SUPPORTING INFORMATION

Lipid-Linked Oligosaccharides in Membranes Sample Conformations that Facilitate Binding to Oligosaccharyltransferase

Nathan R. Kern, Hui Sun Lee, Emilia L. Wu, Soohyung Park, Kenno Vanommeslaeghe, Alexander D. MacKerell, Jr., Jeffery B. Klauda, Sunhwan Jo, and Wonpil Im

llo_dol.inp

```
* CHARMM input file to generate a eukaryotic LLO molecule:
* Glc3-Man9-GlcNAc2-PP-Dolichol
*

! read topology and parameter files

! carbohydrates
open read card unit 10 name top_all36_carb.rtf
read  rtf card unit 10

open read card unit 20 name par_all36_carb.prm
read para card unit 20 flex

! lipids
open read card unit 10 name top_all36_lipid.rtf
read  rtf card unit 10 append

open read card unit 20 name par_all36_lipid.prm
read para card unit 20 append flex

! CGENFF
open read card unit 10 name top_all36_cgenff.rtf
read  rtf card unit 10 append

open read card unit 20 name par_all36_cgenff.prm
read para card unit 20 append flex

! LLO-realted toppar (this file requires CGENFF atom types)
stream toppar_all36_llo.str


read sequence card
* Oligosaccharide chain
*
 15
!     1      2      3      4      5      6      7      8      9     10
 AGLCNA BGLCNA   BMAN   AMAN   AMAN           ! 5-sugar core
   AMAN   AMAN                                ! right (1-6) branch 1
   AMAN   AMAN                                ! right (1-6) branch 2
   AMAN   AMAN   AGLC   AGLC   AGLC           ! left  (1-3) branch

generate osac first none last none setup warn

patch 14bb osac  1  osac  2 setup warn    ! 14 AGLC
patch 14bb osac  2  osac  3 setup warn    !     | (a1-2)
patch 16ab osac  3  osac  4 setup warn    ! 13 AGLC
                                          !     | (a1-3)
patch 16ab osac  4  osac  5 setup warn    ! 12 AGLC
patch 12aa osac  5  osac  6 setup warn    !     | (a1-3)
                                          ! 11 AMAN       8 AMAN      6 AMAN
patch 13ab osac  4  osac  7 setup warn    !     | (a1-2)     | (a1-2)    | (a1-2)
patch 12aa osac  7  osac  8 setup warn    ! 10 AMAN       7 AMAN      5 AMAN
                                          !     |           |          /
patch 13ab osac  3  osac  9 setup warn    !     | (a1-2)     \ (a1-3) / (a1-6)
patch 12aa osac  9  osac 10 setup warn    !     |             \      /
patch 12aa osac 10  osac 11 setup warn    !  9 AMAN            4 AMAN
patch 13ab osac 11  osac 12 setup warn    !       \            /
patch 13ab osac 12  osac 13 setup warn    !        \ (a1-3)   / (a1-6)
patch 12ab osac 13  osac 14 setup warn    !         \        /
                                          !          3 BMAN
autogenerate angle dihe                   !             | (b1-4)
                                          !          2 BGLCNA
                                          !             | (b1-4)
                                          !          1 AGLCNA
                                          !             |
read sequence card                        !             |
*  dol-p-p                                !        DOL-PP
*                                         !
 1
 DL19PP

generate dolp first none last none setup warn
```

```
patch lllo osac 1 dolp 1 setup warn
autogenerate angle dihe

rename segid llo sele segid OSAC end
join   llo dolp renumber

ic para
ic edit
   dihe LLO  3 O5  LLO  3 C1  LLO  2 O4  LLO  2 C4  -81.720    ! from DB charmm input
   dihe LLO  3 C1  LLO  2 O4  LLO  2 C4  LLO  2 C3  131.780    ! 5 -> 7; 6 -> 5; 7 -> 9; 8 -> 10
   dihe LLO  2 O5  LLO  2 C1  LLO  1 O4  LLO  1 C4  -80.530
   dihe LLO  2 C1  LLO  1 O4  LLO  1 C4  LLO  1 C3   99.077
   dihe LLO  5 O5  LLO  5 C1  LLO  4 O6  LLO  4 C6   64.292
   dihe LLO  5 C1  LLO  4 O6  LLO  4 C6  LLO  4 C5 -177.657
   dihe LLO  4 O6  LLO  4 C6  LLO  4 C5  LLO  4 C4   55.887
   dihe LLO  4 O5  LLO  4 C1  LLO  3 O6  LLO  3 C6   65.555
   dihe LLO  4 C1  LLO  3 O6  LLO  3 C6  LLO  3 C5 -166.077
   dihe LLO  3 O6  LLO  3 C6  LLO  3 C5  LLO  3 C4   62.219
   dihe LLO  7 O5  LLO  7 C1  LLO  4 O3  LLO  4 C3   87.042
   dihe LLO  7 C1  LLO  4 O3  LLO  4 C3  LLO  4 C2  -96.428
   dihe LLO  9 C1  LLO  3 O3  LLO  3 C3  LLO  3 C2  -89.112
   dihe LLO  9 O5  LLO  9 C1  LLO  3 O3  LLO  3 C3   78.977
   dihe LLO 10 C1  LLO  9 O2  LLO  9 C2  LLO  9 C1  -83.494
   dihe LLO 10 O5  LLO 10 C1  LLO  9 O2  LLO  9 C2   84.619
end
ic seed 1 c1 1 c2 1 c3
ic build

energy

! orient LLO along Z
coor stat
coor orient sele resname dl19pp .or. resname undpp end
coor stat   sele resname dl19pp .or. resname undpp end
coor stat

open write unit 20 card name llo_dol.psf
write psf  unit 20 card

open write unit 20 card name llo_dol.pdb
write coor unit 20 pdb

open write unit 20 card name llo_dol.crd
write coor unit 20 card

stop
```

