Published in final edited form as:

J Mol Biol. 2007 April 13; 367(5): 1523–1534.

Sugar Binding in Lactose Permease: Anomeric State of a Disaccharide Influences Binding Structure

Jeffery B. Klauda^{*} and Bernard R. Brooks

Laboratory of Computational Biology, National Institutes of Health, Bldg. 50, 50 South Drive, Bethesda, MD 20892

Abstract

Lactose permease in *E. coli* (LacY) transports both anomeric states of disaccharides but has greater affinity for α -sugars. Molecular dynamics (MD) simulations are used to probe the protein-sugar interactions, binding structures, and global protein motions in response to sugar binding by investigating LacY (the experimental mutant and wild-type) embedded in a fully hydrated lipid bilayer. A total of twelve MD simulations of 20–25ns each with $\beta(\alpha)$ -D-galactopyranosyl-(1,1)- β -D-galactopyranoside ($\beta\beta$ -(Galp)₂) and $\alpha\beta$ -(Galp)₂ result in binding conformational families that depend on the anomeric state of the sugar. Both sugars strongly interact with Glu-126 and $\alpha\beta$ -(Galp)₂ has a greater affinity to this residue. Binding conformations are also seen that involve protein residues not observed in the crystal structure, as well as those involved in the proton translocation (Phe-118, Asn-119, Asn-240, His-322, Glu-325, and Tyr-350). Common to nearly all protein-sugar structures, water acts as a hydrogen bond bridge between the disaccharide and protein. The average binding energy is more attractive for $\alpha\beta$ -(Galp)₂ than $\beta\beta$ -(Galp)₂, i.e., -10.7 ± 0.7 and -3.1 ± 1.0 kcal/mol, respectively. Of the twelve helices in LacY, Helix-IV is the least stable with $\beta\beta$ -(Galp)₂ binding resulting in larger distortion than $\alpha\beta$ -(Galp)₂.

Keywords

Lactose Permease; Sugar Binding; Disaccharides; Molecular Dynamics; Membrane Proteins

Introduction

Membrane transport proteins play significant roles in human physiology, drug transport, and diseases, such as cystic fibrosis, but detailed knowledge about their structure and transport mechanisms has been limited partially due to the difficulty in protein crystallization for x-ray diffraction.¹ Although membrane transport proteins are grouped into many families, the major facilitator superfamily (MFS) is very diverse in terms of substrate transport^{2,3} with over 2000 sequenced members. The MFS proteins can transport many small to moderately sized molecules of biological importance, such as drugs and other hydrophobic substances, sugars and their derivatives, organic and inorganic anions, organocations, siderophore iron complexes, and peptide or peptide-like compounds.³ These proteins can function by uniport (one molecule is transported), cation symport (two molecules are transported in the same direction), and cation or solute antiport (two molecules are transported in the opposite direction).²

^{*}To whom correspondence should be addressed: klauda@helix.nih.gov, Phone: (301) 496-9510, Fax: (301) 402-3404

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The atomic-level structure has been determined for only three MFS proteins (lactose permease,⁴ α -glycerophosphate:phosphate antiporter,⁵ and EmrD multidrug transporter⁶) and only in the state open to the cytoplasm, i.e., the inward facing structure. However, the common topology for most proteins of the MFS is twelve transmembrane α -helices, which was determined by experiment and structure prediction.^{2,7} The first six helices are typically similar in sequence to the last forming a two domain structure.³ Lactose permease (LacY) of *E. coli* (a 417 residue protein, Fig. 1) is an important structural and transport model for the MFS proteins¹ because of the common topology and transport mechanisms of MFS proteins. As such, LacY (a cation symporter of disaccharides^{1,8}) is representative of all transmembrane proteins in Archaea to the mammalian central nervous system that act as cation symporters.

The 3.5 Å structure of a mutated LacY with the bound lactose homolog, β -D-galactopyranosyl-1-thiol- β -D-galactopyranoside (TDG), has been determined by Abramson et al.⁴ This is a crystal structure for the protein mutant with Gly replacement of Cys¹⁵⁴, referred to here as LacY-C154G, which prevents the translocation of the disaccharide and is essentially frozen in the inward facing state.⁹

The structure of the LacY-C154G was obtained with only one bound disaccharide, but biochemical experiments on other sugars have suggested some similarities and differences in sugar binding and transport. A galactopyranosyl ring of D-configuration with a free 6-hydroxyl group, i.e., not attached to another ring, is required for disaccharide transport in LacY.¹⁰ Moreover, 4- and 3-hydroxyl are important in galactoside recognition of LacY¹¹ and the order of importance to binding affinity is 4-hydroxyl >> 6-hydroxyl > 3-hydroxyl.¹² The protein specificity is independent of the anomeric state of the sugar, but the affinity of α -anomeric states of disaccharides is greater than β .^{11,13} The atomic-level description for the cause of the anomeric affinity is unknown. However, these experiments suggest that either the binding structure is different for the two stereochemical states of a disaccharide or the protein-sugar interaction is greater for the α -anomeric state. The main focus of this paper is to describe the anomeric binding of disaccharides in LacY using molecular simulations.

Simulation techniques have only been used in two studies of LacY to determine helix stability¹⁴ and protein structural changes.¹⁵ Bennett et al.¹⁴ used molecular dynamics (MD) to study the stability of the transmembrane helices in a solvent, but each helix was simulated separately. Yin et al.¹⁵ were the first to perform simulations with LacY-wt embedded in a lipid bilayer, where structural changes due to proton translocation from Glu-325 to Glu-269 were studied. This work consisted of a total of only three simulations of 10 ns each and found a relatively fast and global structural change with proton translocation. MD simulations of other protein-sugar complexes have been limited to implicit water models,¹⁶ protein docking,¹⁷ or solvation in water.¹⁸

This work is believed to be the first to investigate how subtle changes in stereochemistry influence the binding of a substrate to a protein. LacY is simulated in a fully hydrated lipid bilayer (Fig. 1a). Although this adds complexity to the simulations, it mimics the protein environment in the cytoplasmic membrane of *E. coli* and is essential in determining the importance of water in sugar binding. Molecular dynamics simulations of the disaccharide β (α)-D-galactopyranosyl-(1,1)- β -D-galactopyranoside, (Galp)₂, in two anomeric forms ($\alpha\beta$ and $\beta\beta$) are used to determine the cause of the anomeric binding affinity. For reference, the structure of $\beta\beta$ -(Galp)₂ is shown in Fig. 2. MD simulations are performed with the experiment mutant (LacY-C154G) and the wild-type protein (LacY-wt). The initial system used for this study is shown in Fig. 1, where the protein structure is inward facing (towards the cytoplasm).⁴ Details of the disaccharide initial placement is given in the Methods section, but the sugar is free of any constraints during the MD simulations. Since the protein is open to the aqueous phase, water fills the lumen of the protein (Fig. 1b) but the closed end of the protein prevents water

transport across the membrane. Twelve MD simulations with three initial sugar configurations are run for a total of 20-25ns each to obtain a detailed description of disaccharide binding in LacY.

