JOURNAL OF
EMERGING INVESTIGATORS

Article

The sweetened actualities of neural membrane proteins: A computational structural analysis

Radha B. Chauhan[1], Ruben T. Almaraz[2]
[1] Sacramento Country Day School, Sacramento, California
[2] Cosumnes River College, Sacramento, California

SUMMARY

In nature, every cell is coated by an array of complex molecules known as glycans, which are also known as sugars. Their intricacy makes them difficult to analyze and as a consequence little is known about their function. However, in the last decade, compounding evidence points to glycans participating in critical roles in every cell type, including neurons. In fact, altered glycosylation has been correlated to neural disorders such as Alzheimer's and Parkinson's disease. To learn more about these molecules, we investigated two neural membrane glycoproteins: the type-A γ-aminobutyric acid ($GABA_A$) and the human 5-hydroxytryptamine 2B ($5\text{-}HT_{2B}$) receptors by computational methods. We examined if the glycan structures could physically interfere with the receptor active site. To answer this question, we modeled different glycans on each protein. The model shows that the active sites are heavily glycosylated and all or several glycans are either directly on or bracketing the active site. Furthermore, molecular dynamics simulations demonstrated these molecules to be highly dynamic and in the case of $5\text{-}HT_{2B}$ receptor, capable of covering the entire active site region. Therefore, validating our hypothesis that at least physically, glycans partakes with the active site functions. Future experimental work needs to be done to reveal how the absence or structure changes in these molecules relate to disease. This information can be used for future neurological disease treatments and possible cures.

INTRODUCTION

Sugars are mostly known as sweeteners, but they are an essential class of complex molecules, especially for brain cells, depending on sugars as an energy source and as participants in prominent roles serving as cell communications and cell adhesion. The nervous system is highly complex, coordinating all bodily activities such as collecting and imparting information from the brain to the spine, processing the information, and releasing signals to all regions of the body. The nervous system is made of neurons, unique cells composed of a cell body, an axon, and dendrites. From the aforementioned list, the axon and dendrites play central roles in cell communication **(Figure 1A–B)**. Neurons receive and transmit substances known as neurotransmitters, across an intercellular space called the synapse, a region between the axon's terminal and dendrites of adjacent neurons **(Figure 1C)**. This communication is a multistep process, participating small ions to large complex molecules. Some of these molecules have been well characterized while the function of others is partial or completely unknown; however, it is known that the proper function of neuron-neuron communication is vital to the organism, and inhibition or dysfunctions of this process can lead to a series of neurological disorders or death (1).

Neurological disorders pose a large burden on our society. Every year, about 100 million Americans are affected by at least one type of neurological disorder (2,3). A number expected to rise as Alzheimer's, Multiple sclerosis, and Parkinson's disease among others are becoming more prevalent in our society (4). In the last few years, different branches of science have begun to explore novel frontiers to understand, treat, and possibly cure these atrocious diseases.

An area of science that entered the fight against neurological diseases is Glycobiology, the field of science that deals with the synthesis, structural, and functional understanding of glycans, also known as carbohydrates or sugars. Carbohydrate moieties can be covalently attached to lipids and proteins in a process known as glycosylation, one of the most universal and essential protein modifications (5). Glycans are ubiquitous and there are diverse types of glycosylation in nature. In eukaryotes, the best-studied being attachment of glycans to asparagine (*N*-glycosylation) and serine/threonine (*O*-glycosylation) residues (5). Glycans are an essential part of our macromolecules repertoire and as such, they should be studied.

An additional motivation for this study is that although nucleic acids and proteins have been reported in thousands of studies, there remains a dearth of information on the structural and functional impact of glycans in all areas of science. Especially at the high school and undergraduate level, the role of glycans in neurology as in any other field, is, for the most part, non-existent in their science curriculum. Together, the importance of glycosylation in neurology and the scarcity of information at the basic academic levels were compelling motivations for this study.