llo_und.inp

```
* CHARMM input file to generate a bacterial LLO molecule:
* Glc1-GalNAc5-Bac1-PP-Undecaprenol
*

! read topology and parameter files
! carbohydrates
open read card unit 10 name top_all36_carb.rtf
read  rtf card unit 10

open read card unit 20 name par_all36_carb.prm
read para card unit 20 flex

! lipids
open read card unit 10 name top_all36_lipid.rtf
read  rtf card unit 10 append

open read card unit 20 name par_all36_lipid.prm
read para card unit 20 append flex

! CGENFF
open read card unit 10 name top_all36_cgenff.rtf
read  rtf card unit 10 append

open read card unit 20 name par_all36_cgenff.prm
read para card unit 20 append flex

! LLO-realted toppar (this file requires CGENFF atom types)
stream toppar_all36_llo.str

read sequence card
* Oligosaccharide chain
*
 15
!     1      2      3      4      5      6      7      8      9     10
   ABAC   AGALNA AGALNA AGALNA AGALNA AGALNA
   BGLC

generate osac first none last none setup warn

patch 13ab osac  1  osac  2 setup warn    !     6 AGALNA
patch 14aa osac  2  osac  3 setup warn    !       | (a1-4)
patch 14aa osac  3  osac  4 setup warn    !     5 AGALNA      7 BGLC
patch 14aa osac  4  osac  5 setup warn    !          \          /
patch 14aa osac  5  osac  6 setup warn    !           \(a1-4)  / (b1-3)
                                          !            \      /
patch 13bb osac  4  osac  7 setup warn    !            4 AGALNA
                                          !               | (a1-4)
autogenerate angle dihe                   !            3 AGALNA
                                          !               | (a1-4)
                                          !            2 AGALNA
                                          !               | (a1-3)
                                          !            1 ABAC (bacillosamine)
read sequence card                        !               |
*  und-p-p                                !               |
*                                         !         UND-PP
 1                                        !
 UNDPP

generate dolp first none last none setup warn

patch lllo osac 1 dolp 1 setup warn
autogenerate angle dihe

rename segid llo sele segid OSAC end
join   llo dolp renumber

ic para
ic edit
dihe LLO 1 C1 LLO 8 O22 LLO 8 P2 LLO 8 O12 -90.00
end
ic seed 1 c1 1 c2 1 c3
ic build
coor print
```

```
energy

! orient LLO along Z
coor stat
coor orient sele resname dl19pp .or. resname undpp end
coor stat   sele resname dl19pp .or. resname undpp end
coor stat

open write unit 20 card name llo_und.psf
write psf  unit 20 card

open write unit 20 card name llo_und.pdb
write coor unit 20 pdb

open write unit 20 card name llo_und.crd
write coor unit 20 card

stop
```

OST Docking Procedure

The following docking procedure was done using CHARMM (1). 1. We defined a set of grid points to locate the pyrophosphate (PP) of the bacterial LLO within the PglB catalytic site. 2. For each LLO conformation taken every 0.3 ns from the Und_DOPC simulations, its PP COM was translated to a grid point, retaining its *Z* coordinate. We then performed rigid-body translational (±3 Å along the *Z*-axis) and rotational (±10° along *X* and *Y* by 2.5° rotation as well as 360° along *Z* by 10° rotation) searches to consider thermal motions of both LLO and PglB. This procedure was repeated for all the grid points for each LLO conformation. 3. To identify optimal docking poses of G1Gn5B1-PP-Und, we applied a set of constraints based on the proposed N-glycosylation mechanism of Pg1B: (i) the distance between the acceptor Asn N and sugar residue 1 C1 atoms ≤ 4 Å (**Figure 9C**), (ii) the distance between M^{2+} and P2 (phosphorus to Bac sugar resid) ≤ 6 Å, (iii) the distance between Arg-275 (OST) and P1 (phosphorus to the Und chain) ≤ 6 Å, and (iv) the angle formed by Asn N, 1 C1, and O22 (oxygen linking P1 to Bac) < 90°. Docking poses that did not satisfy these criteria were rejected and a best pose was chosen based on the lowest number of bad contacts between PglB and G1Gn5B1-PP among the remaining poses. The bad contact is defined by a cutoff distance of 2.5 Å between any heavy atoms from PglB and G1Gn5B1-PP. Note that we did not include the Und chain in checking for bad contacts because of its conformational flexibilities in the absence of PglB. Nonetheless, molecular docking of a large ligand like G1Gn5B1-PP is still challenging due to its large number of degrees of freedom. In our docking algorithm, we fixed the conformations of LLO and PglB, which introduced unavoidable bad contacts between them. Therefore, we allowed some structural tolerance in choosing a set of optimal docking poses by setting the maximum number of bad contacts to 20, which can be easily relieved by short minimization.

Table S1. Simulation system information.

System Name	# Lipids	# Water	# Ions	# Total Atoms	System Size ($Å^3$)
Dol_DLPC	98	6,408	17 K^+, 15 Cl^-	30,210	53 × 53 × 101
Dol_DMPC	98	5,714	16 K^+, 14 Cl^-	29,302	54 × 54 × 94
Dol_DOPC	98	5,527	16 K^+, 14 Cl^-	30,477	57 × 57 × 94
Und_DLPC	98	4,988	13 K^+, 11 Cl^-	25,718	53 × 53 × 85
Und_DMPC	98	5,033	14 K^+, 12 Cl^-	27,031	54 × 54 × 87
Und_DOPC	98	5,969	18 K^+, 16 Cl^-	32,031	57 × 57 × 89

