Characterisation of Glycosylation Mutants in *Drosophila melanogaster*

Matthew Joseph Walker

MSc by Research

University of York

Biology

December 2015

Abstract

Glycosylation is a fundamental process in cellular life, conferring structural and functional properties to proteins and lipids. In order for glycosylation to take place properly, a series of carefully regulated steps must be carried out in specific organelles of the cell. The Golgi apparatus is one of the key organelles where this post-translational modification is carefully controlled by a series of steps in a process known as vesicle trafficking. This multi-layered mechanism has many key phases by which vesicles loaded with cargo are transported from one membrane compartment to another. Glycosylation enzymes make up this cargo and the distribution of these proteins throughout the cisternal compartments of the Golgi is crucial for proper glycosylation to take place. One of the steps of vesicle trafficking is tethering, a process regulated by various proteins including the Conserved Oligomeric Golgi (COG) complex. Defects in subunits of the COG complex leads to perturbations in glycosylation which consequently can result in alterations of cellular functional homeostasis.

In this project we utilise *Drosophila melanogaster* (*Dm*) as a model organism to try and understand more about the role of COG in vesicle trafficking and glycosylation both *in vitro* and *in vivo*. To achieve these aims, using COG mutants, we have carried out *N*- and O-linked glycan profiling and flight test analysis in an effort to connect COG defects *in vivo* to altered glycan patterns *in vitro*. Also using a yeast-two-hybrid (Y2H) approach we looked at COG-Rab interactions to observe how well conserved mammalian-invertebrate trafficking interactions are. What we found are not only functional implications of COG in determining *Drosophila* glycan synthesis and flight ability, but also evidence of evolutionary conservation during interactions associated with membrane trafficking events.

Contents

List of Figures

List of Tables

Acknowledgements

Declaration page

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this or any other University. All sources are acknowledged as References.

1. Introduction

Glycobiology can be defined as the study of carbohydrates in cellular life, accounting for their structure, biosynthesis and functional implications in nature. Included in this field are glycans which are oligosaccharide structures covalently bound to proteins and lipids during glycosylation. However unlike other polymeric structures in biology such as proteins and nucleic acids, our understanding of glycans remains limited in comparison. This is attributable to the fact that glycans lack a template sequence or code to confer structural properties. Consequently the resultant agglomeration of glycan structures in a cell or organism (glycome) is highly heterogeneous meaning sensitive experimental approaches are essential to elucidate glycan structure and function.

Glycosylation is one of the more prevalent types of post-translational modification in eukaryotes with roughly one fifth of all proteins in the swiss-prot database found to be glycosylated, a large proportion of these existing in eukaryotic systems [1]. It is therefore not surprising that glycosylation is a fundamental process in assigning functional specificity to molecules. Glycan-specific roles are prevalent in a multitude of scenarios, for instance by influencing cell surface signalling or determining immunogenicity during immune responses [2, 3]. Such roles can be perturbed following defects in normal glycan processing which can consequently result in forms of disease. Faulty glycosylation has been implicated in a variety of conditions most notably Congenital disorders of glycosylation (CDG's) which exhibit highly pleiotropic phenotypes [4]. Moreover other diseases have been associated with improper glycosylation such as Alzheimer's [5] and cancer [6]. However despite much already being known about glycan functions in biology, due to the complexity of the field, more remains to be unveiled. The elucidation of glycan roles have and continue to be achieved following studies on the Golgi apparatus, the organelle where glycosylation chiefly takes place.

Studies in glycosylation have frequently been carried out in mammalian cell types using *in vitro* experimental approaches. Although this provides further insight into glycosylation, the use of *in vivo* systems aids to compile evidence for glycan roles at a multicellular, whole organism level. A key function of glycans is mediating cell-tissue interactions, a feature which cannot be recapitulated using simple cell culture based approaches. Identification of glycan-specific roles have been uncovered using mutation-based manipulations of glycosylation pathways in model organisms such as fruit flies [7] and nematode worms [8]. Such model systems are a useful tool for glycosylation studies in comparison with humans due to their high sequence homology at the DNA and protein level, as well as their simple life cycles and multicellular composition.

1.1 Glycosylation

Proteins and lipids are the molecules which specifically undergo glycosylation, the former being the subject of this project. A variety of different types of protein glycosylation exist, the most common being *N*-and O-linked. *N*-linked glycosylation differs from O-linked in that *N*-linked glycan processing initially occurs in the Endoplasmic reticulum (ER) before continuing in the Golgi apparatus, conversely O-linked takes place exclusively in the Golgi. Secondly *N*-linked glycosylation involves a linkage between a hydroxyl group of an *N*-acetylglucosamine (GlcNAc) carbohydrate molecule and an amide nitrogen atom of an asparagine amino acid residue after recognition of an Asn-X-Ser/Thr consensus sequence, X being any amino acid besides proline [9]. By contrast O-linked glycosylation is the attachment of a carbohydrate to the oxygen atom of a serine or threonine amino acid residue with site specific sequence recognition more poorly characterised.

N-linked glycosylation is a multi-step procedure which begins following the formation of a lipid-carbohydrate structure, specifically a dolichol-linked *N*-acetylglucosamine in the ER. This then becomes the basis for subsequent monosaccharide addition initially on the cytoplasmic face of the ER, followed by within the ER luminal face until an oligosaccharide precursor structure is synthesised. An oligosaccharyltransferase enzyme is responsible for transferring the precursor to an asparagine residue of a polypeptide chain residing in the ER after consensus sequence recognition [10]. Once the precursor is covalently linked to the amino acid, upon correct protein folding, 3 glucose residues are trimmed from the glycan by glucosidase enzymes [10]. This is a cue for protein transit from the ER to the Golgi where further glycan processing occurs. Once fully synthesised, *N*-glycans can be structurally sub-divided into 3 main types based on the monosaccharide compositions of the glycan in question. All 3 forms share a common core structure of 2 GlcNAc and 3 mannose residues (**Figure 1A**) with the high mannose *N*-glycan form only possessing recurring mannose units following this core structure. Conversely a complex *N*-glycan is made up of various different carbohydrate residues as well as a hybrid *N*-glycan which is a combination of the previous forms (**Figure 1A**). There is a vast diversity of glycan structures in terms of their structure leading to a great extent of heterogeneity. The examples shown in **Figure 1A** are just to highlight the different structural subtypes *N*-glycans are classified into, in reality there are many more glycan structures which fit into the category of high mannose, hybrid or complex.

By contrast the mechanism for O-linked glycosylation is not so well understood. Unlike *N*-linked glycans with a precursor core, O-glycan cores are much more heterogeneous as they can be composed of various different carbohydrate residues. O-glycans are classified based on the first monosaccharide residue attachment such as O-GalNAc (*N*-acetylgalactosamine)or O-Fuc (Fucose) which are subsequently elongated and modified to form an array of glycan structures (**Figure 1B**). O-glycans are prevalent on various molecules including mucins, proteoglycans and collagen [11, 12].
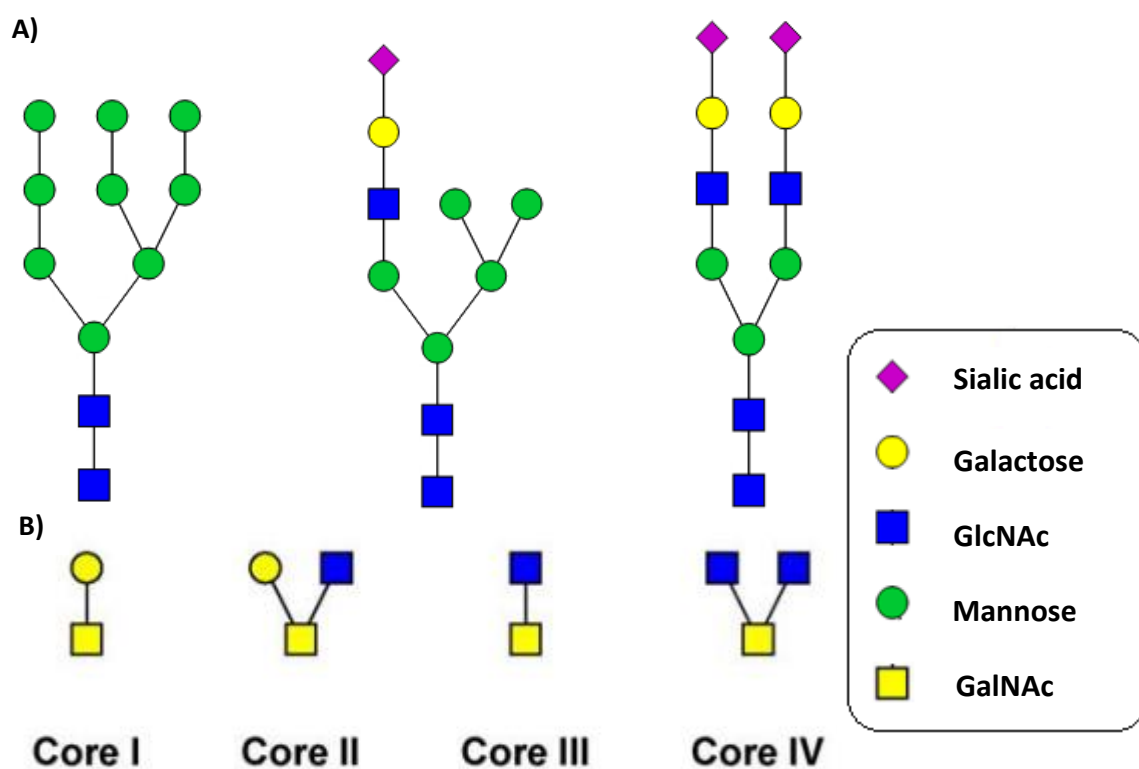
Figure 1. *N*-glycan and O-glycan structures: A) *N*-glycan structural subtype examples, from left to right; High mannose, hybrid and complex forms. B) Mucin-type O-linked glycan cores in mammals [38]

1.2 Golgi vesicle trafficking and The Conserved Oligomeric Golgi (COG) complex

Glycosylation chiefly takes place in the Golgi apparatus, an organelle which can be structurally described as an assemblage of flattened membranes termed cisternae. The cisternal compartments residing throughout the Golgi are organised in a specific manner based on the internal compositions of each. The earliest Golgi cisterna is known as the *cis* Golgi which is followed by the *medial* then *trans* Golgi ending at the latest compartment the *trans*-Golgi network. Residing within each cisternal layer are Golgi resident proteins, particularly in the interest of glycosylation, these include glycosylation enzymes [13].

A broad array of glycosylation enzymes have been identified, all of which can be categorised into two main groups; glycosidases and glycosyltransferases. The former is responsible for glycan trimming and the latter that of monosaccharide addition. Specific glycosylation enzymes are localised to particular cisternal compartments as they are involved at different stages of glycan processing. For instance enzymes found localised to the *medial* Golgi are involved at intermediate stages of glycan processing such as α1, 3-1,6 mannosidase II [14] , whereas those found residing at the *trans* Golgi like α2, 6 sialyltransferase process glycans at a much later stage [14]. Therefore it is the non-uniform distribution of glycosylation enzymes throughout the Golgi which determines subsequent glycosylation patterns and finalised glycan structures.

Various models have been proposed to suggest the movement of resident protein cargo through the Golgi which subsequently determines the distribution of glycosylation enzymes throughout the Golgi. A widely accepted model of such movement is cisternal maturation which depicts that proteins destined for transport to the Golgi, upon correct folding, are packaged into transport vesicles coated in a COPII (Coat Protein II) protein coat which targets vesicles from the ER to the *cis* Golgi face in an anterograde fashion [15]. These vesicles subsequently fuse with retrograde travelling vesicles coated in a COPI (Coat Protein I) protein coat which contain recycled *cis* Golgi resident proteins from an older *cis* Golgi. This results in the formation of a new *cis* Golgi subsequently allowing for *cis* Golgi specific glycosylation modifications to occur [15]. This cisterna will progressively mature later forming *medial, trans* and *trans*-Golgi network compartments via the same mechanism of retrograde recycling. At which point proteins will be fully glycosylated at the *trans*-Golgi network and are packaged into vesicles destined for functional roles elsewhere. It should be stressed that cisternal maturation is a highly dynamic process with new cisternae constantly forming due to recycling of Golgi resident proteins. The cisternal maturation model has evidence to support the claim for this mechanism of vesicle trafficking in the Golgi, including observations of enriched Golgi resident proteins in COPI vesicles over newly synthesised proteins [16] which contradicts the two-way vesicle transport system proposed by the vesicular transport model.

A process termed vesicle trafficking generates the non-uniform distribution of glycosylation enzymes between cisternae by maintaining cargo sorting into vesicles as well as vesicle targeting itself. Vesicle trafficking can be subdivided into different phases; budding, transport, tethering and fusion. Budding initiates trafficking via the release of a vesicle from a donor membrane compartment followed by transport away. Following this comes tethering which involves the capture of a vesicle from a target membrane and subsequently mediating its migration to a target membrane compartment. Finally fusion occurs where stable SNARE (Soluble NSF Attachment Protein Receptor) complexes form tightly, connecting the vesicle to the membrane. This allows for vesicle contents to be incorporated into the target membrane while a SNARE complex comprised of a membrane bound t-SNARE (target-Soluble NSF Attachment Protein Receptor) and vesicle bound v-SNARE (vesicle-Soluble NSF Attachment Protein Receptor) ensures energy provision for fusion. During Golgi vesicle trafficking a plethora of proteins are involved in mediating each step such as the coiled-coil proteins GMAP-210 (Golgi Microtubule-Associated Protein 210) and GM130 (Golgi Matrix Protein 130) involved in tethering at the ER-Golgi interface [17]. One of the main families of proteins involved in regulating all steps of vesicle trafficking are the Rab GTPases. These proteins cycle between inactive guanosine diphosphate (GDP) and active guanosine triphosphate (GTP) bound forms [60]. Rabs will sequentially interact with various effectors allowing them to mediate vesicle trafficking steps from budding to fusion. The sequential events facilitating Rab membrane recruitment and activation firstly involve Rab escort proteins (REPs) which function to deliver Rabs to their appropriate membrane destination [60]. Following Rab delivery Guanine nucleotide exchange factors (GEFs) trigger the exchange of GDP for GTP thereby causing Rab activation [60]. GTP hydrolysis is accelerated through GTPase activating proteins (GAPs) before the Rab is sequestered into the cytosol through recognition by GDP dissociation inhibitors (GDIs) for the cycle to repeat again [60]. These, amongst other players, are critically important in the correct cisternal compartmentalisation of glycosylation enzymes to ensure proper glycosylation takes place **(Figure 2A).**
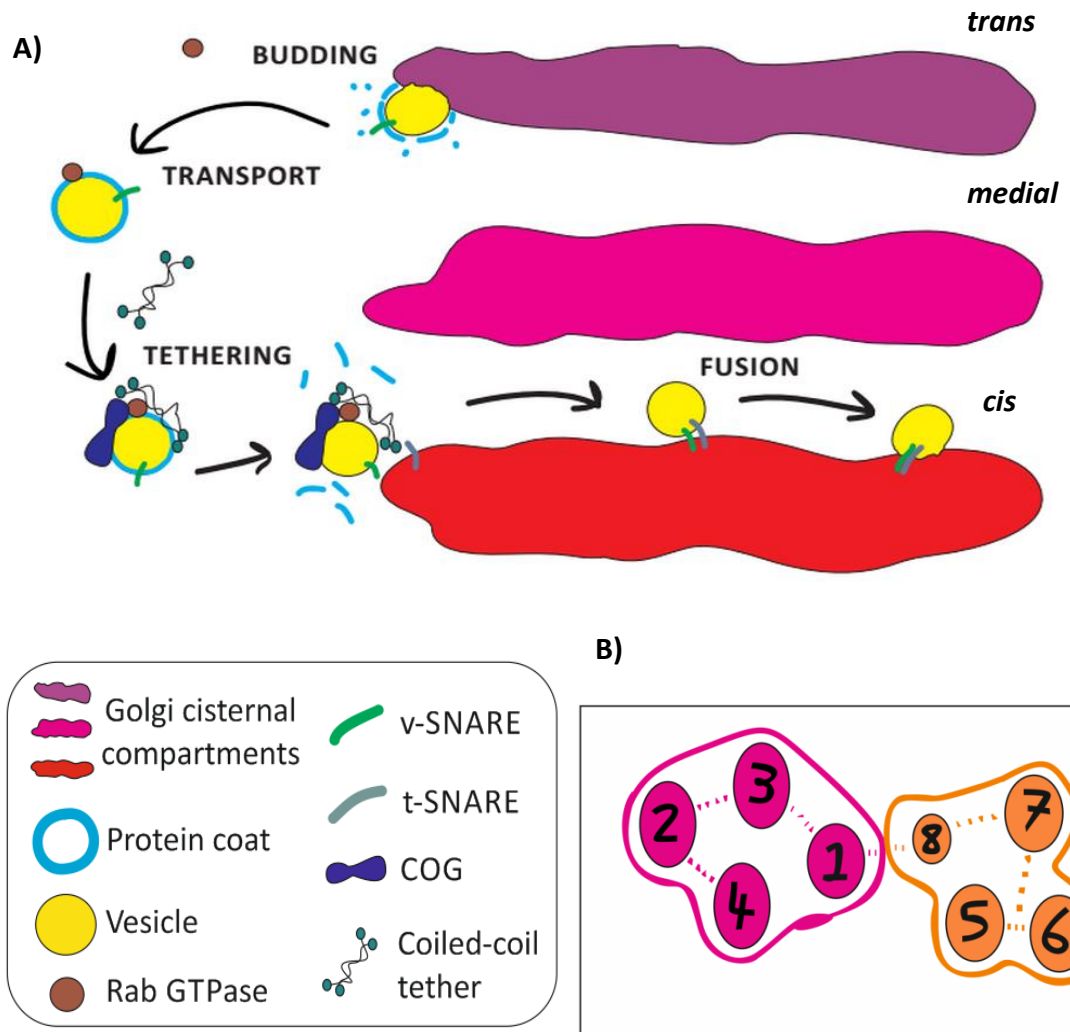
Figure 2. Schematic of budding, transport, tethering and fusion during intra-Golgi vesicle trafficking and Conserved Oligomeric Golgi (COG) structure: A) Golgi vesicle trafficking steps. B) Conserved Oligomeric Golgi (COG) complex structure; lobe A (pink) lobe B (orange)

Rab GTPases and coiled-coiled tethers have been implicated to work in concert with various other proteins in order to achieve vesicle tethering. The Conserved Oligomeric Golgi (COG) complex is a multi-subunit tethering complex (MTC) localised to the Golgi apparatus shown to be involved in tethering [23]. Structurally COG is a hetero-octameric, bi-lobed complex with COG subunits 1-4 forming lobe A and 5-8 comprising lobe B, subunits Cog1 and Cog8 form an interconnected dimer between each lobe [23] **(Figure 2B)**. Lobe A is proposed to function at early stages of glycan processing compared to lobe B at a later stage based on findings of defective lobe A and B subunits affecting levels of different Golgi resident proteins respectively. Levels of early glycosylation enzyme mannosidase II were affected in lobe A defective cells [21], whereas evidence suggests levels of late enzymes galactosyl and sialyltransferase are sensitive to lobe B defects following observations of reduced galactosylation and sialylation in Cog7 mutant patients [22].