For this study, we focused on *N*-glycosylation which has been known to play roles in protein folding, trafficking, stability, and regulation of many cellular activities. Recently, *N*-glycosylation has been suggested to be vital for normal brain functions, including memory and learning (6). At the same time, abnormal *N*-glycosylation has been correlated

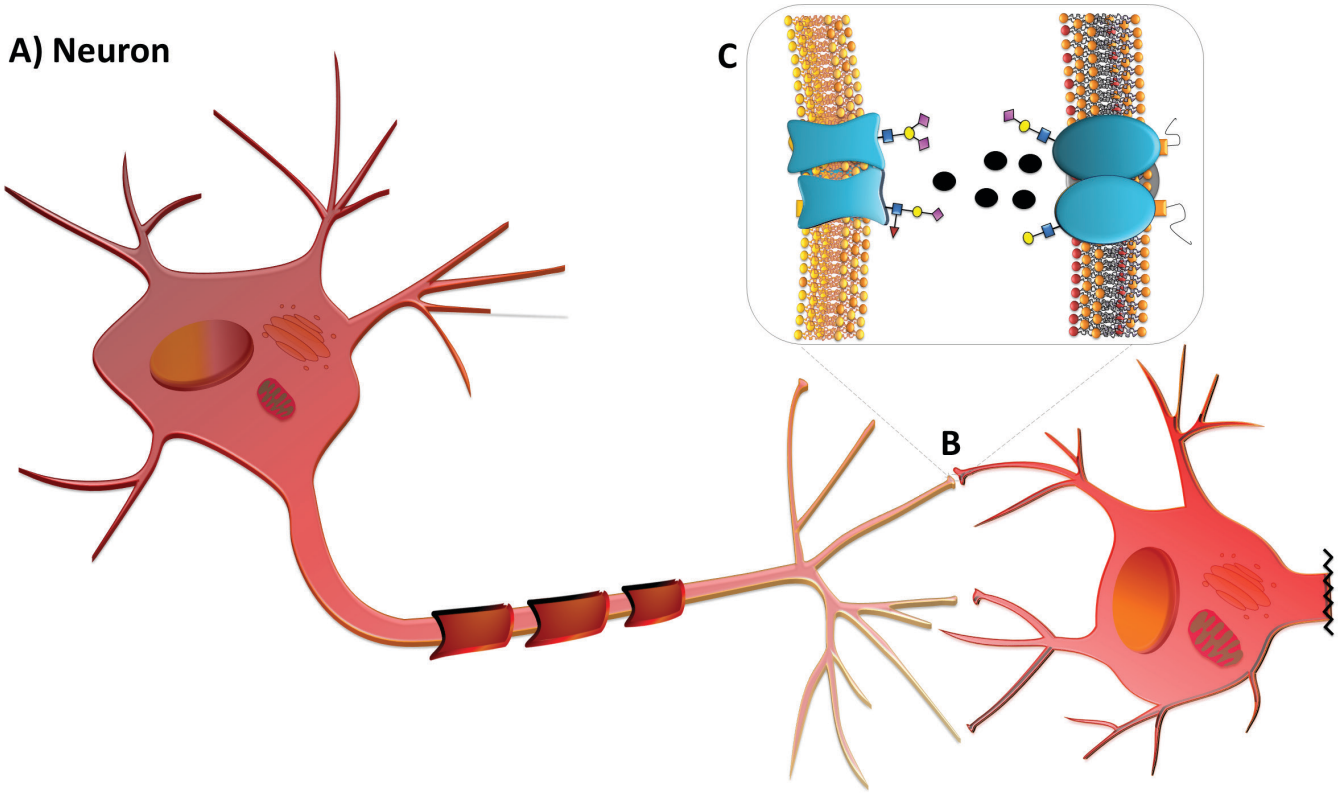

Figure 1: Neuron cell model demonstrating synapsis. A) Illustration of major neuronal cell components: (left to right) neuronal main body, dendrites, axon, myelin sheaths, axon terminals, and synapse junctions. **B)** Depiction of a synapse, a region where a synapse takes place between a dendrite and an axon terminal. **C)** Representing some of the main molecular components of a synaptic apparatus: Transmembrane protein (celeste blue), glycans (red, yellow, and purple geometric figures connected to the transmembrane protein), and neurotransmitters (black circles).

to brain pathology. For example, in an Alzheimer's disease study, specific abnormal glycan structures on the amyloid precursor protein (APP) were observed in the studied Alzheimer's disease patients as compared to healthy people. Furthermore, the same study suggests that *N*-glycosylation plays roles in the sorting, secretion, and transportation of APP and other synaptic glycoproteins (7). In another study, it was found that all schizophrenia patients studied had a decrease in the expression of FUT8, the enzyme responsible to α-1,6-fucosylate *N*-glycans (8).

