan tree (Figure S4) ([PDF](#))

Topologies, input structures, and local elevation biasing potentials for all dimers listed in [Table 1](#) ([ZIP](#))

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [chris.oostenbrink@boku.ac.at](mailto:chris.oostenbrink@boku.ac.at)

### ORCID

Chris Oostenbrink: [0000-0002-4232-2556](https://orcid.org/0000-0002-4232-2556)

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was supported by the Austrian Science Fund in the context of the Doctoral Program BioToP (Grant W 1224). The authors thank Maria Pechlaner for help with programming the gromos++ program `fit_ener` used to filter glycan structures based on their energies.

## ■ REFERENCES

- (1) Mechref, Y.; Novotny, V. M. In *Handbook of glycomics*; Cummings, R., Pierce, J. M., Eds.; Academic Press: San Diego, 2010; pp 1–43.
- (2) Lütteke, T.; von der Lieth, C.-W. PDB-Care (PDB Carbohydrate Residue Check): a Program to Support Annotation of Complex Carbohydrate Structures in PDB Files. *BMC Bioinf.* **2004**, *5*, 69.
- (3) Momany, F. A.; Willett, J. L.; Schnupf, U. Molecular Dynamics Simulations of a Dyclic-DP-240 Amylose Fragment in a Periodic Cell: Glass Transition Temperature and Water Diffusion. *Carbohydr. Polym.* **2009**, *78*, 978–986.
- (4) 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.
- (5) Guvench, O.; Hatcher, E.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D. CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *J. Chem. Theory Comput.* **2009**, *5*, 2353–2370.
- (6) Hansen, H. S.; Hünenberger, P. H. A Reoptimized GROMOS Force Field for Hexopyranose-Based Carbohydrates Accounting for the Relative Free Energies of Ring Conformers, Anomers, Epimers, Hydroxymethyl Rotamers, and Glycosidic Linkage Conformers. *J. Comput. Chem.* **2011**, *32*, 998–1032.
- (7) Pol-Fachin, L.; Rusu, V. H.; Verli, H.; Lins, R. D. GROMOS 53A6 GLYC, an Improved GROMOS Force Field for Hexopyranose-based Carbohydrates. *J. Chem. Theory Comput.* **2012**, *8*, 4681–4690.
- (8) Rao, V. S. R.; Qasba, P. K.; Balaji, P. V.; Chandrasekaran, R. *Conformation of Carbohydrates*; Harwood Academic Publishers: Amsterdam, The Netherlands, 1998.
- (9) Arnold, J. N.; Radcliffe, C. M.; Wormald, M. R.; Royle, L.; Harvey, D. J.; Crispin, M.; Dwek, R. A.; Sim, R. B.; Rudd, P. M. The Glycosylation of Human Serum IgD and IgE and the Accessibility of Identified Oligomannose Structures for Interaction with Mannan-Binding Lectin. *J. Immunol.* **2004**, *173*, 6831–6840.
- (10) Montero-Morales, L.; Maresch, D.; Castilho, A.; Turupcu, A.; Ilieva, K. M.; Crescioli, S.; Karagiannis, S. N.; Lupinek, C.; Oostenbrink, C.; Altmann, F.; Steinkellner, H. Recombinant Plant-Derived Human IgE Glycoproteomics. *J. Proteomics* **2017**, *161*, 81–87.
- (11) Fernandes, C. L.; Sachett, L. G.; Pol-Fachin, L.; Verli, H. GROMOS96 43a1 Performance in Predicting Oligosaccharide Conformational Ensembles within Glycoproteins. *Carbohydr. Res.* **2010**, *345*, 663–671.
- (12) 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.
- (13) Yang, M.; Huang, J.; MacKerell, A. D. Enhanced Conformational Sampling Using Replica Exchange with Concurrent Solute Scaling and Hamiltonian Biasing Realized in One Dimension. *J. Chem. Theory Comput.* **2015**, *11*, 2855–2867.
- (14) Re, S.; Miyashita, N.; Yamaguchi, Y.; Sugita, Y. Structural Diversity and Changes in Conformational Equilibria of Biantennary Complex-Type N-Glycans in Water Revealed by Replica-Exchange Molecular Dynamics Simulation. *Biophys. J.* **2011**, *101*, L44–L46.
- (15) Yamaguchi, T.; Sakae, Y.; Zhang, Y.; Yamamoto, S.; Okamoto, Y.; Kato, K. Exploration of Conformational Spaces of High-Mannose-Type Oligosaccharides by an NMR-Validated Simulation. *Angew. Chem., Int. Ed.* **2014**, *53*, 10941–10944.
- (16) 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.
- (17) Huber, T.; Torda, A. E.; van Gunsteren, W. F. Local Elevation: A Method for Improving the Searching Properties of Molecular Dynamics Simulation. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 695–708.
- (18) Kannan, S.; Zacharias, M. Enhanced Sampling of Peptide and Protein Conformations Using Replica Exchange Simulations with a