The implication of COG in vesicle tethering has been largely elucidated based on findings of COG subunit interactions with a variety of proteins known to function in vesicle trafficking **(Figure 3)**.
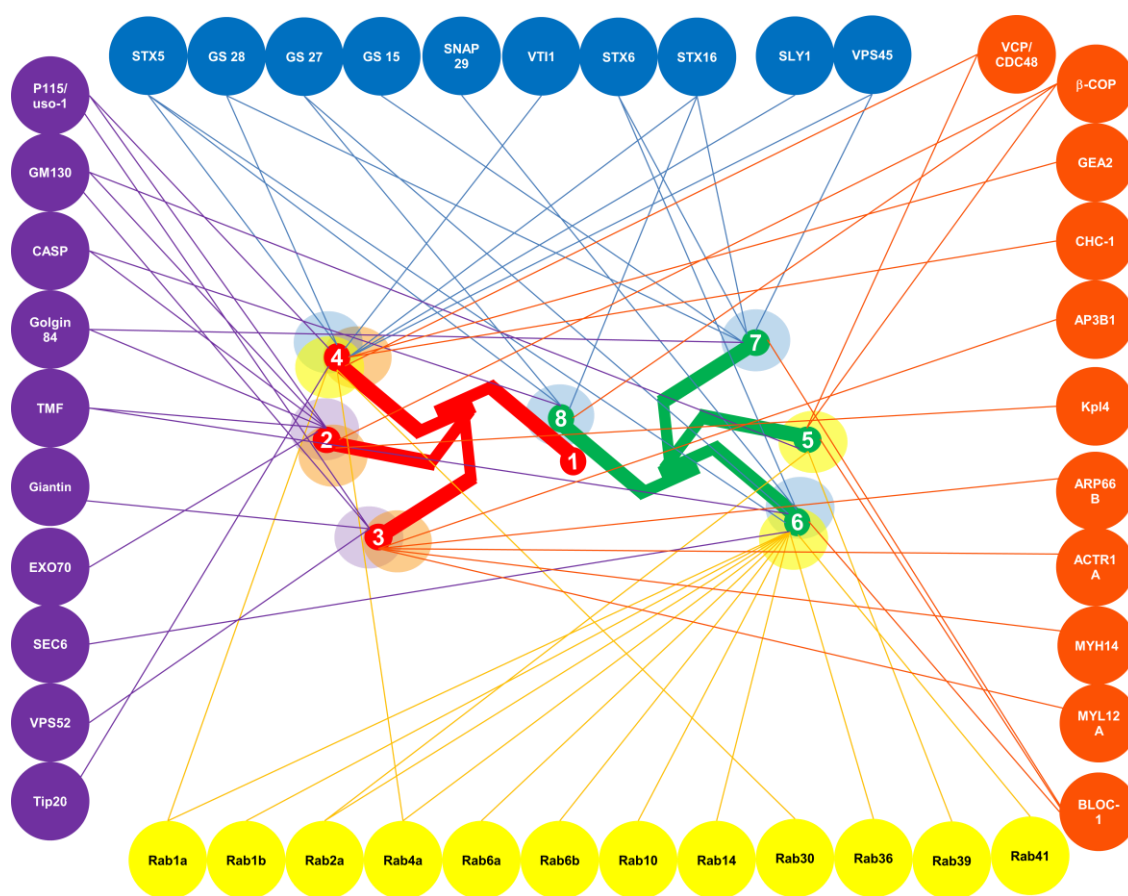
Figure 3. Mammalian COG subunit interactions with vesicle trafficking proteins: Rab GTPases (yellow), SNAREs (blue), tethers (purple), coat and motor proteins (red) [23]

Based on reported interactions, various models have been suggested to operate sequentially as to how COG may assemble with other proteins in aid of vesicle trafficking. The docking station assembly model depicts COG as an orchestrator of protein assembly at an acceptor membrane, including Rab GTPases and coiled-coil tethers. In agreement with this model are observations of Cog4 recruiting COG subunits, *cis*-Golgi Rabs and the coiled-coil tether p115 to STX(Syntaxin)5 membranes, as well as Cog8 recruitment of lobe B components and *trans*-Golgi Rab6 to STX16 membranes [23]. The SNARE stabilisation model suggests a successive step in that COG later acts to mediate stable SNARE complex assembly through multi-pronged binding to pre-formed t- and/or v- SNAREs. In support of this model are observations of COG-deprived cells causing a decrease in steady state levels of Golgi-operating SNARE complexes as well as COG interactions with SNARE-associated proteins such as Cog4 with STX5 partner protein Sly1 to stabilise SNARE complexes [24]. A further model highlights another part for COG to play in this puzzle as a direct facilitator of vesicle tethering by bridging the space between a vesicle and acceptor membrane

through simultaneous interactions with Rab GTPases, coiled-coil tethers and SNARE proteins. Indeed supporting this suggestion are yeast and mammalian cells deficient in COG subunits which were found to accumulate non-tethered vesicles composed of recycled Golgi components [25]. Additionally there is evidence of an immobilised COG complex directly binding to intra-Golgi vesicles [27] as well as observed COG subunit interactions to COPI coat components [26]. Formation of a bridge-like connection between vesicles and membrane compartments could be explainable by the structure of COG. Indeed electron micrographs have shown COG displays a tentacular-like morphology [27] **(Figure 4)** which is explainable by the fact that COG subunits are composed of elongated helical bundle domains, permitting extended tentacle protrusions [20]. Such an elongated structure potentially allows for extended connections between COG subunits and various trafficking proteins thereby directly mediating vesicle tethering. The successive COG interactions described with multiple trafficking partners have shed light onto the roles of COG in vesicle trafficking; further studies into these aspects will certainly substantiate evidence for the mechanisms proposed by the aforementioned models.
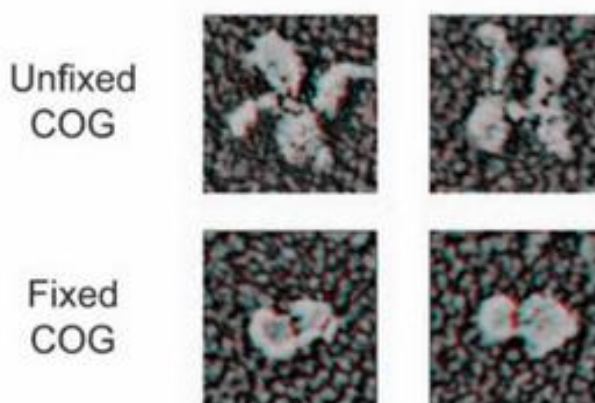
Figure 4. Electron micrographs of fixed and unfixed COG conformations: Native COG purified from bovine brain, Fixed compact COG structure upon glutaraldehyde fixation of 10 µg/ml COG solution mixed with 7% glutaraldehyde (bottom), unfixed COG where glutaralyehyde fixation was omitted displaying tentacular protrusions (top) [37]

1.3 *Drosophila melanogaster;* a model organism for glycosylation

Utilising model organisms as a molecular tool to study biological effects at a whole organism level provides a distinct advantage over using *in vitro* approaches. This is because specific mutations could impact an organism in a variety of different ways which may easily be overlooked when analysing localised effects *in vitro.* Whereas *in vivo* experimentation provides evidence on how a particular deviation from the norm via mutation could impact an organisms behaviour. With glycosylation being a rather complex procedure having implications in cell-cell and cell-tissue interactions, using *Drosophila melanogaster* as a model organism provides a simple, manipulative system in which to look at effects of mutations on glycosylation homeostasis and *Drosophila* behaviour. This offers a distinct advantage over using simple cell culture based methods which overlook organism-wide effects, allowing for *in vivo* roles of glycans to be analysed.

Drosophila melanogaster has and continues to be one of the most widely used model organisms in biological research. This is a useful system for many reasons, including their short generation time and lifespan, high fecundity, and simple genome. Indeed *Drosophila* share high sequence homology with humans at both the DNA and protein level including around 75% known human disease related genes conserved in the *Drosophila* genome [28]. Moreover many proteins key to regulating vesicle trafficking are conserved between mammalian and fruit fly systems including subunits of the COG complex and Rab GTPases.

Although there are many similarities between *Drosophila* glycosylation and that of mammalian systems, it is worth noting differences between the two. Despite *Drosophila* having both hybrid and complex *N-glycans*, these are relatively rare and they predominantly produce high mannose and small high mannose oligosaccharide forms known as paucimannose glycans. This is attributable to the fact that invertebrates lack many of the glycosylation enzymes known to process more complex forms of *N-glycans* in mammalian systems. A key enzyme implicated in the lack of complex and hybrid forms of *N*-glycans in *Drosophila* is a *β-N*-acetylglucosaminidase named fused lobes (Fld) which functions to remove terminal GlcNAc residues [29]. In vertebrates additional GlcNAc residues are maintained on the *N*-glycan chain allowing for elongation of elaborate branching structures from the trimannosyl core; whereas in *Drosophila* due to GlcNAc cleavage such elongation cannot frequently occur. Indeed mutations in Fld lead to more hybrid and complex *N*-glycan forms in *Drosophila* which were more representative of those seen in vertebrates [29]. Another difference between vertebrate and invertebrate *N*-glycans is the possibility for fucosylation linkages to comprise a modified glycan core. Vertebrate *N*-glycan cores have the potential for a 1-6 fucosylated linkage on the core GlcNAc of the glycan. However in invertebrate systems there is not only the capacity for 1-6 linkages, but also a 1-3 core fucosylated linkage **(Figure 5)**.

Mucin type O-glycans have also been found to be functionally implicated in *Drosophila,* for example RNAi of individual genes encoding enzymes in the *pgant* family of GalNAc transferases were found to cause cellular morphological and functional changes, including *pgant*3, *pgant6* or *pgant7* effects on secretion [55]. Also RNAi of *pgant3* and *pgant6* induced changes of Golgi morphology suggesting a key role for O-glycosylation in maintaining proper secretory apparatus structure and function [55]. Moreover, in *Drosophila* RNAi of other members within the *pgant* enzyme family have been shown to cause lethality *in vivo* implicating them in viability [56].
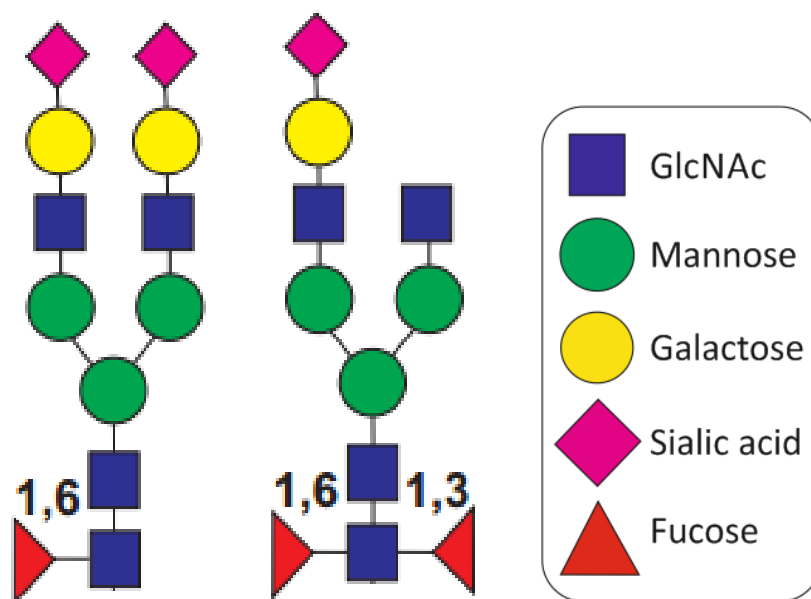
Figure 5. Core-fucosylated *N*-glycans: Mammalian-like glycan with 1-6 linkage (left), *Drosophila*-like glycan with 1-6 and 1-3 linkage (right)

1.3.1 Glycan function in *Drosophila* muscle

Proper glycosylation is important for muscular development and contraction in eukaryotic life. An instance of *N*-glycan significance in *Drosophila* muscle is shown where the *N*-glycan processing enzyme in *Drosophila* MGAT1 (UDP-GlcNAc:α-3-D-mannoside-β1,2-N-acetylglucosaminyl-transferase I) which has the human ortholog GlcNAc-transferase I, was found to reduce neurotransmitter release and perturb normal synaptogenesis at the neuromuscular junction when expressed as a null mutant form [33]. Moreover O-glycosylation has been show to impact on *Drosophila* muscle where mesodermal cells defective in O-glycosylation resulted in an impaired larval muscular system [39].

Preliminary data has indicated COG subunit mutation has an impact on *Drosophila* flight ability **(Figure 6)**. Notable results show significant defects in lines where mutant forms of the *Drosophila* homolog of the COG subunit Cog5 known as *four way stop (Fws)* were expressed as homozygotes

with either Cog1 and Cog2 mutations also present or the Cog1 mutation present alone. However balancer chromosomes are present in many of the COG mutant genotypes and therefore some of the defects observed could be attributable to the affect of the balancer rather than the COG mutation. Therefore in order to substantiate the evidence for a COG specific affect on *Drosophila* flight ability controls of COG mutants without balancer chromosomes present should be analysed.
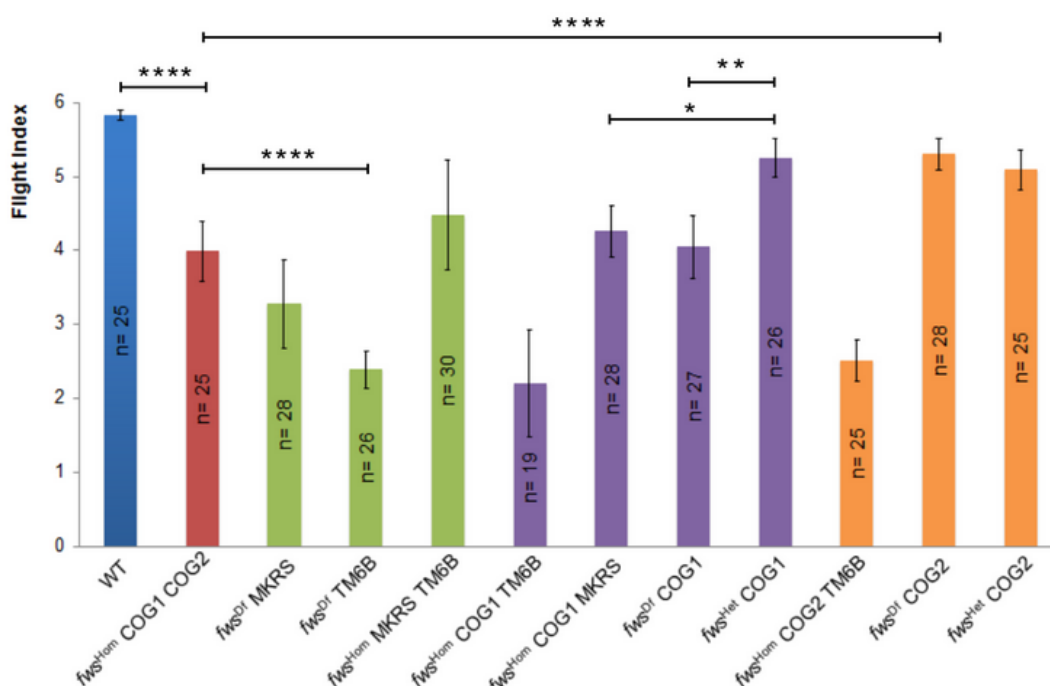
Figure 6. Flight testing of COG mutant flies: 3-5 day old flies released into a clear box with a light source above, scoring based on direction of flight upwards (6), horizontal (4), downwards (2), none (1). *Fws = four way stop* (*Drosophila* Cog5 homolog), Hom = homozygous, Het = heterozygous, Df = *fws* deletion, COG1/COG2 = COG1/COG2 mutant, TM6B and MKRS = balancer chromosomes. One-way ANOVA used for analysis of statistical significance, asterix represent p values upon sample comparison, * = $p<0.05$, ** = $p<0.01$, **** = $p<0.0001$, error bars represent standard deviation (Eric Silva, Ungar lab, unbublished data, 2013)

1.4 Project outline and aims

The questions I set out to ask from the project were three-fold; firstly I wanted to address the query of whether mutations in COG subunits affected *N*- and O-linked glycan presence and abundance in *Drosophila* muscle specific cells. To set out answering this question I have been carrying out *N*- and O-linked glycan profiling using thoraces isolated from a *Drosophila* line with an absence of protein subunit Cog3 in muscle specific cells, a mutation which renders the organism completely flightless, as well as wild type line controls. For *N*-glycan analysis I have been using a novel enzyme PNGaseAr to cleave glycans with *Drosophila*-like fucosylation linkages from glycoproteins which other enzymes are incapable of achieving. I have been able to optimise a working protocol for *N*-glycan release for use with PNGaseAr after initially finding the enzyme was not compatible with the previous protocol and subsequently used this on *Drosophila* samples.

The next question I wanted to answer was whether COG mutations are detrimental enough to protein glycosylation in *Drosophila* muscle to cause observable flight defects *in vivo*. To address this I have carried out various fly line crosses to generate COG mutant genotypes to test for their flight ability. Findings correlate with previous results suggesting certain COG subunits more than others play a role in regulating the processes which underpin flight muscle contraction. This aim ties in with the *N*- and *O*-linked glycan analysis as any flight defects observed could then be attributed to differences in observed glycan profiles of COG mutant *Drosophila* strains. Linking these aims could provide further evidence to implicate COG in mediating the vesicle trafficking essential for glycosylation and therefore subsequently affect the specific glycan structures synthesised which may be important for muscle-specific roles.

The final question I wanted to address was whether COG-Rab protein-protein interactions are conserved between mammalian and *Drosophila* systems. This would certainly provide evidence for how effective *Drosophila* can be as a model organism for glycosylation studies. To answer the question I have used a yeast-two-hybrid approach to look at COG subunit interactions with Rab-GTPases in *Drosophila*. Observations seem to suggest evolutionary conservation between various COG-Rab interactions as well as unveiling some interactions specific to the *Drosophila* system.

Overall from the project I aim to highlight a role for COG in *Drosophila* protein glycosylation and subsequently muscle functionality, as well as provide evidence for conservation between glycosylation pathways in mammalian and invertebrate systems.

2. Materials and Methods

2.1 Yeast-two-hybrid COG constructs

Primers were designed against *Drosophila* COG subunit cDNA sequences with restriction enzyme cut sites present (**Table 1**). PCR products were generated using 30 cycles with 30 second elongation times per cycle and a final elongation time of 12 minutes. PCR products were PCR purified (QIAGEN) before restriction digest and ligation reactions. Plasmid DNA was amplified was in competent DH5α *E.coli* before DNA extraction (QIAGEN miniprep) and Sanger sequencing. Plasmid constructs are shown in **Table 2**.