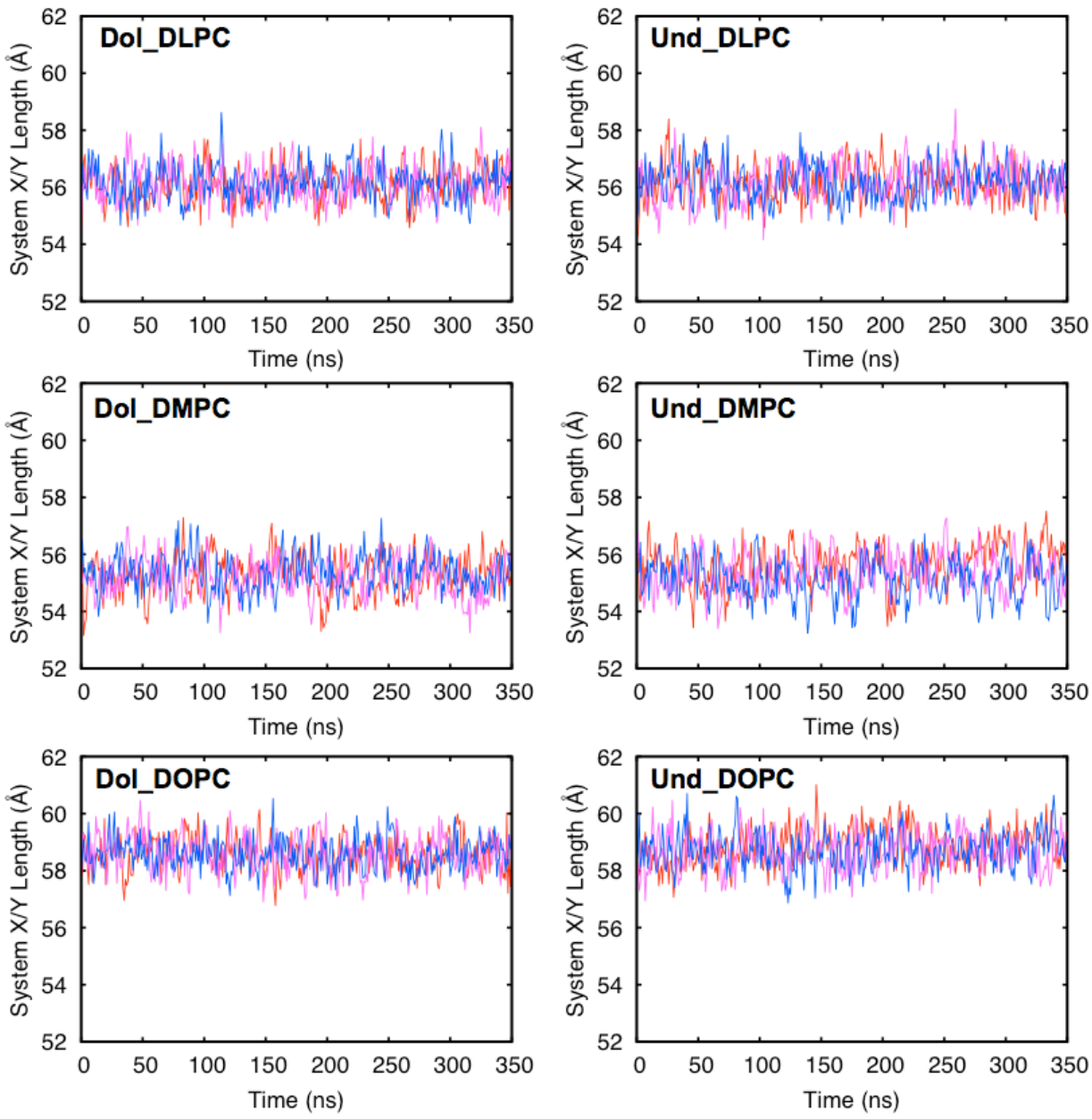

Figure S1. Times-series of the X/Y length in each system with different colors representing data from the three replicate simulations of each system.

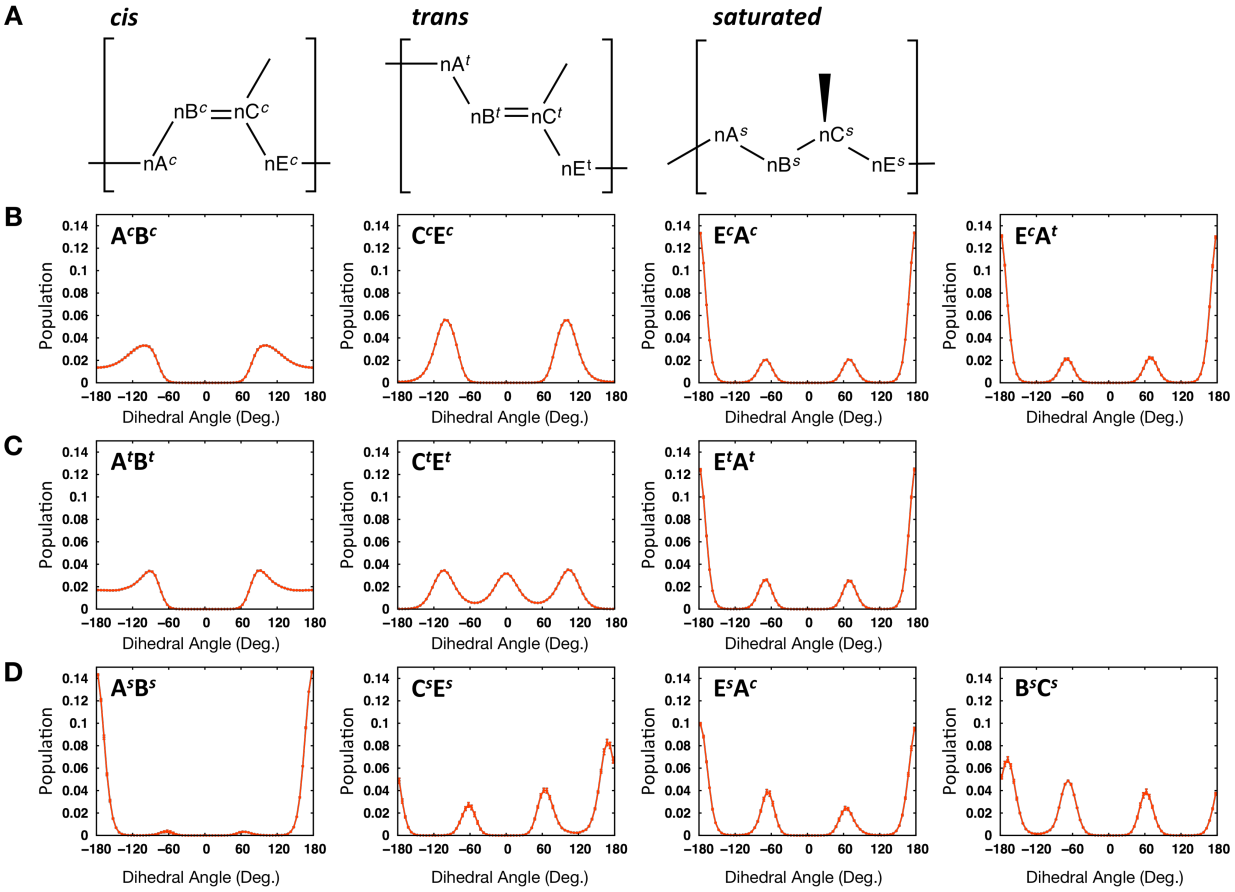

Figure S2. (A) Structure of the *n*-th *cis*-, *trans*-, and saturated isoprenyl units. Main chain carbon atoms are labeled as shown in the figure to define dihedral angles. Dihedral angles were categorized according to the bond about which the dihedral rotates. The following atoms were used for each dihedral definition in unsaturated isoprenyl units: C(n-1)E–CnA–CnB–CnC (dihedral AB), CnA–CnB–CnC–CnE (dihedral BC), CnB–CnC–CnE–C(n+1)A (dihedral CE), and CnC–CnE–C(n+1)A–C(n+1)B (dihedral EA). Dihedrals are defined similarly for isoprenyl units directly attached to diphosphate (saturated for Dol, unsaturated for Und), except that dihedral AB uses the atoms O11–C1A–C1B–C1C. (B) The dihedral angle distributions for *cis*-isoprenyl units: A^cB^c, C^cE^c, E^cA^c, and E^cA^t dihedrals. (C) The dihedral angle distributions for *trans*-isoprenyl units: A^tB^t, C^tE^t, and E^tA^t dihedrals. (D) The dihedral angle distributions for the saturated isoprenyl units: A^sB^s, C^sE^s, E^sA^c, and B^sC^s dihedrals (only from Dol). The distributions are the averages over all the simulations. The standard errors are shown as the error bars.

