THE LECTIN *DOLICHOS BIFLORUS* AGGLUTININ AS A MARKER FOR EARLY MURINE EMBRYONIC STEM CELL DIFFERENTIATION AND CHARACTERIZATION OF THE CARBOHYDRATE EPITOPES INVOLVED IN LECTIN BINDING

by

LORI ANNE HILL NEVES

(Under the Direction of Michael Pierce)

ABSTRACT

Cell surface markers are important tools used to characterize and separate a mixed cell population. Existing cell surface markers used to identify murine embryonic stem cells (mESCs) such as stage specific embryonic antigens (SSEA-1, SSEA-3, SSEA-4) and Forssman antigen are limited because they are unable to define the pluripotent state of mESCs, and are not reliable indicators of early differentiation. In a previous study, a panel of lectins was used to identify glycan cell surface markers that undergo changes in early differentiation. *Dolichos biflorus* agglutinin (DBA) recognizes the glycan epitope α-N-acetylgalactosamine (GalNAc). These glycan epitopes decline drastically upon the first definable step of differentiation, the change from mESCs to primitive ectoderm, and therefore the loss of GalNAc epitopes is the earliest cell surface change that can be assigned to mESCs. In addition, it has been determined that these GalNAc epitopes are present on glycoproteins rather than glycolipids, and are O-linked.

INDEX WORDS: *Dolichos biflorus* agglutinin, DBA, embryonic stem cells, lectin precipitation, lectin affinity chromatography

THE LECTIN *DOLICHOS BIFLORUS* AGGLUTININ AS A MARKER FOR EARLY MURINE EMBRYONIC STEM CELL DIFFERENTIATION AND CHARACTERIZATION OF THE CARBOHYDRATE EPITOPES INVOLVED IN LECTIN BINDING

by

LORI ANNE HILL NEVES

B.S, Furman University, 2003

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTERS OF SCIENCE

ATHENS, GEORGIA

2007

THE LECTIN *DOLICHOS BIFLORUS* AGGLUTININ AS A MARKER FOR EARLY MURINE EMBRYONIC STEM CELL DIFFERENTIATION AND CHARACTERIZATION OF THE CARBOHYDRATE EPITOPES INVOLVED IN LECTIN BINDING

by

LORI ANNE HILL NEVES

Major Professor: Michael Pierce

Committee: Lance Wells
Kelley Moremen
Stephen Dalton

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2007

DEDICATION

This work is dedicated to my husband, James Edward Neves, Jr. He always believed in me, even when I did not believe in myself. Without his constant and unwavering help and support, all of this would have never been possible, and I will strive for the rest of my life to support him in the same way. I love you, dear.

ACKNOWLEDGEMENTS

I would first like to thank my parents, Justin and Karen Hill, and my grandparents, John and Dee Nelson, for their support through out my entire educational career. Without you, I would have never learned to love the pursuit of knowledge, and I greatly appreciate all the sacrifices you went through to ensure that I had the best education possible. I can only hope to be as loving and supportive to my children and grandchildren as you have been to me. I would also like to thank my brother, Brad Hill. You have always pushed me to be a better person, both in life and in academics. If not for your faith in my abilities, as well as your academic competition, I would not be the scientist that I am today.

I would also like to thank Dr. Gerardo Alvarez-Manilla, who has always believed in me and pushed me to achieve great things, both in the laboratory and in life. Also, I would like to thank Dr. Lance Wells, who has never questioned my abilities and has always supported and encouraged my career goals.

Finally, I would like to thank Nicole Warren and Heather Strachan. Without the constant love and support you two have given me over the years I would not be where or who I am today. You are the best friends anyone could ever hope to have, and I love you dearly.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

The ontogeny of multicellular organisms and the activation of various cell types are associated with changes in glycan structures, as cell surface glycolipids and glycoproteins are drastically altered during differentiation [1]. Alterations in glycan structures are thought to participate in cell adhesion, receptor activation, cell differentiation, and tissue morphogenesis. Because the early embryo is composed of cells that must adhere to other embryonic cells, factors that modulate this adherence have an important role in embryonic development [2].

Lectin refers to a group of heterogeneous molecules that are not genetically or structurally related, their only commonality is that they bind carbohydrates in a reversible fashion without enzymatic activity. They are ubiquitous in nature and are found in animals, plants, bacteria, and even viruses. They are mainly involved in recognition functions and play important roles in embryogenesis, cancer, inflammation, immune response, and fertilization [3]. The fact that lectins promote a variety of cell-cell adhesion events suggests the possibility that cell-surface carbohydrates play an integral role in the adherence of embryonic cells. A variety of cell surface markers have been identified that are key tools in the separation and characterization of stem cells at different stages of differentiation. Here I discuss the involvement of glycans and lectins in embryonic development from fertilization to postimplantation embryogenesis.

Murine oocyte fertilization may require glycan recognition by lectins. In mice, O-glycans, which are linked to the protein via a Serine (Ser) or Threonine (Thr), have been shown to be essential for sperm-egg binding, since removal of O-glycans by chemical means destroys

this interaction [4, 5]. Initially, it was thought that the α1-3 Galactose (Gal) terminus was important for this interaction, but α1-3Gal deficient mice do not exhibit reduced sperm-egg interactions or reduced fertility [6]. β1-4Gal was thought to have a function in the oocyte acrosome reaction initiated by sperm binding [7], and α1-3 glycosylated O-glycans were thought to be necessary for high affinity sperm binding [8]. However, in a recent study it was shown that neither Gal or GalNAc on N- or O-glycans are essential receptors for sperm binding on the mouse zona pellucia, the extracellular matrix which surrounds the oocyte and the ovulated eggs [9].

After fertilization, the zygote divides several times as it travels through the oviduct to the uterus, where it implants on the uterine epithelial wall. This stage prior to implantation is termed preimplantation development. It is characterized by cell division and embryo size constraints, which produce daughter cells that have lost totipotency, and by cell-cell interactions that are unique to the parent cell. During the morula stage of embryogenesis, the embryos have approximately 16 cells and will be undergoing compaction, which results in cell flattening and redistribution of the cortical elements [2], an event that will be discussed below. At the last stage of preimplantation, the embryo contains approximately 64 cells. At this point the embryo consists of an inner cell mass, which will become the embryo, and an outer layer of trophoblast cells, which will become the placenta.

The first stage-specific embryonic antigens (SSEAs) are produced during the preimplantation phase, as the embryonic genome becomes active. Most SSEAs were found to be glycan structures that may exist on cell surface proteins or lipids, and their expression pattern changes markedly during early embryogenesis. SSEA-1, or Lewis-X antigen (Figure 1.1), which is found on 8-cell embryos, is lost completely by the 64-cell stage. After implantation it is found

in the embryonic ectoderm and the visceral endoderm. Previously, SSEA-1 had been thought to contribute to embryonic compaction [10] which begins at the 8-cell stage and ends at the 32-cell stage. However, Kudo et.al. (2004) showed that SSEA-1 expression is not essential for *in vivo* embryogenesis by creating a mouse knockout of Fut9, a key enzyme involved in SSEA-1 synthesis [11]. Even with the absence of SSEA-1 due to the knockout, the mice develop normally and are fertile.

Most glycolipids involved in embryogenesis are of the globo series (containing Galα1-4Galβ1-4Glcβ1-ceramide), which includes SSEA-3, SSEA-4, and Forssman antigen (FA) (Figure 1.2.1, 1.2.2, 1.2.3). SSEA-3 and SSEA-4 expression is lost by the blastocyst stage, while FA is found as early as the 8-cell stage and is maintained in the inner cell mass [12]. The functions of these glycans in embryogenesis, however, are unknown.

The next stage in embryonic development is compaction, which can be regulated by glycoprotein bound polylactosamine-type carbohydrates [2]. There are two types of polylactosamine structures. The i-type are unbranched and are expressed in fetal erythrocytes in the zygote stage, and begin to appear in parietal and visceral endoderm of preimplantation embryos (Figure 1.3.1) [13]. The branched I-type polylactosamines are expressed in adult erythrocytes, and therefore are either absent from the embryo or expressed at low levels (Figure 1.3.2). Embryoglycan, a glycoprotein bound polylactosamine-type large carbohydrate which is found in early embryonic cells, carries several cell surface markers that are expressed both in teratocarcinoma and murine embryos, such as SSEA-1 [14]. In addition, some enzymes involved in the synthesis of N-glycans decrease in activity from fertilization onward until they are activated in the blastocyst stage [15]. N-glycans are linked to the protein via an N-

aceytlglucosamine (GlcNAc), which is linked to an Asparagine (Asn) at a sequence of Asn-X-Ser/Thr.

After implantation of the blastocyst stage embryo into the uterine epithelium, a variety of morphogenic and molecular processes occur that are reflected in changes in glycosylation. During implantation, the trophoblast cells link the embryo and the uterine epithelium to form the placenta. This is associated with an increase in the levels of polylactosamine [13]. After gastrulation, or the formation of the three germ layers (ectoderm, endoderm, and mesoderm), i-antigen expression is widespread, as is sialylation of these structures. I-type polylactosamine is found in the endoderm and the mesoderm, while sialyl-I structures are found in the ectoderm. SSEA-3 and SSEA-4 reappear on the endoderm during postimplantation, and FA is found on the primitive ectoderm [16]. Cell movement during gastrulation correlates with changes in the extracellular matrix, which is rich in hyaluronan and proteoglycans. These glycans may promote trafficking and differentiation events. Proteoglycans accumulate in areas of morphological changes, such as organ and limb bud development. However, there are few reagents available to detect changes in glycosaminoglycans *in situ* [2].

In earlier studies, the mouse teratocarcinoma was used as a model to study cell surface glycans during embryogenesis [16]. Teratocarcinoma is a tumor composed of all the three germ layers and malignant stem cells, called embryonal carcinoma (EC) cells [13]. EC cells are similar to the pluripotent cells of early embryos, as many of the cell surface markers that are expressed in EC cells are also found in early embryos [14].

In a previous study, Muramatsu (2000) used [H^3]-fucose labeled glycopeptides obtained by protease digestion of either plasma membranes or whole EC cells to determine that fucose containing glycopeptides are present in EC cells, and decrease in number upon differentiation

[16]. Further biochemical experiments excluded the possibility that the glycopeptides are from glycosaminoglycans, mucin-type glycopeptides, or glycolipids. They were found instead to be glycoprotein-bound carbohydrates that are altered upon differentiation [16].

The glycosylation events during early mouse embryogenesis and teratocarcinoma progression have been studied using several lectins and carbohydrate binding proteins. One lectin of interest in the search for cell surface markers is the seed lectin *Dolichos biflorus* agglutinin (DBA) from the leguminous plant. DBA is a blood group A+H substance binding lectin that recognizes the non-reducing end of a terminal α1-3GalNAc [3]. It is a tetrameric glycoprotein with a MW of 110 kDa, and consists of 2 types of subunits, I and II (Figure 1.4). Subunit II is a post-translationally modified version of subunit I, with 12 residues removed from the C-terminus. DBA has an extreme preference for GalNAc over Gal when compared to other GalNAc binding lectins [3]. This is due to an amino acid substitution in its monosaccharide binding site which replaces an aromatic residue typically found in other Gal binding lectins, such as Peanut Agglutinin (PNA), with an aliphatic residue (Leu127) [3]. This Leu127 allows favorable interactions between the N-acetyl group of the sugar and the Glycine103/Tyrosine104/ Tyrptophan132 (Gly103/Tyr104/Trp132) triad in the lectin, which compensates for the replacement of the aromatic residue in position 127. Only GalNAc can compensate for this substitution, resulting in a lower affinity for Gal (Figure 1.5) [3].

