The Effects of Glycosylation on γ-Sarcoglycan

Allye Gardner

Dr. Gabriel A. Cook's Laboratory

Oklahoma State University

Abstract

Glycoproteins are the underappreciated stars in the world of science. These proteins – specifically, membrane-bound glycoproteins – are known to have been difficult to study, for they present many hurdles: they are complex, hydrophobic, and are stable in only highly specific environments. However, investigations into their structure and function are paramount, for glycoproteins are essential in the biological world. Glycoproteins serve as enzymes and hormones, and aid in clotting and locomotion. While there have been several studies of membrane glycoproteins, most of these have focused on their unique attribute: their association with sugar groups. It has recently been shown that it is possible to use the enzyme N-glycosyltransferase to attach sugars to proteins providing a method for *in vitro* glycosylation (5). Using a membrane-mimetic environment, we are attempting to utilize these methods to glycosylate a truncated form of γ-Sarcoglycan ($GSCG_{tm}$), a protein implicated in the disease etiology of Duchenne Muscular Dystrophy. These studies are aimed at shedding light on the effects of glycosylation on the structure and dynamics of $GSCG_{tm}$, in order to provide information that can be used in future treatment modalities. For these studies, milligram amounts of pure protein are required. The work presented here highlights our successful efforts to recombinantly express and purify this membrane protein.

Introduction

Membrane proteins are abundant within cell membranes. They serve, among many things, to provide a cell structure, as a channel for ions to pass through, and to trigger intracellular signaling pathways (3). **Figure 1** depicts different representations of membrane proteins. However, given their importance, little is none about them, for they are hard to work

with. While it is possible to know which portion of a protein passes through the membrane thanks to the amino acid sequence, the folded three-dimensional structures of most membrane proteins have yet to be determined (1). These membrane proteins are often subjects of glycosylation, which occurs when a sugar is added to the protein (1). **Figure 2** shows a membrane protein that has been glycosylated. Though these proteins prove difficult to work with, the goal of this project is to tackle the unknown and to determine the structure of the membrane-bound glycoprotein γ-Sarcoglycan (GSCG). To do so, we will obtain a purified sample of GSCG and determine its structure by nuclear magnetic resonance (NMR).

Sarcoglycan is a subunit of the Dystrophin glycoprotein complex. This complex is found in the plasma membrane of cells and plays a role in the stabilization of muscle cells (4). **Figure 3** represents the Dystrophin complex and shows the Sarcoglycan subunits. As shown, multiple forms of Sarcoglycan are found within the complex; when the Sarcoglycan subunits are absent, it can lead to a leaky plasma membrane. It has been found that muscular dystrophy can result from genetic mutations to genes encoding GSCG (4). Therefore, the importance of GSCG cannot be understated.

While the structure and exact role of GSCG is not known, many features of the protein are. The gene encoding GSCG is found on chromosome 13 and the protein is believed to play a role in stabilizing the plasma membrane of skeletal and cardiac muscles (4). Cardiomyopathy and muscular dystrophy can result from a recessive mutation of the gene encoding GSCG. The most common mutation is Δ521-T. This frameshift mutation leads to a stop codon, which produces a truncated, unstable protein (4). Determining the structure of GSCG, as well as observing the effects of glycosylation on the protein, will provide information on how to best treat those affected by this mutation.

The truncated version of GSCG I worked with, $GSCG_{tm}$, is approximately 10kDa; this engineered protein contains the necessary glycosylation site, and also includes the transmembrane domain. Removing portions of the extracellular and intracellular domain will facilitate observing any change in the molecular weight once glycosylation is performed. This truncated version proved difficult to work with, leading to some issues within the purification process. The hope is to someday attach a sugar by *in vitro* glycosylation to the then purified $GSCG_{tm}$. Following this glycosylation, additional NMR will be conducted, allowing one to see the changes in structure the added sugar entails.

Methodology

Four 1L growths of expression *E. coli* in either LB or ^{15}N-labeled media are grown with the antibiotic carbenicillin from a starter and overnight growth. IPTG is later added for protein expression. To remove impurities, lysis is performed, utilizing three different buffers. The cells are sonicated and the cell lysis is centrifuged after administering Resuspension I and II buffers. The precipitate, containing the inclusion bodies, is dissolved in binding buffer made with the denaturant guanidine hydrochloride. The solubilized protein is dialyzed to remove the denaturant and associated lipids. Following this the protein is frozen and lyophilized. The target protein, $GSCG_{tm}$, is expressed with a HIS_6-TRPΔLE fusion sequence followed by a methionine residue that allows for the cleavage of the fusion protein from the target protein. Cleavage by cyanogen bromide is done by dissolving the protein in 70% formic acid and adding dry cyanogen bromide. After three hours, the cleavage solution is dialyzed to neutralize its pH. The dialyzed sample is then frozen and lyophilized. The cleaved product can further be separated by FPLC to isolate pure $GSCG_{tm}$. These steps are depicted in **Figure 4**.