To learn more about this phenomenon, we explored the structural glycobiology, by computational methods, of two human neural membrane proteins: the human type-A γ-aminobutyric acid receptor ($GABA_AR$) and the Serotonin or 5-hydroxytryptamine 2B ($5\text{-}HT_{2B}$) receptor. $GABA_AR$ is the principal mediator of rapid inhibitory synaptic transmission in the human brain (3). A decline in $GABA_AR$ signaling prompts hyperactive neurological disorders such as insomnia, anxiety, and epilepsy (3). The $5\text{-}HT_{2B}$ receptor is a serotonin transporter that has been associated with various physiological and pathophysiological roles such as anxiety and sleep regulation (9). The three-dimensional (3D) protein structure of these two proteins has been solved by X-ray crystallography, as well as the location of the *N*-glycans for $GABA_AR$ but there is no experimental data on the *N*-glycosylation on the $5\text{-}HT_{2B}$ receptor. Moreover, the size of the glycans provided on the $GABA_AR$ crystal structure are unrealistically small due to glycosylase treatment, albeit necessary for the crystallization process (3). Often, structural glycobiology data has been obtained as the byproduct of protein structure analysis. Even when the main focus of the study is glycan structural analysis, the techniques often used were meant for protein or nucleic acids, resulting in generic glycan analysis. For example, most glycosylation inhibitions are nonspecific, affecting hundreds to thousands of glycoproteins (10). For example, mutagenesis is a technique used to investigate protein function, and when applied to glycol-analysis, one or several amino acids are mutated to completely block the glycosylation process on that amino acid site. This not only eliminates the control of expression and size of the glycan in question to attenuate the possible function but also now exposed a region of the protein that was previously shielded by the glycan making it impossible to decipher the functional mechanism. Similarly, the analysis done in $GABA_AR$ did not provide information about the glycan structures nor their functions, but the limited information helped us verified their presence and glycosylation sites which gave us a starting point for this project (11). By using an *N*-glycan prediction algorithm such as GlyPro, we can verify its accuracy in predicting the $GABA_AR$ *N*-glycosylation and expand on it by computationally modeling more realistic

glycan structures on the receptor. We can also further expand the project to another neural membrane protein such as the 5-HT_{2B} receptor.

We examined if the glycan structures could substantiate an *N*-glycan functional role on $GABA_AR$ and in a similar manner in the 5-HT_{2B} receptor. We expected to find, at least in the case of $GABA_AR$, one or more of the glycans significantly concealing the active site, by its size, its flexibility, and mobility, or a combination of the aforementioned. For the 5-HT_{2B} receptor, we expected fewer glycosylation sites albeit more significant because of the greater glycan/protein mass ratio. Our hypothesis was tested by computational structurally modeling. The molecular dynamics (MD) simulations further support our hypothesis.

RESULTS

The Created Models for $GABA_A$ and 5-HT_{2B} Receptors Portray *N*-glycans Bracketing Their Active Sites

The effect of glycosylating on the 5-HT_{2B} receptor and the human type-A γ-aminobutyric acid receptor ($GABA_AR$)