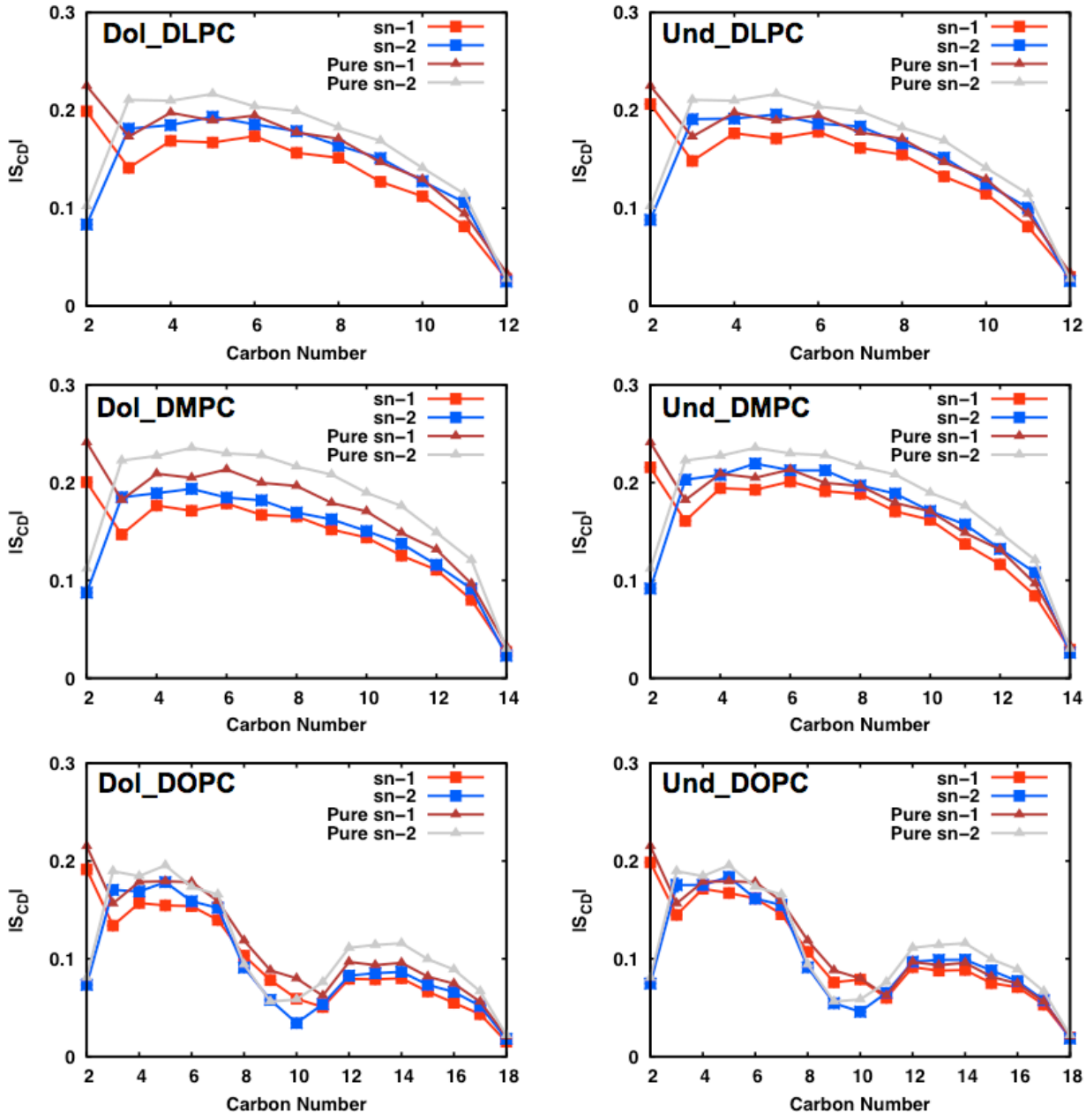

Figure S3. Deuterium order parameters ($|S_{CD}|$) of the *sn*-1 and *sn*-2 chains of DLPC, DMPC, and DOPC in each LLO system, which are compared with those of the corresponding pure bilayers. Only lipid atoms within 10 Å of each LLO center in *XY* were included for $|S_{CD}|$ calculations in the LLO systems. The $|S_{CD}|$ values are the averages of the three independent replicates, and the standard errors are smaller than the symbol sizes.