Alterations to Methodology

As anticipated, this membrane-bound glycoprotein has been challenging to work with. Over the years we have approached it from a variety of different angles. To begin, we originally used high performance liquid chromatography (HPLC). After noticing HPLC was not providing great separation between peaks – which means we were not isolating $GSCG_{tm}$ – we decided to transition to fast purification liquid chromatography (FPLC). From there, we began to run our protein through two size exclusion columns in series, with the hope of furthering separating $GSCG_{tm}$. These results proved promising, but were still not quite where we wanted them to be. Upon the addition of 100mmol NaCl to the running buffer, we noticed that our FPLC graph looked better, with improved peak separation. Currently, we use double FPLC columns, combined with 100mmol NaCl in the running buffer, to purify $GSCG_{tm}$.

While the changes to the FPLC protocol might be the most drastic, other small alterations have also occurred. Following lysis, we began to add the denaturant guanidine hydrochloride, in order to remove lipid impurities. This step has led to what we believe to be more pure $GSCG_{tm}$. The most recent change in protocol, however, comes in the form of a nickel column. Because we believe our protein is aggregating early on in FPLC, we decided to try to purify our protein with an additional step. The six histidines found on our fusion sequence, HIS_6-TRPΔLE, makes it possible for us to run a nickel column. This column binds only our protein, while other impurities within the solution are washed away. We can then collect pure $GSCG_{tm}$. While we are hopeful with this approach, no results have yet been recorded, for it is a recently added step.

Results

We were able to successfully express $GSCG_{tm}$ by recombinant methods using a fusion expression in *E. coli* for both unlabeled and ^{15}N-labeled protein, as shown in **Figure 5**. The amino acid sequence, shown in **Figure 6**, was used to determine the transmembrane portion of $GSCG_{tm}$. This transmembrane domain is illustrated by **Figure 7**. The protein was then purified using chemical cleavage and size exclusion FPLC. Our results indicate that $GSCG_{tm}$ may be aggregating in the FPLC sample; the question remains how to best separate and purify $GSCG_{tm}$. However, we were able to achieve better separation by FPLC when utilizing two columns in series and adding 100mmol NaCl to the running buffer. **Figure 8** shows these results. The purified protein will then be used to conduct *in vitro* glycosylation in order to determine the effects of glycosylation on the structure and dynamics of $GSCG_{tm}$.

Conclusions

The ability to breathe, sit, eat, or walk is commonly taken for granted by those who have fully functioning muscular systems. For those with muscular dystrophy, these thoughtless tasks are prevented due to muscle instability. In fact, by adolescence, many of those affected by muscular dystrophy will be wheelchair bound. To this date, there is no cure for muscular dystrophy, a disorder that affects roughly every 1 in 3500 males (6). However, it is known that the disorder is caused by mutations to proteins found within the dystrophin complex. By focusing my research on $GSCG_{tm}$, I worked to establish beneficial information for those affected by muscular dystrophy. GSCG is a biologically important protein, one whose structures and dynamics will hopefully be determined through an extension of my research, as well as the effects of glycosylation on a truncated form of GSCG.