Peptide Backbone Biasing-Potential. *Proteins: Struct., Funct., Genet.* **2007**, *66*, 697–706.

(19) Hansen, H. S.; Daura, X.; Hünenberger, P. H. Enhanced Conformational Sampling in Molecular Dynamics Simulations of Solvated Peptides: Fragment-Based Local Elevation Umbrella Sampling. *J. Chem. Theory Comput.* **2010**, *6*, 2598–2621.

(20) Schmid, N.; Christ, C. D.; Christen, M.; Eichenberger, A. P.; Van Gunsteren, W. F. Architecture, Implementation and Parallelisation of the GROMOS Software for Biomolecular Simulation. *Comput. Phys. Commun.* **2012**, *183*, 890–903.

(21) Pol-Fachin, L.; Verli, H.; Lins, R. D. Extension and Validation of the GROMOS 53A6(GLYC) Parameter Set for Glycoproteins. *J. Comput. Chem.* **2014**, *35*, 2087–2095.

(22) Crispin, M.; Yu, X.; Bowden, T. A. Crystal Structure of Sialylated IgG Fc: Implications for the Mechanism of Intravenous Immunoglobulin Therapy. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, E3544–E3546.

(23) Lütteke, T.; Bohne-Lang, A.; Loss, A.; Goetz, T.; Frank, M.; von der Lieth, C. W. GLYCOSCIENCES.de: An Internet Portal to Support Glycomics and Glycobiology Research. *Glycobiology* **2006**, *16*, 71R–81R.

(24) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F. J. H. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, 1981; p 331–342.

(25) 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.

(26) Heinz, T. N.; Van Gunsteren, W. F.; Hünenberger, P. H. Comparison of Four Methods to Compute the Dielectric Permittivity of Liquids from Molecular Dynamics Simulations. *J. Chem. Phys.* **2001**, *115*, 1125–1136.

(27) Jo, S.; Im, W. Glycan Fragment Database: A Database of PDB-Based Glycan 3D Structures. *Nucleic Acids Res.* **2013**, *41*, D470–D474.

(28) Frank, M.; Lütteke, T.; von der Lieth, C. W. GlycoMapsDB: A Database of the Accessible Conformational Space of Glycosidic Linkages. *Nucleic Acids Res.* **2007**, *35*, 287–290.

(29) Stencel-Baerenwald, J. E.; Reiss, K.; Reiter, D. M.; Stehle, T.; Dermody, T. S. The Sweet Spot: Defining Virus-Sialic Acid Interactions. *Nat. Rev. Microbiol.* **2014**, *12*, 739–749.

(30) Priyadarzini, T. R. K.; Selvin, J. F. A.; Gromiha, M. M.; Fukui, K.; Veluraja, K. Theoretical Investigation on the Binding Specificity of Sialyldisaccharides with Hemagglutinins of influenza A Virus by Molecular Dynamics Simulations. *J. Biol. Chem.* **2012**, *287*, 34547–34557.

(31) Ligabue-Braun, R.; Sachett, L. G.; Pol-Fachin, L.; Verli, H. V. The Calcium Goes Meow: Effects of Ions and Glycosylation on Fel d 1, the Major Cat Allergen. *PLoS One* **2015**, *10*, 1–19.

(32) Michon, F.; Brisson, J. R.; Jennings, H. J. Conformational Differences Between Linear  $\alpha(2\rightarrow8)$ -linked Homosialooligosaccharides and the Epitope of the Group B Meningococcal Polysaccharide. *Biochemistry* **1987**, *26*, 8399–405.

