Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Christopher Earl Cutler Date

Role of *Cosmc* and Tn Antigen in T Cell Biology and the O-glycoproteome

By

Christopher Earl Cutler

Doctor of Philosophy

Graduate Division of Biological and Biomedical Science

Immunology and Molecular Pathogenesis

Richard D. Cummings, Ph.D. Advisor	John Altman, Ph.D. Committee Member
Sean R. Stowell, Ph.D. Advisor	Brian Evavold, Ph.D. Committee Member
	Judith Fridovich-Keil, Ph.D. Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Role of *Cosmc* and Tn Antigen in T Cell Biology and the O-glycoproteome

By

Christopher Cutler

B. S. Brigham Young University, 2010

Advisor: Richard D. Cummings, Ph.D.

Advisor: Sean Stowell, Ph.D.

An Abstract of

A dissertation submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In Immunology and Molecular Pathogenesis

Graduate Division of Biological and Biomedical Science

2019

Abstract

Role of *Cosmc* and Tn Antigen in T Cell Biology and the O-glycoproteome

By

Christopher Cutler

T lymphocytes are indispensable to the adaptive immune system and proper execution of their role requires appropriate glycosylation. Despite evidence that O-glycosylation of serine and threonine residues in glycoproteins is dynamically and temporally regulated during T cell maturation and activation, little is known about how specifically O-glycosylation impacts T cells or even which proteins on T cells are modified with O-glycans. Here, I present my findings that extended O-glycans beyond the Tn antigen are important for T cell homing to secondary lymph organs and survival in the periphery. These findings underscore the importance of O-glycosylation in regulating T cell biology and emphasize the need for ongoing attention to glycosylation in studies of T cell development and function. Additionally, I present a method for sensitive identification of Tn-expressing glycosylated O-glycoproteins, providing proof of concept for future targeted proteomics to identify proteins with O-glycans, which will allow for the exploration of the physiological effects of glycosylation.

Role of Cosmc and Tn Antigen in T Cell Biology and the O-glycoproteome

By

Christopher E Cutler

B.S., Brigham Young University, 2010

Advisor: Richard D. Cummings, Ph.D.

A dissertation submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In Graduate Division of Biological and Biomedical Science

Immunology and Molecular Pathogenesis

2019

Table of Contents

List of tables

List of figures

List of abbreviations

293F	HEK293 freestyle cells
AMR	Ashwell-Morell receptor
APC	antigen presenting cell
BSM	bovine submaxillary mucin
C1GalT1	Core 1 β3-galactosyltransferase
CBC	complete blood count
Cosmc	C1GalT1 specific molecular chaperone
DN	CD4-CD8- double negative
DP	CD4+CD8+ double positive
EHCKO	endothelial and hematopoietic conditional Cosmc knock out
ER	endoplasmic reticulum
ETD	electron transfer dissociation
FBS	fetal bovine serum
Fuc	fucose
Gal	galactose
GalNAc	N-acetylgalactosamine
GALT	gut associated lymphoid tissue
Glc	glucose
GlcNAc	N-acetylglucosamine
GnT5	Beta 1,6 N-acetylglucosaminltransferase-V
HCD	high energy collision dissociation
HEV	high endothelial venules
HPA	*Helix pomatia* agglutinin
Hsp	heat shock protein
IL	interleuikin
KO	knock out
LacNAc	N-acetylloctosamine
LC-MS	liquid chromatography linked mass spectrometry
MALT	mucous membrane associated lymphoid tissue
MHC	major histocompatibility complex
MLN	mesenteric lymph node
MS	mass spectrometry
NCAM	neural cell adhesion molecule
Neu5Ac	5-acetylneuraminic acid
NKT	natural killer T cell
PLN	peripheral lymph node
PNA	peanut agglutinin
PS	phosphatidyl serine
PSGL-1	P-selectin glycoprotein ligand-1