Figures

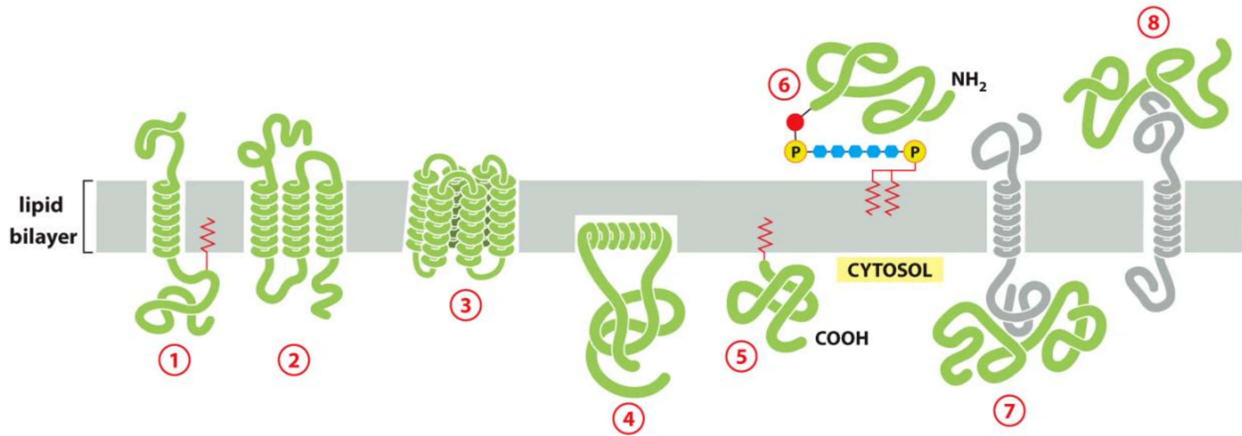

Figure 1. Different forms of membrane proteins (2). GSCGtm most resembles the version labeled "1," for it contains a single transmembrane domain.

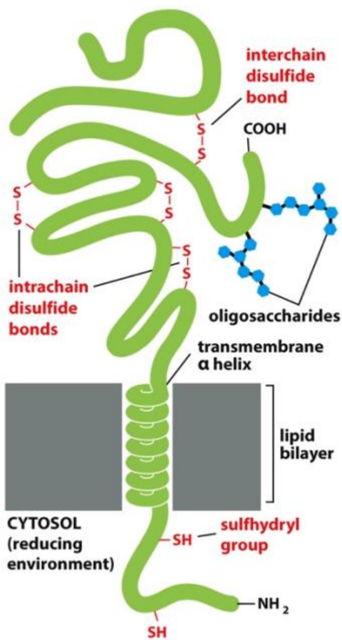

Figure 2. Membrane protein that has been glycosylated, indicated by blue oligosaccharides (2). $GSCG_{tm}$, however, only has one glycosylation site, which will be utilized to add a single glucose monomer.

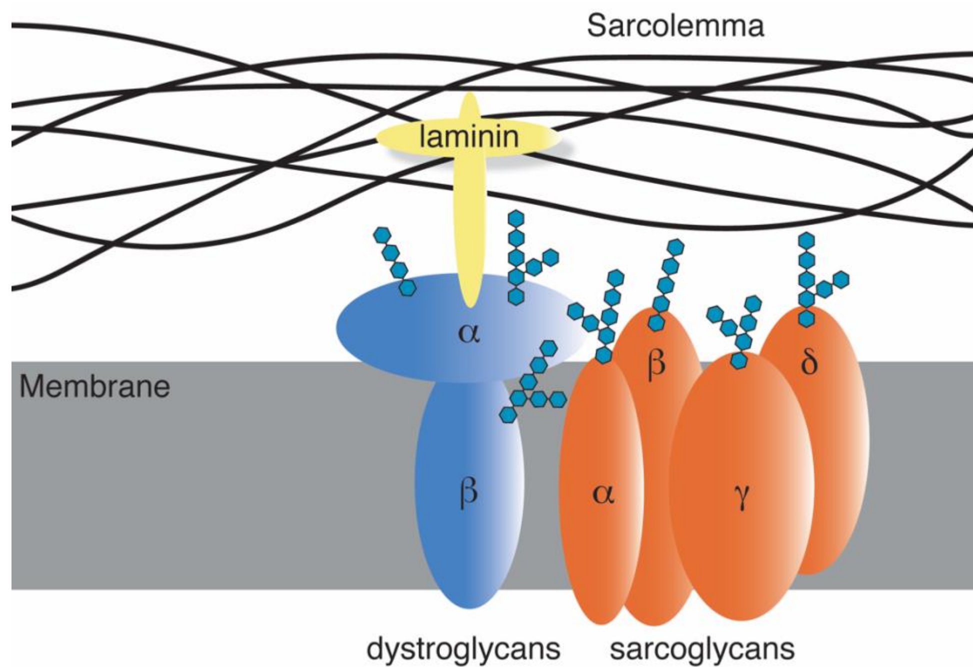

Figure 3. The Dystrophin complex. GSCG can be seen in orange and clearly represents a membrane protein. The complex is shown to be interacting with the sarcolemma of muscle cells.