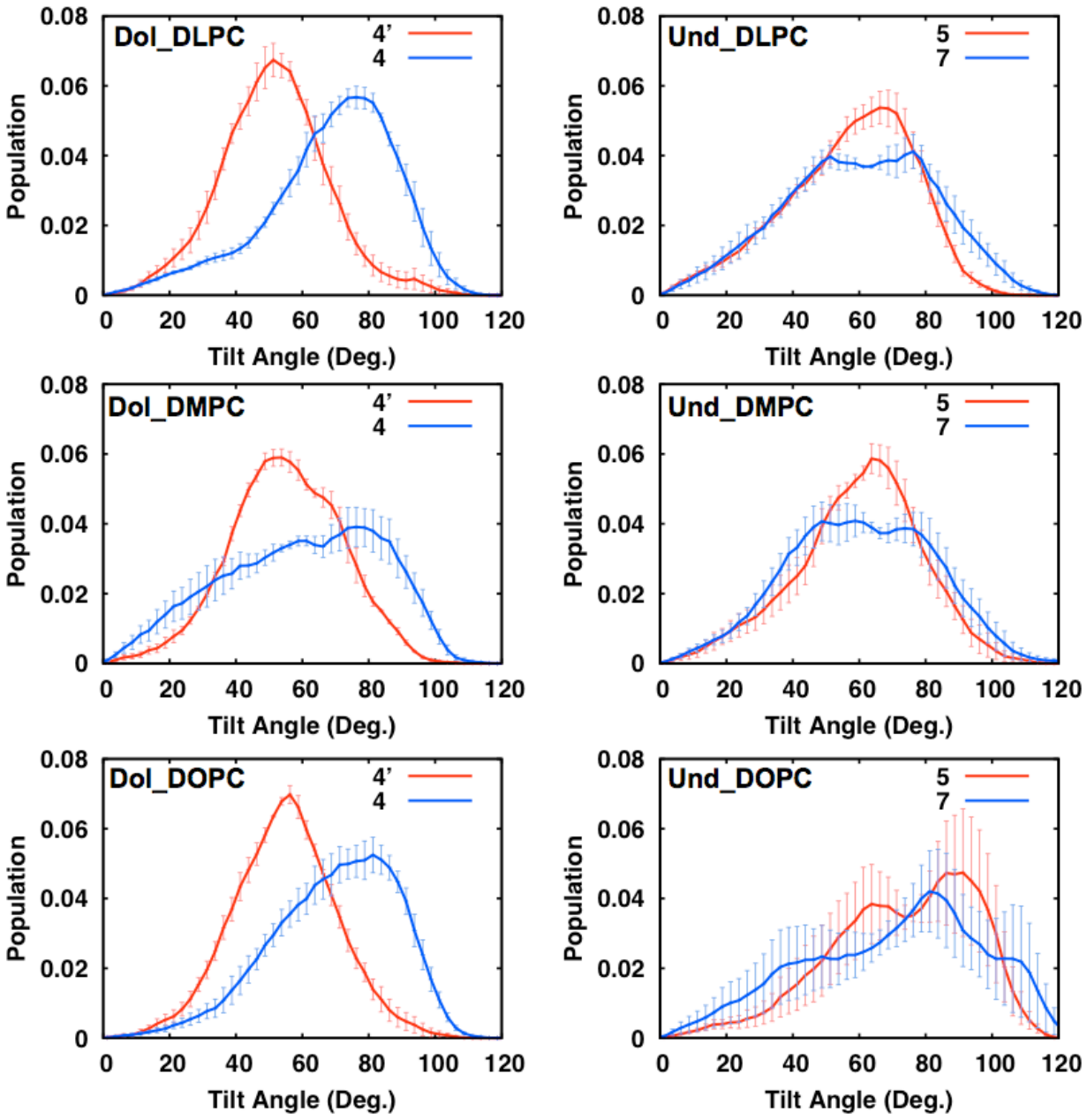

Figure S4. Distributions of tilt angles of the short (red) and long (blue) branches of the oligosaccharide with respect to the *Z*-axis in each system. The distributions are the averages of the three independent replicates, and the standard errors are also shown as the error bars. For G3M9Gn2, the eukaryotic LLO's oligosaccharide, the long-branch is defined by a vector from the C1 carbon of residue 1 GlcNAc and the C4 carbon of residue 4 Man and the short-branch by a vector from residue 1 C1 and residue 4' C4 (see **Figure 1A** for residue numbering). For G1Gn5B1, the bacterial LLO's oligosaccharide, the two representative vectors are defined by residue 1 C1 and residue 5 C4 (red) and by residue 1 C1 and residue 7 C4 (blue) (see **Figure 1B** for residue numbering).

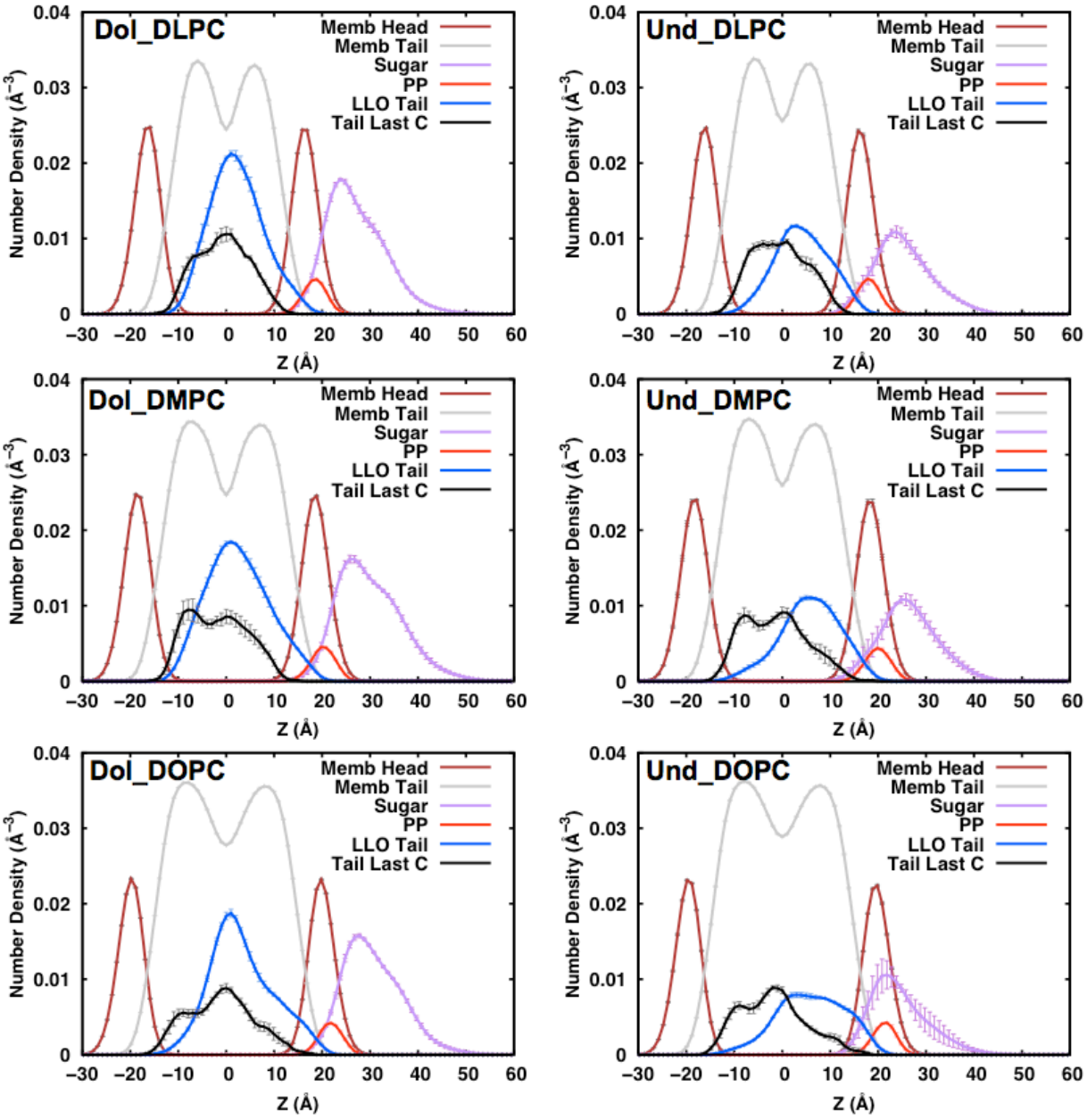

Figure S5. Density profiles of key components in each system: membrane head groups (brown), lipid acyl chains (grey), isoprenoid Dol/Und chain (blue), oligosaccharide (purple), pyrophosphate linkage (red), and the last carbon atom of Dol/Und chain (black). To aid viewing, the oligosaccharide distribution has been scaled by a factor of 5 and the pyrophosphate linkage, dolichol chain, and Dol/Und's last carbon distributions by 10. The profiles are the averages of the three independent replicates, and the standard errors are also shown as the error bars.