Primer name	Sequence	Annealing temp (°C)
*Dm*Cog4Y2Hf	GAGATCGAATTCATGAGTGTGCTGGAACA	51.0°C
*Dm*Cog4Y2Hr	ATCGATGGATCCCTAAAGTTGTAGCCGCTTAATG	51.7°C
*Dm*Cog5Y2Hf	AAAGGGGGATCCATGGTGACTGGAGACCCG	56.7°C
*Dm*Cog5Y2Hr	AACTGCAGGTCGACTTATGGAAGTGCCTTTAAGGCC	55.3°C
*Dm*Cog6Y2Hf	GAG ATC GAA TTC ATG AGC TCG ACG CAG G	54.8°C
*Dm*Cog6Y2Hr	AAGTC CAG ATC GAT CTA AGT GGA GGT CAG GAT GTG	54.6°C

Table 1. Primers used for *dm*COG plasmid construct cloning

Plasmid name	Vector	Additional information
*Dm*Cog4AD	pGAD-C1	AmpR, LEU2 yeast selection, cloned with BamHI and EcoRI
*Dm*Cog5AD	pGAD-C1	AmpR, LEU2 yeast selection, cloned with BamHI and SalI
*Dm*Cog6AD	pGAD-C1	AmpR, LEU2 yeast selection, cloned with EcoRI and ClaI
*Dm*Rab1BD (Q-L)	pGBDU-C1	AmpR, URA3 yeast selection, donated by Rita Sinka
*Dm*Rab2BD	pGBDU-C1	AmpR, URA3 yeast selection, donated by Rita Sinka
*Dm*Rab3BD (Q-L)	pGBDU-C1	AmpR, URA3 yeast selection, donated by Rita Sinka
*Dm*Rab4BD (Q-L)	pGBDU-C1	AmpR, URA3 yeast selection, donated by Rita Sinka
*Dm*Rab6BD (Q-L)	pGBDU-C1	AmpR, URA3 yeast selection, donated by Rita Sinka
*Dm*Rab10BD (Q-L)	pGBDU-C1	AmpR, URA3 yeast selection, donated by Rita Sinka
*Dm*Rab30BD (Q-L)	pGBDU-C1	AmpR, URA3 yeast selection, donated by Rita Sinka
*Dm*Rab39BD (Q-L)	pGBDU-C1	AmpR, URA3 yeast selection, donated by Rita Sinka

Table 2. Plasmid constructs cloned and used for yeast-two-hybrid

2.1.1 Yeast-two-hybrid

GAL4 Activation domain (AD) COG constructs and Rab GAL4 DNA Binding domain (BD) plasmids (donated by Rita Sinka, University of Szeged) were co-transformed into the AH109 HIS reporter yeast strain [44] and grown on selective plates lacking Leucine and Uracil, transformants were re-streaked onto plates lacking Leucine and Uracil. Colonies of yeast from re-streaks were mixed and resuspended in autoclaved water in a 96 well plate. Alongside this a 20-fold dilution of each yeast sample was aliquotted aside the initial resuspension. An inoculating manifold replicator/frogger was soaked in ethanol and flamed before being allowed to cool next to a Bunsen flame. Once cooled the frogger was dipped into wells containing the yeast samples and quickly transferred onto plates lacking Leucine, Uracil and Histidine/Adenine. Plates were incubated at 30°C and left for 1 week to grow.

2.2 PNGase deglycosylation assays

Proteins were denatured by heating to 95°C for 5 minutes in the presence of denature solution (0.2% SDS, 100mM 2-mercaptoethanol). Samples were cooled on ice and briefly centrifuged at 14,000rpm before addition of TritonX-100 (TX-100) to a concentration of 2.25%. PNGaseF and PNGaseAr enzymes were added and samples were incubated at 37°C. Sample buffer (5% glycerol, 50mM Tris 50mM DTT (Dithiothreitol), 1% SDS, 0.74mM bromphenol blue) was added to samples before heating at 95°C for 5 minutes. Proteins were run on 10% and 15% SDS-PAGE gels and then Fairbanks coomassie stained[45].

2.3 *Drosophila* husbandry

Drosophila maintenance and crosses were all performed using a modified semi- defined media (7.5% agar, 41.8% cornmeal, 9.9% brewers yeast, 33.6% sucrose, 7.2% propionic acid) supplemented with anti-fungal agents at 25°C. Crosses were conducted by collecting 10 virgin females and 10 males and adding these to a media vial. These were left for 2 days at 25°C before removing adult flies; vials were left for a further 5 days and then checked morning and evening for progeny genotypes. *Drosophila* stock genotypes are shown in **Table 2**.

Stock Genotype
WT (Canton S)
WT (White$^-$)
$\dfrac{fws\text{p}}{fws\text{p}};\dfrac{\text{Cog2}}{\text{TM6B}}$
$\dfrac{fws\text{p}}{fws\text{p}};\dfrac{\text{Cog1}}{\text{TM6B}}$
$\dfrac{fws\text{p}}{\text{CyO}}$
$\dfrac{fws\text{p}}{fws\text{p}}$
$\dfrac{\text{Cog2}}{\text{TM6B}}$
$\dfrac{\text{Cog1}}{\text{TM6B}}$
Dmef2-GAL4
Dmef2-GAL4
Cog3RNAi-UAS
$\dfrac{fws-\text{GFP}}{\text{CyO}}$

Table 3. *Drosophila* stock lines; $\frac{fws-\text{GFP}}{\text{CyO}}$ line donated (Rita Sinka, University of Szeged), Cog3RNAi line was acquired from the Vienna Drosophila RNAi Center (VDRC) and the Dmef2 lines were provided by John Sparrow (University of York) [59]

2.3.1 *Drosophila* dissections

100 3-5 day old flies of genotypes $\frac{Dmef2-GAL4}{Cog3RNAi-UAS}$, $\frac{Dmef2-GAL4}{WT}$, $\frac{Cog3RNAi-UAS}{WT}$ were anesthetised using CO_2 and their thoraxes were dissected using a razor blade. Thoraces were kept at -20°C during dissections and were stored at -80°C after dissections.

2.4 Cell lysis/protein denaturation

Drosophila thoraces were lysed using a pellet pestle motor in 200µl lysis buffer (4% SDS, 100mM Tris/HCl pH 7.6, 100mM dithiothreitol). In the case of RNAseB and Fetuin proteins they were diluted 1:10 in lysis buffer. Samples were incubated at 95°C for 5 minutes before centrifugation at 14,000rpm for 5 minutes. Lysate supernatants were collected and transferred to fresh microfuge tubes.

2.4.1 Filter-aided *N*-glycan separation (FANGS)

Samples were diluted in Urea solution (8M in 100 mM Tris/HCl pH 8.5) 1:10 using a sample: Urea solution ratio. 400µl of this solution was transferred to an ultrafiltration device (Milipore: Amicon ultra centrifugal 30K 0.5ml tubes) and centrifuged at 15,000xg, 10 minutes at a time; flow through was discarded after each spin. Following transfer of the total solution volume the device was washed once with 250µl Urea solution, 15,000xg, 10 minutes. The samples were then treated with 300µl 50mM iodoacetamide made up in Urea solution and incubated at room temperature in the dark for 15 minutes, the solution was subsequently span through at 15,000xg and flow through removed. 2 more washes in 250µl Urea solution, 15,000xg, 10 minutes were carried out before 3 washes with 250µl 20mM Ammonium Acetate pH5.5, 15,000xg, 10 minutes. The retained proteins in the device were then resuspended in a 100µl volume containing 20mM Ammonium Acetate pH5.5 and PNGaseF or PNGaseAr enzyme. The tubes were subsequently parafilmed tightly with fresh collection tubes added and incubated at 37°C overnight.

2.4.2 *N*-glycan collection

Tubes were centrifuged at 15,000xg, 10 minutes, followed by 2X 250µl HPLC (High Performance Lipid Chromatography) H_2O washes at 15,000xg, 10 minutes. Flow through was then transferred to clean, flamed glass tubes in which samples were dried using a vacuum centrifuge. After drying 2X 300µl HPLC H_2O washes were dried down also before storage at -20°C

2.4.3 O-glycan release and collection

Fresh collection tubes were added to the devices before the addition of 300μl 25% Ammonium hydroxide to the retained proteins in the device following *N*-glycan elution. The sample-containing tubes were subsequently parafilmed tightly and incubated at 45°C overnight. Following incubation the devices were centrifuged at 15,000xg, 10 minutes, before washing the device twice using 150μl HPLC H_2O at 15,000xg, 10 minutes. This volume was then transferred to clean, flamed glass tubes in which samples were dried using a vacuum centrifuge. After drying 2X 300μl HPLC H_2O washes were dried down also before storage at -20°C

2.4.4 Permethylation

Samples in glass tubes were re-dissolved in 20 drops dimethyl sulfoxide (DMSO) using a glass Pasteur pipette before immediately adding 2 microspatulas of ground sodium hydroxide per sample. Samples were then manually agitated for 10 seconds prior to sequences of iodomethane addition: Firstly 10 drops were added, samples manually mixed and 10 minutes elapsed, this was repeated a second time before adding 20 drops and a 20 minute wait. Following this the reaction was quenched by adding 1ml of 100mg/ml sodium thiosulphate and immediately after adding 1 ml dichloromethane. Samples were then vortexed for 10 seconds and centrifuged for 15 seconds to form an emulsion composed of two layers. Upper aqueous layers were removed and replaced with 1ml HPLC H_2O to perform a series of washes using 10 second vortex and 15 second centrifugation steps as before. This was repeated several times (5+) until the organic layer was clear before being dried down in the vacuum centrifuge and stored at -20°C prior to mass spectrometry analysis.

2.4.5 *N*-glycan sample spotting

Dried down samples were re-dissolved in 20μl methanol and 2μl of solute was added to methanol soaked microcentrifuge tubes. This was mixed with a 1:1 and 1:2 sample:matrix ratio using 20mg/ml 2,5-dihydroxybenzoic acid (DHB) dissolved in 30% HPLC H_2O, 70% methanol. Following this 1μl 500mM sodium nitrate was added to the sample/matrix solutions and 2μl of each was spotted out onto the plate and allowed to dry. Just before samples had completely dried 0.2μl ethanol was added to crystals to allow for re-crystallisation to occur.

2.4.6 O-glycan sample spotting

Dried down samples were re-dissolved in 20µl acetonitrile and 2µl of solute was added to microcentrifuge tubes. This was mixed with a 1:1 and 1:2 sample:matrix ratio using 10mg/ml 2,5-dihydroxybenzoic acid (DHB) dissolved in 50% acetonitrile. Following this 2µl of each was spotted out onto the plate and allowed to dry. Just before samples had completely dried 0.2µl ethanol was added to crystals to allow for re-crystallisation to occur.

2.4.7 MALDI mass spectrometry analysis

Sample spots were analysed using a Bruker Daltonics SolariX FTMS (Fourier transform mass spectrometry) mass spectrometer. Calibration was carried out using a peptide mix of 6 known peptide masses with an m/z detection range set between 80-4000. Shots fired were 500 and laser power was initially set to 30% before increasing in 10% increments to improve peak resolution. Once peak resolution was acceptable 5 spectra were added to the sum per spot with 500 shots each, before saving and analysing using the DataAnalysis programme. Glycans were assigned to peaks which matched known glycan masses within a mass range of 0.1 daltons or less. Known glycan masses were determined using the ExPASy GlycoMod online tool. In order to determine peaks which were more likely to be definite structures, a signal to noise ratio of 3 was set and peaks were manually assigned by going through the DataAnalysis spectra and locating peaks within a 0.1 dalton range of glycan monoisotopic masses with isotope patterns of at least 2 naturally occurring isotopes. In order to calculate overall glycan percentage abundances the sum intensity for all spectrum glycan peaks including monoisotopic and other isotopes was calculated. Then for individual glycan percentage abundances the intensity for each individual glycan was calculated as a percentage from the overall sum of glycan intensities in the spectrum.

2.5. Flight testing

3-5 day old flies were tested separately by inserting 5 flies at a time into the centre of a clear box with a light source above and tape segmenting the box into different regions. The flies were scored based on the region of the box in which they landed in. Each group of 5 flies were tested 5 times each before discarding and using a fresh group, 35 flies per genotype were assessed for their flight ability. Flight testing was carried out in the box shown in **Figure 7**. Statistics were done using SigmaPlot programme with a One Way Analysis of Variance (ANOVA). Dunn's method of multiple comparisons versus the control group (WT) was used across all samples.
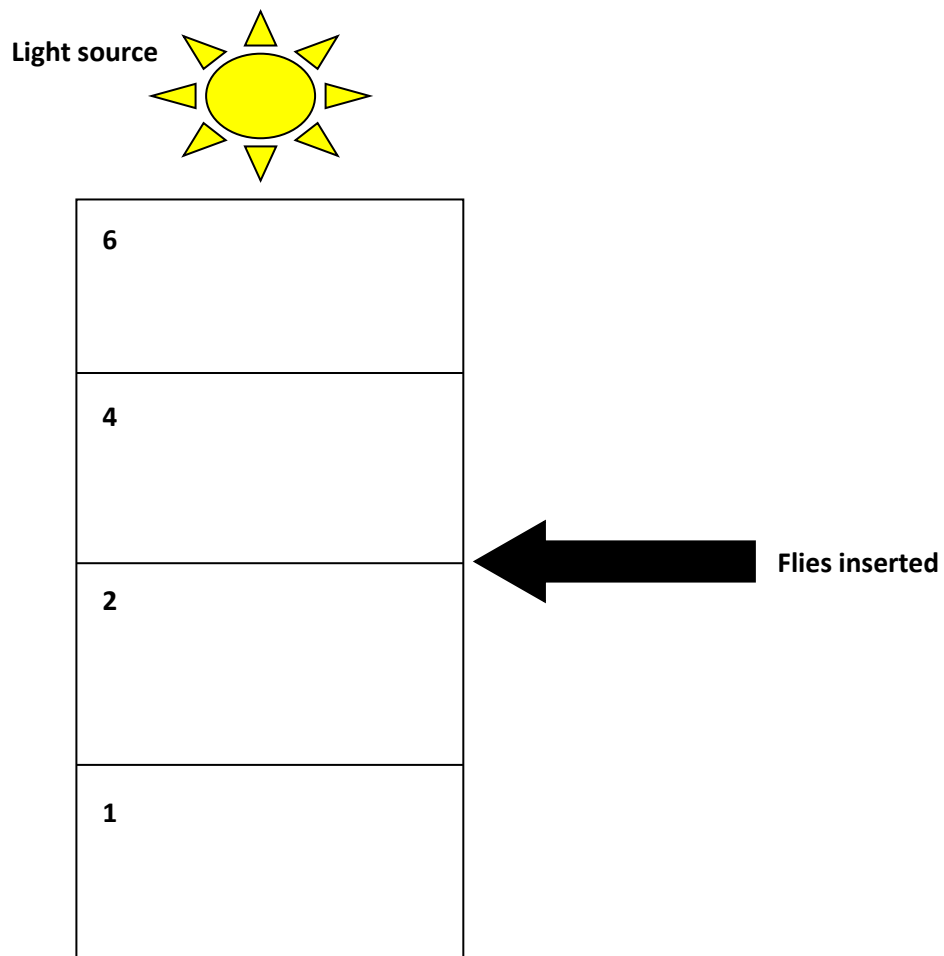

Figure 7. Box used for flight test analysis: Flight scoring = upwards (6), horizontal (4), downwards (2), none (1) With light source above, flies inserted in the centre of the box. Light source shone above and flies inserted into the centre of the box

3. Results

3.1 PNGaseAr vs PNGaseF: Types of *N*-glycan cleavage

The fact that some *Drosophila N*-glycans have 1,3 core fucosylated linkages as part of their structure poses a challenge for removal of these glycans from *Drosophila* glycoproteins. This is because conventionally PNGaseF is used as an amidase to cleave *N*-glycans and this enzyme cannot cleave 1,3 fucosylated *N*-glycans and is limited to cleaving mammalian-like 1,6 fucosylated *N*-glycans. New England Biolabs (NEB) provided us with an aliquot of a novel enzyme named PNGaseAr which has specificity for cleaving 1,3 fucosylated *Drosophila*-like *N*-glycans as well as 1,6 fucosylation linkages. This enzyme therefore has potential to cleave *Drosophila* specific *N*-glycans in order for these to be analysed, a function which PNGaseF would be incapable of carrying out. I began by characterising the PNGaseAr enzyme to find out its capabilities and functional capacity.

Firstly in order to assess which types of *N*-glycans the PNGaseAr enzyme was capable of cleaving in comparison to PNGaseF, deglycosylation assays were carried out on the high mannose glycoprotein RNAseB at varied pH levels. pH was varied as not only may this affect the enzymatic activity, but also has implications for downstream glycan analysis in terms of using a compatible buffer for mass spectrometry. **Figure 8a** shows deglycosylation reactions carried out at two different pH conditions (pH 8 and pH 5.5) each with 3 different concentrations of enzyme. At pH 8 for Ar1 and Ar0.3 units there is almost no deglycosylation of RNAseB, whereas at pH5.5 for Ar1 and Ar0.3 units the ratio of glycosylated to deglycosylated RNAseB shifts favouring more deglycosylation. This indicates a more stringent dependence on pH for the PNGaseAr enzyme to function unlike the PNGaseF enzyme which fully deglycosylates RNAseB at a pH of 5.5 or 8. The next step with this information in mind was to investigate a potential buffer which could be used at pH5.5 for downstream glycan analysis using PNGaseAr.