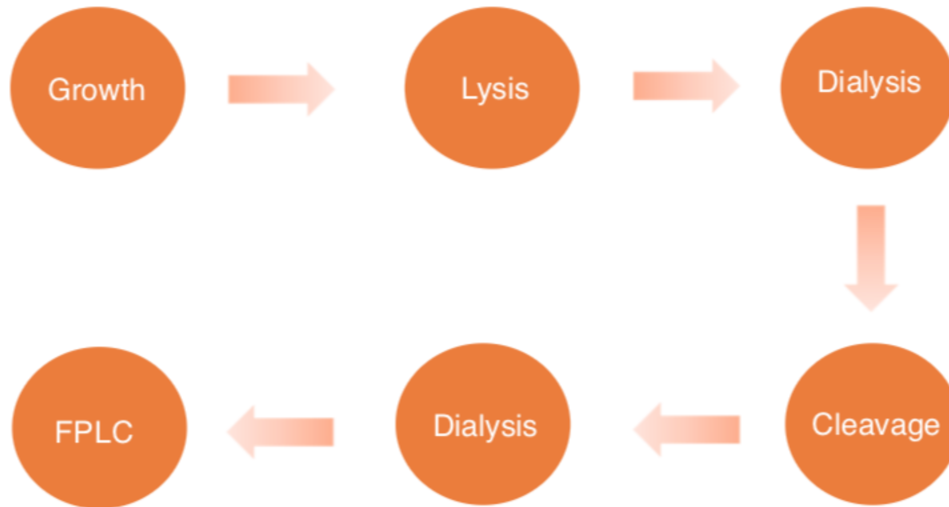

Figure 4. Visual representation of the purification methodology of $GSCG_{tm}$.

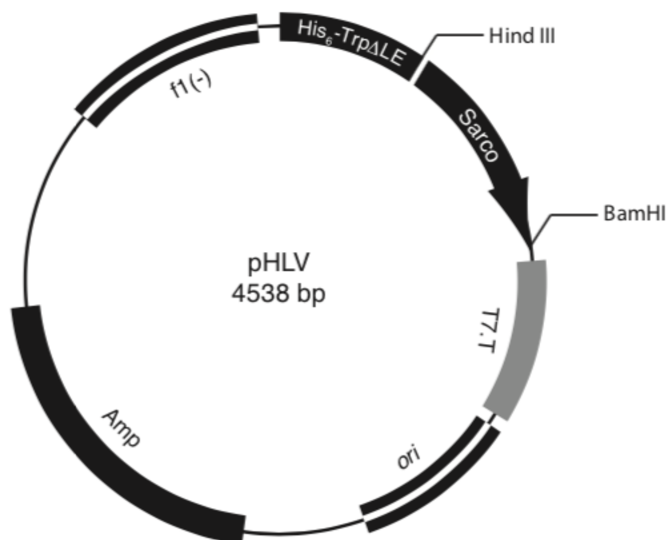

Figure 5. Expression vector map for the $GSCG_{tm}$ construct. The DNA sequence of $GSCG_{tm}$ was inserted within the pHLV plasmid between the HinDIII and BamHI restriction sites. The His-tagged TrpΔLE fusion partner is contained within the plasmid as well. For chemical cleavage, an N-terminal methionine and two C-terminal stop codons were added to the DNA insert.

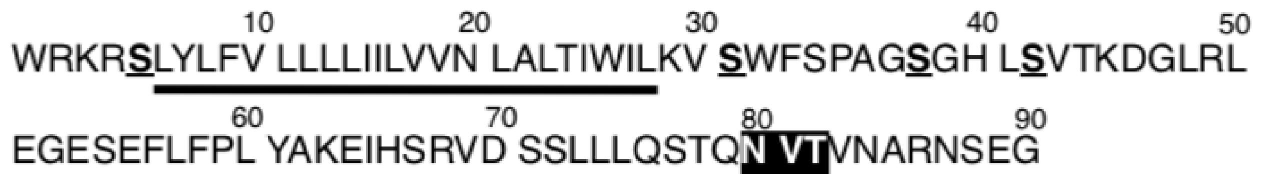

Figure 6. Amino acid sequence of truncated GSCGtm. The glycosylation site, made up of the sequence NVT, is highlighted. Serine substitutions for cysteine and methionine residues are shown underlined and bolded. The transmembrane domain is shown by a solid bar below the sequence.