Previous studies have shown that DBA receptors are expressed in unfertilized ova, and that this expression declines during pre-implantation embryogenesis [13]. Embigin, a glycosylated core protein expressed in extraembryonic endoderm during early post implantation embryos, has a DBA binding site [16]. The expression of this epitope tends to decline and becomes too low to detect by day 10 of embryogenesis [16]. In addition, antibodies against

DBA binding glycoproteins from teratocarcinoma cells have been used to detect the high molecular weight glycoprotein brushin, which is expressed in the primitive endoderm, the visceral endoderm, and the renal tubular brush border [13]. Other antibodies were also able to detect HBP-44, a heparin-binding protein, which binds members of the LDL receptor family and acts as a chaperone [16].

Another lectin that has been used is *Helix pomatia* agglutinin (HPA). HPA is also an A+H blood group terminal α-GalNAc binding lectin, which is found in the albumin gland of the roman snail and has the ability to aggregate bacteria in order to protect the snail from infection [17]. In addition, HPA has the capability to agglutinate blood group A red cells, but not B or O [17]. A previous study utilized HPA affinity chromatography to show that HPA receptors are expressed on the surface of EC cells, but are lost upon differentiation with retinoic acid (RA) [18]. In addition, they observed a 3-4 fold decrease in Forssman glycolipid in the RA differentiated EC cells [18]. This data reinforced the thought that differentiation of EC cells is accompanied by a decline in the presence of terminal α-GalNAc epitopes.

In mammal species, pluripotent embryonic stem cells (ESCs) are derived from the inner cell mass (ICM) of blastocyst-stage embryos. When cultured over extensive periods of time under appropriate conditions ESCs retain many of the characteristics associated with pluripotent cells of the ICM, including the capacity to generate the three embryonic germ lineages as well as the extraembryonic tissues that support development. This pluripotency of ESCs provides the basis for replicating a wide variety of somatic and extraembryonic tissues [19, 20]. Elucidating the molecular mechanisms for stem cell differentiation is perhaps one of the most critical areas of contemporary research in developmental biology because of its potential therapeutic applications in the treatment of diseases and injuries. The molecular mechanisms that regulate

differentiation of stem cells into specific cell lineages are just beginning to be revealed and a huge effort is being mounted to develop *in vitro* systems that direct the differentiation of ESCs into specific lineages that could be used for drug development, cell replacement therapeutics and gene delivery [21].

Although more that 20 human ESCs have been developed and approved for NIH-funded experimentation in the United States, for ethical and practical reasons, murine ESC (mESC) lines have provided for several years an *in vitro* model for understanding their differentiation into multiple tissues [20]. In mESCs, the presence in the culture media of the leukemia inhibitory factor (LIF), a cytokine of the interleukin (IL-6) family, stimulates mESC renewal and suppresses differentiation [22]. When LIF is removed from the media, mESCs spontaneously differentiate into spheroid colonies, which are termed embroid bodies (EBs) because they recapitulate certain aspects of early embryogenesis with the appearance of lineage-specific regions of differentiation similar to that found in the embryo [23]. Several cell surface markers such as CD9, alkaline phosphatases (AP) and the glycan antigens SSEA-1 and FA have been found to be preferentially expressed during the mESC stage [21, 24] however the expression of these epitopes is greatly reduced or absent in EBs.

Unlike mESCs, which can be maintained in a pluripotent state by the presence of LIF in the media, human ESCs (hESCs) are more difficult both to maintain in the pluripotent state and to induce into differentiation [25]. hESCs have been routinely cultured on mouse embryonic feeder layers (MEFs), laminin, fibronectin, and human serum in MEF cultured media [20, 26-28]. In order to induce differentiation, the cells must be removed enzymatically from the feeder layer and re-cultured as EBs. Thus, standard culture conditions for hESCs are limited by the

need for transfer of the cells between two culture systems, which can cause substantial variability of the culture systems [25].

A recent study has addressed this problem by attempting to create a single culture system that would allow cells to change between differentiation states in the same culture setting [25]. This procedure utilizes a hyalruonic acid (HA) hydrogel [25]. Hydrogel capsules have a high water content, which promotes cell viability, and are structurally and mechanistically similar to the extracellular matrix of many tissues [29]. HA is a nonsulfated linear polysaccharide of (1-β-4) glucuronic acid and (1-β-3) N-acetylglucosamine [30]. HA supports hESCs *in vitro* because it suppresses gene expression, signaling, proliferation, motility, adhesion, and morphogenesis of hESCs [30]. It was determined that HA-hydrogels allow hESCs to maintain the pluripotent state in culture and still allows differentiation to EBs [25].

While there are a variety of cell surface markers used to define the differentiation state of ESCs, currently available markers are unable to define the pluripotent state and identify differentiation commitments. ESC and ectoderm primitive layer (EPL) cells are closely related cell populations that previously could only be discriminated by transcript analysis. All pathways of differentiation toward one of the three germ layers must first go through this cell type, and therefore it is the first definable step in differentiation [24]. Nash et. al. (2007) addressed this problem by screening a panel of 18 lectins to determine if any epitopes are elevated on the surface of mESCs at various stages of differentiation. They found that the lectin DBA recognized cell surface epitopes in mESC, and these epitopes significantly decline upon the first step of differentiation, the change from ESC to EPL [24]. Therefore, DBA, as a cell surface marker, is able to discriminate between ESC and EPL cells, and could serve as an indicator of developmental pluripotency.

In this study we developed methods to determine which cell surface macromolecule contains the glycan that results in DBA binding. We show that some of the binding to DBA in mESC is caused by protein and not lipid, and initial analysis suggests that these proteins are O-linked rather than N-linked. We have been able to support previous results suggesting that DBA could act as a cell surface marker for the first definable step in mESC differentiation [24]. If future studies with human ESC demonstrate the same DBA reactivity to cells during early differentiation, these techniques could be utilized in a variety of therapeutic methods.

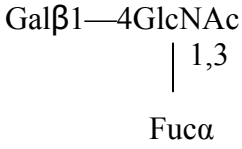


Figure 1.1: The determinant structure of the cell surface marker SSEA-1, or Lewis X. This is expressed in 8-cell stage embryos, but is lost by the 64-cell stage. It can also be found in the embryonic ectoderm and visceral endoderm after implantation [31].
Gal=Galactose GlcNAc=N-acetylglucosamine Fuc=Fucose

1.2.1 GalNAcβ1—3Galα—4Gal

1.2.2 NeuAcα2—3Galβ1—3GalNAc

1.2.3 GalNAcα1—3GalNAcβ1—3Gal

Figure 1.2.1, 1.2.2, 1.2.3: The determinant structures of (1.2.1) SSEA-3, (1.2.2) SSEA-4, and (1.2.3) Forssman antigen. These cell surface markers belong to the globo series of glycolipids. SSEA-3 and SSEA-4 expression is gone by the blastocyst stage. Forssman antigen can be found as early as the 8-cell stage and is later maintained in the inner cell mass [31]. GalNAc=N-acetylgalactosamine Gal=Galactose NeuAc=Neurominic acid

1.3.1 Galβ1—4GlcNAcβ1—3Galβ1—4GlcNAcβ1—3Gal

1.3.2 Galβ1—4GlcNAcβ1

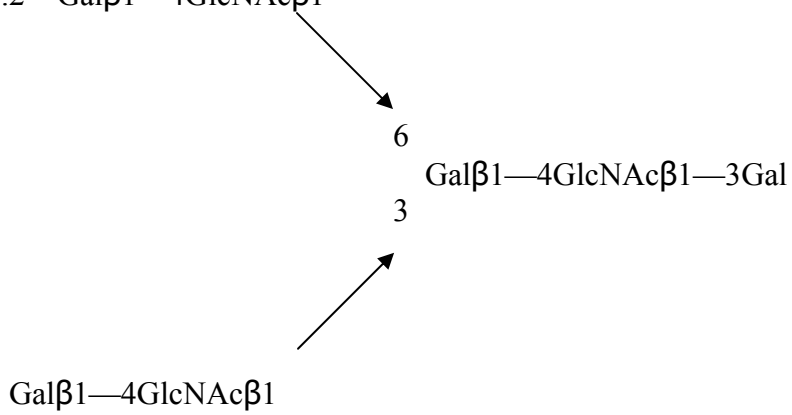


Figure 1.3.1, 1.3.2: The determinant structures of (1.3.1) the unbranched i-type polylactosamine and (1.3.2) the branched I-type polylactosamine. Polylactosamine-type structures are involved in the regulation of compaction. The i-type are expressed in fetal erythrocytes, while the I-type are expressed in adult erythrocytes and are absent from the embryo [13].
Gal=Galactose GlcNAc=N-acetylglucosamine

Figure 1.4: Overall structure of *Dolichos biflorus* agglutinin in complex with adenine. DBA consists of two different subunits, subunit I (displayed in pink and gold), and subunit II (displayed in cyan and purple). Subunit I and II form dimers (shown in pink/cyan and gold/purple), and two dimers form the tetrameric protein shown. Adenine is shown in spacefill cpk. Ca^{2+} (green) and Mn^{2+} (RedOrange) are present in each carbohydrate binding site and are required for GalNAc binding. Monosaccharide binding sites are indicated with stars. Structures were manipulated in RasMol from Protein Data Bank (PDB) accession number 1bjq.