Results

Binding of $\beta\beta$ -(Galp)₂ in LacY-C154G

From the x-ray diffraction of LacY-C154G/TDG,⁴ distances consistent with hydrogen bonding are observed between the protein/disaccharide, i.e., Arg-144/O3 (and possibly O4 on the sugar) and Lys-358/O4' (Fig. 3a). (Hydrogen bonds are classified between a donor and acceptor if the distance is less than 3.2 Å.) The notation used for the disaccharide atoms is described in the caption of Fig. 2. Protein/sugar interactions via a water molecule have been suggested to occur with Glu-126/O6 and Asp-237/O4' based x-ray diffraction density maps, but water could not be resolved at the 3.5 Å resolution.⁴ Initially, three simulations of $\beta\beta$ -(Galp)₂ inserted near the binding site are used to verify the accuracy of our methods and force fields to reproduce experiment. Table 1 lists the total simulation time for the three simulations. The disaccharide in the x-ray crystal structure, TDG, contains galactose moieties linked by a C-S-C linkage, whereas the disaccharide used in our simulations has a C-O-C glycosidic linkage. Although $\beta\beta$ -(Galp)₂ differs from TDG, simulated binding structures agree with the experimental crystal structure as shown in Fig. 3(b-d). This agreement is specific to the formation of hydrogen bonds between the same oxygen atoms on the sugar and the corresponding acceptor or donor amino acid. However, the water-bridging between the protein and sugar is different between the runs and crystal structure. The suggested water-bridging between Glu-126 and O6⁴ is not observed for these crystal-like structures; rather a direct hydrogen bond is made between the disaccharide and amino acid for all three runs. Instead, a water-bridge with Glu-269 occurs in **2** and **3** and with Asp-237 in **1** (simulation runs are defined in bold numbers and described in the Methods section).

For each simulation, a time series for the total number of experimentally observed hydrogen bonds is shown in Fig. 4 (water-bridged hydrogen bonds in the simulations are not included in the counting). From the experimental structure,⁴ there are five possible protein/sugar interactions; two are direct hydrogen bonds between the protein and sugar (Arg-144/O3 and Lys-358/O4'), see Fig. 3a. The three other protein/sugar interactions (Glu-126/O6, Arg-144/O4, and Asp-237/O4') have larger distances and suggest weaker bonding or water-bridging. Since MD simulations at physiological temperature inherently result in more flexible binding than that of the crystal structure at 100K, the three weaker interactions have been including in the hydrogen bond counting for Fig. 4. The amount of crystal-like hydrogen bonds is consistent throughout most of the simulation **1** (1-2 direct hydrogen bonds), where as **2** and **3** have shorter lived states.

In addition to hydrogen bonds, hydrophobic interactions exists between $\beta\beta$ -(Galp)₂ and the indole ring of Trp-151 and the Sδ atom of Met-23 in the simulations and experimentally with TDG.⁴ The distance between the C6 atom on $\beta\beta$ -(Galp)₂ and Sδ atom of Met-23 is consistent with a hydrophobic interaction (< 6Å with an average separation of 4.3Å) for all the crystal-like configurations. Similarly, the ring-ring distance between the disaccharide and Trp-151 indicates a van der Waals contact during these configurations. The three crystal-like conformational families in Fig. 3 have a population of 9%, 9%, and 11% for the total simulation time of **1**, **2**, and **3**, respectively. For the remainder of the simulation, additional binding conformations or transitions between various conformations are observed.

As shown in Fig. 5, during the three simulations runs the number hydrogen bonds between $\beta\beta$ -(Galp)₂ and amino acids near the binding site ranges from 0 to 6. The probability of having one or more hydrogen bonds is greater than 87% and 48% for 2 or more. Many of these sugar-

protein hydrogen bonds are formed with amino acids not observed in the crystal structure (compare Figs. 4 and 5). Examples of two conformational families that differ from the crystal-like structure are shown in Fig. 6. A snapshot of a conformational family that occurs at the end of simulation **2** is shown in Fig. 6a and has a population of 34%. Interactions with amino acids in the crystal structure (Glu-126, Arg-144, and Asp-237) exist but with different oxygens on $\beta\beta$ -(Galp)₂. Moreover, Arg-144 and Asp-237 have water-bridged hydrogen bonds. Asn-119 and Asp-240 also form direct hydrogen bonds with the sugar and were not observed in the crystal structure.⁴ Asp-240 is not directly involved with proton translocation but is located near the site which a proton is transferred.¹ The shorter-lived conformational family (2.3%) in Fig. 6b has interactions with Tyr-350 that has been recently determined to be essential to sugar transport via Cys-scanning.¹⁹ For **3**, Phe-118 interacts with O3 on $\beta\beta$ -(Galp)₂ forming a protein backbone hydrogen bond with the disaccharide.

The final conformation of the $\beta\beta$ -(Galp)₂ simulations is shown for **2** in Fig. 6a and **1** and **3** in Fig. 7. As described previously, the end of **2** consists of a conformation that differs from the crystal structure. The final conformation for **1** maintains the two crystal-like hydrogen bonds: Glu-126/O6 and a water mediated interaction of Lys-358/O4'. The sugar in **3** appears to be leaving the binding site (Fig. 7b, left). Three interactions exist in this final conformation: Glu-269/O6', Glu-269/O4' and a water mediated interaction of Glu-126/O3'. There also is a switch in protein-protein interactions from Arg-144 interacting with Glu-126 to Glu-269.

Glu-126 has the highest propensity to form hydrogen bonds with a sugar and on average occurs 35±9% of the simulation time for the three runs (see Table 1a). Asp-237 has a similar population of sugar hydrogen bonds. Four other residues (Phe-118, Asn-119, Arg-144, and Lys-358) have sugar-amino acid hydrogen bond populations that are on average greater than 10%. Although Glu-269 is important in most of the binding structures of $\beta\beta$ -(Galp)₂, it primarily interacts via a water bridge.

The average binding energy of $\beta\beta$ -(Galp)₂ in LacY-C154G, $\langle\Delta E^{\text{bind}}\rangle$, is listed in Table 1a. In general, there is an inverse relationship with the number of protein-sugar hydrogen bonds and $\langle\Delta E^{\text{bind}}\rangle$, i.e., an increase in the number of hydrogen bonds result in a decrease (more attractive) in the binding energy. For these three runs, the average $\langle\Delta E^{\text{bind}}\rangle$ is -3.8±1.7 kcal/mol.

Binding of $\alpha\beta$ -(Galp)₂ in LacY-C154G

For the three runs of $\alpha\beta$ -(Galp)₂ with LacY-C154G, negligible occurrences of crystal-like structures are found. The only crystal-like hydrogen bond occurs between Glu-126 and O6 on $\alpha\beta$ -(Galp)₂ for **2** and **3**, which occurs with a population of 25%, whereas **1** does not exhibit any crystal-like hydrogen bonds. Although the sugar in **2** has a significant population of hydrogen bonds with Glu-126 (see Table 1a), the sugar is escaping the binding site region towards the open end of the protein.