RAG	recombination activating genes
RTE	recent thymic emigrants
SC	simple cell
SD	standard deviation
SEC-MALS	size exclusion chromatography with multi-angle light scattering
Sia	sialic acid
SLO	secondary lymphoid organ
SP	single positive
sT	sialyl T
ST6GalNAc1	ST6 N-acetylgalactosaminide alpha-2,6-sialyltransferase 1
sTn	sialyl Tn
TACA	tumor associated carbohydrate antigens
TCKO	T cell conditional Cosmc knock out
TCR	T cell receptor
Tfh	T follicular helper cell
TGF-β	tumor growth factor beta
Th	T helper cell
Treg	T regulatory cell
TSP	thymus-seeding progenitors
VVA	*Vicia villosa* agglutinin
WT	wild type
ZPS	zwitterionic polysaccharides

Chapter 1: Background and Significance

Introduction

T cells are required for defense against pathogens and cancer, as well as the avoidance of autoimmunity. T cells depend on glycoproteins on their surface membranes and in secretions for all stages of their differentiation, activation, homing, and immune functions, and appropriate glycosylation is necessary for these proteins to perform their roles correctly. Glycosylation is an enzymatic process whereby proteins in the endoplasmic reticulum and Golgi apparatus acquire monosaccharides and complex sugar chains linked to specific amino acids, e.g. asparagine (N-glycosylation), and serine/threonine residues (O-glycosylation), through enzymatic reactions involving lipid-linked sugars and nucleotide sugar donors. While glycosylation is dynamically regulated both during T cell maturation and during an immune response, the direct role of glycans in T cell biology remains enigmatic. This is exceptional, as O-glycans in membrane glycoproteins are known to regulate trafficking of polymorphonuclear cells, e.g. neutrophils and monocytes, and activated T cells, in inflammatory responses. As a class, O-glycans in particular participate in cell adhesion and T cell signaling, and examining their role will enhance the understanding of T cell biology. Furthermore, identification of the T cell glycoproteins that contain O-glycans is essential for elucidating the mechanisms behind such mechanisms. Doing so requires an advance in methodology to be able to achieve reasonable sample sizes from primary cells.

In my dissertation studies I have examined the role of O-glycosylation in T cell biology and methods for identifying proteins modified by O-glycosylation. This chapter highlights the current understanding of T cell maturation and the effects of glycosylation on that process, as well as the methods for analyzing glycoproteins. First, T cell maturation and function are

reviewed. Second, the types of protein glycosylation, methods of analysis, and links to human pathology are examined. Third, the known roles of glycosylation in T cell biology are described.

1.1 T cells

T lymphocytes are a crucial element of the adaptive immune system. T cells discriminate between self and non-self through the T cell receptor, they direct the immune response through their secreted cytokines and their cell surface signaling, and they directly eliminate virally infected cells (Abbas, Murphy, and Sher 1996; Jankovic et al. 2002; Bromley et al. 2001; Weninger, Manjunath, and von Andrian 2002). The maturation, homing, activation, and effector function of T cells are each carefully regulated processes, the dysregulation of which can result in immune dysfunction.

1.1.1 T cell maturation

T cells develop from bone marrow precursors that migrate or "home" to the thymus, where they undergo a series of highly regulated developmental stages to produce mature naïve T cells which further mature in the periphery, defined as the blood and blood-containing organs (Zuniga-Pflucker and Lenardo 1996; Carlyle and Zuniga-Pflucker 1998). Early thymic progenitors (ETPs) home to the thymus partly via P-selectin, a C-type lectin expressed constitutively by perinatal thymic endothelium that recognizes poorly defined P-selectin ligand glycans on the glycoprotein PSGL-1 of the ETPs (Zlotoff and Bhandoola 2011; Rossi et al. 2005). Once in the thymus they receive signals from thymic stromal cells to induce the recombination activating genes (RAG) and initiate recombination of the T cell receptor (TCR) genetic loci, generating a unique TCR for each cell (Borowski et al. 2002; Bleul and Boehm 2000). Positive selection checkpoints ensure that only cells productively expressing a rearranged TCR that can engage major histocompatibility complex (MHC) survive and proliferate (Marrack and Kappler 1997; Boehmer 1986; Zinkernagel et al. 1978; Zerrahn, Held, and Raulet 1997).