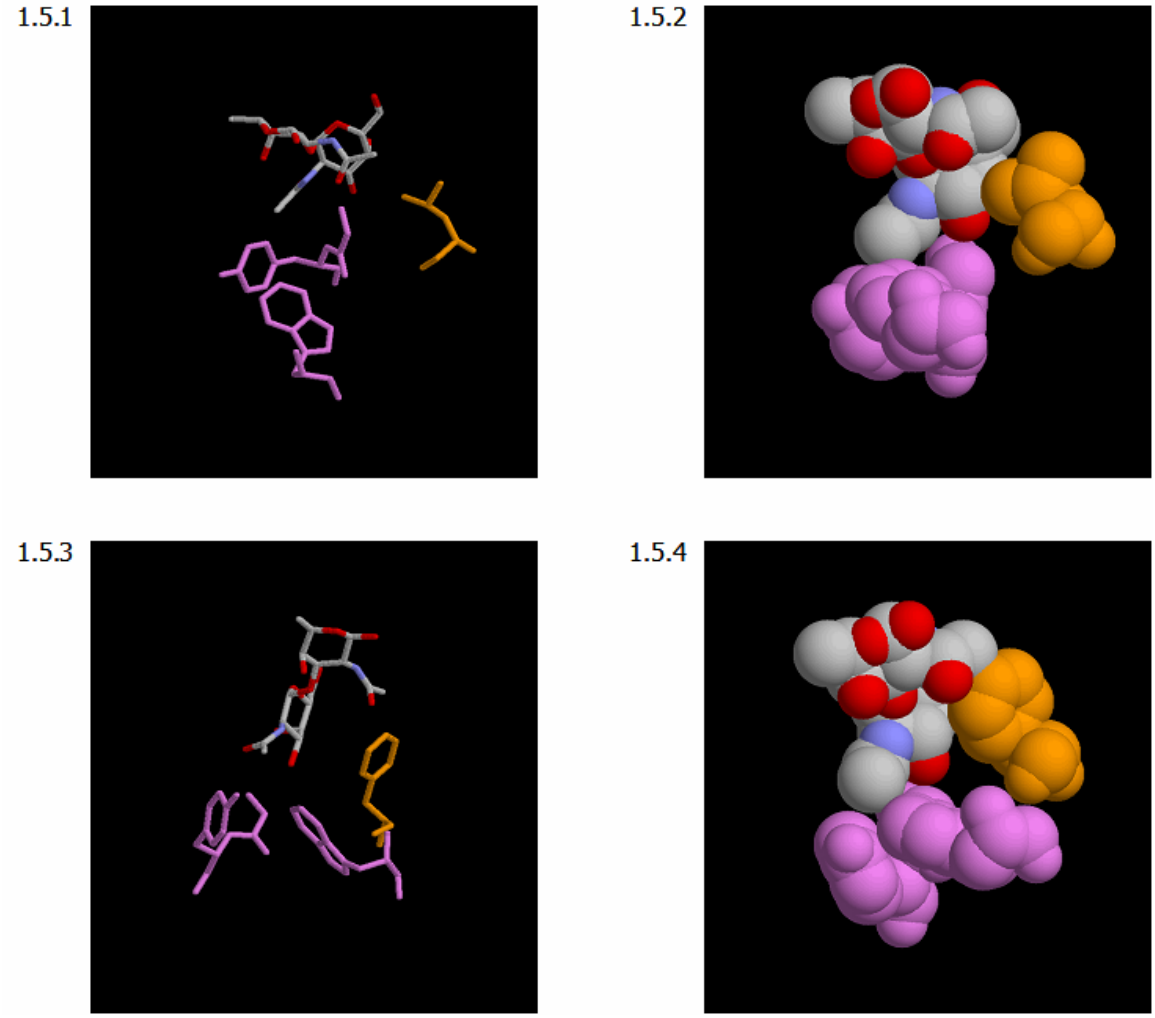


Figure 1.5.1-1.5.4 Structure of *Dolichos biflorus* agglutinin monosaccharide binding site. (1.5.1) Amino acids essential for monosaccharide binding are shown in complex with Forssman antigen (α1,3GalNAcβ1,3GalNAc), which is shown in cpk. The Leu 127 substitution is shown in gold, and the Gly103/Tyr104/Trp132 triad is shown in violet. (1.5.2) Same as (1.5.1) shown in spacefill to illustrate steric interactions between Leu 127 and N-acetyl group of the non-reducing N-acetylgalactosamine. (1.5.3) Monosaccharide binding site is shown with the aromatic residue Phe, which is found in the monosaccharide binding site of Peanut agglutinin (PNA), a Galactose binding lectin. Phe is displayed in gold, the Gly103/Tyr104/Trp132 triad is shown in violet, and Forssman antigen is shown in cpk. (1.5.4) Same as (1.5.3) shown in spacefill to illustrate steric hindrance between the N-acetyl group of N-acetylgalactosamine and the aromatic residue typically found in galactose binding lectins [3]. Structures were manipulated in RasMol from PDB accession number 1lu1.

CHAPTER 2

MATERIALS AND METHODS

Cell Culture and Lysis

R1 mESC were cultured as previously described [24]. Cells were radiolabeled by replacement of growth media with media containing nitrogen-dried [H^3]-glucosamine. Cells were grown in the presence of the label for 48 hours. Cell pellets were prepared as previously described [24]. Whole cell lysates were prepared by resuspending pellets in 2:1 volume of cold lysis buffer (0.2M HEPES, 100mM NaCl, 5mM EDTA, 1% Triton-X-100, 1 protease inhibitor tablet (Roche)) at ~1mg protein per ml. Pellets were then probe-sonicated for 2 minutes at an intensity of 5.0 (12W-15W) and centrifuged at 3500 rpm for 30 minutes at 4ºC. The supernatant was then removed, diluted to 0.1% Triton-X-100, and protein concentration was determined by bicinchoninic acid (BCA) assay (Pierce).

N-acetylgalactosamine synthesis

Galactosamine hydrochloride ($GalNH_2 \cdot HCl$, 5g, 23.19mmol) was crushed into a fine powder and dried overnight in a desicator over P_2O_5. The powder was added to an oven-dried flask under argon, which was then charged with anhydrous methanol (MeOH) (25ml) and stirred for 5 minutes. The slurry was then cooled on an ice bath, and then 0.5M anhydrous sodium methoxide in methanol (46.2ml) was added. After stirring for 10 minutes, acetic anhydride (2.2ml, 23.19mmol) was added slowly. The reaction was stirred for 2 hours until all of the $GalNH_2 \cdot HCl$ dissolved and a fine precipitate of NaCl appeared. The reaction was monitored by TLC eluted in either dichloromethane:MeOH:H_2O (60:40:10 v/v/v) or isopropanol:water:28%

ammonium hydroxide (80:20:10 v/v/v). The TLC plates were visualized using 95% sulfuric acid in methanol followed by charring. The reaction mixture was left at 4ºC overnight to ensure maximum NaCl precipitation. The precipitant was removed by filtration through a fritted glass funnel. The mixture was concentrated in vacuo and co-evaporated with toluene (to ensure complete removal of the acetic acid formed in the second step) to give a colorless foam (3g). Purification was achieved by loading the sugar in minimal amount of ethanol (EtOH) onto a column of Iatrobeads (50g) packed using chloroform. The column was eluted with chloroform:MeOH:H_2O (65:35:5 v/v/v), and 5ml fractions were collected until the product began to elute. After 12-13 more fractions were collected, the eluent was changed to 63:37:5 (v/v/v) and 30 2ml fractions were collected. The eluent was then changed to 60:40:5 (v/v/v). Fractions 45-137 were tested by TLC. The fractions containing the product (54-96) were combined and concentrated in vacuo and then lyophilized overnight to give N-acetylgalactosamine as a white powder. The structure was confirmed by Nuclear Magnetic Resonance (NMR) spectroscopy.

Lectin Coupling

In the first protocol, which was used to prepare a 2ml DBA column, 4-5mg/ml of the lectin (Vector) was dissolved in 0.05M sodium phosphate/0.05M sodium citrate at pH 7. Biosupport Medium (Pierce) (0.25g/ml) and 2ml 0.2M GalNAc were added to the lectin and rocked at 4ºC overnight. The following day, 500µl of 3M ethanolamine was added and rocked for 3 hours to stop the coupling. The resin beads were centrifuged for 10 minutes at 6000 rpm and the supernatant was removed. The beads were washed 3 times with 2ml TBS-B (100mM Tris, 1.5M NaCl, 10mM $CaCl_2$ at pH 8) and the supernatant of each wash was saved for BCA assay (Pierce) to determine the coupling efficiency. The resin was then packed in a 2ml column or saved for lectin precipitations.

The second coupling protocol was used for a 1ml DBA column. All solutions were equilibrated to 4ºC before use. First, 20mg DBA (Vector) was dissolved in 2ml 0.2M GalNAc in Buffer B (100mM MOPS, pH7.5/0.3M NaCl), and 5µl was removed and diluted 10 fold in Buffer B for protein estimation by BCA (Pierce). Then 1ml of NHS activated Sepharose media (GE-Healthcare) was placed in a 15ml tube and centrifuged at 1000rpm for 1 minute and supernatant was removed. The resin was then washed 3 times with 2.5ml of 1mM HCl, followed by 3 washes with 2ml of cold milli-Q-H_2O. The lectin was then added to the washed resin and incubated 4-5 hours on a rotating platform at room temperature. After incubation, the mixture was centrifuged to remove supernatant, and then the resin was washed with 2ml Buffer B and both supernatants were saved for yield estimation. GalNAc was present in this supernatant and the carbonyl of the N-acetyl group could result in chelation of copper, which will produce high background and a false positive reaction. Therefore, 3ml of the supernatant was applied to a DG-10 cartridge that was previously equilibrated with 20ml TBS-B. When eluting the DG-10 cartridge, the first 3ml were discarded and the next 4ml, which contained the desalted protein, were collected. Dilutions of 1/2, 1/5, and 1/10 were prepared for BCA assay to estimate coupling efficiency. 2ml of Buffer D (0.5M ethanolamine, pH 8.3/ 0.5M NaCl) was added to resin and incubated for 2 hours at room temperature on a rotating platform, then the coupled resin was washed with 5ml of TBS-B. The resin was then packed in a 1ml serological pipette and washed with at least 50ml of TBS-B with 0.1% Triton-X-100 for glycoproteins and at least 20ml TBS-B for glycopeptides.

Preparation of Glycopeptides from Stem Cell Pellets

Cell pellets frozen at -80ºC were thawed on ice for 15 minutes, then deionized water (dH_2O) was added to a proportion of 1:5 (tissue:water). The mixture was then probe-sonicated

five times for 10 seconds, each at an intensity of 5.0 (12W-15W). To remove glycolipids, CH-Cl_3:MeOH:dH_2O was added to the lysed aqueous mixture to give a 4:8:3 solvent mixture, followed by a 2 hour incubation at room temperature on a rocking platform. After the incubation period, the mixture was brought to 4:8:5.6 (CH-Cl_3:MeOH:dH_2O), vortexed for 1 minute, and then centrifuged at 7000 x g for 45 minutes to separate the mixture into three phases. The upper aqueous phase, containing glycolipids in a mixture of 27% H_2O and 73% MeOH, and the lower organic phase, containing phospholipids, triglycerides, and non-polar lipids in CH-Cl_3, were removed, leaving the intermediate solid phase which contained the precipitated proteins. This intermediate phase was then washed twice with 2ml of cold 100% acetone, quickly vortexed and then centrifuged for 15 minutes at 7000 x g. Then 4ml of dH_2O was added, solution was sonicated as described previously, and protein solution was transferred to a 15ml Falcon tube after removing a small amount for protein quantitation by BCA assay (Pierce). The delipidated protein mixture was then frozen and lyophilized overnight.

To prepare glycopeptides, the lyophilized powder was resuspended in 4ml of 50mM Tris (pH 7.5), 2M urea and sonicated as described earlier. Proteins were then reduced by addition of DTT in the Tris/urea buffer to a concentration of 5mM and incubation at 50ºC for 45 minutes. Then carboxyamidomethylation was achieved by the addition of iodoacetamide (IAM) in the Tris/urea buffer to a concentration of 15mM and incubation at room temperature, in the dark, for 45 minutes. Proteolysis was then performed on reduced, alkylated sample by addition of 1:50 (trypsin:sample) of TPCK-Trypsin in the Tris/urea buffer and sample was digested overnight at 37ºC.

After proteolysis, the glycopeptides were purified by C_{18} Sep-Pak cartridges. First, 500µl of 5% CH_3COOH was added to bring the pH below 4.0. C_{18} Sep-Pak cartridges were attached to

10ml syringes and activated by eluting successively with 5ml MeOH, 5ml 5% CH_3COOH, 5ml propan-1-ol, and 15ml 5% CH_3COOH. The sample was then directly loaded onto the cartridge and was washed with 30ml of 5% CH_3COOH, and glycopeptides were eluted by the addition of 4ml 20% propan-1-ol in 5% CH_3COOH, followed by 4ml of 40% propan-1-ol in 5% CH_3COOH, and an elution with 100% propan-1-ol, which was collected in a separate tube. The first two eluents were then dried on a speed-vac overnight for lectin affinity chromatography.

Lectin Affinity Chromatography

Dried glycopeptides were resuspended in 5ml TBS-B and applied to a 1ml Con A gravity column (GE-Heathcare) that was equilibrated at room temperature with 10 volumes of TBS-B. The sample was then loaded onto the column and branched glycopeptides were eluted with 10 volumes of TBS-B (Con A I). Biantennary glycopeptides were eluted from the column with 10 volumes of 10mM α-D-methyl-mannopyranoside in TBS-B (Con A II), and high mannose glycopeptides were then eluted with 10 column volumes of 100mM α-D-methyl-mannopyranoside in TBS-B (Con A III). The first 5 fractions of each eluent were pooled and brought up to 5% CH_3COOH for desalting with a 3cc (60mg) Oasis MCX cation exchange cartridge (Waters). After desalting and drying in speed-vac, Con A I eluent was resuspended in 200µl TBS-B and was applied to the 1ml DBA column.