All the binding conformational families of $\beta\beta$ -(Galp)₂ consist of interactions with both domains of the protein. However, **1** contains a binding structure of $\alpha\beta$ -(Galp)₂ that interacts with a single domain (Fig. 8a). The sugar interacts with the first domain (helices I-VI) and has four direct protein/sugar hydrogen bonds Glu-126/O2 (and O3), Arg-144/O2, and Glu-269/O6' and a water mediated interaction with Phe-118/O4. Similar to $\beta\beta$ -(Galp)₂, the Phe-118/O4 hydrogen bond forms with the backbone carbonyl. This binding conformational family has a population of 16.7%.

Three other binding conformational families are observed for $\alpha\beta$ -(Galp)₂ in LacY-C154G and one of these is shown in Fig. 8b. Three protein/sugar hydrogen bonds are found in this family: Asn-119/O6, Asp-237/O2, and Glu-126/O3' (mediated by water). The hydrogen bond with Asp-237/O2 is weak and exchanges with Lys-358/O2 and consequently forms a hydrogen bond

with Asp-237/O3 (not shown). Together this conformational family has a population of 30.6% with 16.5% of the time spent interacting solely with Asp-237/O2. A less common conformational family is observed (5.8%) forming hydrogen bonds with Tyr-350/O3' and Glu-126/O6.

Similar to $\beta\beta$ -(Galp)₂ binding in LacY-C154G, Glu-126 has the highest propensity to form hydrogen bonds with $\alpha\beta$ -(Galp)₂. However, the population of forming such a hydrogen bond is much greater than with the $\beta\beta$ -anomeric state; excluding the escaping sugar run (2), the population is 64±22% (all average calculations henceforth will not include 2). Protein/sugar hydrogen bond populations with Arg-144 and Asp-237 are high for 1 and 3, respectively. Moreover, Asn-119, Glu-269, and Lys-358 all form stable hydrogen bonds with $\alpha\beta$ -(Galp)₂.

The average $\langle\Delta E^{\text{bind}}\rangle$ of $\alpha\beta$ -(Galp)₂ for the two runs that the sugar remain in the binding site is more attractive than $\beta\beta$ -(Galp)₂ (−11.4±1.2 kcal/mol). This is consistent with experimental results that the affinity of α -sugars is greater than β -sugars in LacY.^{11,13} The strength of protein-sugar interactions is greatly reduced for the escaping sugar (2) and thus the $\langle\Delta E^{\text{bind}}\rangle$ increases by nearly 7 kcal/mol.

Binding of (Galp)₂ in LacY-wt

Both anomeric states of (Galp)₂ are simulated with LacY-wt and result in only a single conformational family that is consistent with that measured from x-ray diffraction.⁴ Three protein/sugar binding structures of $\beta\beta$ -(Galp)₂/LacY-wt are similar to those with the LacY-C154G; 1. Asp-237/O3' and Lys-358/O4' (6% of 2), 2. Asp-237/O2'(or O3') and Lys-358/O6' (or O2') (11% of 3), and 3. Gly-126/O2(and O3), Asn-119/O4', and Lys-358/O6' (via water) (5% of 3). The conformational family shown in Fig. 9b (12.5% of 2) varies from other binding conformations and has contacts with Asp-237/O4', Asp-240/O3', Asp-269/O6, and Glu-325/O6'. Glu-325 is protonated in these conformations to conform with the inward facing structure (towards the cytoplasm) and this residue is involved in proton translocation.^{1,8} $\beta\beta$ -(Galp)₂ in 1 quickly escapes the binding site and is not included in the remainder of the population analysis.

The $\alpha\beta$ -(Galp)₂/LacY-wt simulations resulted in only a single conformational family consistent with that measured from x-ray diffraction.⁴ Three protein/sugar binding conformational families of $\alpha\beta$ -(Galp)₂ are obtained that are similar to those with LacY-C154G: 1. Glu-126/O3(or O2), Arg-144/O2, and Glu-269/O6' (7% of 1 and 24% of 3), 2. Glu-126/O3'(or O2') and Asn-119/O2 (41% of 1), and 3. Glu-126/O6, Asn-119/O2(or O3 via water), Asp-237/O4'(or O3'), and Lys-358/O4'(or O3') (5% of 2 same protein residues of 3 $\beta\beta$ -(Galp)₂/LacY-C154G but different sugar oxygens). An additional binding conformational family similar to the crystal (Fig. 3a) forms the following hydrogen bonds is obtained during 2; Glu-126/O6 and an alternating Lys-358/O4'(or O3') (32%) and Asp-237/O4'(or O3') (29%). An alternate binding conformational family shown in Fig. 9a (12% of 3) contains hydrogen bonds with Glu-126/O2(or O3), Asp-240/O3'(via water), Glu-269/O6', and His-322/O4'.

The propensity for $\alpha\beta$ -(Galp)₂ to form hydrogen bonds with Glu-126 in LacY-wt (90% on average) is significantly larger than in LacY-C154G (see Table 1). Similar to LacY-C154G, stable hydrogen bonds are formed with Arg-144, Asn-119, Glu-269, and Lys-358. However, the propensity of forming hydrogen bonds with Asp-237 is reduced with the wild-type protein. The population of forming a hydrogen bond between Glu-126 and $\beta\beta$ -(Galp)₂ is reduced in LacY-wt (22±4%). The average hydrogen bond population between the protein and $\beta\beta$ -(Galp)₂ in LacY-wt is similar to LacY-C154G, i.e., sugar hydrogen bonds with Phe-118, Asn-119, Asp-237, and Lys-358. However, there is a decrease in forming hydrogen bonds with Arg-144 and an increase with Asp-240.

The average $\langle \Delta E^{\text{bind}} \rangle$ is lower for $\alpha\beta$ -(Galp)₂ compared to $\beta\beta$ -(Galp)₂ (-10.3 ± 0.9 and -2.0 ± 0.6 kcal/mol, respectively) in LacY-wt. The difference between the averages $\langle \Delta E^{\text{bind}} \rangle$ of $\alpha\beta$ - and $\beta\beta$ -(Galp)₂ is similar to that of simulations with LacY-C154G.

Protein and Lipid Bilayer Structure

The lipid bilayer maintains a stable liquid crystalline (L_α) phase throughout the MD simulations. An example of the final configuration of lipids is shown in the left column of Fig. 7. The deuterium order parameters (S_{CD}) of the lipid chains are constant with that of the disordered L_α state, i.e., $\langle |S_{\text{CD}}| \rangle$ for chain1 is 0.16 ± 0.01 from MD compared to experiment 0.20 .²⁰ Moreover, the thickness of the lipid bilayer remains stable with an average phosphate distance between leaflets of 35.4 ± 0.2 Å.