Subsequently, negative selection occurs in which cells with a TCR that reacts too strongly with self-peptides presented on the MHC are induced to undergo apoptosis, ensuring that autoreactive cells are eliminated from the T cell repertoire, thus preventing autoimmunity (Sprent 1995; Kishimoto and Sprent 1997; Ashton-Rickardt et al. 1994). The major stages of thymocyte maturation are characterized by changes in expression of the co-receptors CD4 and CD8 (Dzhagalov and Phee 2012; Borowski et al. 2002). ETPs express neither and are termed double negative (DN) cells. Cells expressing a TCR and undergoing positive selection express both coreceptors and are termed double positive (DP) cells. Cells that are undergoing or have successfully passed negative selection downregulate one of the two coreceptors and maintain only CD4 or CD8 on their surface, becoming single positive (SP)(Dzhagalov and Phee 2012).

In addition to changes in coreceptor expression, progression through thymocyte maturation is marked by changes in location in the thymus. ETPs enter the thymus at the cortico-medullary junction and migrate to the cortex, passing through four distinct DN stages, characterized by differential expression of RAGs and various surface markers, before becoming DP cells (Bleul and Boehm 2000; Cosgrove et al. 1992; Ueno et al. 2004). After passing positive selection, the cells divide several times to proliferate, then migrate to the thymic medulla as SP cells, where negative selection takes place (Sprent 1995; Webb and Sprent 1990). Cells that successfully pass the checkpoints of positive and negative selection are authorized to egress from the thymus, mediated by the receptor S1P1R (Hogquist et al. 2015; Dzhagalov and Phee 2012). Thymic maturation takes about two weeks to complete, and only approximately 10% of thymocytes successfully pass both quality control check points and enter the peripheral circulation(Merkenschlager et al. 1997; Shortman et al. 1990; Huesmann et al. 1991).

After egress from the thymus, recent thymic emigrants (RTEs) home to the secondary lymphoid organs (SLO), which include spleen, lymph nodes, mucous membrane associated lymphoid tissue (MALT), and gut associated lymphoid tissue (GALT)(Cyster 2000, 1999). These sites provide growth and survival factors, including the critically important chemokine CCL25, also known as the thymus-expressed chemokine or TECK, recognized by its chemokine receptor CCR9 (Cyster 2000). Once in the SLO, T cells within the T cell area await activation by antigen presenting cells, e.g. B cells in the B cell follicles, upon which activation they act to eliminate their targets as well as direct the development and activation of other immune cells (Itano and Jenkins 2003; Kapsenberg 2003).

1.1.2 Function of mature T cells

T cells fill multiple roles in the adaptive immune system. Thymic selection events produce a pool of naïve T cells that recognize non-self antigens. Once activated, these cells are then able to act to directly eliminate infected cells, to secrete signals that direct other immune cells and determine the type and strength of a response, and to inhibit other immune cells and limit the extent of immune activation.