For glycopeptides, the NHS Sepharose DBA column was run at 4ºC. After equilibration with 20ml TBS-B, the sample was applied and column was washed with 7.5ml TBS-B. Then bound material was eluted with 7.5ml 0.2M GalNAc in TBS-B. The first six 500µl fractions were pooled and desalted on 3cc MCX cation exchange cartridge (Waters).

The Biosupport DBA column was first equilibrated with 20 ml of wash buffer (50mM HEPES, 100mM NaCl, 0.1% Triton-X-100), then 400µg-1mg cell lysate was added to the

column. The column was then washed with 20ml of wash buffer and 40ml of 0.2M α-methyl-mannopyranoside (Pierce) in wash buffer was applied to remove non-specifically bound glycoproteins. Then 20ml of 0.2M GalNAc in wash buffer was applied to elute bound glycoproteins. Fractions were then pooled accordingly and were precipitated with 12.5% TCA for 20 minutes on ice, then were centrifuged for 20 minutes at 20,000 g at 4ºC. The pellets were washed 3 times in 90% acetone before drying in the speed vac. The sample was then resuspended in TBS-B, and was either prepared for SDS-PAGE, or in cases when radiolabeled material was used, fractions were not pooled for TCA precipitation, but 20% of each fraction was removed for scintillation counting.

Solid Phase Extraction by Oasis MCX cartridge

Glycopeptide fractions were brought up to 5% CH_3COOH for desalting with a MCX cartridge. A 3cc Oasis MCX cartridge was attached to a vacuum manifold and activated with 6ml of MeOH followed by 6ml of 5% CH_3COOH. Glycopeptides were then loaded onto the column and washed with 3ml of 5% CH_3COOH, then the column was washed with MeOH to remove any impurities. The desalted glycoproteins were eluted from the cartridge with 3ml 5% ammonium hydroxide (NH_4OH) in 50% MeOH, followed by 5% NH_4OH in 100% MeOH, and the samples were dried overnight in a speed-vac.

Lectin Precipitation

A 600µg sample of total cell lysate was prepared by addition of $CaCl_2$ and $MgCl_2$ to a concentration of 5mM, then volume was brought up to 600µl with lectin precipitation (LP) buffer (10mM HEPES, 100mM NaCl, 0.1% Triton-X-100). The sample was first pre-cleared with 10µl Protein-A-agarose (Santa Cruz Biotech, Inc.) for 30 minutes at 4ºC, followed by centrifugation at 1000 rpm for 1 minute, and then the supernatant was removed. A 55µl aliquot

of pre-washed DBA agarose (1:1 slurry;Vector Laboratories) was then added, and the solution was incubated on a rotating platform at 4ºC overnight. The next day, the agarose beads were washed 5 times with lysis buffer.

Bound proteins were eluted by two methods. In the first, protein was boiled at 100ºC in the presence of Laemmli Sample Buffer with 2-mercaptoethanol (BIO-RAD) for 10 minutes. The sample was then cooled and loaded directly onto SDS or Bis-Tris polyacrylamide gels.

In the second method of elution, the DBA-agarose beads were first incubated with 100µl of 0.2M α-methyl-mannopyranoside in LP buffer five times for 5 minutes each at 37ºC to remove non-specifically bound glycoproteins. This was followed by elution of bound glycoproteins with 100µl of 0.2M GalNAc in LP buffer five times for 5 minutes each at 37ºC. The first 3 fractions of both eluates were pooled and 10% of each was loaded onto a Bis-Tris gel, which was silver stained with mass spectrometry compatible protocol. Bands of interest were excised for in-gel digest with TPCK trypsin (Pierce).

SDS-PAGE, Bis-Tris Gels, and Lectin Blots

Proteins were resolved on pre-cast 4-20% gradient or 7% SDS polyacrylamide gels (BIO-RAD) and were run at 120V for 1 hour. For lectin blots, proteins were transferred to a nitrocellulose membrane (Bio-Rad) using BIO-RAD 10x western transfer buffer with 10% MeOH at 280mA for 1 hour at 4ºC. The membrane was then incubated in blocking buffer (2% bovine serum albumin (BSA) (Sigma) in TBS-A (100mM Tris, 1.5M NaCl at pH 7.6) containing 0.05% Tween-20 (TBS-T) on a rotating platform overnight at 4ºC. For lectin blots, the membrane was then rocked with a 1:2000 dilution of biotinylated DBA (Vector Laboratories) in blocking buffer for 1 hour at room temperature, followed by three 15 minute washes with TBS-T. The membrane was then rocked with a 1:10,000 dilution of streptavidin-conjugated HRP in

blocking buffer for 30 minutes at room temperature, followed by three 15 minute washes with TBS-T. The bands were then detected by enhanced chemiluminescence (ECL) (Perkin Elmer) and developed on Kodak chemiluminescence film.

Proteins were also resolved on a 4-12% gradient Bis-Tris polyacrylamide gel run in MES buffer (Invitrogen) for 35 minutes at 200V. For lectin blots, proteins were transferred onto a nitrocellulose membrane (Bio-Rad) in 20x transfer buffer with 10% MeOH for 1-2 hours at 160mA at room temperature. The lectin blot was then carried out as described for SDS-PAGE.

Silver and Sypro Ruby stain

For silver stain, the protein gel was incubated with 50% EtOH for 30 minutes, followed by a 1 minute pre-treat ($NaS_2O_35H_2O$, 0.2g/L) and three 20 second washes with dH_2O. The gel was then impregnated with stain (0.1g AgNO, 37.5µl HCOH, 50ml dH_2O) for 20 minutes and washed with dH_2O three times for 30 seconds each. Bands were then developed in 60g/L Na_2CO_3, 500µl HCOH, 4mg/L $NaS_2O_35H_2O$ for 10 minutes, and the gel was placed in stop buffer (50% MeOH, 12% CH_3COOH) for 10 minutes. The gel was stored in 50% MeOH, and was then dried with Gel Drying Solution (Invitrogen) overnight.

For a mass spectroscopy compatible silver stain, the gel was fixed (50% EtOH, 10% CH_3COOH) for 30 minutes, followed by a 15 minute incubation with 10% EtOH and three 1 minute washes with milli-Q-H_2O. The gel was then sensitized for 90 seconds in $NaS_2O_35H_2O$, 0.2g/L, and washed 3 times with milli-Q-H_2O. This was followed by impregnation with 0.2g/100ml $AgNO_3$ for 30 minutes at 4ºC and then three washes with milli-Q-H_2O. Bands were developed (Na_2CO_3, 6g/100ml; HCOH, 50µl/100ml; $Na_2S_2O_3H_2O$, 2ml/100ml) for 10 minutes, and development was stopped (6% CH_3COOH) for 10 minutes. Bands of interest were then excised for in-gel digest.

For Sypro Ruby stain, the gel was fixed for 1 hour in 40% MeOH/10% CH_3COOH, and then was incubated over night in 10-20ml Sypro Ruby stain (Invitrogen). Protein bands were then visible under UV light. Sypro Ruby is mass spectroscopy compatible, and it provides the same sensitivity as silver stain techniques with out the background. It is also able to detect proteins of low expression missed by Coomassie Blue and reveals proteins that stain poorly with silver staining.

In-gel Digest

Corresponding sections of the gel from the GalNAc and α-methyl-mannopyranoside elutions were excised and incubated in 30µl of a 1:1 mix of 30mM potassium ferricyanide and 100mM sodium thiosulfate for 10 minutes. The buffer was removed and gel pieces were washed with milli-Q-H_2O 5 times. Gel pieces were then incubated with 200µl of 40mM NH_4HCO_3 for 10 minutes, followed by 10 minutes in 100µl of 100% CH_3CN, and these steps were repeated until the gel pieces were dehydrated. Gel pieces were then reswelled in 50µl of 10mM DTT in 40mM NH_4HCO_3 at 55ºC for 1 hour. After cooling to room temperature, the solution was exchanged with 50µl of 55mM IAM in 40mM NH_4HCO_3 and incubated in the dark for 45 minutes at room temperature. The gel was then dehydrated by repeating a ten minute incubation with 50µl 40mM NH_4HCO_3 followed by a 10 minute incubation with 50µl of 100% CH_3CN three times. The gel pieces were then dried in the speed-vac for 10 minutes and then reswelled on ice in 20µl of 100ng/µl trypsin in 40mM NH_4HCO_3 for 45 minutes. Trypsin was then removed and replaced with 20µl 40mM NH_4HCO_3 and digested overnight at 37ºC. The following day, peptides were extracted from the gel with 50µl of 5% formic acid (HCOOH)/ 50% CH_3CN three times for 10 minutes. Peptides were then dried by speed-vac for 2 hours and then resuspended in 200µl of 40mM NH_4HCO_3 for analysis by LC-MS/MS.

PNGase F Release of N-linked Oligosaccharides

DBA precipitates were boiled in 10mM sodium phosphate buffer pH 7.5, 0.1% SDS, 0.1% v/v 2-mercaptoethanol for 5 minutes then immediately chilled on ice. SDS salts were removed from eluates by addition of KCl to 100mM followed by centrifugation at 4ºC. Clarified eluates were then subjected to N-glycosidase F (NEB) treatment at 37ºC for 2 hours and protein was precipitated by addition of cold 100% EtOH and microcentrifugation. PNGase F, a type of N-glycosidase, releases N-linked glycans by hydrolyzing the glycosidic linkage between the GlcNac of the N-glycan core and the Asn of the glycosylated protein. This reaction cleaves the glycan from the protein and cleaves the protein at the Asn residue. The cleaved peptides are deaminated, converting the Asn to Aspartic acid (Asp) and an increase of 1 atomic mass unit (amu). This mass change, along with the amino acid change can be detected by mass spectroscopy, allowing mapping of N-linked glycosylation sites (Figure 2.1). Digestion of Fetuin (Sigma) was performed in parallel as a positive control for glycosidase F activity. Precipitated protein was then resolved on a 4-12% Bis-Tris gradient polyacrylamide gel (Invitrogen), transferred onto a nitrocelluose membrane (Bio-Rad) and probed with biotinylated DBA (Vector Laboratories) followed by streptavidin-conjugated HRP. Bands were then detected by ECL (Perkin Elmer).

Glycopeptides, after being eluted from each column, desalted, and dried were transferred brought up in 200µl 100mM NH_4CO_3 and dried for enzymatic digestion. The dried glycopeptides were then resuspended in 30µl of ^{18}O labeled water and incubated overnight at 37ºC with 5µg PNGase F. This reaction causes H_2O to be incorporated into the N-terminus of the cleaved glycan and the Asn, resulting in conversion to Asp, increases the cleaved peptide mass by 3 amu (Figure 2.2). The digested samples were then added to 40µl of 0.1% HCOOH

and the PNGase F was removed by microcentrifugation with Microcon centrifugal filter devices (Millipore) at a maximum speed of 14 x g for 20 minutes. The filter was then washed once more with 40µl milli-Q-H_2O, centrifuged, and dried overnight in a speed-vac. Samples were then prepared for LC-MS/MS as previously described [32].