As with RNAseB, deglycosylaton was also carried out on the complex *N*-glycan glycoprotein Fetuin (**Figure 8b**). At all concentrations tested the PNGaseAr enzyme failed to deglycosylate Fetuin at pH 5.5. This is in contrast to the action of PNGaseF which was able to deglycosylate Fetuin at pH 8 and pH 5.5. These results indicate that PNGaseAr is only capable of cleaving high mannose *N*-glycans whereas PNGaseF is capable of cleaving both high mannose and complex forms. In the context of *Drosophila* however this is not such an issue as they predominantly produce the high mannose types of *N*-glycan over complex forms.
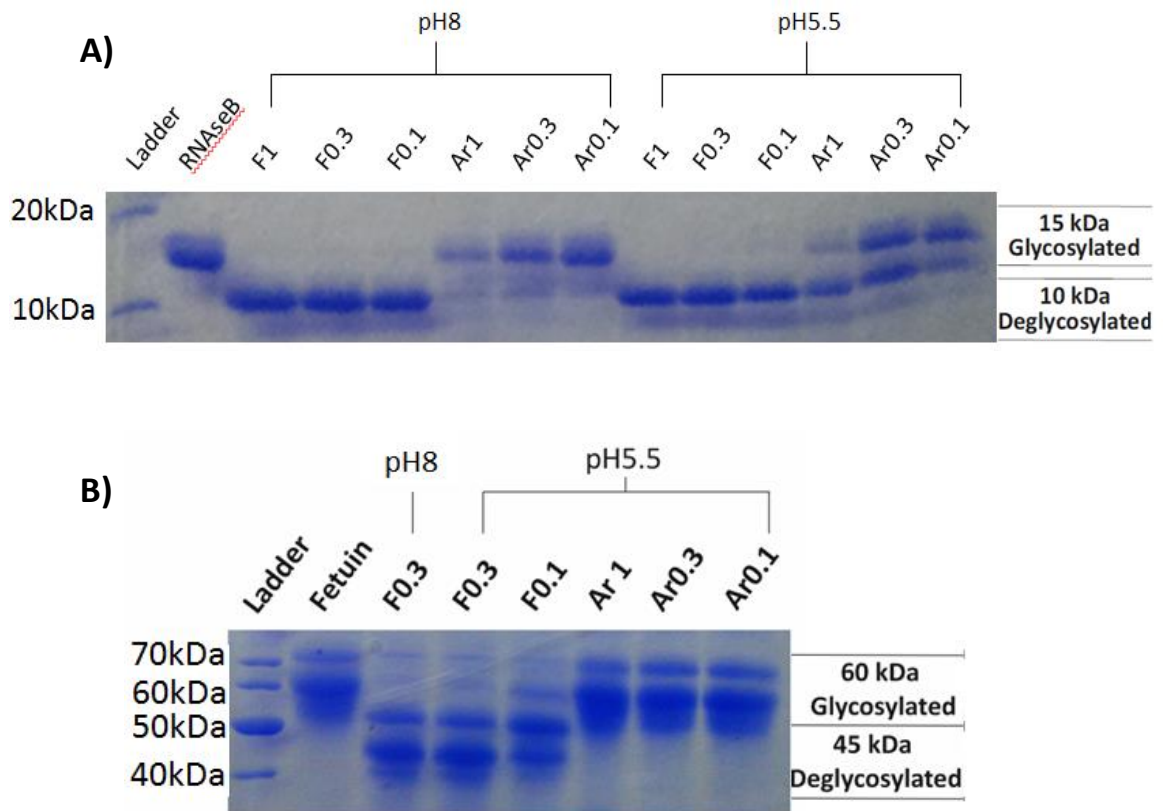
Figure 8. PNGaseAr can cleave high mannose *N*-glycans but not complex glycans: A) Deglycosylation assay on 4µg RNAseB, treated with 1, 0.3, and 0.1 units of PNGaseF and PNGaseAr at pH8 (20mM Ammonium Bicarbonate buffer pH8) and pH5.5 (20mM Ammonium Acetate buffer pH5.5) at 37°C for 5 hours. Samples were run on 15% SDS PAGE gel. 4µg untreated RNAseB was loaded in lane 1 for comparison as well as unstained PageRuler ladder (10-250kDa) B) Deglycosylation assay on 4µg Fetuin, treated with 0.3, and 0.1 units of PNGaseF and 1, 0,3 and 0.1 units PNGaseAr at pH8 (20mM Ammonium Bicarbonate buffer pH8) and pH5.5 (20mM Ammonium Acetate buffer pH5.5) at 37°C for 5 hours. Samples were run on 10% SDS PAGE gel. 4µg untreated Fetuin was loaded in lane 1 for comparison as well as unstained PageRuler ladder (10-250kDa)

3.2 Ammonium acetate functionality in FANGS procedure

To find out if a 20mM Ammonium Acetate pH5.5 buffer was compatible with downstream mass spectrometry analysis FANGS was carried out using PNGaseF and RNAseB. The results in **Figure 9** show *N*-glycan mass peaks corresponding to those of RNAseB suggesting that the buffer is compatible with the procedure and can therefore be used with the PNGaseAr enzyme. The peaks shown in **Figure 9** correspond to different mass high mannose *N*-glycans which are present on different *N*-glycosylation sites of the RNAseB protein. These *N*-glycan structures make up the glycosylated mass of RNAseB shown in **Figure 8**, once these glycans have been enzymatically cleaved the specific masses of each glycan structure can be visualised by mass spectrometry. This is also an indication of enzymatic functionality in the 20mM Ammonium acetate pH5.5 buffer as the mass peaks of cleaved RNAseB glycans can be seen, whereas no peaks would be visible had there been no enzymatic activity. This procedure was repeated with PNGaseAr using the same amount of enzyme; however no RNAseB *N*-glycan peaks were shown suggesting an inhibition or lack of enzymatic activity during the protocol.
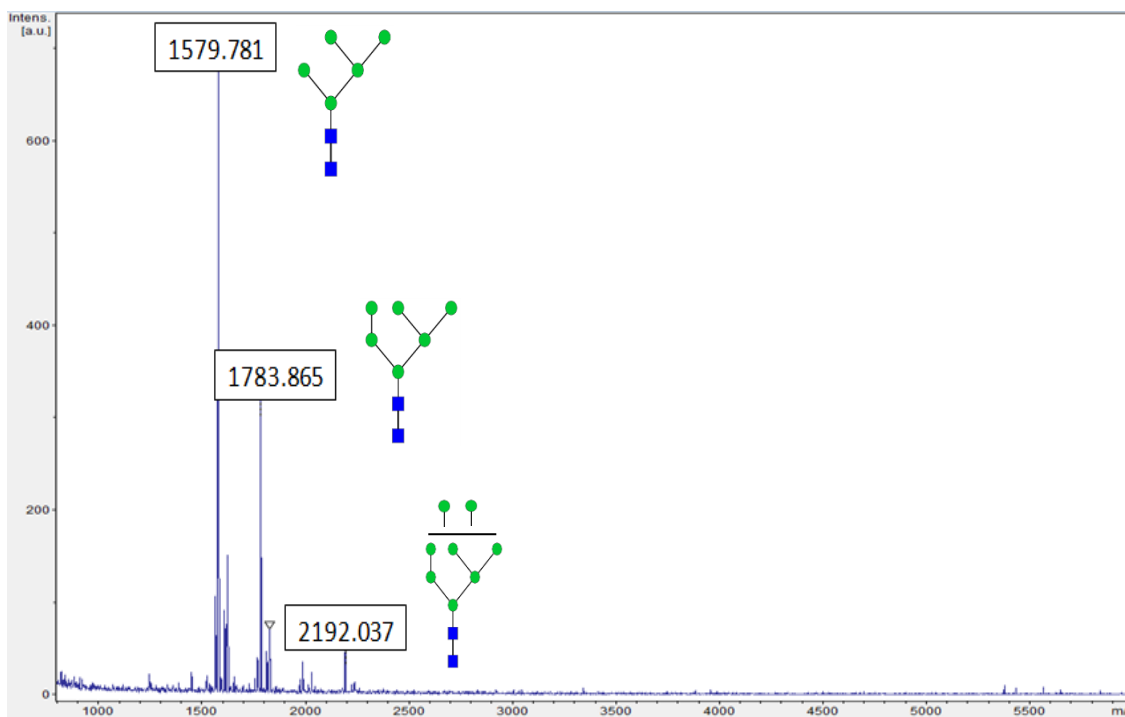
Figure 9. 20mM Ammonium Acetate is a compatible buffer with the FANGS procedure: Matrix-assisted laser desorption/ionization (MALDI) mass spectra following protein denaturation and FANGS on 50µg RNAseB using 8 units PNGaseF with a 16 hour incubation period at 37°C and subsequent permethylation.

3.3 PNGaseAr vs PNGaseF: *N*-glycan cleavage efficiency

There was the possibility that the PNGaseAr enzyme lacked efficiency at *N*-glycan cleavage when compared to PNGaseF which was why no glycans were seen via mass spectrometry on RNAseB cleaved *N*-glycans. Therefore to assess this, different units of PNGaseAr and PNGaseF were used in deglycosylation assays of RNAseB over different time courses. The data in **Figure 10** shows that PNGaseAr is roughly 100 times less efficient at *N*-glycan cleavage than PNGaseF over a 16 hour time course. This can be seen when comparing PNGaseAr units of 0.3 and PNGaseF units of 0.003, or PNGaseAr units of 0.1 and PNGaseF units of 0.001; when compared similar deglycosylation patterns can be observed. This implies that higher concentrations of PNGaseAr would need to be used in order to achieve deglycosylation comparable to that of PNGaseF. Another observation from this result is that PNGaseAr continues to cleave RNAseB glycans when given a longer incubation time. This can be seen when comparing 0.1 PNGaseAr units after 16 hours with that at 40 hours where a clear increase in deglycosylation can be seen at 40 hours. This implies that the enzyme does not lose activity or denature over time and in fact continues to cleave which is a promising sign for its use in the FANGS procedure.
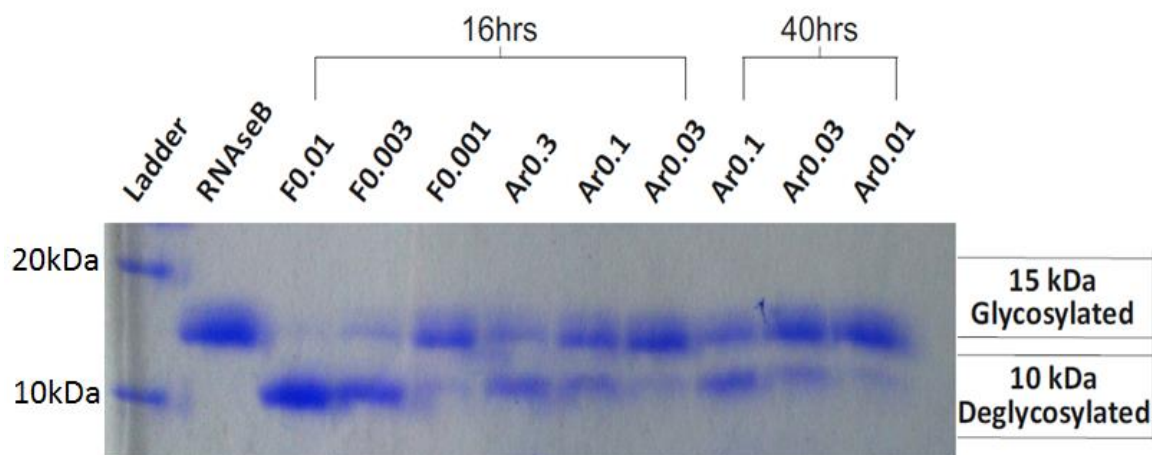
Figure 10. PNGaseAr is ~100x less efficient at *N*-glycan cleavage than PNGaseF but continues to cleave given prolonged incubation times: Deglycosylation assay on 4µg RNAseB, treated with 0.01, 0.003, and 0.001 units of PNGaseF and 0.3, 0.1, 0.03, and 0.01 units PNGaseAr. Samples were incubated for either 16 hours or 40 hours at 37°C. Samples subsequently were run on 15% SDS PAGE gel 4µg untreated RNAseB was loaded in lane 1 for comparison as well as unstained PageRuler ladder (10-250kDa)

3.4 PNGaseF FANGS unit requirements

Having established the efficiency difference between the two enzymes, observing the minimum amount of PNGaseF required to cleave N-glycans in FANGS and subsequently still observe glycan peaks via mass spectrometry would give an indication as to how little PNGaseAr may need to be used to produce a similar result. Due to the small aliquot of PNGaseAr readily available, FANGS experiments were carried out using PNGaseF with decreasing unit amounts of the enzyme to cleave *N*-glycans from the same amount of RNAseB protein. FANGS was carried out on 50μg RNAseB using 8, 4, 2 and 1 units PNGaseF to see where the cut-off point may be for observable readouts of *N*-glycan signal following PNGaseF cleavage. The data shown in **Figure 11** is reminiscent of that seen in the spectra using 8 units and 4 units PNGaseF unlike that of **Figure 12** which differs. The data from these figures suggests that using 2 units or above PNGaseF is sufficient to produce clear *N*-glycan peaks following FANGS on the amount of RNAseB used. However when down to using 1 unit PNGaseF the signal from the spectrum begins to deteriorate indicating that using amounts below this may be insufficient to release enough *N*-glycan to be observed via mass spectrometry.
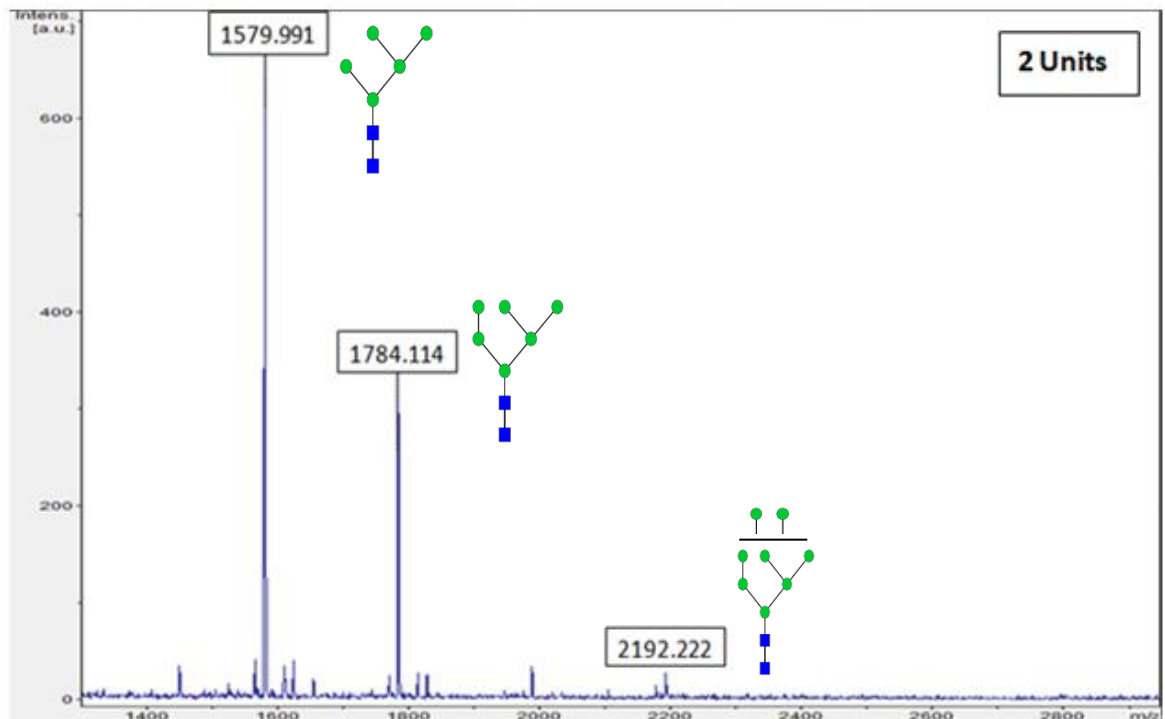
Figure 11. 2 units PNGaseF is sufficient to produce accurate *N*-glycan readouts via mass spectrometry following FANGS: MALDI mass spectra following protein denaturation and FANGS on 50μg RNAseB using 2 units PNGaseF with a 16 hour incubation period at 37°C and subsequent permethylation.
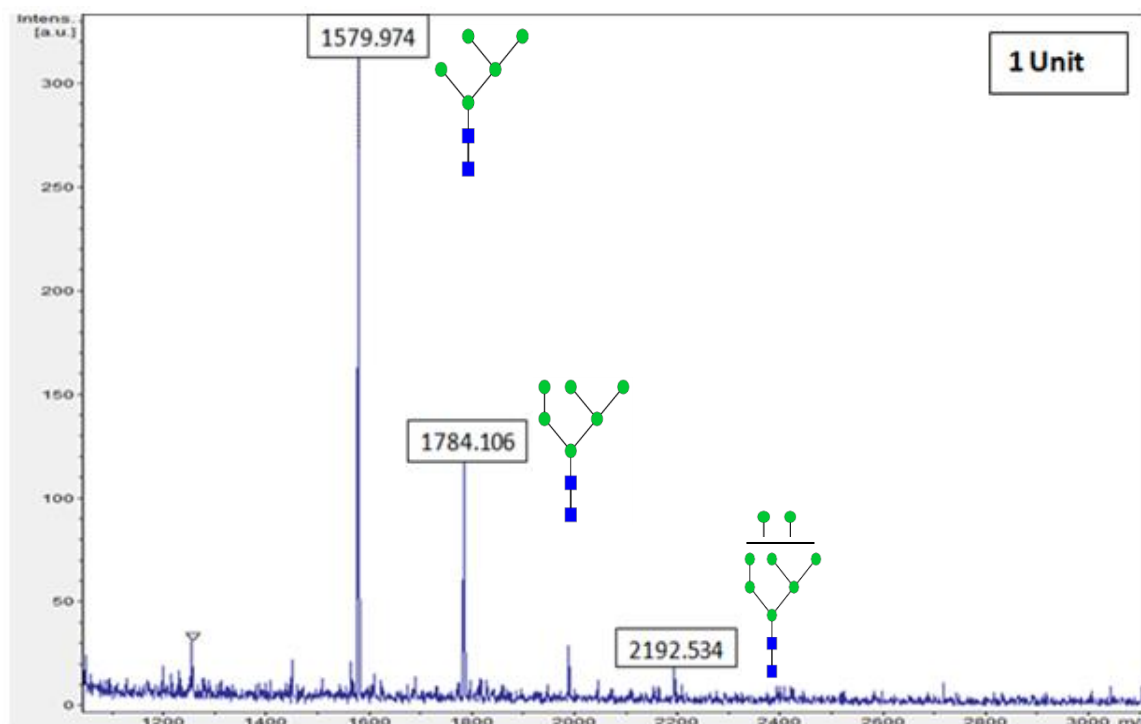
Figure 12. 1 unit PNGaseF causes signal decreased signal to noise ratio of *N*-glycan readouts via mass spectrometry following FANGS: MALDI mass spectra following protein denaturation and FANGS on 50µg RNAseB using 1 unit PNGaseF with a 16 hour incubation period at 37°C and subsequent permethylation.

3.5 PNGase SDS sensitivity

In order to assess what could be causing PNGaseAr inhibition during the FANGS procedure, different elements of the protocol were removed or altered to see if this had an effect on PNGaseAr cleavage of *N*-glycans. A few factors were tested such as the removal of the alkylating agent iodoacetamide from the procedure and also changing the reaction vessel for the cleavage from the ultrafiltration device to a microfuge tube. These factors however showed minimal effects if any on PNGaseAr activity and were therefore not pursued further. The results in **Figure 13** show the effect of SDS on PNGaseF and PNGaseAr activity using a 100-fold difference between the concentrations of the enzymes to normalise the activities of each based on the results in **Figure 10**. The effect of SDS on PNGaseAr activity was investigated as this is a component of the lysis buffer used prior to FANGS and it is known that SDS causes an inhibition of PNGaseF activity [46] and therefore could also impact PNGaseAr function. **Figure 13** shows that in the F0.001 section between 0.03-0.27% SDS high to full levels of deglycosylation occur before an increase to 0.81 and 2.43% SDS where deglycosylation is inhibited. In the Ar0.1 section high to full levels of deglycosylation do not occur until 0.09% SDS which is most likely due to the PNGaseF enzyme being actually more than 100X higher in cleavage efficiency than PNGaseAr. Between 0.09 and

0.81% SDS, PNGaseAr is capable of complete deglycosylation before an increase to 2.43% where inhibition occurs. These results indicate firstly that PNGaseAr is better at cleaving *N*-glycans from more native protein forms than PNGaseF shown by comparing the two enzymes at 0.09% SDS where PNGaseAr showed slightly greater deglycosylation. Also the data shows that PNGaseAr is less sensitive to SDS than PNGaseF which can be seen when comparing the enzymes at 0.81% SDS where PNGaseAr can still fully deglycosylate whereas PNGaseF activity is inhibited. Finally it is evident that PNGaseAr is sensitive to SDS inhibition albeit at a higher concentration than PNGaseF sensitivity which is shown at a 2.43% SDS concentration. This was therefore evidence that SDS could be effecting the PNGaseAr activity in FANGS consequently resulting in no or insufficient *N*-glycan cleavage to be visualised by mass spectrometry; so this was subsequently pursued further.
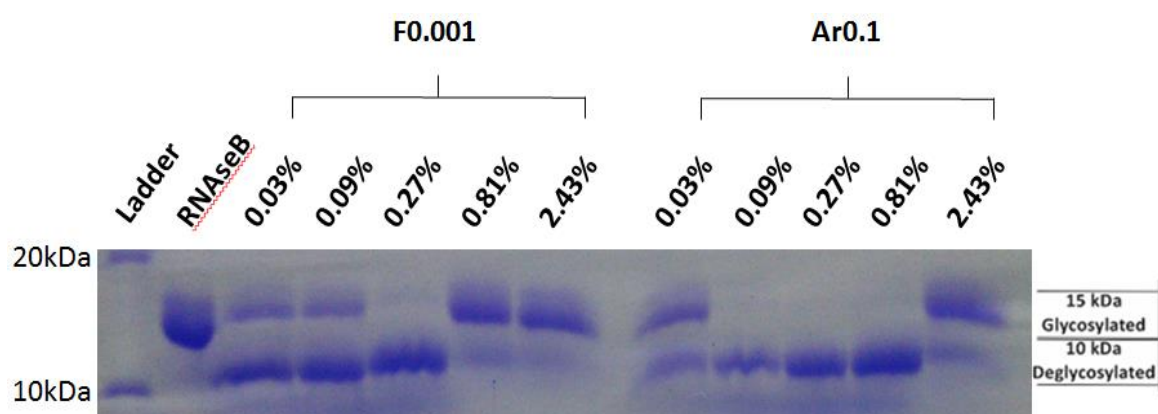
Figure 13. PNGaseAr is sensitive to SDS: Deglycosylation assay on 4μg RNAseB, treated with 0.001 units PNGaseF and 0.1 units PNGaseAr. Denature solution was composed of increasing concentrations of SDS to give final reaction volume concentrations of 0.03%, 0.09%, 0.27%, 0.81% and 2.43%. Samples were incubated for 16 hours at 37°C. Samples subsequently were run on 15% SDS PAGE gel. 4μg untreated RNAseB was loaded in lane 1 for comparison as well as unstained PageRuler ladder (10-250kDa)