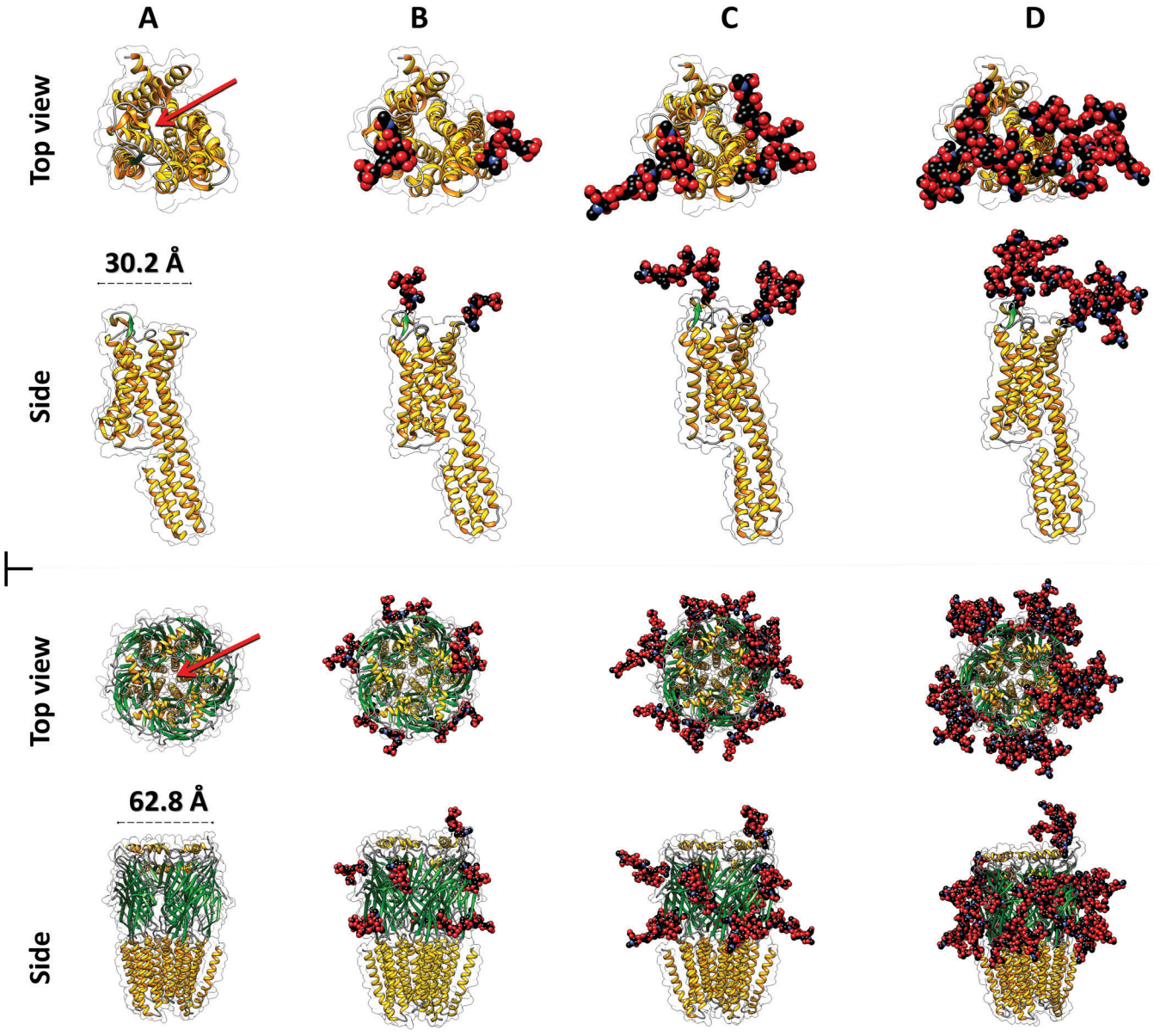

Figure 2: Neural unglycosylated glycoproteins and *N*-glycosylated with basic, hybrid, and complex *N*-glycans. The top panel shows the top and side view of the unglycosylated 5-HT_{2B} receptor **(A)** and glycosylated receptor with basic **(B)**, hybrid **(C)**, and complex **(D)** *N*-glycans. The lower panel shows the top and side view of the unglycosylated $GABA_A$ receptor **(A)** and glycosylated receptor with basic **(B)**, hybrid **(C)**, and complex **(D)** *N*-glycans. Glycans are shown as spheres with carbon atoms in black, oxygen atoms in red, and nitrogen atoms in blue. Hydrogen atoms were omitted for simplicity. The protein portion is shown as a cartoon model where helixes are shown in yellow and beta sheets in green. The active site average distance is also shown for each receptor.