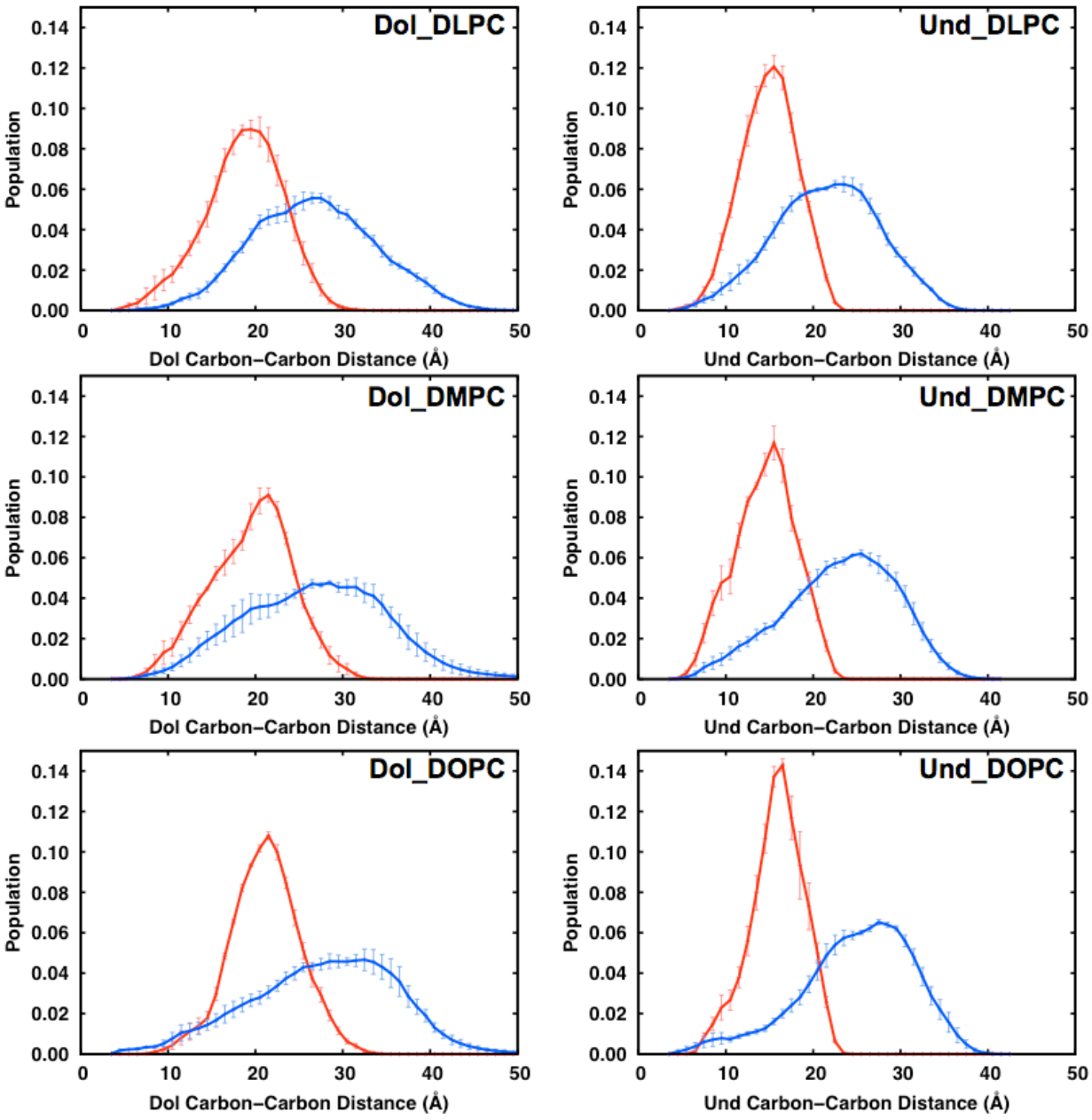

Figure S6. Distributions of the distances of C1A-C9E (red) and C1A-C19E (blue) in the Dol tail and C1A-C6E (red) and C1A-C11E (blue) in the Und tail in the DLPC, DMPC, and DOPC bilayers. For the Dol tail, C1A is the first carbon of Dol, C9E the carbon at the end of 9th isoprenyl unit, and C19E the carbon at the end of 19th isoprenyl unit. For the Und tail, C1A is the first carbon of Und, C9E the carbon at the end of 6th isoprenyl unit, and C19E the carbon at the end of 11th isoprenyl unit.