(33) Brisson, J. R.; Baumann, H.; Imberty, A.; Perez, S.; Jennings, H. J. Helical Epitope of the Group B Meningococcal  $\alpha(2\rightarrow8)$ -linked Sialic Acid Polysaccharide. *Biochemistry* **1992**, *31*, 4996–5004.

(34) Säwén, E.; Massad, T.; Landersjö, C.; Damberg, P.; Widmalm, G. Population Distribution of Flexible Molecules from Maximum Entropy Analysis Using Different Priors as Background Information: Application to the  $\phi$ ,  $\psi$ -Conformational Space of the  $\alpha(1\rightarrow2)$ -linked Mannose Disaccharide Present in N- and O-linked Glycoproteins. *Org. Biomol. Chem.* **2010**, *8*, 3684.

(35) Corzana, F.; Busto, J. H.; Jiménez-Osés, G.; De Luis, M. G.; Asensio, J. L.; Jiménez-Barbero, J.; Peregrina, J. M.; Avenoza, A. Serine versus Threonine Glycosylation: The Methyl Group Causes a Drastic Alteration on the Carbohydrate Orientation and on the Surrounding Water Shell. *J. Am. Chem. Soc.* **2007**, *129*, 9458–9467.

(36) Patel, D. S.; Pendrill, R.; Mallajosyula, S. S.; Widmalm, G.; MacKerell, A. D. Conformational Properties of  $\alpha$ - or  $\beta$ -(1 $\rightarrow$ 6)-Linked

Oligosaccharides: Hamiltonian Replica Exchange MD Simulations and NMR Experiments. *J. Phys. Chem. B* **2014**, *118*, 2851–2871.

(37) Franca, E. F.; Lins, R. D.; Freitas, L. C. G.; Straatsma, T. P. Characterization of Chitin and Chitosan Molecular Structure in Aqueous Solution. *J. Chem. Theory Comput.* **2008**, *4*, 2141–2149.

(38) Franca, E. F.; Freitas, L. C. G.; Lins, R. D. Chitosan Molecular Structure as a Function of N-Acetylation. *Biopolymers* **2011**, *95*, 448–460.

(39) Chen, C.; Colley, K. J. Minimal Structural and Glycosylation Requirements for ST6Gal I Activity and Trafficking. *Glycobiology* **2000**, *10*, 531–583.

(40) Fast, D. G.; Jamieson, J. C.; McCaffrey, G. The Role of the Carbohydrate Chains of Gal beta-1,4-GlcNAc alpha 2,6-Sialyltransferase for Enzyme Activity. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1993**, *1202*, 325–330.

(41) Kuhn, B.; Benz, J.; Greif, M.; Engel, A. M.; Sobek, H.; Rudolph, M. G. The Structure of Human  $\alpha$ -2,6-Sialyltransferase Reveals the Binding Mode of Complex Glycans. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 1826–1838.

(42) Shade, K.-T.; Anthony, R. Antibody Glycosylation and Inflammation. *Antibodies* **2013**, *2*, 392–414.

(43) Arnold, J. N.; Wormald, M. R.; Sim, R. B.; Rudd, P. M.; Dwek, R. A. The Impact of Glycosylation on the Biological Function and Structure of Human Immunoglobulins. *Annu. Rev. Immunol.* **2007**, *25*, 21–50.

(44) Barb, A. W.; Prestegard, J. H. NMR Analysis Demonstrates Immunoglobulin G N-glycans are Accessible and Dynamic. *Nat. Chem. Biol.* **2011**, *7*, 147–153.

(45) Wormald, M. R.; Rudd, P. M.; Harvey, D. J.; Chang, S. C.; Scragg, I. G.; Dwek, R. A. Variations in Oligosaccharide-Protein Interactions in Immunoglobulin G Determine the Site-Specific Glycosylation Profiles and Modulate the Dynamic Motion of the Fc Oligosaccharides. *Biochemistry* **1997**, *36*, 1370–1380.

(46) Yamaguchi, Y.; Kato, K.; Shindo, M.; Aoki, S.; Furusho, K.; Koga, K.; Takahashi, N.; Arata, Y.; Shimada, I. Dynamics of the Carbohydrate Chains Attached to the Fc Portion of Immunoglobulin G as Studied by NMR Spectroscopy Assisted by Selective <sup>13</sup>C Labeling of the Glycans. *J. Biomol. NMR* **1998**, *12*, 385–394.