Molecule	Average Length (Å)
Basic *N*-glycan Asn-FGA	21.4
Hybrid *N*-glycan Asn-FGA	32.2
Complex *N*-glycan Asn-FGA	42.9
$GABA_AR$ ASC-NG	30.7
$5\text{-}HT_{2B}$ ASC-NG	15.1

Table 1: Average glycans length and distance to the receptor active site.
Asn-FGA, Asparagine to the Farthest Glycan Atom
ASC-NG, Active Site Center to the Nearest Glycan

were examined through the manipulation of their 3D protein structures. The $5\text{-}HT_{2B}$ and the $GABA_AR$ 3D structure coordinates were obtained from the protein database (PDB) and submitted to the Glyprot server which generated the glycosylated models (9). Two glycosylated sites were predicted at asparagine 204 and 354 for $5\text{-}HT_{2B}$. The three different *N*-glycan types may be possible at some given time or/and set of conditions. Due to the heterogeneous nature of glycosylation, it may be impossible to accurately predict the actual identity of the *N*-glycans. Hence a representative model of each type was selected: a basic, a hybrid, and a complex *N*-glycan structure as suggested by the Glyprot server, to generate a total of three modeled glycoproteins with each glycan type (12). The created glycoproteins were energy-optimized using Scigress. Chimera was implemented to visualize and analyze the three modeled glycoproteins as well as the unglycosylated protein (**Figure 2A–D**, upper panel). Note that we call the active site the top portion of the glycoprotein, the region parallel to the cell surface and most likely to interact with the environment but lateral sites could also be active sites, albeit less likely. The center of the active site is shown by a red arrow, **(Figure 2)**. For the $5\text{-}HT_{2B}$ receptor, the average distance to the closest predicted glycans was measured to be approximately 15.1 Å. Average distances for each glycan were taken from the amino acid to the terminal monosaccharide to estimate the reach of the glycans. After optimization, the average length for the basic, hybrid, and complex glycan was found to be 21.4, 32.2, and 42.9 Å respectively **(Table 1)**. All *N*-glycans modeled here have the potential to reach the active site center.

As with the $5\text{-}HT_{2B}$ receptor, the $GABA_AR$ protein 3D coordinates were obtained from the PDB database and submitted to the GlyProt server to glycosylate the most probable *N*-linked sites (3). Ten potential *N*-glycosylation sites were predicted at asparagine 80, and 149 on each of the five subunits. An additional glycan was predicted at asparagine 8 on the first subunit similar to the published structure (3). The modeled glycoproteins, the basic, the hybrid, and the complex *N*-glycosylated glycoprotein as well as the unglycosylated protein were visualized (**Figure 2A–D**, lower panel). Measurements were taken and it was found that the distance between the active site center to the nearest *N*-glycans, located at subunit A, is about 30.7 Å, within reach of a complex *N*-glycan.

MD Simulations Reveal that the Dynamic Motion of *N*-glycans can Cover a Large Surface Area and are not Sequestered to One Region

To explore the *N*-glycans flexibility and potential of interfering with the active site of each receptor, we performed MD simulations using our hybrid glycans glycoprotein, under physiological temperature and a dielectric constant for water. The hybrid glycan model was used since it represents a median size glycoprotein in this study. Due to the large size of these glycoproteins, the protein portion was held fixed and only the glycan portion was free to move during the simulation. The simulations – which are most clearly evident in **Videos 1 and 2** – show how highly flexible these molecules are. Implementation of Chimera provided the means to visualize time intervals of the MD simulation trajectories. The glycan was colored differently at each time frame: 0.1 (tan), 0.2 (blue), 0.4 (purple), 0.6 (spring green), 0.8 (salmon) and 1.0 (red) nanoseconds (ns) **(Figure 3)**. Overlaid glycoproteins at different simulation times are shown to better illustrate their dynamics. The lack of monochrome glycans on the overlap figure illustrates the dynamic behavior of these structures. For these receptors, based on the predicted glycan location, size, and their dynamic behavior, it is probable that glycosylation is of functional significance to the receptor.