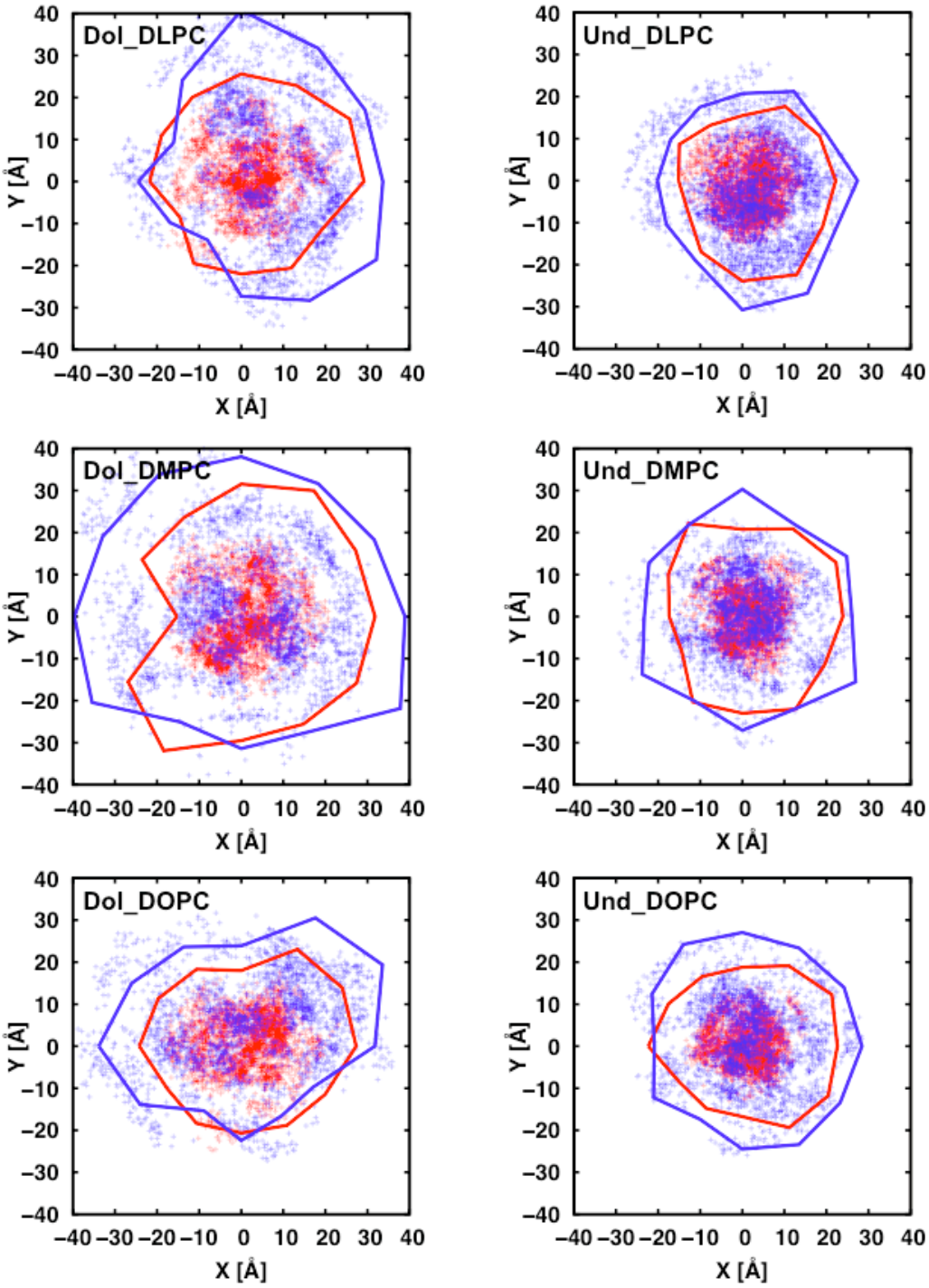

Figure S7. Distributions of the *XY* locations of C9E (red) and C19E (blue) of the Dol tail and C6E (red) and C11E (blue) of the Und tail. C1A in the Dol/Und tail is always positioned at $X = Y = 0$, and the long branch of the oligosaccharide (i.e., the 1C1-14C1 vector for the eukaryotic LLO and the 1C1-6C1 vector for the bacterial LLO) is aligned along the positive *X*-axis. See the **Figure S4** caption for atom naming and the **Figure 5** captions for the solid lines.

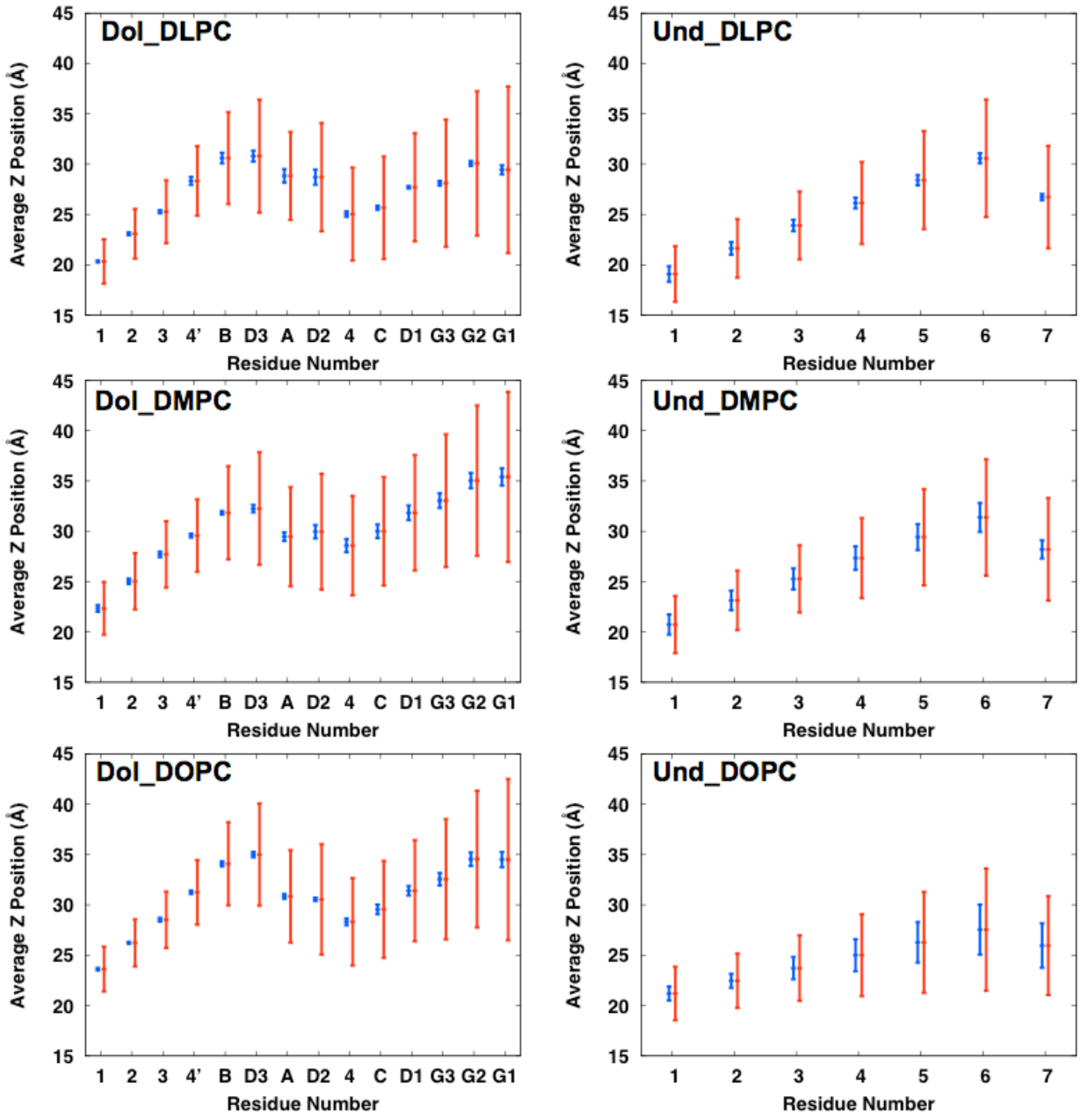

Figure S8. Average Z distance of the center of mass of each sugar residue from the bilayer center (i.e., Z = 0) in each system: the blue bars for the standard errors from the three independent simulations and the red bars for the standard deviations within each simulation (see the residue names in **Figure 1**).