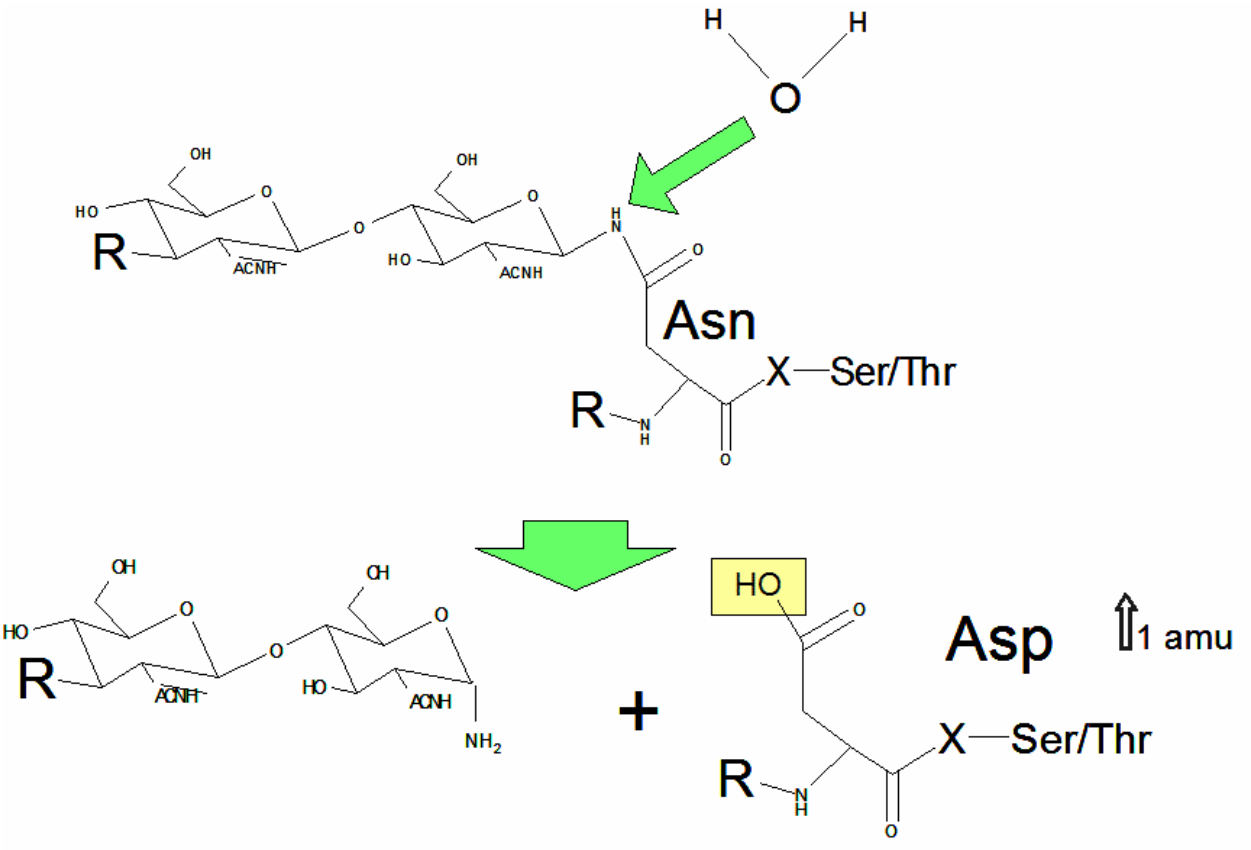


Figure 2.1: Schematic of PNGase F digestion. Showing the glycosidic bond of N-glycans that is cleaved by PNGase F. This reaction results in an increase of the truncated peptide by 1 amu due to the Asn to Asp conversion.

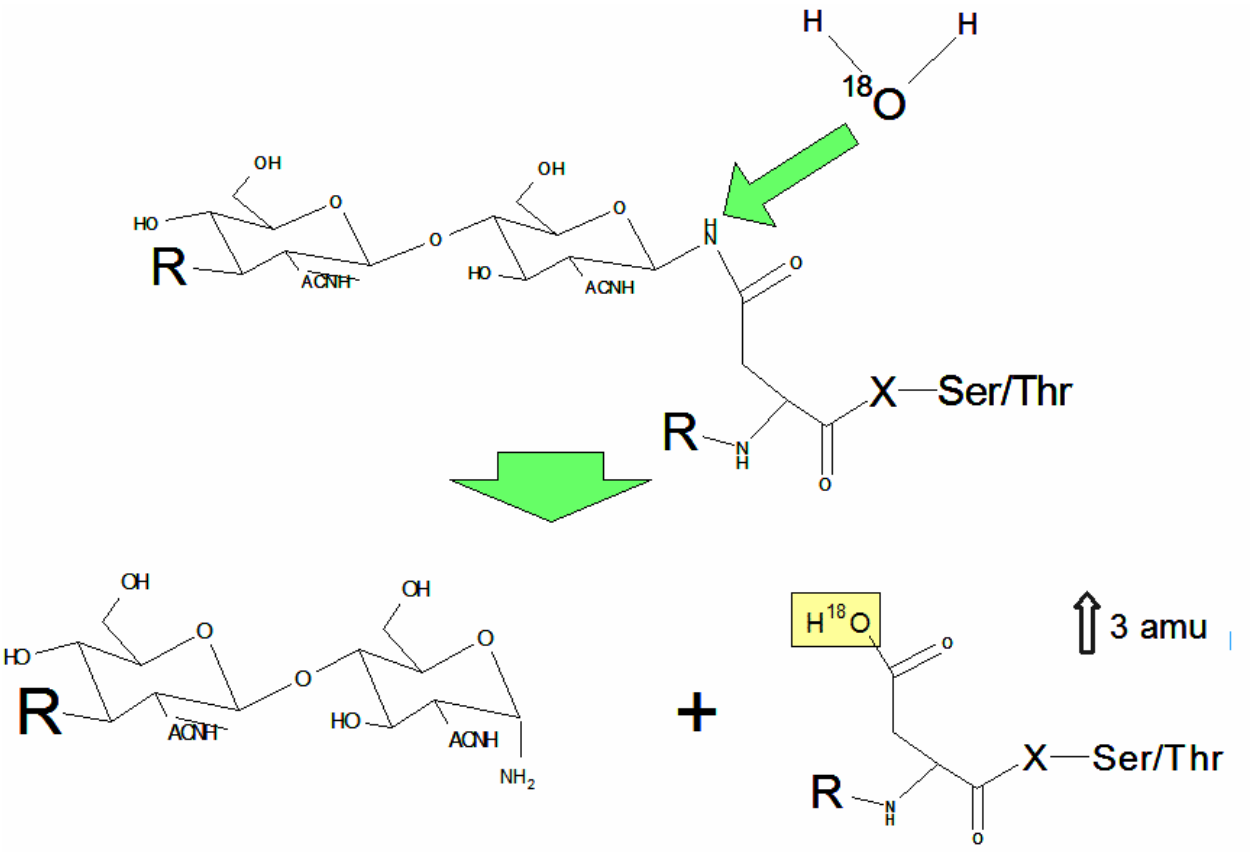


Figure 2.2: Schematic of PNGase F digestion with ^{18}O water. Shows how, when ^{18}O water is used in a PNGase F digestion, there in an increase of 3 amu due to both the higher density of the resulting hydroxyl group that converts the Asn to Asp.

CHAPTER 3

RESULTS

DBA reactivity is due to glycoproteins and not only glycolipids

A DBA lectin blot was performed with 10μg total cell lysate to determine if any DBA reactivity in mESC was due to glycoproteins. About 8 clear protein bands were visible on the lectin blot (Figure 3.1). By denaturing the proteins with SDS and β-ME, the glycan epitopes were fully exposed, allowing DBA to bind more efficiently. Therefore, it was determined that at least some of the DBA reactivity in mESC was due to glycoproteins.

Lectin Coupling Efficiency and Affinity Chromatography

The coupling efficiency was 95% for the Biosupport DBA column and 90% for the NHS Sepharose DBA column. Protein elution for the Biosupport DBA column was determined by using [H^3]-Glucosamine labeled mESC. It was determined that the majority of non-specifically bound proteins were eluted in the first 6 α-methyl-mannopyranoside fractions, with 3% of the total protein eluting in these fractions, and specifically bound proteins were eluted in the first 3 GalNAc fractions, with 1.86% of total protein eluting in these fractions (Figure 3.2.1,3.2.2). It was expected that proteins and peptides would elute in a similar fashion for the NHS Sepharose DBA column.

LC-MS/MS of DBA bound glycopeptides digested with PNGase F

Analysis of glycopeptides applied to the NHS Sepharose DBA column by PNGase F release revealed 5 N-glycosylated peptides in mESC (Table 3.1). We are confident that these represent specific DBA binding because they were found only in the samples applied to the DBA column, while none were found when the lectin Lotus was used. If the peptides were non-

specifically interacting with the agarose, they would have been seen in both lectin samples. Of these proteins, laminin and prosaposin contain only N-linked glycans. Fibronectin, however, contains O-linked glycans when found in both amniotic fluid and in serum. These glycans are thought to play a role in protecting the glycoprotein from protealytic attack at its sensitive regions [33]. The peptides with no N-glycosylations that were retained in the DBA column are rich in Asp and Glu (Table 3.2), which would cause them to be retarded in the column due to their ionic content and the tendency of affinity columns to also act as ion exchange columns [34]. In addition, these proteins may contain O-linked GalNAc epitopes, which would cause them to bind to the DBA column, though it is unlikely because the majorities are cytosolic proteins. It should be noted that this experiment was performed only once. Future studies will attempt to verify this data.

Lectin Precipitation is more efficient at isolating specifically bound proteins than DBA lectin affinity chromatography

The efficiency of the Biosupport DBA column and the DBA lectin precipitation was determined by performing SDS-PAGE, silver stains, and lectin blots with DBA on eluates from both techniques and comparing the results to the DBA lectin blot of the total cell lysate. In the DBA column eluate, 3 protein bands were visible on a silver stained gel, while the lectin blot of the precipitated boiled eluate contained most protein bands found in the total cell lysate lectin blot (Figure 3.3.1, 3.3.2). This data suggests that lectin precipitation may be a more efficient means of isolating proteins that bind to DBA than the Biosupport DBA column. However, the boiled eluent from the lectin precipitation still has a high likelihood of containing non-specifically bound proteins.

In order to reduce the appearance of non-specifically bound proteins in the lectin precipitation, the DBA-agarose beads were eluted with 0.2M GalNAc at 37ºC rather than boiling. It has been previously shown that 0.2M GalNAc is sufficient to release DBA bound glycans [3]. Each sample was resolved on a 4-12% Bis-Tris gradient polyacrylamide gel, and proteins were visualized by Sypro Ruby staining and lectin precipitation with DBA (Figure 3.4.1, 3.4.2). It was determined that elution with 0.2M GalNAc is a more efficient way of protein purification than boiling, as elution with GalNAc reduces the appearance of non-specific bands. Therefore, we are confident that the protein bands found in the GalNAc eluate contain glycans that bind specifically to DBA.

Glycan structures recognized by DBA are O-linked

As an initial characterization of DBA epitopes, it was considered whether the glycan structures recognized by this lectin were N- or O-linked. This was addressed by digesting denatured DBA precipitated protein with PNGase F, an enzyme that cleaves N-linked, but not O-linked glycan structures. The proteins must be denatured with SDS prior to digestion to allow PNGase F complete access to all N-linked sites [35]. After digestion, samples were resolved on a 4-12% Bis-Tris gradient gel and a DBA lectin blot was performed in addition to a Sypro Ruby stain of the gel. This analysis showed that the DBA precipitated proteins were insensitive to digestion, where as Fetuin, a serum protein with well characterized N-linked structures, was sensitive to PNGase F treatment (Figure 3.5.1,3.5.2). It was concluded that glycans containing the DBA epitope are likely to be O-linked.