DISCUSSION

Glycosylation is structurally significant in neural communication. Many neurological disorders remain largely untreatable, in part because important aspects of the neuron's communication system, such as its glycobiology are still unknown. Without a clear understanding of the glycobiology and aberrations in glycosylation associated with neurological disease cases, it is difficult to target them diagnostically or therapeutically. In this study, we used a computational approach to address some of these issues, first by modeling the three different types of *N*-glycans on two neural receptors. The two receptors studied here, $GABA_A$ and $5\text{-}HT_{2B}$, are heavily *N*-glycosylated. As most cell membrane surface proteins are glycosylated, these results were not surprising (5). What was novel is the size of the glycans relative to the proteins, as even the smallest

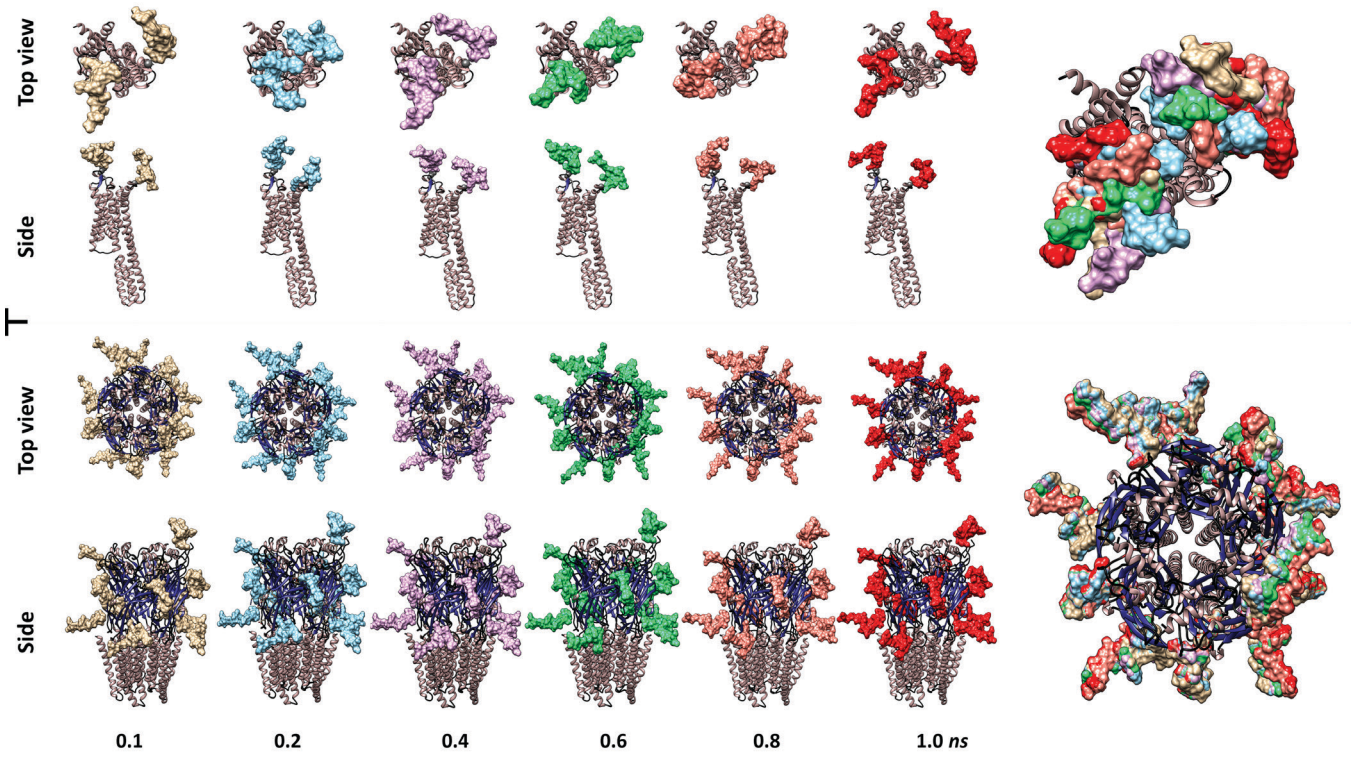

Figure 3: Simulated movement of two neural glycoproteins. Structures of the two membrane glycoproteins at specific times during the MD trajectory. The upper panel shows the side and top view for 5-HT_{2B} and the lower panel shows the top and side view for the $GABA_AR$ modeled glycoproteins at 0.1 (tan), 0.2 (blue), 0.4 (purple), 0.6 (spring green), 0.8 (salmon) and 1.0 (red) nanoseconds (ns). The top view of the aforementioned structures superimposed to better illustrate the movement of the same structures at different times is shown on the right side.