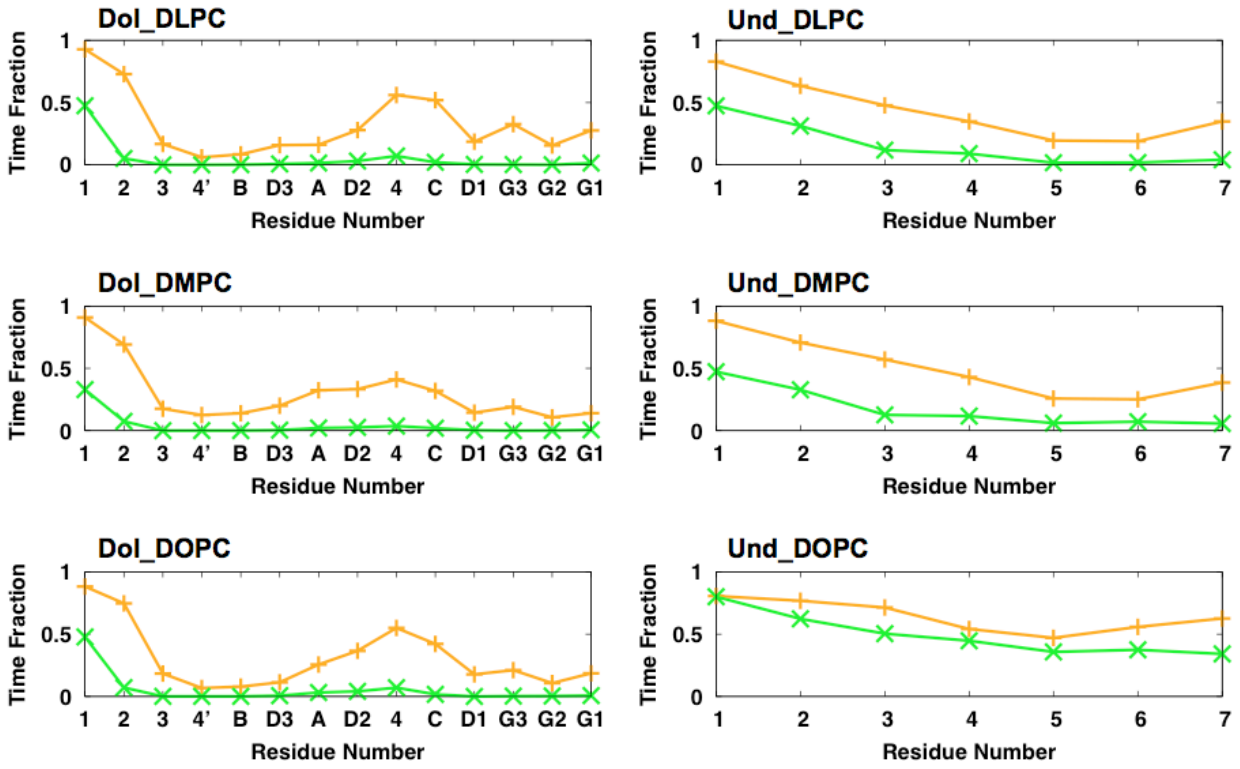

Figure S9. Time fraction of each sugar residue in contact with lipid head groups (orange) or tails (green) in each system. The contact is counted when any heavy atom of each sugar residue is within 4.5 Å from any heavy atom from the head groups or the tails. The values are the averages of the three independent replicates, and the standard errors are smaller than the symbol sizes for most cases.

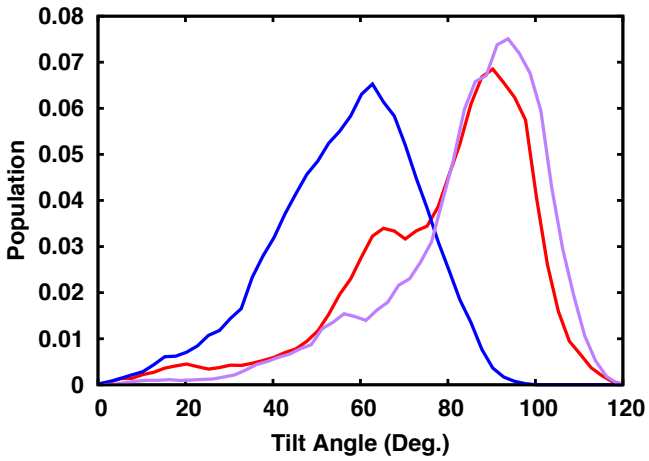

Figure S10. Distributions of tilt angles of the vector from oligosaccharide residue 1 C1 and residue 5 C4 with respect to the *Z*-axis in three independent Und_DOPC systems (with different colors). Their average distribution with the standard errors is shown in **Figure S4** Und_DOPC (red).

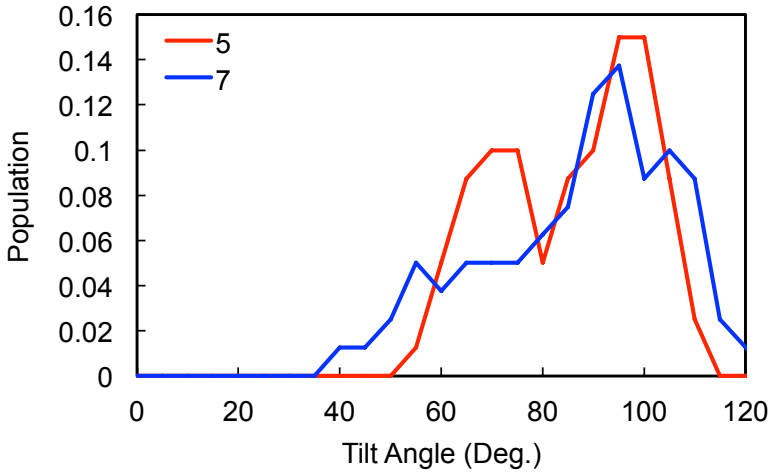

Figure S11. Distributions of tilt angles of the bacterial LLO's oligosaccharide (in DOPC) docked into the OST PglB. The two representative vectors are defined by residue 1 C1 and residue 5 C4 and by residue 1 C1 and residue 7 C4. For this plot, only LLO trajectories with successful docking poses were used.

Reference:

1. Brooks, B. R., C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus. 2009. CHARMM: The biomolecular simulation program. J. Comput. Chem. 30:1545-1614.