3.6 Impact of SDS during FANGS

After observing the sensitivity of PNGaseAr to SDS from a deglycosylation assay, taking this into the FANGS procedure was the next step. Lower concentrations of SDS were used in the lysis buffer at the start of the protocol and compared to the 4% SDS used previously. The results in **Figure 14a** show that following FANGS when using SDS concentrations of lysis buffer at 2% or 1% when compared to the 4% used previously, a marked improvement in PNGaseF cleavage of RNAseB N-glycans was seen. However the samples were also treated with TX-100 as this functions to counteract any SDS which may be present in the reaction to cause an inhibition of enzyme activity. In **Figure 14b** TX-100 was not added to the samples given 4% and 2% SDS concentrations in the lysis buffer and this resulted in no deglycosylation of RNAseB by PNGaseF suggesting that TX-100 is needed to counteract residual SDS activity ('NO FANGS' samples were used as a positive control for RNAseB *N*-glycan cleavage using the parameters from previous assays such as those used in **Figure 10** to show glycosylated and deglycosylated bands for comparison). This was not an ideal situation as TX-100 is a detergent which interferes with downstream mass spectrometry analysis and therefore is not desirable to have in the reaction. In order to remove the requirement for TX-100 addition the residual SDS present during the PNGase addition stage of FANGS needed to be removed more effectively.
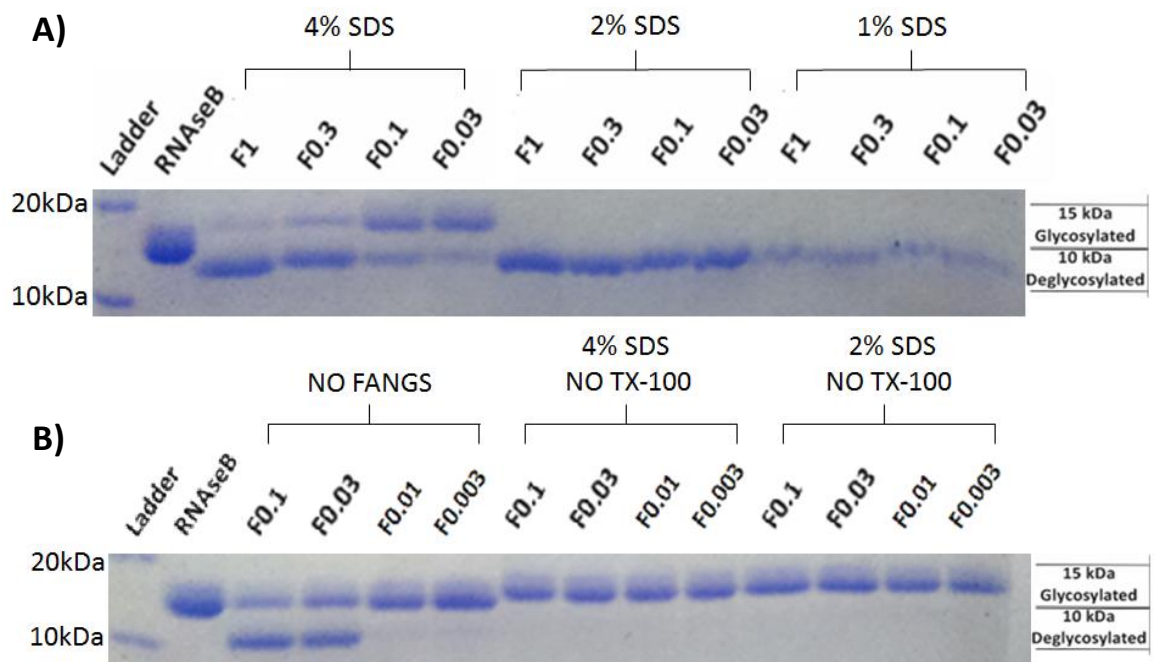
Figure 14. Reduced SDS improved PNGaseF *N*-glycan cleavage during FANGS but requires TX-100 addition to counteract residual SDS activity in order to function: A) Deglycosylation assay on 4µg RNAseB with 40µg having gone through protein denaturation and FANGS, treated with lysis buffer composed of 4%, 2% and 1% SDS. Samples treated with 1, 0.3, 0.1 and 0.03 units PNGaseF and incubated for 16 hours at 37°C. Samples subsequently were run on 15% SDS PAGE gel. 4µg untreated RNAseB was loaded in lane 1 for comparison as well as unstained PageRuler ladder (10-250kDa) B) Deglycosylation assay on 4µg RNAseB with 40µg having gone through protein denaturation and FANGS, treated with lysis buffer composed of 4% and 2% SDS. Samples treated with 0.1, 0.03, 0.01 and 0.003 units PNGaseF and were not treated with TX-100. 'NO FANGS' samples were used as a control deglycosylation assay which did not contain protein having gone through FANGS and had 0.2% SDS and TX-100 treatment. Samples incubated for 16 hours at 37°C and subsequently were run on 15% SDS PAGE gel. 4µg untreated RNAseB was loaded in lane 1 for comparison as well as unstained PageRuler ladder (10-250kDa)

3.7 Effective removal of SDS from FANGS

A possible explanation for the difficulty in removing SDS effectively from the FANGS procedure was that SDS micelles form as the SDS concentration goes above the critical micelle concentration (CMC) for SDS. This was an issue as the micelles are larger than individual SDS molecules and would consequently be retained above in the FANGS ultrafiltration device being too large to go through the pores into the flow through. Hence the PNGase enzymatic inhibition was potentially being caused by residual SDS micelles retained above the filter during the *N*-glycan cleavage. To attempt to resolve this issue during the dilution in Urea solution following protein denaturation a 1:20 dilution of the sample in Urea solution was done instead of 1:10, this was also done using lysis buffer containing 2% SDS rather than 4%. The result in **Figure 15** shows that PNGaseF and PNGaseAr fully deglycosylate RNAseB *N*-glycans following use of this modified FANGS procedure and even results in greater deglycosylation than the control samples which had not gone through FANGS. What can be inferred from this data is that the SDS has been removed more effectively eliminating the requirement for TX-100 addition based on the PNGase enzymes cleaving more effectively.
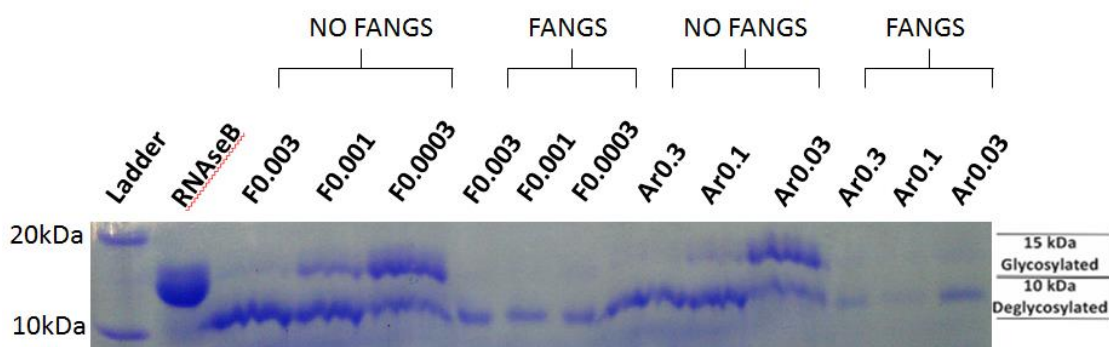
Figure 15. 1:20 dilution of 2% SDS containing lysis volume with Urea solution effectively removes SDS from FANGS. Deglycosylation assay on 4μg RNAseB with 40μg having gone through protein denaturation and FANGS, treated with lysis buffer composed of 2% SDS. Samples treated with 0.003, 0.001 and 0.0003 units PNGaseF and 0.3, 0.1 and 0.03 units PNGaseAr. 'NO FANGS' samples were used as a control deglycosylation assay which did not contain protein having gone through FANGS and had 0.2% SDS and TX-100 treatment. Samples were incubated for 16 hours at 37°C and subsequently were run on a 15% SDS PAGE gel. 4μg untreated RNAseB was loaded in lane 1 for comparison as well as unstained PageRuler ladder (10-250kDa)

3.8 Using optimised FANGS procedure with PNGaseAr

To finalise the optimisation of the FANGS procedure for use with PNGaseAr, 3 samples were used for FANGS. One as a negative control using the original FANGS protocol with 0.003 units PNGaseF and the other two with the procedure used on samples which underwent FANGS in **Figure 15** but here treated with 0.003 units PNGaseF and 1 unit PNGaseAr. The spectra in **Figure 16a** shows as expected no *N*-glycan peaks from RNAseB following use of the original FANGS procedure with 0.003 units PNGaseF. **Figure 16b** shows that with the same amount of PNGaseF but using the optimised protocol *N*-glycan peaks are apparent on the spectrum. And finally in **Figure 16c** it is clear that the procedure has been successfully optimised for use with PNGaseAr as *N*-glycan peaks are shown from the samples treated with 1 unit PNGaseAr using the optimised FANGS procedure. The procedure was now ready for use with *Drosophila melanogaster* samples.
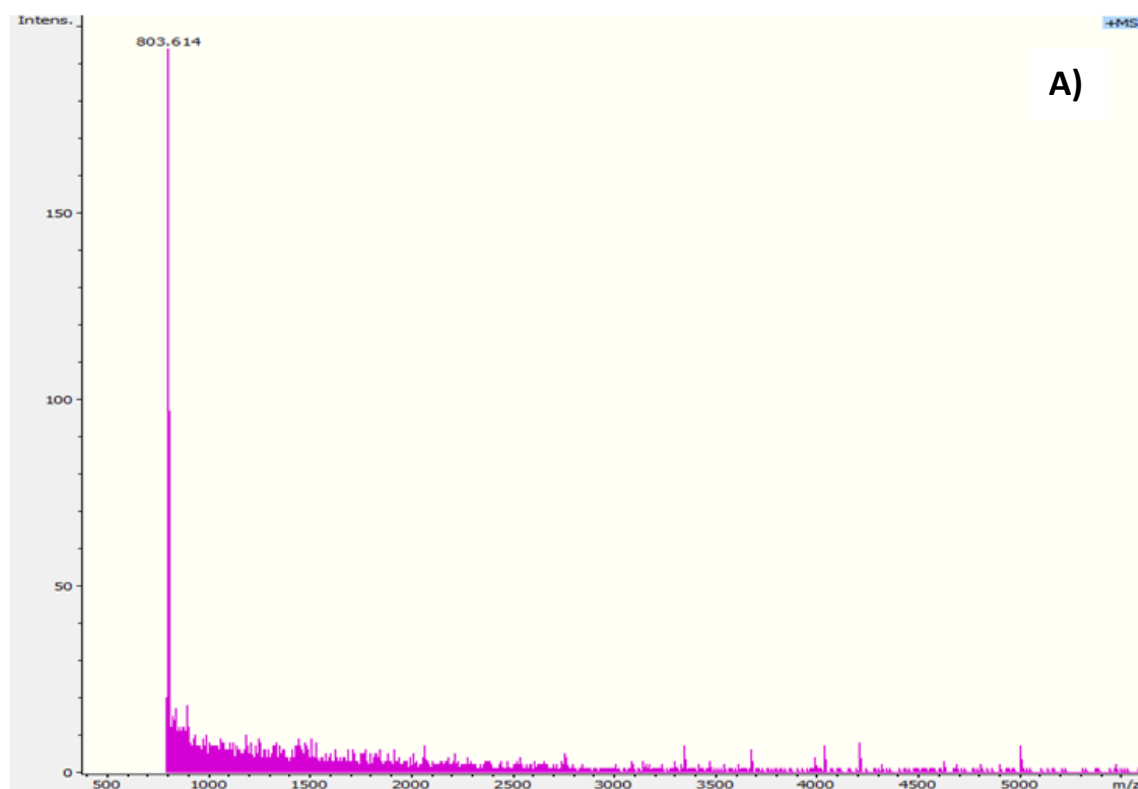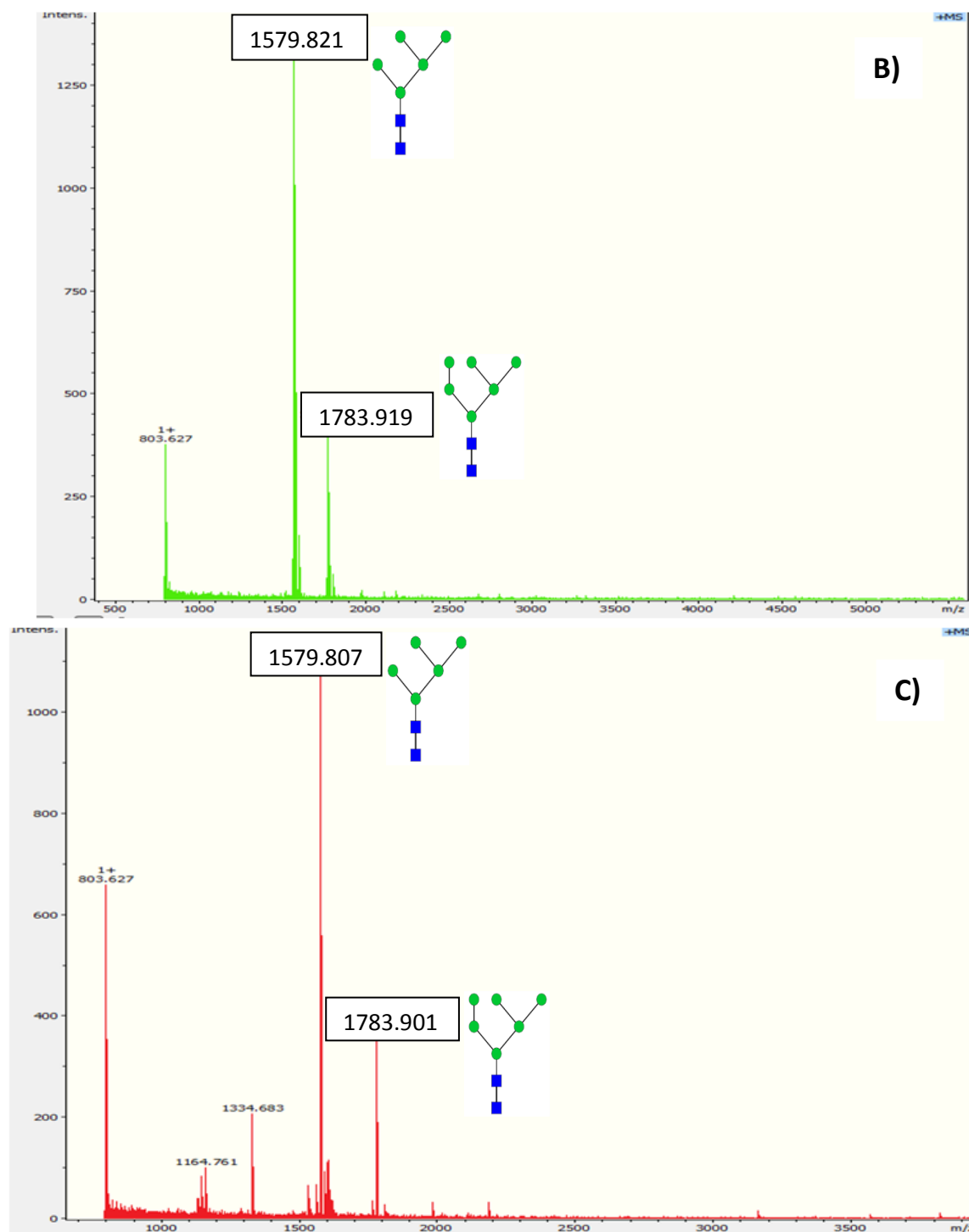
Figure 16. *N*-glycan cleavage during FANGS is significantly improved for both PNGaseF and PNGaseAr when using the optimised procedure over the original: A) MALDI mass spectra following protein denaturation and original FANGS procedure on 50µg RNAseB using 0.003 units PNGaseF with a 16 hour incubation period at 37°C and subsequent permethylation. B) MALDI mass spectra following protein denaturation and optimised FANGS procedure on 50µg RNAseB using 0.003 units PNGaseF with a 16 hour incubation period at 37°C and subsequent permethylation. C) MALDI mass spectra following protein denaturation and optimised FANGS procedure on 50µg RNAseB using 1 unit PNGaseAr with a 16 hour incubation period at 37°C and subsequent permethylation.

3.9 *Drosophila N*-glycan profiling

In order to assess how the loss of COG subunit Cog3 in muscle cells may impact on *Drosophila* glycosylation, *N*-glycan analysis was carried out. With the FANGS protocol optimised for use with the PNGaseAr enzyme, this was applied to *Drosophila* COG mutant line $\frac{Dmef2-GAL4}{Cog3RNAi-UAS}$ and wild type control lines $\frac{Dmef2-GAL4}{WT}$ and $\frac{Cog3RNAi-UAS}{WT}$ (**Table 4**). Both PNGaseF (4 units) and PNGaseAr (3units) were applied to the samples during FANGS.

The results in **Figure 17** show that predominantly the *N*-glycans detected were high mannose or paucimannose which is expected in *Drosophila*. Most of the glycans did not differ largely between the mutant and wild type lines; however there were a few differences. Interestingly the glycan $HexNAc_3Hex_3Fuc_3NeuAc_1$ showed an increased abundance in the mutant compared to wild type lines. Additionally this glycan structure was most likely cleaved by PNGaseAr due to the multiple fucosylation linkages in the structure suggesting a unique *Drosophila*-like glycan was observed due to use of PNGaseAr. Additionally this *N*-glycan structure was not previously seen in *N*-glycan profiling when using PNGaseF for glycan cleavage (Personal communication, Daniel Ungar). Other differences include an increase in Hex8 high mannose glycan in the mutant line as well as no presence in the mutant of $HexNAc_3Hex_5$.

Carrying out tandem mass spectrometry on these structures would certainly shed more light onto the monosaccharide linkages of each structure. Also more repeats of each sample may unveil more differences in the glycan abundances between the Cog3 mutant and wild type lines. The glycans which differed between the mutant and wild type flies may have functional roles in *Drosophila* flight muscle. This is because with the Cog3 mutant used in this experiment having a flightless phenotype, differences in glycan abundances could be attributable to defective COG localisation of glycosylation enzymes at the Golgi consequently resulting in improper glycosylation of proteins important for muscular function. Hence the functionality of these proteins may be dependent on specific glycosylation patterns which are altered, potentially impacting cell-cell or cell-tissue interactions which are crucial for muscle contractile homeostasis.