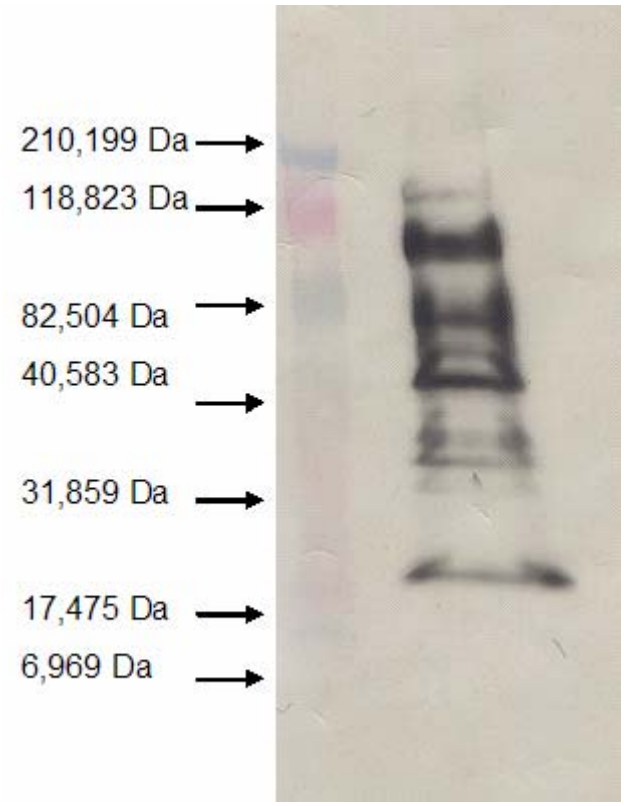
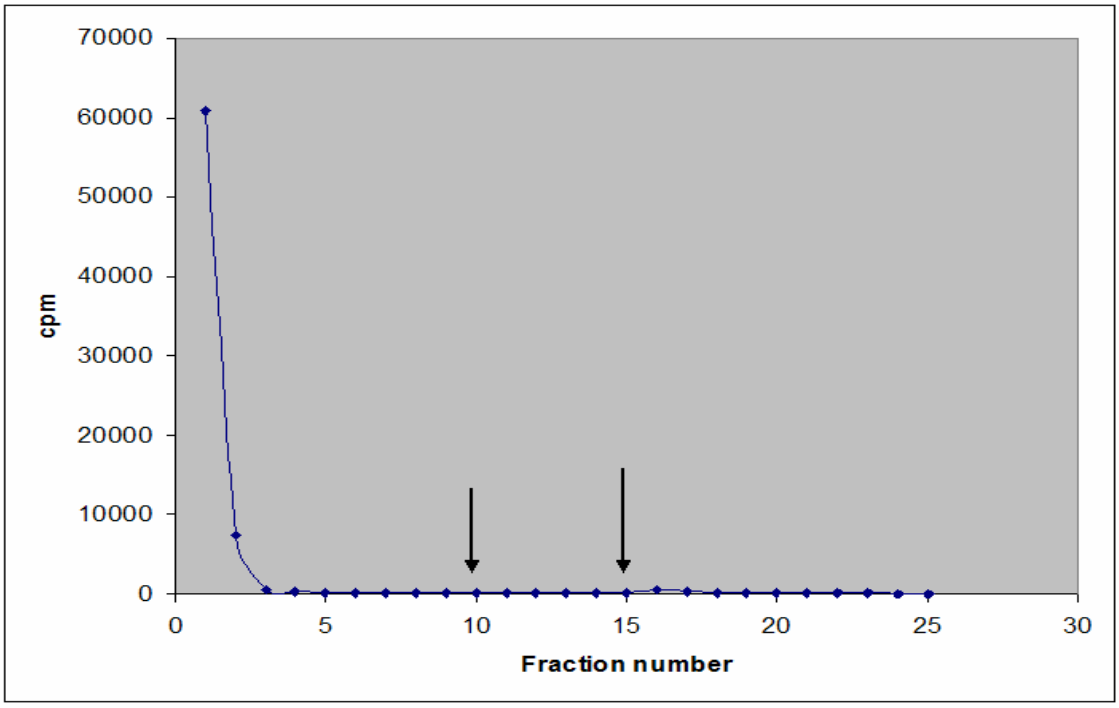


Figure 3.1: DBA lectin blot on whole cell lysate. Whole R1 mESC lysate was resolved on a 4-20% SDS polyacrylamide gel. Proteins were transferred to a nitrocellulose membrane and a lectin blot was performed as described in Materials and Methods. The membrane was exposed to film for 1 minute before development. Around 8 protein bands were present, with molecular weights of 117kDa, 104kDa, 88kDa, 68kDa, 55kDa, 49kDa, 42kDa, and 20kDa.

Figure 3.2.1,3.2.2: [H^3]-glucosamine labeled protein elution from DBA affinity column. mESC lysate labeled with [H^3]-glucosamine was applied to the Biosupport DBA column and elution of protein with various haptens was determined by scintillation counting. The counts per minute (cpm) was compared to the corresponding fraction collected. Fraction 1 is the sample flow-through. Fractions 2-9 are column washes with wash buffer (see Materials and Methods). The arrow at fraction 10 indicates the addition of 0.2M α-methyl-mannopyranoside, and the arrow at fraction 15 indicates the addition of 0.2M GalNAc. (3.2.1) is an overview of all cpm eluted from column. (3.2.2) is a smaller scale view of the cpm eluted with α-methyl-mannopyranoside and GalNAc.

3.2.1


70000
60000
50000
40000
30000
20000
10000
0
cpm
0
5
10
15
20
25
30
Fraction number

3.2.2.

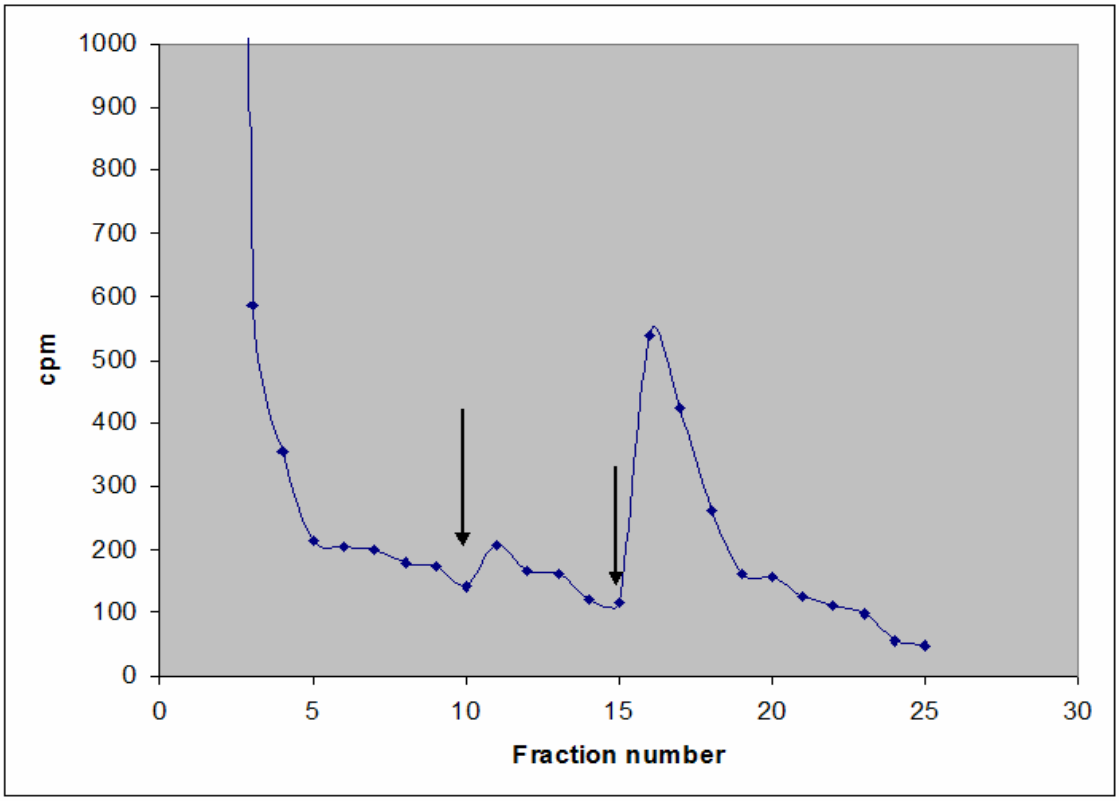

1000
900
800
700
600
500
400
300
200
100
0
cpm
0
5
10
15
20
25
30
Fraction number

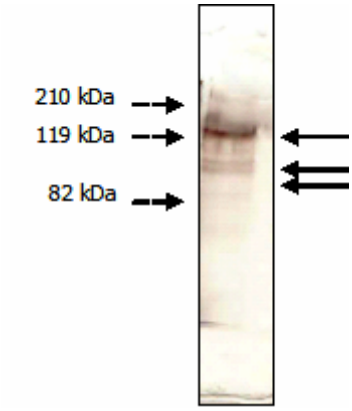
Table 3.1: Proteins identified from bound fractions in 1ml DBA column by LC-MS/MS. In peptide sequence, J represents the site of the N-glycosylation, @ represents the presence of Cys, S-S represents the presence of a disulfide bond, and GaseFO18 refers to the 3 amu increase due to the O^{18} water utilized in the PNGase F reaction. It should be noted that this experiment was performed only once.

Corrected peptide sequence	Mascot score	Protein ID
TLAGEJQTALEIEELNR +GaseFO18	118.66	laminin, gamma 1 [Mus musculus]
FLGNQVLSYGQJLSFSFR +GaseFO18	118.66	laminin, gamma 1 [Mus musculus]
FSELIVNJATEELLVK +GaseFO18	91	prosaposin [Mus musculus]
EVTEGGVVTVJ@SLQEEKPPIFFK +S-S; GaseFO18	57.09	platelet/endothelial cell adhesion molecule 1 isoform 2 [Mus musculus]
DQ@IVDDITYNVJDTFHK +S-S; GaseFO18	39	fibronectin 1 [Mus musculus]
TLSSGJ@TLNVPAKNSYR +GaseFO18	31.43	PREDICTED: similar to RASD family, member 2 [Mus musculus]

Table 3.2: A representation of some of the proteins identified from the bound fractions in the 1ml DBA column that do not contain N-glycosylations. @ represents the presence of Cys, S-S represents the presence of a disulfide bond. There is no mass change from the ^{18}O water because as there are no N-glycans present, PNGase F did not hydrolyze the peptide which results in the addition of a hydroxyl group containing ^{18}O. It is unlikely that these were retained in the column due to O-glycans containing a terminal α-GalNAc because they are cytosolic proteins. It should be noted that this experiment was performed only once.