glycan has the potential to occupy a large surface area **(Figure 2)**. Furthermore, the location of a few glycans, in the case of $GABA_AR$, or all, in the case of 5-HT_{2B}, relatively close to the active site suggests a critical functional role in cell communication. Even the glycans away from the active site may play vital roles too. A relevant study demonstrated that removing the β2 *N*149 *N*-glycan on $GABA_AR$, a lateral glycan, resulted in the receptor having reduced sensitivity to GABA (3). In the case of the 5-HT_{2B} receptor, our study suggests the protein is heavily glycosylated coveting most of its active site, and the MD simulation demonstrated the hybrid *N*-glycans to be highly flexible further shielding the active site. This is probably more so with larger or more complex glycans than the one modeled here. However, it is beyond the scope of these studies to perform simulations of all glycoproteins modeled. Glycosylation of proteins is known to be highly heterogeneous; it is also possible that a portion of the protein is unglycosylated — all good reasons to further investigate the glycobiology of these and all neural receptors.

Together, the modeled glycoproteins and MD simulations, show the *N*-glycans as highly flexible and capable of covering a significant area on the active site. This suggests that it is reasonable and logical to speculate that the conformational and chemical nature of the *N*-glycans on these membrane proteins has the potential to modulate the interaction of these two neural membrane proteins and all neural glycoproteins and begs the questions as to how they do it? Answering these questions can lead to future potential therapeutic targets to eradicate mental diseases and improve mental health.

MATERIALS AND METHODS

General Computational Modeling Approaches

Computational modeling was done with the UCSF Chimera version 1.15 program, and the Scigress version 3.1.1 software package (Scigress FJ 2.6, 2008-2021 Fujitsu Limited) (13). All computations were performed on a Cyberpower computer equipped with an Intel® Core™ i7-3970X Extreme Edition Six-Core 3.50 GHz, >20% overclock processor, and 32 GB of RAM (CyberPowerPC Inc, CA). Starting protein 3D structure coordinates were extracted from the PDB database (http://www.rcsb.org/). Optimization of the structure complexes was done by first using the Newton conjugate gradient procedure for 600 cycles followed by the steepest descent procedure for 300 cycles to locate the energy minimum with a dielectric of 78.4. The following interactions were optimized: bond stretch, bond angle, dihedral angle, improper torsion, torsion stretch, van der Waals, electrostatics, and hydrogen bonds. Optimization continued until the energy change was less than 0.00100 kcal/mol or after the molecule was updated for 900 cycles.

Glycoprotein Modeling

Briefly, the structure of the $GABA_AR$ (PDB ID: 4cof) and the $5\text{-}HT_{2B}$ (PDB ID: 5tud) receptor were submitted to the GlyProt server for the prediction of the *N*-glycans sites and generation of a glycosylated protein (http://www.glycosciences.de/). GlyProt works by calculating the probabilities of physicochemical properties such as mass, accessible surface, and radius of gyration (12). A representative *N*-glycan from the three main groups; basic, hybrid, and complex *N*-glycans, was selected based on the options offered by the Glyprot server. The glycosylated membrane protein complexes were energy optimized with the MM3 force field from the Scigress software package. UCSF Chimera v 1.15 (https://www.cgl.ucsf.edu/chimera/) was implemented to visualize and analyzed the created models (13). Measurements of the glycans' length and distances between the closest *N*-glycan to the center of the active site, were done with the Structural Analysis functions on Chimera.

Molecular Dynamics (MD) Simulations

The glycoprotein structure complexes previously energy-optimized were used as a geometrical starting point for all MD simulations. MD simulations were carried out at the physiological temperature of 310.0 K with a dielectric of 78.4. The equilibration time was set to 1.0 picoseconds (ps) and sampling time to 1000 ps with the protein portion locked in place. The *N*-glycans as well as the asparagine to which they were attached, were freely flexible. The time step was 2.0 femtoseconds (fs) and was sampled every 200 steps. Chimera was utilized to visualize the glycoproteins at 0.1 (tan), 0.2 (blue), 0.4 (purple), 0.6 (spring green), 0.8 (salmon) and 1.0 (red) ns. All extracted glycoproteins at different times were overlaid to further appreciate the volume covered and mobility of the *N*-glycans.