Figure 17. Absence of COG3 in *Drosophila* muscle reveals differences in *N*-glycan profile in comparison to wild type lines: *N*-glycan percentage abundances from MALDI MS (Mass Spectrometry) analysis following cell lysis and optimised FANGS procedure on 100 flies per genotype performed in triplicate using 3 units PNGaseAr and 4 units PNGaseF with a 16 hour incubation period at 37°C and subsequent permethylation.

3.10 *Drosophila* O-glycan profiling

To complement the findings in **Figure 17** regarding altered glycosylation shown by differences in *N*-glycan abundance between the Cog3 mutant and wild type strain, after *Drosophila N*-glycan release and collection, O-glycans were sequentially released. This gave a broader range of glycan species which is attributable to the fact that O-glycans have heterogeneous core structures opening up more possibilities for glycan structures than *N*-glycans.

Figure 18 shows the O-glycan percentages in the Cog3 mutant compared to the wild type controls and unveils differences in glycan abundance. Glycans which showed increased abundance in the Cog3 mutant compared to the wild types were Xyl_2NeuAc_1 $HexA_1$, $Hex_1HexNAc_2NeuAc_1$, Hex_1Fuc_2KDN (Deaminoneuraminic acid)$_1$ and Fuc_1Xyl_3. Also there were some glycan structures which had reduced abundances in the Cog3 mutant when compared to wild types which were $HexNAc_3$ and $Hex_1Fuc_1NeuAc_1$. It is worth noting that some of the glycan structures when analysed using the ExPASy GlycoMod tool were of a similar mass to various other structures. Therefore glycans were selected based on reported *Drosophila* glycosylation enzymes in the literature which would permit certain structures to be synthesised [47, 48].

In the case of these glycans tandem mass spectrometry would help to reveal the monosaccharide linkages of the glycans in question. Overall both **Figure 17** and **Figure 18** suggest that the loss of COG subunit Cog3 has an impact on the glycosylation machinery causing differences in the abundance of certain glycan structures in comparison to the wild type.
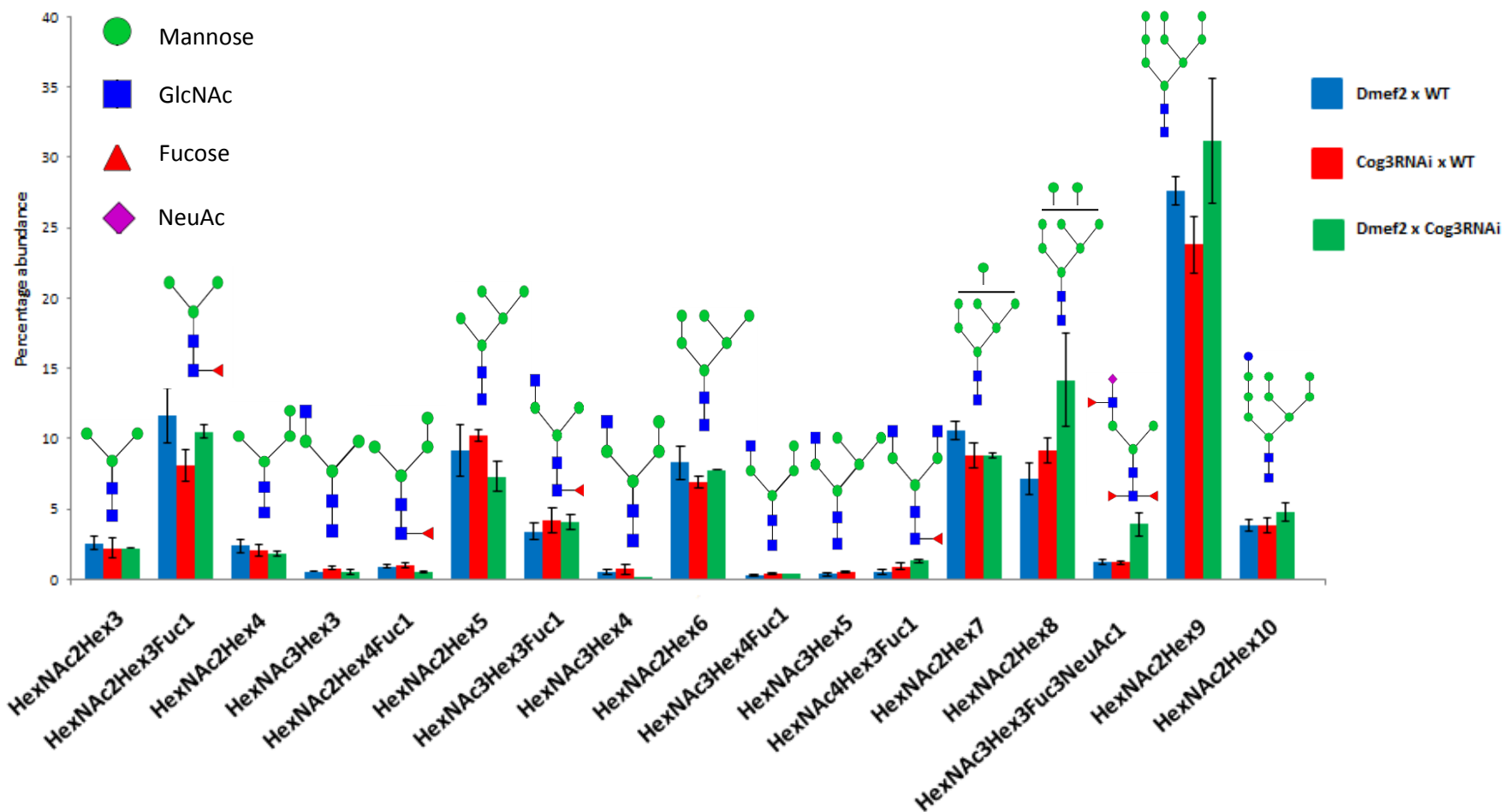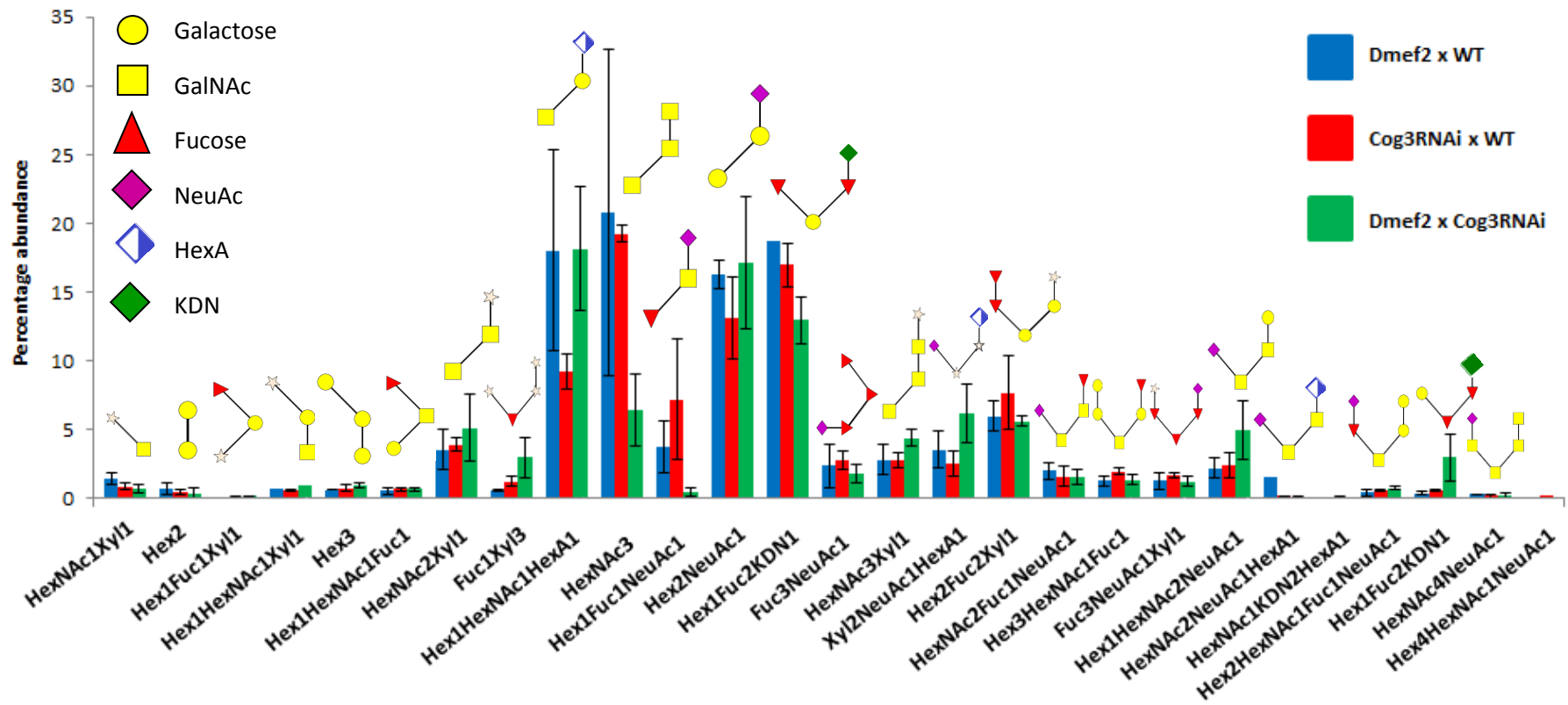
Figure 18. Absence of COG3 in *Drosophila* muscle reveals differences in O-glycan profile in comparison to wild type lines: O-glycan percentage abundances from MALDI MS analysis following cell lysis and optimised FANGS procedure on 100 flies per genotype in triplicate with a 16 hour incubation period at 45°C and subsequent permethylation.

Flight Test		
Genotype 1	**Genotype 2**	**Progeny genotype**
WT (Canton S)	WT (White[-])	WT
$\frac{fwsp}{fwsp};\frac{Cog1}{TM6B}$	$\frac{fwsp}{fwsp};\frac{Cog2}{TM6B}$	$\frac{fwsp}{fwsp};\frac{Cog1}{Cog2}$
$\frac{fwsp}{fwsp};\frac{Cog2}{TM6B}$	$\frac{fwsp}{CyO};\frac{+}{+}$	$\frac{fwsp}{fwsp};\frac{Cog2}{+}$
$\frac{fwsp}{fwsp};\frac{Cog1}{TM6B}$	$\frac{fwsp}{CyO};\frac{+}{+}$	$\frac{fwsp}{fwsp};\frac{Cog1}{+}$
$\frac{fwsp}{CyO};\frac{+}{+}$	$\frac{fwsp}{CyO};\frac{+}{+}$	$\frac{fwsp}{fwsp};\frac{+}{+}$
$\frac{fwsp}{CyO};\frac{+}{+}$	WT (Canton S)	$\frac{fwsp}{+};\frac{+}{+}$
$\frac{+}{+};\frac{Cog2}{TM6B}$	WT (Canton S)	$\frac{+}{+};\frac{Cog2}{+}$
$\frac{+}{+};\frac{Cog1}{TM6B}$	WT (Canton S)	$\frac{+}{+};\frac{Cog1}{+}$
$\frac{+}{+};\frac{Cog1}{TM6B}$	$\frac{+}{+};\frac{Cog2}{TM6B}$	$\frac{+}{+};\frac{Cog1}{Cog2}$
$\frac{fws-GFP}{CyO};\frac{+}{+}$	$\frac{fwsp}{fwsp};\frac{+}{+}$	$\frac{fws-GFP}{fwsp};\frac{+}{+}$
Glycan profiling		
Genotype 1	**Genotype 2**	**Progeny genotype**
Dmef2-GAL4	Cog3RNAi-UAS	$\frac{Dmef2-GAL4}{Cog3RNAi-UAS}$
Dmef2-GAL4	WT (Canton S)	$\frac{Dmef2-GAL4}{WT}$
Cog3RNAi-UAS	WT (Canton S)	$\frac{Cog3RNAi-UAS}{WT}$

Table 4. *Drosophila* crosses between 'Genotype 1' and 'Genotype 2' to generate 'Progeny genotype' which were used in either the flight testing assay (Flight test section) or for *N-* and *O-*linked glycan profiling (Glycan profiling section) ($\frac{fws-GFP}{fwsp}$ and $\frac{fws-GFP}{CyO}$ lines donated by Rita Sinka, University of Szeged) Cog3RNAi line was acquired from the Vienna Drosophila RNAi Center (VDRC) and the Dmef2 lines were provided by John Sparrow (University of York)[59]

3.11 Flight testing

In order to assess how defects in COG subunits may impact *Drosophila* flight muscle, flight testing was carried out on various COG mutant lines and compared to that of wild type flight; flight test crosses shown in **Table 4**. Regarding the data shown by Eric in **Figure 6**, although this shows flight defects in certain COG mutant lines, balancer chromosomes were present along with COG mutations. This means that the defect in flight cannot at this stage be conclusive as a COG specific effect as the balancer presence may be responsible for the defects observed. Therefore to tackle this problem fly lines were generated and tested with only COG mutations to remove the possibility of balancer chromosome defects.

Flight test data is shown in **Figure 18** and shows results which correlate with those seen in **Figure 6** showing that defects were indeed COG specific. Firstly it is worth noting that in all lines containing a $\frac{fwsp}{fwsp}$ mutation that showed a significant difference in flight had an observable hovering phenotype during the flight testing, something which other lines did not show. Specifically the phenotype shown was an extended period of *Drosophila* flight following their release into the box where the flies did not land on the sides of the box and maintained hovering flight, often gradually decreasing in flight index before eventually landing. *Drosophila* lines $\frac{fwsp}{fwsp};\frac{\text{Cog1}}{\text{Cog2}}, \frac{fwsp}{fwsp};\frac{\text{Cog1}}{+}, \frac{fwsp}{fwsp}, \frac{\text{Cog1}}{+}, \frac{\text{Cog1}}{\text{Cog2}}$ all showed defective flight and a significant difference in flight ability compared to wild type. These samples all have in common either a homozygous loss of Cog5 function or a lower gene dosage of lobe A subunits of the COG complex. In correlation with the data in **Figure 6** is the $\frac{fwsp}{fwsp};\frac{\text{Cog1}}{\text{Cog2}}$ line which showed a flight index of around 4 indicating horizontal flight. Moreover the $\frac{fwsp}{fwsp};\frac{\text{Cog1}}{+}$ line also showed a similar defect to the $\frac{fwsp}{fwsp};\frac{\text{Cog1}}{\text{MKRS}}$ line in **Figure 6** implying that the defect in flight is specific to the COG mutations and not the balancer presence. Also the $\frac{fwsp}{fwsp};\frac{\text{Cog2}}{+}$ line showed no significant difference to wild type flight which agrees with the flight index shown in lines $\frac{fwsp}{+};\frac{\text{Cog2}}{+}$ and $\frac{fwsp}{\text{Df}};\frac{\text{Cog2}}{+}$ in **Figure 6**. Lines $\frac{fwsp}{fwsp}, \frac{\text{Cog1}}{+}$ and $\frac{\text{Cog1}}{\text{Cog2}}$ showed a significant difference to wild type flight unlike the $\frac{fwsp}{fwsp};\frac{\text{Cog2}}{+}, \frac{\text{Cog2}}{+}$ and $\frac{fwsp}{+}$ lines which agrees with the suggestion from the data in **Figure 6** that the Cog2 mutation causes a suppression of the $\frac{fwsp}{fwsp}$ phenotype. This data implies also that the $\frac{fwsp}{fwsp}$ mutation causes defective flight and slightly less so does the $\frac{\text{Cog1}}{+}$ mutation; a combination of the mutations in the $\frac{fwsp}{fwsp};\frac{\text{Cog1}}{+}$ line also maintains the flight defect, showing an additive effect on the phenotype. Finally the $\frac{fwsp-fws:\text{GFP}}{fwsp}$ line which rescues the *fws^p* mutation to a wild type *fws* protein was able to rescue the flight defect to a wild type index. These results imply that certain COG subunit mutations have more of an impact on *Drosophila* flight ability than others. It seems that subunits

which have an impact on flight ability are Cog5 and Cog1 with Cog2 having no impact on flight ability. The effect of the Cog5 mutant on flight index can be further shown by the introduction of the rescue line which restores the WT phenotype by having a genotype which is effectively a $\frac{fwsp}{+}$ line with a *fws^p* mutation being replaced with a *fws*-GFP wild type Cog5 gene. These defects highlight potential roles in regulating glycosylation of proteins important for flight muscle functionality for COG subunits in both lobe A and lobe B of the COG complex. This along with the glycan profiling data of the Cog3 mutant line of *Drosophila* highlight that it only requires defects in individual subunits of the COG complex to disrupt the glycosylation machinery enough to alter glycosylation patterns and also impact on organism function, in this case muscular contraction.
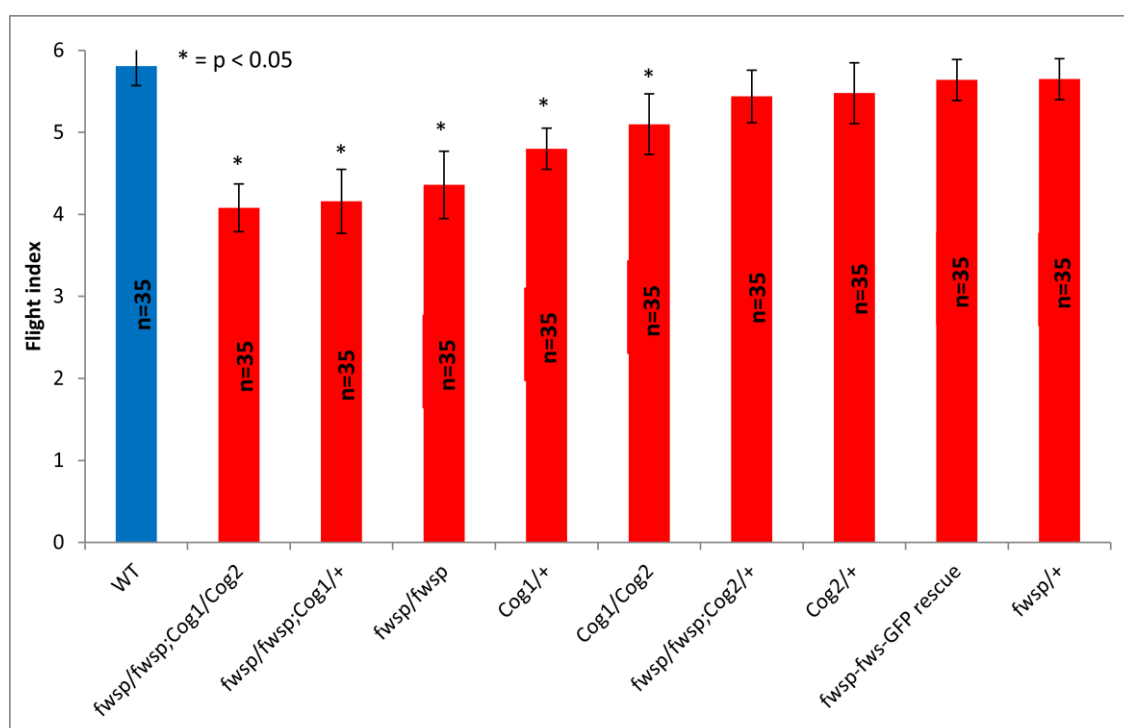
Figure 19. Mutations in *fws* and Cog1 *Drosophila* subunits cause flight defects and Cog2 mutation suppresses the *fws* mutant phenotype: 3-5 day old flies released into a clear box with a light source above, scoring based on direction of flight upwards (6), horizontal (4), downwards (2), none (1).*fws = four way stop* (*Drosophila* Cog5 homolog), *fws^p = four way stop* p element disruption, + = wild type, COG1/COG2 = COG1/COG2 mutant. 35 flies were used for each individual flight test genotype. Significance was tested based on a one way ANOVA on ranks using the WT line as the control group. P values <0.05 indicates results significantly different from the WT score. Error bars represent standard deviation.