As a measure of protein structural change from the x-ray crystal structure (1PV7),⁴ the root-mean squared deviation (RMSD) of the backbone atoms is calculated for each of the simulations. Overall the RMSD of the protein is 2.97 ± 0.14 Å, but the RMSD for the all helices except H-IV is relatively low (0.90 ± 0.03 Å), Fig. 10. A RMSD of 2–4 Å is typical for MD simulations with proteins of similar size.^{18,21–23} The C154G mutant has little effect on overall backbone structure, i.e., the overall RMSD is 3.03 ± 0.23 Å and 2.90 ± 0.19 Å for LacY-C154G and LacY-wt, respectively. Moreover, the RMSD for the helix with the mutated residue (H-V) are statistically equivalent.

The average RMSD, %helicity, and angle between the two domains (θ_{domain}) of LacY is listed in Table 2 for each simulation. A residue is considered to be in an α -helix conformation when $\varphi = -57 \pm 40^\circ$ and $\psi = 47 \pm 40^\circ$. Only two simulations have a final average RMSD of greater than 4 Å and subsequently have a larger $\langle \theta_{\text{domain}} \rangle_f$ compared to the other simulations. Embedding LacY into a POPE bilayer results in a $\langle \theta_{\text{domain}} \rangle$ that is less than the experimental structure (Table 2). Overall, the %helicity remains stable for all the simulations and identical to the value from experiment. The protein structure remains stable through the course of these 20–25 ns simulations because there is little change in these three protein structural measures.

The anomeric state of the sugar influences the local structure of LacY, see Fig. 10. The overall RMSD is higher for the $\alpha\beta$ -(Galp)₂ simulations (3.16 ± 0.24 Å) compared to $\beta\beta$ -(Galp)₂ (2.76 ± 0.11 Å). The binding of $\alpha\beta$ -(Galp)₂ increases the RMSD from the crystal structure for the second domain. This change in structure is not the result of helix deformation, rather the flexibility of the loops regions account for the larger RMSD. The binding of the two anomeric states of the disaccharide have no appreciable effect on the helix structure except H-IV, see Fig. 10. The RMSD of H-IV with $\beta\beta$ -(Galp)₂ is higher (1.99 ± 0.15 Å) than that with $\alpha\beta$ -(Galp)₂ (1.57 ± 0.06 Å). This helix contains three of the important residues to protein-sugar binding (Phe-118, Asn-119, and Glu-126). From Fig. 11, $\beta\beta$ -(Galp)₂ distorts the end of H-IV that is closest to Glu-126.

Salt bridges and cross-amino acid hydrogen bonds form readily near the binding site and proton translocation site.^{1,24,25} The salt bridge of Asp-237/Lys-358 and hydrogen bond between Tyr-236/Arg-302 are stable throughout the MD simulations (populations of greater than 85%). Only **3** with LacY-C154G/ $\beta\beta$ -(Galp)₂ resulted in a significant population (>10%) of the hydrogen bond pair Glu-269/Arg-144 (46%). The more common pairing is between Glu-126 and Arg-144. From the crystal structure,⁴ Arg-144 may be at an average position between Glu-126 and Glu-269 suggesting interactions with both glutamic acid residues. The hydrogen bond structure between Tyr-236/His-322 important for proton translocation^{4,8} populates these simulations on average 14%. Hydrogen bonds between the C α on Gly-150 and –147 and the backbone of Phe-20 and –16 suggested by Ermolova et al.²⁶ are only rarely seen in our simulations (<1% on average). Consistent with both escaping sugar simulations is an increase population of the Asp-240/Lys-319 salt bridge, i.e., $18.6 \pm 3.8\%$ versus $5.6 \pm 1.1\%$.

Discussion

The initial simulations with $\beta\beta$ -(Galp)₂/LacY-C154G demonstrate that our methods can reproduce sugar binding conformations similar to that measured experimentally.⁴ The structure of LacY-C154G⁴ is with a sugar that contains a sulfur glycosidic linker (TDG) and as a result the sugar can extend about 0.5Å farther than the more native-like sugar $\beta\beta$ -(Galp)₂ with an oxygen linker. Therefore, the ability and affinity to form cross-domain interactions is increased for TDG and one should not expect direct agreement with our simulations. Nevertheless, our binding results are quite similar to that of x-ray diffraction, see Fig. 3. Consistent with most binding structures from our simulations are water-mediated protein-sugar interactions, which has been suggested previously by Abramson et al.⁴ On occasion, the bridging water is the only contact the sugar has with one of the domains, see Fig. 3b.

In addition to crystal-like conformations, two additional conformational families are found for $\beta\beta$ -(Galp)₂/LacY-C154G, see Fig. 6. These contain additional interactions with residues not observed experimentally in the crystal structure: Phe-118, Asn-119, Asp-240, and Tyr-350. However, the latter residue has been recently found to be important in sugar transport.¹⁹

An important result of this work is that the anomeric state of the sugar influences binding conformations, i.e., binding with $\alpha\beta$ -(Galp)₂ results in different conformations compared to the crystal structure with $\beta\beta$ -TDG. Protein-sugar hydrogen bonds can exist with similar amino acids as the $\beta\beta$ -(Galp)₂ but with differing sugar oxygens. Only **2** of $\alpha\beta$ -(Galp)₂/LacY-wt forms a binding conformational family that is similar to the crystal structure. $\alpha\beta$ -(Galp)₂ is also able to form a binding conformational family that interacts solely with the first domain of LacY, e.g., Fig. 8a, **3** of $\alpha\beta$ -(Galp)₂/LacY-C154G, and **1** of $\alpha\beta$ -(Galp)₂/LacY-wt.

The single point mutation of LacY also influences the binding of (Galp)₂ (Galp)₂. The binding of (Galp)₂ in LacY-wt results in protein-sugar interactions with amino acids involved in proton translocation, i.e., Asp-240, His-322, and Glu-325. Moreover, there is more sugar binding conformational families in LacY-wt compared to the experimental mutant LacY-C154G, which implies the wild-type protein promotes conformational flexibility. This is supported by recent isothermal titration calorimetry experiments by Nie et al.²⁷ with p-nitrophenyl α -D-galactopyranoside (NPG). Sugar binding in LacY-wt was found to have an increase in entropy compared to LacY-C154G, which suggests larger conformational flexibility in sugar binding for the wild-type protein.

The existence of several conformational families for sugar binding in MD simulations is not surprising. These simulations inherently include more flexibility because the sugar and protein dynamics are explicitly included, where as x-ray crystallography is used to obtain a single structure. Since the x-ray crystallography structure was obtained at low temperatures (100K), the flexibility of the protein and sugar is small. Artificial structural averaging can also exist in these experimental structures, e.g., Arg-144 is between Glu-126 and Glu-269 (Fig. 3a) and there appears to be a dynamic hydrogen bond switch between arginine and one glutamic acid. In most of our simulations Arg-144 forms a hydrogen bond with Glu-126, but it can also interact with Glu-269 (Fig. 6b and 7b). The advantage of using MD simulations is the ability to directly observe the inherent dynamics and flexibility of substrate binding and protein motions.