Corrected peptide sequence	Mascot Score	Protein ID
VNQIGSVTESLQA@K +S-S	115.88	PREDICTED: similar to T21B10.2b [Mus musculus]
VTIAQGGVLPNIQAVLLPK	97.79	histone 3, H2a [Mus musculus]
SLTNDWEEHLAVK	87.4	heat shock protein 1, alpha [Mus musculus]
DAGTIAGLNVLR	86.91	heat shock protein 8 [Mus musculus]
DATNVGDEGGFAPNILENK	83.21	enolase 1, alpha non-neuron [Mus musculus]
FNVWDTAGQEK	81.86	RAS-like, family 2, locus 9 [Mus musculus]
GVVDSEDLPLNISR	87.4	heat shock protein 1, alpha [Mus musculus]
TLTIVDTGIGMTK	87.4	heat shock protein 1, alpha [Mus musculus]
ISGLIYEETR	72.3	histone 1, H4c [Mus musculus]
LAPDYDALDVANK	71.85	PREDICTED: similar to 60S ribosomal protein L23a [Mus musculus]
IGGIGTVPVGR	70.21	eukaryotic translation elongation factor 1 alpha 1 [Mus musculus]
TSGPPVSELITK	70.01	histone 1, H1e [Mus musculus]
TVTAMDVVYALK	72.3	histone 1, H4c [Mus musculus]
HDVVFLITK	68.58	clathrin, heavy polypeptide (Hc) [Mus musculus]
EAFQEALAAAGDK	68.28	thioredoxin 1 [Mus musculus]

3.3.1



3.3.2

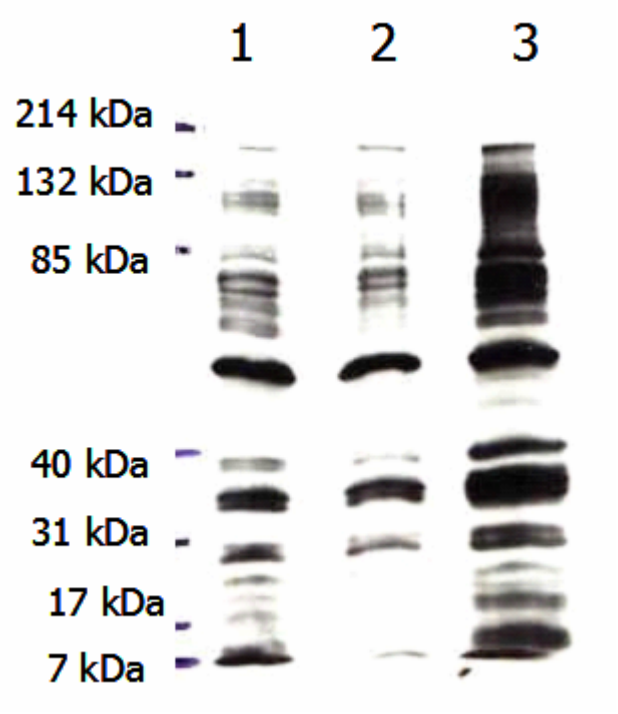


Figure 3.3.1, 3.3.2: Comparison of the efficiency of DBA affinity chromatography and lectin precipitation. (3.3.1) GalNAc eluates from the 2ml DBA column were resolved on a 7% SDS gel and silver stained as described in Materials and Methods. About 3 protein bands were visible at around 132 (Band 1), 111 (Band 2), and 103 (Band 3) kDa. (3.3.2) Boiled eluates from DBA lectin precipitation were resolved on a 7% SDS gel and lectin blot was performed as described in Materials and Methods. The membrane was exposed to the film for 30 seconds before development. Lane 1: Whole mESC lysate. Lane 2: Wash of DBA-agarose beads after lectin precipitation. Lane 3: Eluting removed by boiling. Though 3.3.1 is a representation of total protein present in the Biosupport column GalNAc eluent and 3.3.2 is only showing protein reactive to DBA in the lectin precipitation boil, it appears that there is more recovered protein in the lectin precipitation, possibly due to the removal of the TCA precipitation required for Biosupport column fractions. The unbound material found in lane 2 is thought to be similar to lanes 1 and 3 because the sample was saturated with proteins containing the epitope for DBA binding. This could be resolved by using less total cell lysate or by using more DBA-agarose.

3.4.1 3.4.2

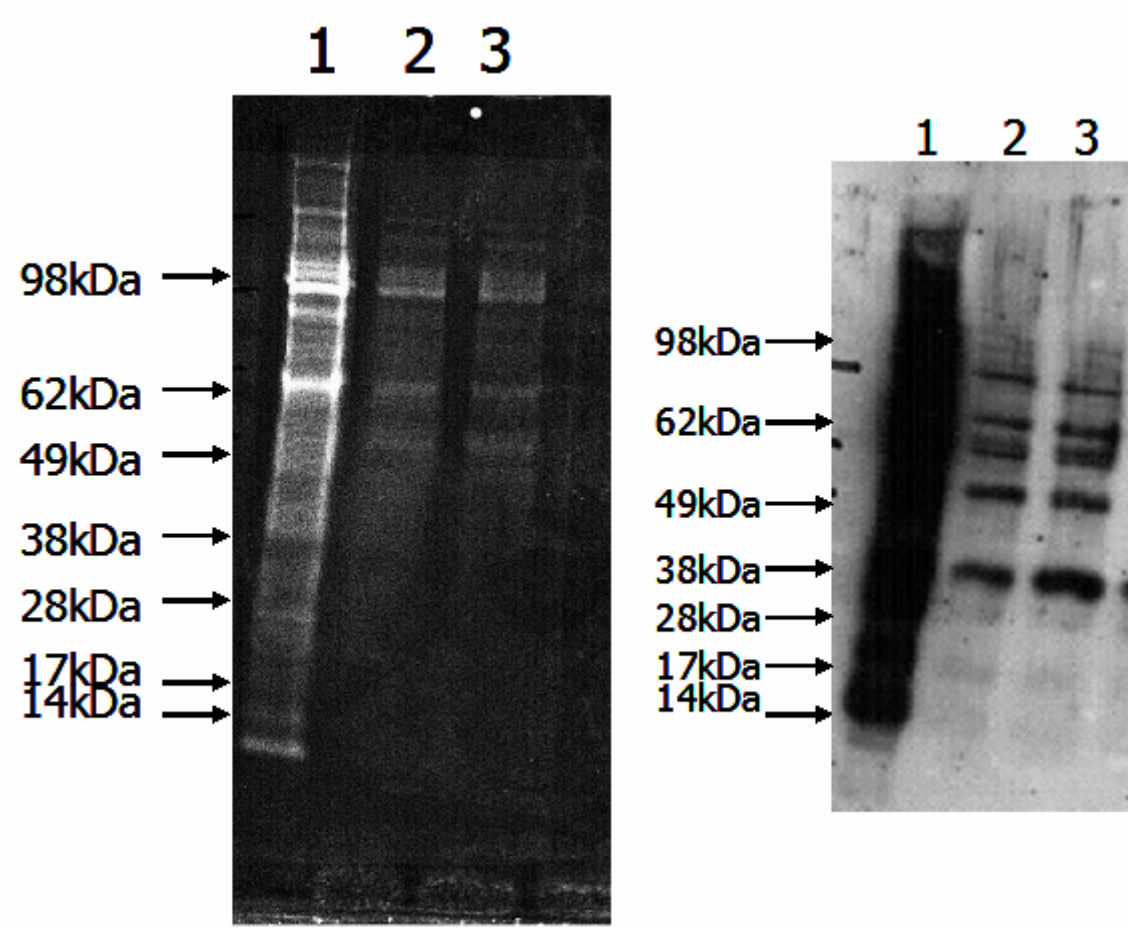


Figure 3.4.1, 3.4.2: Comparison of the efficiency of releasing bound glycoproteins from DBA lectin precipitation by boiling and elution with GalNAc. Proteins were resolved on a 4-12% Bis-Tris gel, and (3.4.1) Sypro stain and (3.4.2) lectin blot was performed as described in Materials and Methods. The membrane was exposed to the film for 15 seconds prior to development. Lane 1: Boiled eluent. Lane 2: First GalNAc eluent. Lane 3: Second GalNAc eluent.

3.5.1 3.5.2

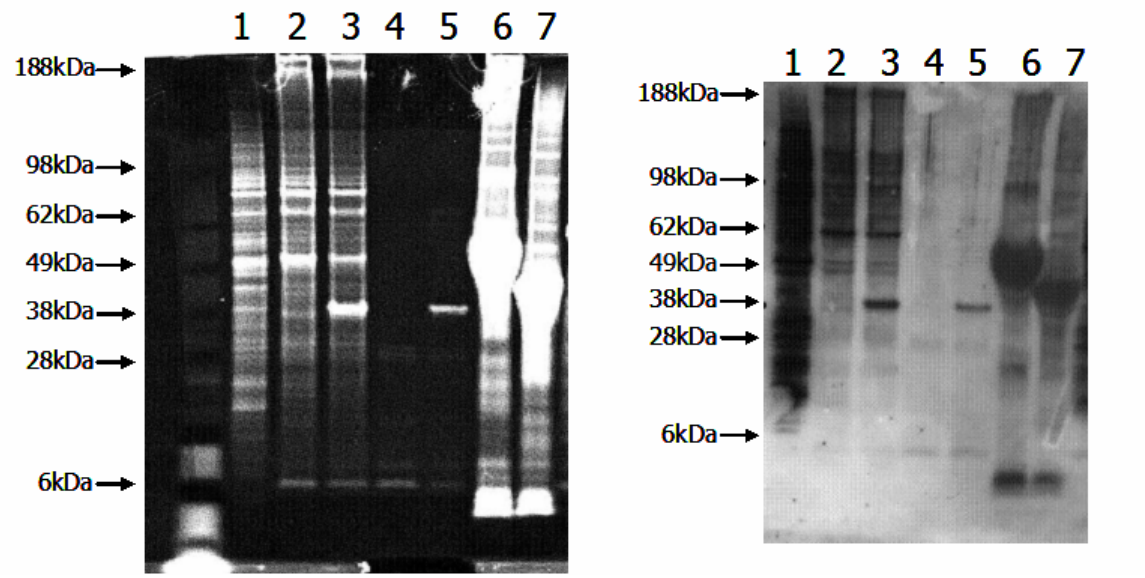


Figure 3.5.1, 3.5.2: Glycan analysis of DBA epitopes by N-glycanase treatment. Proteins were resolved in a 4-12% Bis-Tris gel and (3.5.1) Sypro stain and (3.5.2) lectin blot were performed as described in Materials and Methods. The membrane was exposed to the film for 30 seconds prior to development. Lane 1: mESC lysate. Lane 2: Lectin precipitation (LP) not treated with PNGase F. Lane 3: LP treated with PNGase F. Lane 4: Boiled eluent from untreated DBA agarose beads not treated with PNGase F. Lane 5: Boiled eluent from untreated DBA agarose beads treated with PNGase F. Lane 6: Fetuin not treated with PNGase F. Lane 7: Fetuin treated with PNGase F. The 38kDa band found in lanes 3 and 5 represent the enzyme PNGase F. When comparing lanes 6 and 7, the band shift of fetuin from ~49kDa to ~38kDa is apparent. This shows the sensitivity of Fetuin to PNGase F. No such band shifts are apparent in lanes 2 and 3

CHAPTER 4

DISCUSSION

Nash et. al. suggested that the lectin DBA could be a reliable marker for the pluripotent state of mESC, since DBA is able to discriminate between mESC and primitive ectoderm (EPL), which have lost the ability to colonize blastocyst stage embryos [24]. In addition, DBA could serve as a higher resolution tool for the analysis of stages of differentiation than the previously used markers such as SSEA1, CD9, and FA. Even though the sugar recognized by DBA is also found on FA, DBA epitopes decrease long before those recognized by the FA antibody. This could be due to differences in the affinity of lectins and antibodies for their targets, as the association constant of DBA for GalNAc ligands is around 10^3 mol^{-1} while the association constant for an antibody can range between 10^5 and 10^{12} mol^{-1} [3]. Therefore, DBA should be a more sensitive detector during early stages of differentiation than the FA antibody.

Here we have shown that some of the DBA epitopes are present on glycoproteins (Figure 3.2), which was unexpected as the FA is found only on glycolipids. If the FA antibody and DBA recognize the same epitope then it would be the first instance in which FA is found to be a part of a glycoprotein and not only glycolipid [24].

Glycoproteins were purified by two methods, lectin affinity chromatography (LAC) and lectin precipitation (LP). By LAC with a Biosupport DBA column using [H^3]-glucosamine labeled mESCs it was found that glycoproteins were in fact binding to the DBA column and could be eluted by the addition of 0.2M GalNAc (Figure 3.1.1, 3.1.2). It was then determined that, while both are reasonable methods to purify glycoproteins containing the DBA epitope, LP

was a more efficient method at isolating and eluting DBA bound glycoproteins than LAC with a Biosupport DBA column containing a total of ~9mg of DBA bound to sepharose (Figure 3.3.1, 3.3.2). In addition, by eluting the LP with GalNAc we were able to reduce the appearance of non-specifically bound proteins compared to boiling the LP (Figure 3.4.1, 3.4.2). Affinity chromatography can be susceptible to non-specific interactions, which can be caused by a variety of factors. The use of glycoproteins rather than glycopeptides in this experiment increases the possibility of the formation of hydrophobic aggregates. Also, the nature of the sepharose used in affinity chromatography can create a weak ion exchange column, which would retard negatively charged particles in the column [34].