3.12 Protein-protein interactions of COG and Rab GTPases

As a method of looking for possible interactions between *Drosophila* COG subunits and *Drosophila* Rab GTPases, a yeast-two-hybrid approach was used. This approach has the advantage of being able to investigate many different potential interactions simultaneously. The reasoning behind using the COG subunits and Rab GTPases chosen for this series of yeast-two-hybrid assays was based on the reported interactions between them in mammalian systems. The image shown in **Figure 3** summarises these interactions, what we wanted to see was how well conserved these interactions would be in *Drosophila* to evaluate the usefulness of this system for such glycosylation studies.

Firstly in **Figure 19** potential interactions were analysed between various *Drosophila* Rab GTPases with the *Drosophila* Cog4 subunit. These results show that *Drosophila* Cog4 interacts with *Drosophila* Rabs 1, 2, 4, 10, 30, and 39, three of which are conserved interactions observed in mammalian systems between Rab1, 4 and 30 (**Figure 3**). This can be deduced from the growth of yeast colonies on the plates lacking histidine (-Leu-Ura-His) as only a protein-protein interaction would permit the production of histidine, subsequently allowing the yeast to grow. The premise behind this being a protein-protein interaction brings the GAL4AD and GAL4BD in close proximity as these domains are fused to the proteins of interest, and this allows for GAL4 to function as a transcriptional activator and drive transcription of a reporter gene responsible for encoding a protein crucial for yeast growth, in this case one involved in histidine biosynthesis. No protein-protein interaction would give the opposite effect in that the GAL4AD and GAL4BD would not be in close enough proximity to allow GAL4 to function and therefore yeast would not grow on plates lacking histidine.

The remaining interactions novel Cog-Rab interactions which are specific to *Drosophila*. Interactions were not strong enough to grow on the more stringent –Leu-Ura-Ade plate but could on the less stringent –Leu-Ura-His plates.

	DIL		DIL		DIL		DIL
mCog4 X mCog2	mCog4 X mCog2	pGAD x pGBDU	pGAD x pGBDU	pGBDU x Cog4	pGBDU x Cog4		
pGAD x Rab1 Q-L	pGAD x Rab1 Q-L	pGAD x Rab2	pGAD x Rab2	pGAD x Rab3 Q-L	pGAD x Rab3 Q-L	pGAD x Rab4 Q-L	pGAD x Rab4 Q-L
pGAD x Rab6 Q-L	pGAD x Rab6 Q-L	pGAD x Rab10 Q-L	pGAD x Rab10 Q-L	pGAD x Rab30 Q-L	pGAD x Rab30 Q-L	pGAD x Rab39 Q-L	pGAD x Rab39 Q-L
Cog4 x Rab1 Q-L	Cog4 x Rab1 Q-L	Cog4 x Rab2	Cog4 x Rab2	Cog4 x Rab3 Q-L	Cog4 x Rab3 Q-L	Cog4 x Rab4 Q-L	Cog4 x Rab4 Q-L
Cog4 x Rab6 Q-L	Cog4 x Rab6 Q-L	Cog4 x Rab10 Q-L	Cog4 x Rab10 Q-L	Cog4 x Rab30 Q-L	Cog4 x Rab30 Q-L	Cog4 x Rab39 Q-L	Cog4 x Rab39 Q-L

-Leu-Ura   **-Leu-Ura-Ade**   **-Leu-Ura-His**

Figure 20. *Drosophila* Cog4 interacts with *Drosophila* Rabs 1, 2, 4, 10, 30 and 39: *Drosophila* Cog4 was crossed with *Drosophila* Rabs 1, 2, 3, 4, 6, 10, 30, 39 and plated on control plate –Leu-Ura and selection plates –Leu-Ura-Ade and –Leu-Ura-His. Rab3 cross used as negative control along with crosses with empty vectors pGAD and pGBDU as well as positive control mCog4xmCog2 of which a strong mammalian interaction is known. DIL = Dilution, Q-L = GTP-locked active Rab GTPase.

The same Rab GTPases were also crossed with *Drosophila* Cog5 to look for interactions between the Rab GTPases and the Cog5 subunit of *Drosophila* (**Figure 20**). The results indicate that *Drosophila* Cog5 interacts with Rab4, 10, 30 and 39, of which the interaction with Rab39 is also seen in mammalian systems (**Figure 3**). The interactions with Rab4, 10 and 30 are therefore interactions specific to *Drosophila*.

	DIL		DIL		DIL		DIL
mCog4 X mCog2	mCog4 X mCog2	pGAD x pGBDU	pGAD x pGBDU	pGBDU x Cog5	pGBDU x Cog5		
pGAD x Rab1 Q-L	pGAD x Rab1 Q-L	pGAD x Rab2	pGAD x Rab2	pGAD x Rab3 Q-L	pGAD x Rab3 Q-L	pGAD x Rab4 Q-L	pGAD x Rab4 Q-L
pGAD x Rab6 Q-L	pGAD x Rab6 Q-L	pGAD x Rab10 Q-L	pGAD x Rab10 Q-L	pGAD x Rab30 Q-L	pGAD x Rab30 Q-L	pGAD x Rab39 Q-L	pGAD x Rab39 Q-L
Cog5 x Rab1 Q-L	Cog5 x Rab1 Q-L	Cog5 x Rab2	Cog5 x Rab2	Cog5 x Rab3 Q-L	Cog5 x Rab3 Q-L	Cog5 x Rab4 Q-L	Cog5 x Rab4 Q-L
Cog5 x Rab6 Q-L	Cog5 x Rab6 Q-L	Cog5 x Rab10 Q-L	Cog5 x Rab10 Q-L	Cog5 x Rab30 Q-L	Cog5 x Rab30 Q-L	Cog5 x Rab39 Q-L	Cog5 x Rab39 Q-L

-Leu-Ura **-Leu-Ura-Ade** **-Leu-Ura-His**

Figure 21. *Drosophila* Cog5 interacts with *Drosophila* Rabs 4, 10, 30 and 39: *Drosophila* Cog5 was crossed with *Drosophila* Rabs 1, 2, 3, 4, 6, 10, 30, 39 and plated on control plate –Leu-Ura and selection plates –Leu-Ura-Ade and –Leu-Ura-His. Rab3 cross used as negative control along with crosses with empty vectors pGAD and pGBDU as well as positive control mCog4xmCog2 of which a strong mammalian interaction is known. DIL = Dilution, Q-L = GTP-locked active Rab GTPase.

Finally the crosses were repeated with same Rab GTPases and with *Drosophila* Cog6 to look for interactions between the Rab GTPases and the Cog6 subunit of *Drosophila* (**Figure 21**). The results suggest that *Drosophila* Cog6 interacts with Rab10, 30 and 39, of which the interaction with Rab10 is also seen in mammalian systems (**Figure 3**). The interactions with Rab30 and 39 are therefore interactions specific to *Drosophila*.

	DIL		DIL		DIL		DIL
mCog4 X mCog2	mCog4 X mCog2	pGAD x pGBDU	pGAD x pGBDU	pGBDU x Cog5	pGBDU x Cog5		
pGAD x Rab1 Q-L	pGAD x Rab1 Q-L	pGAD x Rab2	pGAD x Rab2	pGAD x Rab3 Q-L	pGAD x Rab3 Q-L	pGAD x Rab4 Q-L	pGAD x Rab4 Q-L
pGAD x Rab6 Q-L	pGAD x Rab6 Q-L	pGAD x Rab10 Q-L	pGAD x Rab10 Q-L	pGAD x Rab30 Q-L	pGAD x Rab30 Q-L	pGAD x Rab39 Q-L	pGAD x Rab39 Q-L
Cog5 x Rab1 Q-L	Cog5 x Rab1 Q-L	Cog5 x Rab2	Cog5 x Rab2	Cog5 x Rab3 Q-L	Cog5 x Rab3 Q-L	Cog5 x Rab4 Q-L	Cog5 x Rab4 Q-L
Cog5 x Rab6 Q-L	Cog5 x Rab6 Q-L	Cog5 x Rab10 Q-L	Cog5 x Rab10 Q-L	Cog5 x Rab30 Q-L	Cog5 x Rab30 Q-L	Cog5 x Rab39 Q-L	Cog5 x Rab39 Q-L

-Leu-Ura **-Leu-Ura-Ade** **-Leu-Ura-His**

Figure 22. *Drosophila* Cog6 interacts with *Drosophila* Rabs 10, 30 and 39: *Drosophila* Cog6 was crossed with *Drosophila* Rabs 1, 2, 3, 4, 6, 10, 30, 39 and plated on control plate –Leu-Ura and selection plates –Leu-Ura-Ade and –Leu-Ura-His. Rab3 cross used as negative control along with crosses with empty vectors pGAD and pGBDU as well as positive control mCog4xmCog2 of which a strong mammalian interaction is known. DIL = Dilution, Q-L = GTP-locked active Rab GTPase.

4. Discussion

Overall there are many things to take away from the results obtained in this project. Firstly the work done using PNGaseAr has characterised the efficiency and functionality of this enzyme in comparison to the generally used PNGaseF for *N*-glycan cleavage assays. As well as this not only has the FANGS protocol been optimised for use with PNGaseAr but PNGaseF also. Using these enzymes on *Drosophila* lacking Cog3 expression in muscle cells has shown differences in *N*- and O-linked glycan profile in comparison to wild type flies. Furthermore in the flight testing analysis mutations in certain COG subunits has unveiled defects in flight ability compared to others. This can be correlated with differences in the *N*- and O-linked glycan profiles of the Cog3 mutant *Drosophila* line potentially attributing certain glycan structures to having functional relevance in *Drosophila* flight muscle contractile apparatus. And lastly the yeast-two-hybrid analysis has unveiled not only conservation with interactions observed in mammalian systems, but also novel interactions between COG subunits and Rab GTPase binding partners which are unique to the *Drosophila* system.

The optimisation of the FANGS protocol has permitted the use of PNGaseAr in the FANGS procedure which has useful potential for future studies on *Drosophila N*-glycans. Indeed another study using the PNGaseAr enzyme has shown similarities to my data in that they observed only high mannose forms of *N*-glycans [49] as well as observing glycans with more than one fucose addition [49] like the $GlcNAc_3Hex_3Fuc_3NeuAc_1$ glycan seen in **Figure 17**. They did not observe glycans with sialyation, however for their study only cell lines were used [49] as opposed to segments of whole *Drosophila* as I used with the thoraces containing many cell types, of which some may have specifically sialylated glycans. But not only that, with its observed enhancement of PNGaseF activity due to the effectiveness of SDS removal from the procedure, this could be used in *N*-glycan studies for other systems also. Using the published FANGS protocol [40] although useful for removing the vast majority of *N*-glycans from samples and generating glycan profiles via mass spectrometry, this could be potentially improved by using the protocol optimised in this project. Indeed another group found an alternate way of optimising FANGS through using a coupled method with individuality normalization when labelling with glycan hydrazide tags (INLIGHT) [50]. However in order to directly compare my optimisation of FANGS with the optimisation they carried out, a comparison between the original FANGS procedure and the optimised version I have developed should be carried out using cell samples rather than purified protein as I have. The more effective removal of SDS will reduce inhibition of the PNGase enzymes which could allow for glycans in lower abundance which would not be cleaved and/or detected by mass spectrometry to appear when they wouldn't using the previously established FANGS protocol.

Although the PNGaseAr enzyme is less efficient and has less *N*-glycan cleavage capacity than PNGaseF enzyme, it appears on the *N*-glycan profile that a $Hex_3HexNAc_3Fuc_3NeuAc_1$ glycan has been cleaved which could only be achieved by the PNGaseAr enzyme given the fucosylation linkages of the structure. This shows that the optimisation of the PNGaseAr enzyme for use in FANGS on the *Drosophila* samples was not in vein unveiling an increase in abundance of this glycan in the Cog3 mutant line compared to wild type. Indeed previous studies have shown the importance of sialylation in *Drosophila* muscle [41] leading to the possibility that this observed increase is having an impact on the flight muscle contractile apparatus and consequently flight ability. Not only have observations been made regarding the impact of *N*-glycosylation on *Drosophila* muscle through MGAT1 null effects on neuromuscular junction synaptogenesis [33], but also its role in *Drosophila* neural transmission in that *N*-glycan sialylation has been found to control neural excitability [51]. Moreover an increase of the $Hex_8HexNAc_2$ glycan was observed in the mutant along with an increase of $Hex_3HexNAc_3Fuc_3NeuAc_1$. Alterations in sialylated *N*-glycans is of interest due to its reported roles in both muscular and neurological environments. There may be correlation with changes in sialyation and the observations of defective flight in particular *Drosophila* COG mutant lines. As well as clear defects in *Drosophila* flight ability and therefore muscular contraction, they also displayed a landing defect by hovering for extended periods of time (as described in Results) possibly suggesting a neurological perturbation causing delays in landing times.

The O-glycan profile unveiled some interesting alterations in glycan abundance also with an observed increase in glycans such as $Hex_1HexNAc_2NeuAc_1$ and $Hex_1Deoxyhexose_2HexA_1$ as well as decreases in structures $HexNAc_3$ and $Hex_1Deoxyhexose_1NeuAc_1$. The O-glycan data is comparable to the N-glycan profiling in that differences in glycan abundances were also observed in sialylated glycans. As with the *N*-glycan results this may point towards highlighting the role of sialylation in cellular events responsible for muscular contraction. Such deviations from wild type glycosylation patterns may point to mislocalisation of glycoslyation enzymes due to a lack of Cog3 mediated vesicle tethering. Indeed it has been shown previously that a deficiency of the Cog7 subunit in human fibroblasts leads to altered recycling of Golgi proteins [52]. It has been shown previously that faulty O-glycosylation also has an impact on *Drosophila* muscle through impaired mucin type O-glycosylation leading to muscle weakening and progressive degeneration [42]. Indeed the decrease in abundance of the $HexNAc_3$ glycan is of interest to muscular function in that blistered wings has previously been observed following *pgant3* mutation of RNAi, a phenotype of integrin-mediated cell interactions which are critical in many diverse processes [57] including muscular function [58]. The fact that *pgant* enzymes are GalNAc transferases means observations of differences in the $HexNAc_3$ structure could be attributable to the mislocalisation of certain *pgant* enzymes between Golgi cisternae causing glycan processing perturbations and an

underproduction of the HexNAc$_3$ glycan. This glycan structure may therefore be important as a cell surface glycan which mediates cell-cell or cell-tissue interactions involved in muscular contraction and consequently *pgant* mislocalisation results in impaired muscular function.

More glycan structures both *N*- and O-linked may be unveiled through multiple repeats of the *Drosophila* glycan analysis as in this data set some glycans only appeared in one or two samples. Therefore more repeats could certainly improve the reproducibility of the data and perhaps lead to more glycan alterations between mutant COG and wild type lines. Moreover tandem mass spectrometry on the *N*- and O-linked glycan structures of which differences were observed between mutant and wild type would allow elucidation of the specific monosaccharide linkages which make up the glycans in question. Overall this data shows that COG has an impact on the glycosylation machinery at the Golgi in *Drosophila* and impacts the *N*- and O-linked glycans produced in muscle specific cells.

Flight test analysis on the various COG mutant lines of *Drosophila* revealed certain COG subunit defects have a greater impact on flight muscle than others. The data suggests a Cog2 mutation suppression of the $\frac{fwsp}{fwsp}$ defect in flight, reverting the flight index to a wild type-like score. Furthermore it seems that the $\frac{fwsp}{fwsp}$ mutation is the predominant cause of a defect in flight as well as the $\frac{Cog1}{+}$ mutation having some impact also. Moreover the rescue of the *fws*ᵖ mutation leading to a rescue of the flight defect furthers the suggestion that the *fws* COG subunit plays a key role in regulating glycosylation of proteins involved in muscular contraction. What is useful about this data is the fact that there are no balancer chromosomes present in any of the genotypes meaning that any defect in flight can be directly correlated with mutations in COG subunits. The balancer chromosomes can have a significant impact in flight ability as shown in **Figure 6** where the TM6B balancer presence in fly lines caused a severe defect in flight compared to lines which did not have this in their genotype. The mechanism underlying the defect in flight caused by the mutations in COG subunits is likely due to compromised interactions between components of the extracellular matrix and muscle through glycosylated proteins such as dystroglycan, a consequence of which is muscular dystrophy [43]. Defects in COG subunits may lead to mislocalised glycosylation enzymes consequently impairing the glycosylation machinery and thereby improperly glycosylating proteins important for interactions key to muscular function. This data can be correlated with differences in the *N*- and O-linked glycan profiles of the Cog3 mutant *Drosophila* line potentially attributing certain glycan structures to having functional relevance in *Drosophila* flight muscle contractile apparatus. In combination with the glycan profiling results there is certainly evidence here for a role of COG in mediating the glycosylation

apparatus in *Drosophila.* Perturbations in particular COG subunits will lead to differences in glycan synthesis which translates to functional defects *in vivo* in the form of flight ability.

Finally from the data obtained in the yeast-two-hybrid analysis it is evident that conservation exists between the *Drosophila* and mammalian systems in terms of their COG-Rab interactions. Levels of conservation are expected due to the high levels of sequence homology between the protein sequences of COG subunits and Rabs in *Drosophila* and mammals. Indeed it would be unsurprising that COG subunits would interact with Rab GTPases in *Drosophila* with there being observations in flies of Rab GTPases interacting with other vesicle tethering proteins such as Rab30 interactions with coiled-coil tether proteins dGCC88, dGolgin-97 and dGolgin-245 [54]. Cog4, 5 and 6 in *Drosophila* showed interactions with Rab30 and further interaction studies would certainly shed light onto which other potential molecular players are involved in COG-mediated vesicle tethering. Indeed other studies have been carried out looking at *Drosophila* COG-Rab interactions [54] which in correlation with my results showed interactions between Cog4 and Rab1, 2, 30 and 39, as well as between Cog6 and Rab30. However they also showed Cog5 and Cog6 interactions with Rab6 and Rab2 which I did not observe. I also in contrast to their results showed an interaction between Cog4, 5 and 6 with Rab10 and 39 of which Cog5-Rab39 and Cog6-Rab10 interactions are conserved between the mammalian system. A possibility for some of the differences observed between the sets of data could be the use of the yeast-two-hybrid system rather than pull-down experiments, therefore in order to validate my results; further protein-protein interaction assays should be conducted.