VIDEOS

Molecular Dynamics Simulations for Glycoproteins.

Video 1: $5\text{-}HT_{2B}$
URL: https://youtu.be/yOTIecuUzRQ
Molecular Dynamics (MD) Simulations for the 5-HT2B (PDB ID: 5tud) glycoproteins. Sampling time of 1000 ps with the protein portion locked in place. The *N*-glycans as well as the asparagine to which they were attached, were freely flexible.

Video 2: $GABA_AR$
URL: https://youtu.be/IF7vqyiO0q0
Molecular Dynamics (MD) Simulations for the $GABA_AR$ (PDB ID: 4cof) glycoproteins. Sampling time of 1000 ps with the protein portion locked in place. The *N*-glycans as well as the asparagine to which they were attached, were freely flexible.

ACKNOWLEDGMENTS

We would like to thank Dr. Jerry L. Hedrick from the University of California Davis, for allowing us to use the Scigress software. The authors gratefully acknowledge the JEI staff for their help in the preparation of this manuscript.

Received: February 20, 2021
Accepted: April 4, 2022
Published: November 3, 2022

REFERENCES

1. Mary Ann Clark, M.D., Jung Choi. (2018), *Biology 2e*. 2e ed. OpenStax, Houston, Texas.
2. Gooch, C.L., Pracht, E. and Borenstein, A.R. (2017) The burden of neurological disease in the United States: A summary report and call to action. *Ann Neurol*, **81,** 479-484.
3. Miller, P.S. and Aricescu, A.R. (2014) Crystal structure of a human GABAA receptor. *Nature*, **512,** 270-275.
4. Espay, A.J., Aybek, S., Carson, A., Edwards, M.J., Goldstein, L.H., Hallett, M., LaFaver, K., LaFrance, W.C., Jr., Lang, A.E., Nicholson, T. *et al.* (2018) Current Concepts in Diagnosis and Treatment of Functional Neurological Disorders. *JAMA Neurol*, **75,** 1132-1141.
5. Reily, C., Stewart, T.J., Renfrow, M.B. and Novak, J. (2019) Glycosylation in health and disease. *Nat Rev Nephrol*, **15,** 346-366.
6. Scott, H. and Panin, V.M. (2014) N-glycosylation in regulation of the nervous system. *Adv Neurobiol*, **9,** 367-394.
7. Schedin-Weiss, S., Winblad, B. and Tjernberg, L.O. (2014) The role of protein glycosylation in Alzheimer disease. *FEBS J*, **281,** 46-62.
8. Mueller, T.M., Yates, S.D., Haroutunian, V. and Meador-Woodruff, J.H. (2017) Altered fucosyltransferase expression in the superior temporal gyrus of elderly patients with schizophrenia. *Schizophr Res*, **182,** 66-73.
9. Ishchenko, A., Wacker, D., Kapoor, M., Zhang, A., Han, G.W., Basu, S., Patel, N., Messerschmidt, M., Weierstall, U., Liu, W. et al. (2017) Structural insights into the extracellular recognition of the human serotonin 2B receptor by an antibody. *Proc Natl Acad Sci U S A,* **114,** 8223-8228.
10. Esko, J.D. and Bertozzi, C.R. (2009) In nd, Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W. and Etzler, M. E. (eds.), *Essentials of Glycobiology*, Cold Spring Harbor (NY).
11. Martinez-Maza, R., Poyatos, I., Lopez-Corcuera, B., E, N.u., Gimenez, C., Zafra, F. and Aragon, C. (2001) The role of N-glycosylation in transport to the plasma membrane and sorting of the neuronal glycine transporter GLYT2. *J Biol Chem,* **276,** 2168-2173.
12. Bohne-Lang, A. and von der Lieth, C.-W. (2005) GlyProt: in silico glycosylation of proteins. *Nucleic Acids Research*, **33,** W214-W219.

13. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem,* **25,** 1605-1612.