Separate from the individual conformational families, there are some distinct differences and similarities in specific amino acid-sugar hydrogen bonds for the two anomeric states of (Galp)₂. On average (excluding the two escaped sugar runs) five residues form hydrogen bonds at least 10% of the simulated time for both anomeric states, i.e., Glu-126, Asn-119, Arg-144, Asp-237, and Lys-358. The protein residue that forms the most stable hydrogen bonds with the disaccharides is Glu-126. This result and the 17±5% population of hydrogen bonds with

Arg-144 are consistent with the experimental observation that Glu-126 and Arg-144 are irreplaceable in sugar binding.²⁴ However, there are several differences in the protein-sugar interactions for $\alpha\beta$ - and $\beta\beta$ -(Galp)₂. The affinity of Glu-126 to the sugar is dramatically greater for $\alpha\beta$ -(Galp)₂, i.e., $80\pm 10\%$ versus $30\pm 14\%$, and also an increased affinity for $\alpha\beta$ -(Galp)₂ in LacY-wt. In addition, Glu-269 appears to play a more important role in $\alpha\beta$ -(Galp)₂ binding with an average population of $22\pm 10\%$, whereas Phe-118 is more important with $\beta\beta$ -(Galp)₂ ($13\pm 6\%$). Together these results indicate possible mutations in residues that will influence anomeric binding.

These changes in conformational families and hydrogen bond forming propensity of the two anomeric states of the sugar also influence the $\langle\Delta E^{\text{bind}}\rangle$ and ultimately its binding affinity. The $\langle\Delta E^{\text{bind}}\rangle$ for $\alpha\beta$ - and $\beta\beta$ -(Galp)₂ is -10.7 ± 0.7 and -3.1 ± 1.0 kcal/mol, respectively. This difference in $\langle\Delta E^{\text{bind}}\rangle$ is consistent with experimental results that the affinity of α -sugars is greater than β -sugars in LacY.^{11,13} There is little dependence in the relative affinity of the two anomeric states of the sugar to LacY-C154G and LacY-wt (Table 1). The binding energy is inversely proportional to the number of hydrogen bonds between the sugar and protein. The binding conformations with the most attractive binding energy have a total of 4-6 direct protein-sugar hydrogen bonds. For example, the conformational family in Fig. 9b (2 of LacY-wt/ $\beta\beta$ -(Galp)₂) has a $\langle\Delta E^{\text{bind}}\rangle$ of -10.6 ± 3.1 kcal/mol

In general, the structural changes of the protein from the x-ray crystal structure are predominately in the loop regions of the protein. Changes in loops and interactions with the lipid bilayer result in a more compact structure compared to experiment, i.e., the angle between the two domains of LacY is reduced from 59° to $40\pm 3^\circ$ for experiment and simulation, respectively. The binding of $\beta\beta$ -(Galp)₂ distorts H-IV to a greater extent than $\alpha\beta$ -(Galp)₂, see Figs. 10 and 11. This helix is the most flexible and agrees well with the experimental and simulation helix stability results of Bennett et al.¹⁴ The cause for this distortion is likely the increased backbone interaction of $\beta\beta$ -(Galp)₂ with Phe-118 and possibly interactions with the other two residues on this helix (Asn-119 and Glu-126) that are important to binding.

The sugar escapes the binding site for two of the simulations (2 of LacY-C154G/ $\alpha\beta$ -(Galp)₂ and 1 of LacY-wt/ $\beta\beta$ -(Galp)₂). The cause for this escape from the binding site is the result of a stable Asp-240/Lys-319 salt bridge that prevents the sugar from forming favorable interactions with either Asp-237 or Glu-269. For the escaped sugar runs the Asp-240/Lys-319 salt bridge populates the simulation $18.6\pm 3.8\%$, while the other runs only $5.6\pm 1.1\%$. A high propensity of forming hydrogen bonds with either Asp-237 or Glu-269 is consistent with the sugar-binding simulations, where as the escaped sugars do not form hydrogen bonds with these residues (see Table 1). This escaping behavior appears to be temperature dependent because a Asp-240/Lys-319 salt bridge exists in the measured structure.⁴

In summary, the conformational families of $\beta\beta$ -(Galp)₂ in LacY-C154G are consistent with a x-ray diffraction structure that contains a similar 1-1 glycosidic linked sugar. However, other binding structures that are thermally stable have sugar-protein interactions with Phe-118, Asn-119, Asn-240, His-322, Glu-325, and Tyr-350. Glu-126 is important in sugar binding of both anomeric states of (Galp)₂, but interacts more strongly with the $\alpha\beta$ -anomeric state. Consequently, $\langle\Delta E^{\text{bind}}\rangle$ of $\alpha\beta$ -(Galp)₂ is more favorable than $\beta\beta$ -(Galp)₂ by 7.6 kcal/mol. H-IV is the least stable of the twelve helices and $\beta\beta$ -(Galp)₂ distorts this helix to a greater extent. Collectively, these simulation results demonstrate that the anomeric state of the disaccharide influences not only binding structures and affinity to the protein but also the structure of the protein itself. In general, these results also have implications to the importance of subtle changes in substrate structure on binding conformations and protein structure.

Methods

For all the MD simulations, the CHARMM²⁸ and NAMD²⁹ programs are used. The molecular building, simulations with complex periodic boundary conditions, and data analysis are performed with CHARMM. The production runs use NAMD because it is more efficient in its parallelization across multiple processors. The CHARMM family of force fields are used to describe the atomic interactions of the lipids,^{30,31} protein,³² and substrate.³³

Since the head group composition of lipids in *E. coli* is 70-80% phosphatidylethanolamine (PE),³⁴ the palmitoyloleoylphosphatidylethanolamine (POPE) lipid bilayer will be used to model the plasma membrane. POPE lipid bilayer is also important because the topology and function of LacY is exclusively dependent on the existence of the PE moiety.³⁵⁻³⁷ A previously equilibrated and hydrated POPE bilayer³⁸ is used for the initial lipid structure and the experimental surface area of 65.2 Å²/lipid²⁰ is fixed. A cylindrical cut of lipids ($r \approx 20$ Å) is removed from the center of the POPE bilayer, such that LacY (1PV64) can be inserted into the membrane. Additional lipids that are in energetically unfavorable contact with LacY are removed resulting in 272 total POPE lipids. TIP3P waters^{39,40} from previous bulk simulations are then added on the top and bottom of the lipid bilayer such that sufficient water is present to prevent any artificial protein-protein and lipid-lipid interactions due to periodicity. The thickness of the water slab is 22 Å above and below the phosphate for the top and bottom leaflet with a total of 16,668 waters. Since the LacY structure is charged (+8e), eight chlorine atoms are required to maintain electroneutrality. The total number of atoms in our simulations is 90,685.