 All 3 COG subunits crossed with a variety of Rabs showed at least one conserved interaction when compared to mammalian interactions. However interactions between COG subunits and Rab partners which have not been observed in mammalian studies are most likely attributable to differences in the protein structures which are specific to the *Drosophila* sequences. These have unveiled novel COG-Rab interactions which are unique to the *Drosophila* system and therefore they most likely have some differences in vesicle trafficking regulation in comparison to mammals. However having said that the conservation between some interactions highlights the usefulness of *Drosophila* to study as a model organism for vesicle tethering events and glycosylation patterns as a result of glycosylation enzyme trafficking.

Overall this project has highlighted the role of COG in regulating the glycosylation machinery and shown the importance of this mediation in *Drosophila* muscular contraction. Moreover it has highlighted evolutionary conservation of interactions between COG subunits and Rab GTPases indicating the usefulness of *Drosophila* as a model system for glycosylation studies. Further studies on the impact of COG on *Drosophila* glycosylation and its subunit interactions with

different partners of vesicle trafficking can only substantiate our understanding of this intriguing

protein complex both mechanistically *in vitro* and functionally *in vivo*.

Abbreviations

AD = GAL4 Activation Domain

AmpR = Ampicillin resistant

ANOVA = Analysis of Variance

BD = GAL4 DNA Binding Domain

CDG = Congenital disorders of Glycosylation

CMC = Critical micelle concentration

COG = Conserved Oligomeric Golgi

COPI = Coat protein I

COPII = Coat protein II

CyO = Curly wings

Df = Deficiency

DHB = 2,5-dihydroxybenzoic acid

Dm = Drosophila melanogaster

*D*mef2 = *Drosophila* myocyte enhancer factor-2

DMSO = Dimethyl sulfoxide

DTT = Dithiothreitol

ER = Endoplasmic reticulum

FANGS = Filter-aided *N*-glycan separation

FGF = Fibroblast growth factor

Fld = Fused lobes

FTMS = Fourier transform mass spectrometry

Fuc = Fucose

Fws = four way stop

Fws[p] = *four way stop* p element disruption

GalNAc = *N*-acetylgalactosamine

GAP = GTPase activating protein

GDI = GDP dissociation inhibitor

GDP = Guanosine diphosphate

GEF = Guanine nucleotide exchange factor

GlcNAc = *N*-acetylglucosamine

GM130 = Golgi matrix protein 130

GMAP-210 = Golgi microtubule-associated protein 210

GTP = Guanosine triphosphate

Het = Heterozygous

Hex = Hexose

HexA = Glucoronic acid

HexNAc = *N*-acetylhexosamine

Hom = Homozygous

HPLC = High Performance Lipid Chromatography

INLIGHT = individuality normalization when labelling with glycan hydrazide tags

KDN = deaminated neuraminic acid

LEU2 = Gene encoding Beta-isopropylmalate dehydrogenase, part of leucine biosynthesis pathway

MALDI = Matrix-assisted laser desorption/ionisation

Man = Mannose

mCog = Mammalian Cog

MGAT = UDP-GlcNAc:α-3-D-mannoside-β1,2-N-acetylglucosaminyl-transferase I

MKRS = *Drosophila* balancer

MS = Mass spectrometry

MTC = Multi-subunit tethering complex

NeuAc = Sialic acid

pGAD = plasmid GAL4 Activation Domain

pGBDU = plasmid GAL4 DNA Binding Domain Uracil

PNGaseAr = Peptide *N* Glycosidase Ar

PNGaseF = Peptide *N* Glycosidase F

REP = Rab escort protein

RNAseB = RibonucleaseB

SNARE = Soluble NSF Attachment Protein Receptor

STX = Syntaxin

TM6B = *Drosophila* balancer

t-SNARE = target-Soluble NSF Attachment Protein Receptor

TX-100 = TritonX-100

UAS = Upstream activation sequence

v-SNARE = vesicle-Soluble NSF Attachment Protein Receptor

WT = Wild type

Xyl = Xylose

Y2H = Yeast-two-hybrid

References

1. Khoury, G. Baliban, R. and Floudas, C. (2011). Proteome-wide post-translational modification statistics: Frequency analysis and curation of the Swiss-Prot database. *Science Reports*. 1(90). pp. 1-5

2. Garner, O. and Baum, L. (2008). Galectin–glycan lattices regulate cell-surface glycoprotein organization and signalling. *Biochemical Society Transactions*. 36(6). pp.1472.

3. Gavrilov, B. Rogers, K. Fernandez-Sainz, I. Holinka, L. Borca, M. and Risatti, G. (2011). Effects of glycosylation on antigenicity and immunogenicity of classical swine fever virus envelope proteins. *Virology*. 420(2). pp.135-145.

4. Leroy, J. (2006). Congenital disorders of *N*-glycosylation including diseases associated with O- as well as *N*-glycosylation defects. *Pediatric Research Nature*. 60(6). pp.643-656.

5. Akasaka-Manya, K. Manya, H. Sakurai, Y. Wojczyk, B. Kozutsumi, Y. Saito, Y. Taniguchi, N. Murayama, S. Spitalnik, S. and Endo, T. (2009). Protective effect of N-glycan bisecting GlcNAc residues on -amyloid production in Alzheimer's disease. *Glycobiology*. 20(1). pp. 99-106.

6. Radhakrishnan, P. Dabelsteen, S. Madsen, F. Francavilla, C. Kopp, K. Steentoft, C. Vakhrushev, S. Olsen, J. Hansen, L. Bennett, E. Woetmann, A. Yin, G. Chen, L. Song, H. Bak, M. Hlady, R. Peters, S. Opavsky, R. Thode, C. Qvortrup, K. Schjoldager, K. Clausen, H. Hollingsworth, M. and Wandall, H. (2014). Immature truncated O-glycophenotype of cancer directly induces oncogenic features. *Proceedings of the National Academy of Sciences*. 111(39). pp. 4066-4075.

7. Islam, R. Nakamura, M. Scott, H. Repnikova, E. Carnahan, M. Pandey, D. Caster, C. Khan, S. Zimmermann, T. Zoran, M. and Panin, V. (2013). The role of *Drosophila* cytidine monophosphate-sialic acid synthetase in the nervous system. *Journal of Neuroscience*. 33(30). pp. 12306-12315.

8. Stevens, J. and Spang, A. (2013). *N*-glycosylation is required for secretion and mitosis in *C. elegans*. *Public Library of Science ONE*. 8(5). pp. 63687.

9. Gavel, Y. and Heijne, G. (1990). Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. *Protein Engineering Design and Selection*. 3(5). pp. 433-442.

10. Schwarz, F. and Aebi, M. (2011). Mechanisms and principles of *N*-linked protein glycosylation. *Current Opinion in Structural Biology*. 21(5). pp. 576-582.

11. Fu, J. Wei, B. Wen, T. Johansson, M. Liu, X. Bradford, E. Thomsson, K. McGee, S. Mansour, L. Tong, M. McDaniel, J. Sferra, T. Turner, J. Chen, H. Hansson, G. Braun, J. and Xia, L. (2011). Loss of

intestinal core 1–derived O-glycans causes spontaneous colitis in mice. *Journal of Clinical Investigation*. 121(4). pp. 1657-1666.

12. Tang, G. Ruiz, T. and Mintz, K. (2012). O-polysaccharide glycosylation is required for stability and function of the collagen adhesin EmaA of *Aggregatibacter actinomycetemcomitans*. *Infection and Immunity*. 80(8). pp. 2868-2877.

13. Nilsson, T. Au, C. and Bergeron, J. (2009). Sorting out glycosylation enzymes in the Golgi apparatus. *Federation of European Biochemical Societies Letters*. 583(23). pp. 3764-3769.

14. Rabouille, C. Hui, N. Kieckbusch, R. Berger, EG. Warren, G. Nilsson, T. (1995). Mapping the distribution of Golgi enzymes involved in the construction of complex oligosaccharides. *Journal of Cell Science*. 108 (1). pp. 1617-27.

15. Elsner, M. Hashimoto, H. Nilsson, T. (2003) Cisternal maturation and vesicle transport: Join the band wagon! (Review). *Molecular Membrane Biology.* 20(3). pp. 221-229.

16. Orci, L. Amherdt, M. Ravazzola, M., Perrelet, A. and Rothman, J. (2000). Exclusion of Golgi residents from transport vesicles budding from Golgi cisternae in intact cells. *The Journal of Cell Biology*. 150(6). pp. 1263-1270.

17. Roboti, P. Sato, K. and Lowe, M. (2015). The golgin GMAP-210 is required for efficient membrane trafficking in the early secretory pathway. *Journal of Cell Science*. 128(8). pp. 1595-1606.

18. Sztul, E. and Lupashin, V. (2009) Role of vesicle tethering factors in the ER–Golgi membrane traffic. *Federation of European Biochemical Societies Letters.* 583(23). pp. 3770-3783.

19. Huaqing, C. Reinisch, K. and Ferro-Novick, S. (2007) Coats, tethers, Rabs and SNAREs work together to mediate the intracellular destination of a transport vesicle. *Developmental Cell.* 12(5). pp. 671-682.

20. Richardson, BC. Smith, RD. Ungar, D. Nakamura, A. Jeffrey, PD. Lupashin, VV. Hughson, FM. (2009) Structural basis for a human glycosylation disorder caused, in part, by mutation of the COG4 gene. *Proceedings of the National Academy of Sciences.* 106. pp. 13329–13334.

21. Oka, T. Ungar, D. Hughson, FM. Krieger, M. (2004) The COG and COPI complexes interact to control the abundance of GEARs, a subset of Golgi integral membrane proteins. *Moleculuar Biology of the Cell.* 15. pp. 2423–2435.

22. Wu, X. Steet, RA. Bohorov, O. Bakker, J. Newell, J. Krieger, M. Spaapen, L. Kornfeld, S. Freeze, HH. (2004) Mutation of the COG complex subunit gene COG7 causes a lethal congenital disorder. *Nature Medicine*. 10. pp. 518–523.

23. Willett, R. Kudlyk, T. Pokrovskaya, I. Schönherr, R. Ungar, D. Duden, R. and Lupashin, V. (2013). COG complexes form spatial landmarks for distinct SNARE complexes. *Nature Communications*. 4, pp. 1553.

24. Laufman, O. Kedan, A. Hong, W. and Lev, S. (2009). Direct interaction between the COG complex and the SM protein Sly1 is required for Golgi SNARE pairing. *European Molecular Biology Organisation*. 28(14). pp. 2006-2017.

25. Shestakova, A. Zolov, S. and Lupashin, V. (2005). COG complex-mediated recycling of Golgi glycosyltransferases is essential for normal protein glycosylation. *Traffic*. 7(2). pp. 191-204.

26. Miller, V. Sharma, P. Kudlyk, T. Frost, L. Watson, I. Duden, R. Lowe, M. Lupashin, V. and Ungar, D. (2013). Molecular insights into vesicle tethering at the Golgi by the conserved oligomeric Golgi (COG) complex and the golgin TATA element modulatory factor (TMF). *Journal of Biological Chemistry*. 288(6). pp. 4229-4240.

27. Lees, J. Yip, C. Walz, T. and Hughson, F. (2010). Molecular organization of the COG vesicle tethering complex. *Nature Structural & Molecular Biology*. 17. pp. 1292–1297.

28. Reiter, L. (2001). A systematic analysis of human disease-associated gene sequences in *Drosophila melanogaster*. *Genome Research*. 11(6). pp. 1114-1125.

29. Leonard, R. Rendic, D. Rabouille, C. Wilson, IB. Preat, T. (2006) The *Drosophila fused lobes* gene encodes an *N*-acetylglucosaminidase involved in *N*-glycan processing. *Journal of Biological Chemistry.* 281. pp. 4867–4875

30. Lane, S. Frankino, W. Elekonich, M. and Roberts, S. (2014). The effects of age and lifetime flight behavior on flight capacity in *Drosophila melanogaster*. *Journal of Experimental Biology*. 217(9). pp. 1437-1443.

31. Beltrán-Valero de Bernabé, D. Currier, S. Steinbrecher, A. Celli, J. van Beusekom, E. van der Zwaag, B. Kayserili, H. Merlini, L. Chitayat, D. Dobyns, W. Cormand, B. Lehesjoki, A. Cruces, J. Voit, T. Walsh, C. van Bokhoven, H. and Brunner, H. (2002). Mutations in the O-mannosyltransferase gene POMT1 give rise to the severe neuronal migration disorder Walker-Warburg Syndrome. *The American Journal of Human Genetics*. 71(5). pp. 1033-1043.

32. Nakamura, N. Stalnaker, S. Lyalin, D. Lavrova, O. Wells, L. and Panin, V. (2009). *Drosophila* dystroglycan is a target of O-mannosyltransferase activity of two protein O-mannosyltransferases, rotated abdomen and twisted. *Glycobiology*. 20(3). pp. 381-394.

33. Parkinson, W. Dear, M. Rushton, E. and Broadie, K. (2013). *N*-glycosylation requirements in neuromuscular synaptogenesis. *Development*. 140(24). pp. 4970-4981.

34. Matsumura, K. and Campbell, K. (1994). Dystrophin-glycoprotein complex: Its role in the molecular pathogenesis of muscular dystrophies. *Muscle & Nerve*. 17(1). pp. 2-15.

35. Ungar, D. (2009). Golgi linked protein glycosylation and associated diseases. *Seminars in Cell & Developmental Biology*. 20(7). pp. 762-769.

36. Zolov, S. and Lupashin, V. (2005). Cog3p depletion blocks vesicle- mediated Golgi retrograde trafficking in HeLa cells. *Journal of Cell Biology*. 165(5). pp. 747-759.

37. Ungar, D. (2002). Characterization of a mammalian Golgi-localized protein complex, COG, that is required for normal Golgi morphology and function. *The Journal of Cell Biology*. 157(3). pp. 405-415.

38. Mariappa, D. Sauert, K. Marino, K. Turnock, D. Webster, R. Van Aalten, D. Ferguson, M. and Muller, H. (2011). Protein O-GlcNAcylation is required for fibroblast growth factor signaling in *Drosophila*. *Science*. 4(204). pp. 1-8.

39. Martin-Blanco, E. and Garcia-Bellindo, A. (1996). Mutations in the rotated abdomen locus affect muscle development and reveal an intrinsic asymmetry in Drosophila. *Proceedings of the National Academy of Sciences*. 93. pp. 6048-6052.

40. Abdul Rahman, S. Bergström, E. Watson, C. Wilson, K. Ashford, D. Thomas, J. Ungar, D. and Thomas-Oates, J. (2014). Filter-Aided N -Glycan Separation (FANGS): A convenient sample preparation method for mass spectrometric *N* -glycan profiling. *Journal of Proteome Research*. 13(3). pp. 1167-1176.

41. Repnikova, E. Koles, K. Nakamura, M. Pitts, J. Li, H. Ambavane, A. Zoran, M. and Panin, V. (2010). Sialyltransferase regulates nervous system function in *Drosophila*. *Journal of Neuroscience*. 30(18). pp. 6466-6476.

42. Aoki, K. Porterfield, M. Lee, S. Dong, B. Nguyen, K. McGlamry, K. and Tiemeyer, M. (2008). The diversity of O-linked glycans expressed during *Drosophila melanogaster* development reflects stage- and tissue-specific requirements for cell signaling. *Journal of Biological Chemistry*. 283(44), pp. 30385-30400.

43. Moore, C. and Winder, S. (2010). Dystroglycan versatility in cell adhesion: a tale of multiple motifs. *Cell Communication and Signaling*. 8(1). p.3.

44. James, P. Haliaday, J. & Craig, E. A. (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics*. 144. pp. 1425–1436.

45. Perrett, D. (2003). Book Reviews: The Proteome Revisited. The theory and practice of all relevant electrophoretic steps. *Proteomics*. 3(1). pp. 104-105.

46. Biolabs, N. (2015). *PNGase F | NEB*. [online] Neb.com. Available at: https://www.neb.com/products/p0704-pngase-f [Accessed 13 Dec. 2015].

47. Kim, K. Lawrence, S. Park, J. Pitts, L. Vann, W. Betenbaugh, M. and Palter, K. (2002). Expression of a functional *Drosophila melanogaster N*-acetylneuraminic acid (Neu5Ac) phosphate synthase gene: evidence for endogenous sialic acid biosynthetic ability in insects. *Glycobiology*. 12(2). pp. 73-83.

48. Lee, T. Sethi, M. Leonardi, J. Rana, N. Buettner, F. Haltiwanger, R. Bakker, H. and Jafar-Nejad, H. (2013). Negative regulation of Notch signalling by xylose. *Public Library of Science.* 9(6). pp .e1003547.

49. Mabashi-Asazuma, H. Kuo, C. Khoo, K. and Jarvis, D. (2013). A novel baculovirus vector for the production of nonfucosylated recombinant glycoproteins in insect cells. *Glycobiology*. 24(3). pp. 325-340.

50. Hecht, E. McCord, J. and Muddiman, D. (2015). Definitive screening design optimization of mass spectrometry parameters for sensitive comparison of filter and solid phase extraction purified, INLIGHT plasma *N*-glycans. *Analytical Chemistry*. 87(14). pp. 7305-7312.

51. Repnikova, E. Koles, K. Nakamura, M. Pitts, J. Li, H. Ambavane, A. Zoran, MJ. Panin, VM. (2010) Sialyltransferase regulates nervous system function in Drosophila. *Journal of Neuroscience.* 30. pp. 6466–6476.

52. Steet, R. Kornfeld, S. (2006): COG7-deficient human fibroblasts exhibit altered recycling of Golgi proteins. *Molecular Biology of the Cell.* 17(5). pp. 2312-2321.

53. Haines, N. S. Seabrooke, B. A, Stewart. (2007) Dystroglycan and protein O-mannosyltransferases 1 and 2 are required to maintain integrity of *Drosophila* larval muscles. *Molecular Biology of the Cell.* 18(12). pp. 4721-4730

54. Gillingham, A. Sinka, R. Torres, I. Lilley, K. Munro, S. (2014) Toward a comprehensive map of the effectors of Rab GTPases. *Developmental Cell.* 31(3). pp. 358-373.

55. Zhang, L. and Hagen, K. (2010). Dissecting the biological role of mucin-type O-glycosylation using RNA interference in *Drosophila* cell culture. *Journal of Biological Chemistry*. 285(45). pp. 34477-34484.

56. Tran, D. Zhang, L. Zhang, Y. Tian, E. Earl, L. and Ten Hagen, K. (2011). Multiple members of the UDP-GalNAc: Polypeptide *N* -acetylgalactosaminyltransferase family are essential for viability in *Drosophila*. *Journal of Biological Chemistry*. 287(8). pp. 5243-5252.

57. Zhang, L. Zhang, Y. and Hagen, K. (2008). A mucin-type O-glycosyltransferase modulates cell adhesion during *Drosophila* development. *Journal of Biological Chemistry*. 283(49). pp. 34076-34086.

58. Chanana, B. Graf, R. Koledachkina, T. Pflanz, R. and Vorbrüggen, G. (2007). αPS2 integrin-mediated muscle attachment in Drosophila requires the ECM protein Thrombospondin. *Mechanisms of Development*. 124(6). pp. 463-475.

59. Katzemich, A. Kreiskother, N. Alexandrovich, A. Elliott, C. Schock, F. Leonard, K. Sparrow, J. and Bullard, B. (2012). The function of the M-line protein obscurin in controlling the symmetry of the sarcomere in the flight muscle of Drosophila. *Journal of Cell Science*. 125(14). pp. 3367-3379.

60. Schwartz, S. Cao, C. Pylypenko, O. Rak, A. and Wandinger-Ness, A. (2008). Rab GTPases at a glance. *Journal of Cell Science*. 121(2). pp. 246-246.