We considered whether the glycan structures containing the DBA epitope were N- or O-linked, which was addressed by the digestion of DBA precipitated protein with PNGase F, an enzyme that specifically cleaves N-linked but not O-linked glycans. The glycans containing the DBA epitope were found to be O-linked , as they were not sensitive to digestion with PNGase F, unlike the glycoprotein Fetuin (Figure 3.5.1, 3.5.2). Some of the DBA epitopes may be found on N-linked structures, but the abundance of these proteins may be too low for detection by lectin blot.

We believe these results are unlikely to be due to structural inhibition of PNGase F in murine proteins. The ability of PNGase F to cleave N-linked glycans is inhibited by a core α1,3-fucose (Fuc), which is thought to be found in insects and plants. However, the core α1,6-Fuc found in mammalian cells does not affect the activity of PNGase F [36].

The reactivity of PNGase F and Fetuin to DBA seen on the lectin blot (Figure 3.5.2) was likely caused by the high concentration of both proteins present in the reaction mixtures and loaded onto the gel, which is evident on the Sypro Ruby stain (Figure 3.5.1). In addition, Fetuin

contains O-linked glycans with a terminal αGalNAc, which would also contribute to its reactivity to DBA [37].

Future experiments could include β-elimination and hexosaminidase treatment. By utilizing hexosaminidase, all hexose sugars would be removed. Then, by resolving the digested protein by SDS-PAGE and detecting protein with a DBA lectin blot, it would be expected that proteins would lose binding to DBA, showing no reactivity on the blot. In addition, by staining a protein gel, mass changes due to sugar removal should be evident as band shifts.

With β-elimination, the O-linked glycans would be removed, which should further confirm the result that the glycans which bind to DBA are O-linked. After β-elimination, the proteins could be resolved by SDS-PAGE and detected by a DBA lectin blot. If glycans binding to DBA are in fact O-linked, it would be expected that the proteins would lose reactivity to DBA on the lectin blot.

We wished to determine which cell surface proteins contained the DBA epitope. DBA bound glycopeptides were digested with PNGase F and identified by analysis on LC-MS/MS. A total of 5 glycoproteins were identified from the mESCs (Table 3.1). The main function of the majority of these proteins is cell adhesion, which further suggests that these epitopes are stage specific, and could therefore be used as a means of separating various differentiation states. In the future, β-elimination/Michael Addition (BEMAD) will be used to map the locations of and determine the structures of O-glycans involved in this binding, such as those found in fibronectin [33]. It would also be of interest to determine exactly how these structures or their corresponding proteins are regulated upon cellular differentiation. In addition, protein identification studies should be repeated to ensure the accuracy of the identities mentioned here. Once glycoproteins that are retained in the DBA column are confirmed, specific antibodies for

these proteins could be utilized to isolate individual glycoproteins by immunoprecipitation. By resolving the purified glycoprotein on SDS-PAGE and performing a DBA lectin blot on the purified protein, is could be determined whether these glycoproteins contain glycans which bind specifically to DBA, or if they are retained in the column by non-specific interactions.

This study has further demonstrated the usefulness of DBA as a marker for the pluripotent mESC. DBA epitopes were found to be present on glycoproteins, not only glycolipids, and the identities of these glycoproteins were determined. Initial analysis of the glycans containing the DBA epitope showed that they are O-linked rather than N-linked. The techniques developed here could be useful in a variety of investigations utilizing not only mESC, but possibly human ESC as well. Therefore, DBA could be used as a marker to separate pluripotent stem cells for use in a variety of therapeutic methods.

REFERENCES

1. Arita, Y., et al., *Distinct properties of receptors for Dolichos biflorus agglutinin (DBA) isolated from small intestine of adult mice and endoderm cells of early embryos and teratocarcinomas.* Differentiation, 1985. **28**(3): p. 254-9.

2. Varki, A., Cummings, R., Esko, J., Freeze, H., Hart, G., Marth, J., *Essentials of Glycobiology*, ed. C.o.G. Editors. Vol. 2. 1999, La Jolla, California: CSHL Press.

3. Hamelryck, T.W., et al., *Carbohydrate binding, quaternary structure and a novel hydrophobic binding site in two legume lectin oligomers from Dolichos biflorus.* J Mol Biol, 1999. **286**(4): p. 1161-77.

4. Bleil, J.D. and P.M. Wassarman, *Galactose at the nonreducing terminus of O-linked oligosaccharides of mouse egg zona pellucida glycoprotein ZP3 is essential for the glycoprotein's sperm receptor activity.* Proc Natl Acad Sci U S A, 1988. **85**(18): p. 6778-82.

5. Florman, H.M. and P.M. Wassarman, *O-linked oligosaccharides of mouse egg ZP3 account for its sperm receptor activity.* Cell, 1985. **41**(1): p. 313-24.

6. Thall, A.D., P. Maly, and J.B. Lowe, *Oocyte Gal alpha 1,3Gal epitopes implicated in sperm adhesion to the zona pellucida glycoprotein ZP3 are not required for fertilization in the mouse.* J Biol Chem, 1995. **270**(37): p. 21437-40.

7. Youakim, A., et al., *Overexpressing sperm surface beta 1,4-galactosyltransferase in transgenic mice affects multiple aspects of sperm-egg interactions.* J Cell Biol, 1994. **126**(6): p. 1573-83.

8. Shaper, J., *Murine sperm-zona binding: An alpha-3-fucosyl residue is required for the assemply of a high affinity sperm-binding ligand.* Glycobiology, 1998. **8**: p. A51.

9. Williams, S.A., et al., *Fertilization in mouse does not require terminal galactose or N-acetylglucosamine on the zona pellucida glycans.* J Cell Sci, 2007. **120**(Pt 8): p. 1341-9.

10. Solter, D. and B.B. Knowles, *Monoclonal antibody defining a stage-specific mouse embryonic antigen (SSEA-1).* Proc Natl Acad Sci U S A, 1978. **75**(11): p. 5565-9.

11. Kudo, T., et al., *Normal embryonic and germ cell development in mice lacking alpha 1,3-fucosyltransferase IX (Fut9) which show disappearance of stage-specific embryonic antigen 1.* Mol Cell Biol, 2004. **24**(10): p. 4221-8.

12. Kannagi, R., et al., *Stage-specific embryonic antigens (SSEA-3 and -4) are epitopes of a unique globo-series ganglioside isolated from human teratocarcinoma cells.* Embo J, 1983. **2**(12): p. 2355-61.

13. Muramatsu, T., *Developmentally regulated expression of cell surface carbohydrates during mouse embryogenesis.* J Cell Biochem, 1988. **36**(1): p. 1-14.

14. Ozawa, M. and T. Muramatsu, *The glycoprotein-bound large carbohydrates from embryonal carcinoma cells carry receptors for several lectins recognizing N-acetylgalactosamine and galactose.* J Biochem (Tokyo), 1985. **97**(1): p. 317-24.

15. Armant, D.R., H.A. Kaplan, and W.J. Lennarz, *N-linked glycoprotein biosynthesis in the developing mouse embryo.* Dev Biol, 1986. **113**(1): p. 228-37.

16. Muramatsu, T., *Protein-bound carbohydrates on cell-surface as targets of recognition: an odyssey in understanding them.* Glycoconj J, 2000. **17**(7-9): p. 577-95.

17. Sanchez, J.F., et al., *Biochemical and structural analysis of Helix pomatia agglutinin. A hexameric lectin with a novel fold.* J Biol Chem, 2006. **281**(29): p. 20171-80.

18. Clark, G.F., et al., *Decreased biosynthesis of Forssman glycolipid after retinoic acid-induced differentiation of mouse F9 teratocarcinoma cells. Lectin-affinity chromatography of the glycolipid-derived oligosaccharide.* Carbohydr Res, 1991. **213**: p. 155-68.

19. Reubinoff, B.E., et al., *Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro.* Nat Biotechnol, 2000. **18**(4): p. 399-404.

20. Thomson, J.A., et al., *Embryonic stem cell lines derived from human blastocysts.* Science, 1998. **282**(5391): p. 1145-7.

21. Conley, B.J., et al., *Derivation, propagation and differentiation of human embryonic stem cells.* Int J Biochem Cell Biol, 2004. **36**(4): p. 555-67.

22. Burdon, T., A. Smith, and P. Savatier, *Signalling, cell cycle and pluripotency in embryonic stem cells.* Trends Cell Biol, 2002. **12**(9): p. 432-8.

23. Keller, G.M., *In vitro differentiation of embryonic stem cells.* Curr Opin Cell Biol, 1995. **7**(6): p. 862-9.

24. Nash, R., et al., *The lectin Dolichos biflorus agglutinin recognizes glycan epitopes on the surface of murine embryonic stem cells: a new tool for characterizing pluripotent cells and early differentiation.* Stem Cells, 2007. **25**(4): p. 974-82.

25. Gerecht, S., et al., *Hyaluronic acid hydrogel for controlled self-renewal and differentiation of human embryonic stem cells.* Proc Natl Acad Sci U S A, 2007. **104**(27): p. 11298-303.

26. Richards, M., et al., *Human feeders support prolonged undifferentiated growth of human inner cell masses and embryonic stem cells.* Nat Biotechnol, 2002. **20**(9): p. 933-6.

27. Stojkovic, P., et al., *Human-serum matrix supports undifferentiated growth of human embryonic stem cells.* Stem Cells, 2005. **23**(7): p. 895-902.

28. Xu, C., et al., *Feeder-free growth of undifferentiated human embryonic stem cells.* Nat Biotechnol, 2001. **19**(10): p. 971-4.

29. Yang, S., et al., *The design of scaffolds for use in tissue engineering. Part I. Traditional factors.* Tissue Eng, 2001. **7**(6): p. 679-89.

30. Toole, B.P., *Hyaluronan: from extracellular glue to pericellular cue.* Nat Rev Cancer, 2004. **4**(7): p. 528-39.

31. Jessell, T.M. and J. Dodd, *Structure and expression of differentiation antigens on functional subclasses of primary sensory neurons.* Philos Trans R Soc Lond B Biol Sci, 1985. **308**(1136): p. 271-81.

32. Atwood, J.A., 3rd, et al., *Glycoproteomics of Trypanosoma cruzi trypomastigotes using subcellular fractionation, lectin affinity, and stable isotope labeling.* J Proteome Res, 2006. **5**(12): p. 3376-84.

33. van Rooijen, J.J., et al., *Expression of N-linked sialyl Le(x) determinants and O-glycans in the carbohydrate moiety of human amniotic fluid transferrin during pregnancy.* Glycobiology, 1998. **8**(11): p. 1053-64.

34. Besselink, G. and D. de Korte, *Sephadex-based cell-affinity adsorbents: preparation and performance.* Biotechnol Appl Biochem, 2002. **35**(Pt 1): p. 55-60.

35. Tarentino, A.L. and T.H. Plummer, Jr., *Oligosaccharide accessibility to peptide:N-glycosidase as promoted by protein-unfolding reagents.* J Biol Chem, 1982. **257**(18): p. 10776-80.

36. Tretter, V., F. Altmann, and L. Marz, *Peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase F cannot release glycans with fucose attached alpha 1----3 to the asparagine-linked N-acetylglucosamine residue.* Eur J Biochem, 1991. **199**(3): p. 647-52.

37. Edge, A.S. and R.G. Spiro, *Presence of an O-glycosidically linked hexasaccharide in fetuin.* J Biol Chem, 1987. **262**(33): p. 16135-41.