The initial equilibration is simulated in the constant particle number, pressure, surface area, and temperature ensemble (NPAT) with CHARMM at a temperature of 310.15 K. Currently, NAMD only supports temperature control with Langevin dynamics, temperature coupling, or temperature rescaling. To minimize the perturbation on the lipid bilayer by these methods, the temperature was rescaled every 10ps for the first 0.5ns in equilibration and then unscaled for the remainder of the simulation. The average and standard deviation for the temperature of the twelve runs are listed in Table 1. Consequently, the ensemble for the production runs is NPAH (H, constant enthalpy). The Lennard-Jones interactions are smoothed by a switching function over 10 to 12 Å.²⁸ The extended system formalism is used to maintain the pressure with a barostat.^{41,42} The particle mesh Ewald⁴³ method is used for the long-range (beyond 12 Å) electrostatic contribution to the total energy. All hydrogen atoms are constrained using the SHAKE algorithm⁴⁴ and a time step of 1 fs.

The number of lipids per leaflet in a tetragonal box should be asymmetric with a cone shaped membrane protein like LacY. Simply removing lipids that overlap the protein results in a bilayer that is asymmetric is the sole method currently used in MD simulations of membrane proteins. However, we use a more accurate and rigorous approach that allows for lipid exchange across leaflets to reach a free energy minimum⁴⁵ with P2₁ boundary conditions instead of P1. The number of lipids per leaflet was found to reach equilibrium after 0.5 ns with P2₁ (132 lipids in the leaflet in contact with the open end and 140 in the other leaflet) and P1 boundary conditions are used for the remainder of the simulations for optimal efficiency.

The disaccharide (Galp)₂ is inserted near the experimental binding site with the overlapping waters removed after lipid equilibrium. The center of mass of the disaccharide is initially placed between Glu-126 and Asp-237. Then the sugar is translated and rotated to form the initial three configurations. Consequently, these configurations are not biased to the TDG conformation in the LacY-C154G crystal structure. The three initial positions of the sugar are labeled as follows: (1) initially imposing a disaccharide/Glu-126 distance consistent with a hydrogen bond, (2) initially imposing a hydrogen bond distance for disaccharide/Arg-144, and (3) random

configuration near the binding site with no imposed hydrogen bond distances. The imposing of hydrogen bonds is only done for the model building and the sugars in the equilibrium and productions runs are free of any imposed hydrogen bonds or bias. These initial positions are used for both anomeric states of the sugar and for LacY-wt and LacY-C154G (a total of 12 simulations).

After a 0.5 ns equilibration period for the lipid number per leaflet and insertion of the sugar, the system is simulated for 20-25 ns (see Table 1 for total production run times). Hydrogen bonds are classified between a donor and acceptor if the distance is less than 3.2 Å. $\langle \Delta E^{\text{bind}} \rangle = E^{\text{bind,LacY}} - E^{\text{sol}}$, where $E^{\text{bind,LacY}}$ is the binding energy of the sugar and E^{sol} is that of the sugar in pure water. Two 10 ns simulations of both anomeric states of (Galp)₂ are used to obtain E^{sol} at 310.15 K and 1 bar. The calculated E^{sol} for αβ- and ββ-(Galp)₂ is -76.1 and -82.7 kcal/mol, respectively.

Acknowledgements

This research was supported in part by the Intramural Research Program of the NIH, National Heart, Lung and Blood Institute. Some of the simulation results utilized the high-performance computational capabilities of the Biowulf PC/Linux cluster at the National Institutes of Health, Bethesda, MD. (<http://biowulf.nih.gov>) We would like to thank Dr. H. Ronald Kaback for his thoughtful comments on this work.

References

1. Kaback HR, Sahin-Toth M, Weinglass AB. The kamikaze approach to membrane transport. *Nature Reviews Molecular Cell Biology* 2001;2:610–620.
2. Chang AB, Lin R, Studley WK, Tran CV, Saier MH. Phylogeny as a guide to structure and function of membrane transport proteins (Review). *Molecular Membrane Biology* 2004;21:171–181. [PubMed: 15204625]
3. Saier MH. Families of transmembrane sugar transport proteins. *Molecular Microbiology* 2000;35:699–710. [PubMed: 10692148]
4. Abramson J, Smirnova I, Kasho V, Verner G, Kaback HR, Iwata S. Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* 2003;301:610–615. [PubMed: 12893935]
5. Huang YF, Lemieux MJ, Song JM, Auer M, Wang DN. Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*. *Science* 2003;301:616–620. [PubMed: 12893936]
6. Yin Y, He X, Szewczyk P, Nguyen T, Chang G. Structure of the multidrug transporter EmrD from *Escherichia coli*. *Science* 2006;312:741–744. [PubMed: 16675700]
7. Paulsen IT, Brown MH, Skurray RA. Proton-dependent multidrug efflux systems. *Microbiological Reviews* 1996;60:575. [PubMed: 8987357]
8. Mirza O, Guan L, Verner G, Iwata S, Kaback HR. Structural evidence for induced fit and a mechanism for sugar/H⁺ symport in LacY. *Embo Journal* 2006;25:1177–1183. [PubMed: 16525509]
9. Abramson J, Iwata S, Kaback HR. Lactose permease as a paradigm for membrane transport proteins - (Review). *Molecular Membrane Biology* 2004;21:227–236. [PubMed: 15371012]
10. Sander mann H. β-D-Galactoside Transport In *Escherichia-Coli* - Substrate Recognition. *European Journal Of Biochemistry* 1977;80:507–515. [PubMed: 336372]
11. Sahin-Toth M, Akhoon KM, Runner J, Kaback HR. Ligand recognition by the lactose permease of *Escherichia coli*: Specificity and affinity are defined by distinct structural elements of galactopyranosides. *Biochemistry* 2000;39:5097–5103. [PubMed: 10819976]
12. Sahin-Toth M, Lawrence MC, Nishio T, Kaback HR. The C-4 hydroxyl group of galactopyranosides is the major determinant for ligand recognition by the lactose permease of *Escherichia coli*. *Biochemistry* 2001;40:13015–13019. [PubMed: 11669639]
13. Green AL, Hrodey HA, Brooker RJ. Evidence for structural symmetry and functional asymmetry in the lactose permease of *Escherichia coli*. *Biochemistry* 2003;42:11226–11233. [PubMed: 14503872]
14. Bennett M, D'Rozario R, Sansom MSP, Yeagle PL. Asymmetric stability among the transmembrane helices of lactose permease. *Biochemistry* 2006;45:8088–8095. [PubMed: 16800633]