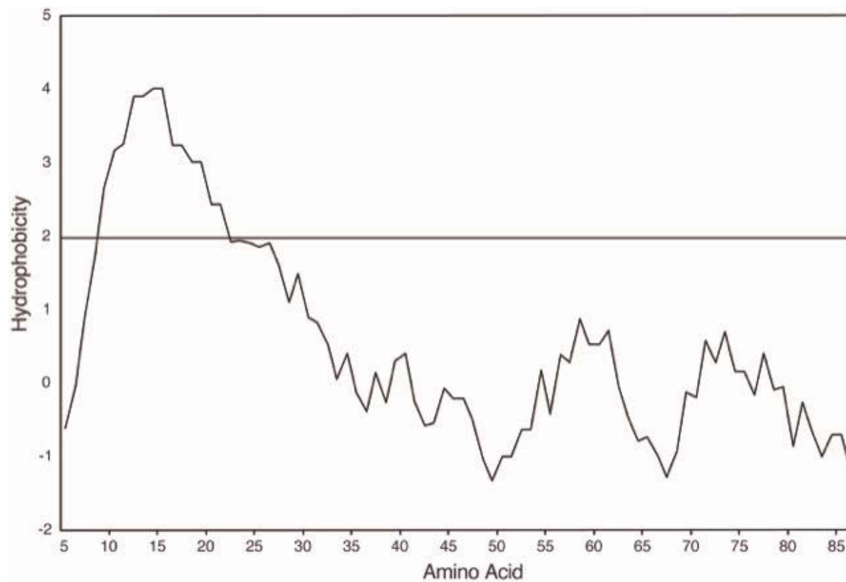

Figure 7. Kyte-Doolite hydropathy plot of GSCGtm. This plot shows that a transmembrane domain exists near the N-terminus of the protein and encompasses approximately 23 amino acids.

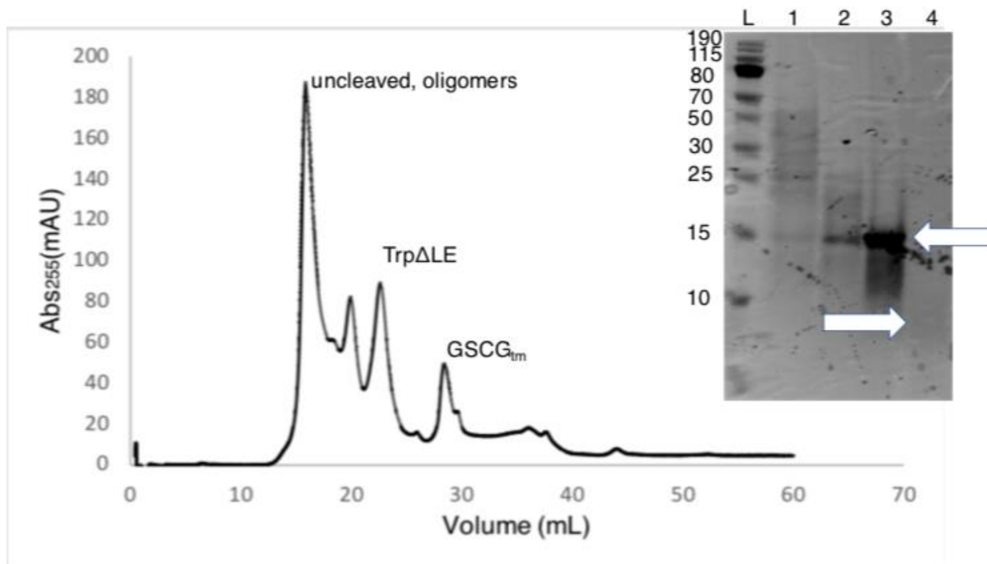

Figure 8. Fast purification liquid chromatography (FPLC) was used to purify the cleaved target protein, followed by an SDS-PAGE gel. Lanes 1-4 correspond to peaks at 16mL, 20mL, 23mL, and 29mL. The third lane corresponds with the size of TrpΔLE while the fourth lane contains a faint band of the purified $GSCG_{tm}$.

References

1. Alberts, B., Johnson, A., & Lewis, J., et. al. (2002). *Molecular biology of the cell* (4th ed.). New York, NY: Garland Science Taylor & Francis.
2. Alberts, B., Johnson, A., & Lewis, J., et al. (2015). *Molecular Biology of the Cell* (6th ed.). New York, NY: Garland Science Taylor & Francis.
3. Lodish, H., Berk, A., Kaiser, C. A., Krieger, M., Bretscher, A., Ploegh, H., . . . Martin, K. C. (2000). *Molecular Cell Biology* (4th ed.). New York: W.H. Freeman and Company.
4. McNally, E. M. (1970, January 01). The Sarcoglycans. Retrieved April 30, 2019, from https://www.ncbi.nlm.nih.gov/books/NBK6317/
5. Schwarz, Flavio et al. "Cytoplasmic *N*-Glycosyltransferase of *Actinobacillus Pleuropneumoniae* Is an Inverting Enzyme and Recognizes the N*X*(S/T) Consensus Sequence." *The Journal of Biological Chemistry* 286.40 (2011): 35267–35274. *PMC*. Web. 23 Jan. 2018.
6. Understanding Duchenne. (2019). Retrieved April 30, 2019, from https://www.duchenne.com/about-duchenne