15. Yin Y, Jensen MO, Tajkhorshid E, Schulten K. Sugar binding and protein conformational changes in lactose permease. *Biophys. J* 2006;91:3972–3985. [PubMed: 16963502]
16. Ravindranathan KP, Gallicchio E, Levy RM. Conformational equilibria and free energy profiles for the allosteric transition of the ribose-binding protein. *J. Mol. Biol* 2005;353:196–210. [PubMed: 16157349]
17. Datta D, Vaidehi N, Floriano WB, Kim KS, Prasadara NV, Goddard WA. Interaction of E-coli outer-membrane protein A with sugars on the receptors of the brain microvascular endothelial cells. *Proteins* 2003;50:213–221. [PubMed: 12486715]
18. Stockner T, Vogel HJ, Tieleman DP. A salt-bridge motif involved in ligand binding and large-scale domain motions of the maltose-binding protein. *Biophys. J* 2005;89:3362–3371. [PubMed: 16143635]
19. Ermolova N, Madhvani RV, Kaback HR. Site-directed alkylation of cysteine replacements in the lactose permease of *Escherichia coli*: Helices I, III, VI, and XI. *Biochemistry* 2006;45:4182–4189. [PubMed: 16566592]
20. Shaikh SR, Brzustowicz MR, Gustafson N, Stillwell W, Wassall SR. Monounsaturated PE does not phase-separate from the lipid raft molecules sphingomyelin and cholesterol: Role for polyunsaturation? *Biochemistry* 2002;41:10593–10602. [PubMed: 12186543]
21. Deol SS, Domene C, Bond PJ, Sansom MSP. Anionic phospholipid interactions with the potassium channel KcsA: Simulation studies. *Biophys. J* 2006;90:822–830. [PubMed: 16272446]
22. Hashido M, Ikeguchi M, Kidera A. Comparative simulations of aquaporin family: AQP1, AQPZ, AQP0 and GlpF. *Febs Letters* 2005;579:5549–5552. [PubMed: 16225876]
23. Henin J, Maigret B, Tarek M, Escricut C, Fourmy D, Chipot C. Probing a model of a GPCR/ligand complex in an explicit membrane environment: The human cholecystokinin-1 receptor. *Biophys. J* 2006;90:1232–1240. [PubMed: 16326901]
24. Frillingos S, Sahin-Toth M, Wu JH, Kaback HR. Cys-scanning mutagenesis: a novel approach to structure-function relationships in polytopic membrane proteins. *FASEB Journal* 1998;12:1281–1299. [PubMed: 9761772]
25. Weinglass A, Whitelegge JP, Faull KF, Kaback HR. Monitoring conformational rearrangements in the substrate-binding site of a membrane transport protein by mass spectrometry. *J. Biol. Chem* 2004;279:41858–41865. [PubMed: 15272008]
26. Ermolova NV, Smirnova IN, Kasho VN, Kaback HR. Interhelical packing modulates conformational flexibility in the lactose permease of *Escherichia coli*. *Biochemistry* 2005;44:7669–7677. [PubMed: 15909981]
27. Nie YL, Smirnova I, Kasho V, Kaback HR. Energetics of ligand-induced conformational flexibility in the lactose permease of *Escherichia coli*. *J. Biol. Chem* 2006;281:35779–35784. [PubMed: 17003033]
28. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM - a Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem* 1983;4:187–217.
29. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. Scalable molecular dynamics with NAMD. *J. Comput. Chem* 2005;26:1781–1802. [PubMed: 16222654]
30. Klauda JB, Brooks BR, MacKerell AD Jr. Venable RM, Pastor RW. An Ab Initio Study on the Torsional Surface of Alkanes and its Effect on Molecular Simulations of Alkanes and a DPPC Bilayer. *J. Phys. Chem. B* 2005;109:5300–5311. [PubMed: 16863197]
31. Klauda JB, Pastor RW, Brooks BR. Adjacent gauche stabilization in linear alkanes: Implications for polymer models and conformational analysis. *J. Phys. Chem. B* 2005;109:15684–15686. [PubMed: 16852989]
32. MacKerell AD Jr. Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kucsera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kucsera J, Yin D, Karplus M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* 1998;102:3586–3616.

33. Kuttel M, Brady JW, Naidoo KJ. Carbohydrate solution simulations: Producing a force field with experimentally consistent primary alcohol rotational frequencies and populations. *J. Comput. Chem* 2002;23:1236–1243. [PubMed: 12210149]
34. Dowhan W. Molecular basis for membrane phospholipid diversity: Why are there so many lipids? *Annu. Rev. Biochem* 1997;66:199–232. [PubMed: 9242906]
35. Xie J, Bogdanov M, Heacock P, Dowhan W. Phosphatidylethanolamine and monoglucosyldiacylglycerol are interchangeable in supporting topogenesis and function of the polytopic membrane protein lactose permease. *J. Biol. Chem* 2006;281:19172–19178. [PubMed: 16698795]
36. Wang XY, Bogdanov M, Dowhan W. Topology of polytopic membrane protein subdomains is dictated by membrane phospholipid composition. *Embo Journal* 2002;21:5673–5681. [PubMed: 12411485]
37. Bogdanov M, Heacock PN, Dowhan W. A polytopic membrane protein displays a reversible topology dependent on membrane lipid composition. *Embo Journal* 2002;21:2107–2116. [PubMed: 11980707]
38. Tieleman DP, Berendsen HJC. A molecular dynamics study of the pores formed by Escherichia coli OmpF porin in a fully hydrated palmitoylcholine bilayer. *Biophys. J* 1998;74:2786–2801. [PubMed: 9635733]
39. Durell SR, Brooks BR, Bennaïm A. Solvent-Induced Forces between Two Hydrophilic Groups. *J. Phys. Chem* 1994;98:2198–2202.
40. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys* 1983;79:926–935.
41. Nosé S, Klein ML. A Study of Solid and Liquid Carbon Tetrafluoride Using the Constant Pressure Molecular-Dynamics Technique. *J. Chem. Phys* 1983;78:6928–6939.
42. Andersen HC. Molecular-Dynamics Simulations at Constant Pressure and/or Temperature. *J. Chem. Phys* 1980;72:2384–2393.
43. Darden T, York D, Pedersen L. Particle Mesh Ewald - an NLog(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys* 1993;98:10089–10092.
44. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-alkanes. *J. Comp. Phys* 1977;23:327–341.
45. Dolan EA, Venable RM, Pastor RW, Brooks BR. Simulations of membranes and other interfacial systems using P2₁ and pc periodic boundary conditions. *Biophys. J* 2002;82:2317–2325. [PubMed: 11964222]

Figure 1.

(a) the initial inward facing configuration of LacY/(Galp)₂ embedded in the POPE bilayer (half of the membrane in front of the protein is omitted for clarity) and (b) lumen of the protein (domain 1-green and domain 2-blue) with sugar and water (red spheres).

Figure 2. β -D-galactopyranosyl-(1,1)- β -D-galactopyranoside ($\beta\beta$ -(Galp)₂) with the atom types labeled. The *prime* character ' is used to distinguish the two rings, e.g., O6 interacting with domain 1 has an equivalent O6' on the second ring with $\beta\beta$ -(Galp)₂. For $\alpha\beta$ -(Galp)₂, the ' is used for the β ring.

Figure 3.
The IPV7 crystal structure⁴ (a) and crystal-like structures for **1** (t=8.1 ns), **2** (t=4.9 ns) and **3** (t=2.3 ns) for (b-d), respectively. Hydrogen bonds between the sugar and protein are shown in green.

Figure 4.
The number of hydrogen bonds during the LacY-C154G/ $\beta\beta$ -(Galp)₂ simulations that conform to those observed in the 1PV7 crystal structure.

Figure 5.

The total number of LacY-C154G/ $\beta\beta$ -(Galp)₂ hydrogen bonds as a function of the simulation time.

Figure 6. Snapshots of binding site structures that differ from 1PV7⁴ for LacY-C154G/ $\beta\beta$ -(Galp)₂ for: (a) **2** at t=20.0 ns and (b) **3** at t=16.5 ns.

Figure 7. Final snapshots of LacY-C154G/ββ-(Galp)₂ for (a) **1** and (b) **3** with the whole system shown on the left and the binding site on the right. The final binding site snapshot for **2** is shown in Fig. 6a.

Figure 8. Snapshots of binding site structures for LacY-C154G/αβ-(Galp)₂ for: (a) **1** at t=19.4 ns and (b) **3** t=13.1 ns.

Figure 10.

The RMSD from 1PV7⁴ for the backbone of LacY-wt and LacY-C154G comparing the influence of the anomeric state of the sugar on protein structure. Helices (H) and domains are based on previous definitions.⁴

Figure 11.

Helix-IV of LacY-wt from 1PV7 (red), end of one LacY-wt/ββ-(Galp)₂ (blue) and LacY-wt/αβ-(Galp)₂ simulation (green). The coordinates are orientated such that the RMS to 1PV7 is minimized.

Table 1

Percentage of time a sugar-protein amino acid forms a hydrogen bond during the sampled period of time for each run. The binding energy (in kcal/mol) of the sugar and average temperature are also given with standard errors in parenthesis, where a) is with LacY-C154G and b) is with LacY-wt.

a)	α -Galp-(1,1)- β -Galp			β -Galp-(1,1)- β -Galp		
	1	2	3	1	2	3
E126	85.8	40.6	42.9	57.7	21.6	32.7
R144	58.5	4.5	1.1	18.4	12.0	30.5
D237	0.1	0.6	43.9	9.3	57.2	25.1
E269	32.9	0.0	0.0	0.1	0.0	18.2
K358	0.5	8.3	30.9	19.0	15.4	24.6
F118	2.8	0.5	1.4	3.0	6.2	15.0
N119	2.6	0.5	19.6	13.6	35.2	0.0
D240	0.0	0.0	0.0	0.0	6.2	0.0
H322	1.9	0.0	0.0	0.1	0.7	0.8
Y350	0.0	1.4	8.0	3.2	4.0	3.4
t_{run} [ns]	20	20	25	20	20	25
$\langle T \rangle$ [K]	310.2 (0.8)	310.8 (0.9)	310.7 (0.9)	310.7 (0.8)	311.5 (0.8)	310.8 (0.8)
$\langle \Delta E^{\text{bind}} \rangle$	-12.6 (1.8)	-4.6 (0.1)	-10.2 (0.9)	-2.0 (0.8)	-7.1 (2.2)	-2.2 (0.5)
b)	α -Galp-(1,1)- β -Galp			β -Galp-(1,1)- β -Galp		
	1	2	3	1	2	3
E126	100.0	77.3	91.8	7.1	18.1	26.7
R144	18.9	6.0	32.2	0.0	3.8	11.3
D237	1.6	38.5	0.9	0.9	33.7	28.5
E269	21.4	0.0	54.4	0.0	15.5	1.5
K358	3.4	37.8	4.5	6.2	12.4	24.1
F118	0.4	0.0	0.0	0.0	23.7	4.2
N119	41.3	18.5	14.1	0.0	30.0	12.4
D240	0.0	0.0	0.0	0.0	16.8	1.4
H322	0.9	1.8	14.6	0.2	4.3	0.4
Y350	0.0	1.2	5.7	4.5	0.6	8.9
t_{run} [ns]	20	20	25	20	20	20
$\langle T \rangle$ [K]	310.2 (0.9)	310.4 (0.9)	310.0 (1.1)	310.9 (1.0)	310.0 (0.9)	310.0 (0.9)
$\langle \Delta E^{\text{bind}} \rangle$	-11.1 (0.3)	-8.6 (1.0)	-11.2 (0.2)	5.4 (0.0)	-2.6 (1.1)	-1.4 (0.1)

Table 2

Average protein and lipid bilayer properties for the twelve MD simulations, where a) is with LacY-C154G and b) is with LacY-wt. The root mean squared deviation (RMSD) is calculated from the x-ray crystal structure, 1PV7, of the LacY-C154G mutant.⁴ $\langle\% \text{ helicity}\rangle$ has a standard error of 1% and $\langle\% \text{ helicity}\rangle > 80\%$ from the x-ray crystal structure. $\langle\theta_{\text{domain}}\rangle$ is the angle between domain 1 and 2 of the protein with a standard error of 0.2° (1PV7 = 59°). $\langle r_{P-P}\rangle$ is the average distance between phosphorous in POPE with a standard error of 0.2 Å. The subscript f refers to the average of the last 0.5 ns and without a subscript the average for the entire simulation.

a)	1	$\alpha\text{-Galp}-(1,1)\text{-}\beta\text{-Galp}$ 2	3	1	$\beta\text{-Galp}-(1,1)\text{-}\beta\text{-Galp}$ 2	3
$\langle\text{RMSD}\rangle$	3.8	2.5	3.6	2.6	2.6	3.2
$\langle\text{RMSD}\rangle_f$	3.6	2.7	4.6	3.3	2.8	3.6
$\langle\% \text{ helicity}\rangle$	84	83	82	82	82	83
$\langle\% \text{ helicity}\rangle_f$	83	82	82	80	82	82
$\langle\theta_{\text{domain}}\rangle$	43±3	35±3	42±6	39±2	38±3	40±3
$\langle\theta_{\text{domain}}\rangle_f$	38	36	50	37	33	34
$\langle r_{P-P}\rangle$ [Å]	35.4	35.4	35.5	35.3	35.5	35.5
$\langle r_{P-P}\rangle_f$ [Å]	35.4	35.6	35.2	35.4	35.3	35.5
b)	1	$\alpha\text{-Galp}-(1,1)\text{-}\beta\text{-Galp}$ 2	3	1	$\beta\text{-Galp}-(1,1)\text{-}\beta\text{-Galp}$ 2	3
$\langle\text{RMSD}\rangle$	3.7	2.6	2.8	2.6	2.8	3.0
$\langle\text{RMSD}\rangle_f$	4.3	2.7	2.8	3.3	3.2	3.3
$\langle\% \text{ helicity}\rangle$	84	82	83	83	82	83
$\langle\% \text{ helicity}\rangle_f$	83	82	83	82	82	83
$\langle\theta_{\text{domain}}\rangle$	46±5	40±3	40±4	41±3	42±3	35±4
$\langle\theta_{\text{domain}}\rangle_f$	47	39	34	36	43	34
$\langle r_{P-P}\rangle$	35.4	35.4	35.4	35.4	35.4	35.4
$\langle r_{P-P}\rangle_f$	35.8	35.2	35.3	35.5	35.3	35.4