1.1.2.1 Naive T cells

Mature naive T cells circulate through the blood and lymph, surveying antigen presenting cells (APCs) such as dendritic cells until they encounter their cognate antigen (Itano and Jenkins 2003; Thery and Amigorena 2001). In the absence of cognate antigen, the half-life of naïve T cells ranges from 1-12 months (Farber, Yudanin, and Restifo 2014; De Boer and Perelson 2013). Three signaling events are crucial to determining what kind of T cell a naïve T cell develops into upon antigen recognition. Dendritic cells process and present antigens on MHC class I and MHC class II complexes (Gatti and Pierre 2003). When a TCR-MHC interaction is of sufficient

strength and duration, strong TCR signaling results, indicating that the T cell has encountered its cognate antigen. If the dendritic cell collected the antigen in the context of inflammation or other immune-activating molecules such as bacterial lipopolysaccharide, the dendritic cell will also provide secondary signals to authorize the T cell for immune activation (Bour-Jordan and Blueston 2002). Tertiary signals from APCs and the tissue microenvironment provide additional information on the context and type of antigen and direct the T cell towards expression of a particular set of effector molecules for a given type of response, determining which subtype of effector T cell it will become (Kapsenberg 2003; Murphy and Reiner 2002). In the absence of secondary or tertiary signals, T cells encountering their cognate antigen become either non-responsive or anti-inflammatory regulatory T cells (Arnold 2002; Fantini et al. 2004). Effector T cells home to sites of inflammation and extravasate via high endothelial venules (von Andrian and Mempel 2003; Rosen 2004). Once in the target tissue, the cells perform their designated function, as described below, producing a mixture of cytokines and other effector molecules to direct other immune cells and/or kill infected target cells. Effector T cells have a short half-life of 28 days (Chu et al. 2016).

1.1.2.2 Effector T cells

Effector functions of T cells are divided first by whether they express CD8 or CD4, which restrict cells to the recognition of antigens presented on class I or class II MHC, respectively, then by the secondary and tertiary cytokine signals the cells receive, directing them to differentiate in response to the immune context of their activation.

Effector CD8 T cells predominantly function as cytotoxic cells, targeting cells that have been infected by a virus and are producing foreign proteins (Harty, Tvinnereim, and White 2000). Recognition of these foreign proteins occurs through presentation of endogenously

produced peptides on MHC-1 molecules, which are expressed on all cell types (Williams, Peh, and Elliott 2002). Engagement of MHC-I with TCR on CD8 T cells provides a survey of the proteins being synthesized in each cell. Effector cytotoxic CD8 T cells that encounter their cognate antigen recognize the target cell as infected and deliver contact signals and a cocktail of secreted molecules to kill the target cell (Barry and Bleackley 2002). Signaling via Fas/FasL protein induces apoptosis, and perforins create holes in the target membrane through which granzymes enter and act to proteolytically degrade targets and induce caspase activation (Russell and Ley 2002; Smyth et al. 2001).

Effector CD4 T helper (Th) cells direct the immune response. Broadly speaking, Th cells provide activating cytokines to other immune cells, including CD8 cytotoxic T cells, B cells, and granulocytes. CD4 T cells can be directed by combinations of cytokines to differentiate into several subtypes of Th cell, including Th1, Th2, Th17, and T follicular helper (T_{FH}) (Santana and Rosenstein 2003). The signals that induce each subtype, together with the transcription factors each subtype expresses, the effector molecules produced, and the general function exerted are summarized in Table 1.1.

In contrast to the Th cells described above, which act to enhance an immune response and direct it towards an appropriate profile of effector cell types, CD4 T regulatory (Treg) cells function to limit immune responses and prevent damage from excessive inflammation (Sakaguchi and Sakaguchi 2005). Unlike other T cells which mature into their final effector phenotype when they encounter their cognate antigens in the periphery, Tregs can be produced in the thymus or the periphery (Fantini et al. 2004; Roncarolo et al. 2001). Tregs are produced in the thymus from cells with an intermediate strength of TCR signaling in response to self-antigens (Arnold 2002). Tregs can also be induced in the periphery by dendritic cells presenting their

cognate antigen together with tumor growth factor-beta (TGF-β) (Roncarolo, Levings, and Traversari 2001). Tregs are characterized by the expression of the transcription factor Foxp3 and the production of anti-inflammatory cytokines IL-10 and TGF- β (Roncarolo et al. 2001). These factors influence CD4 and CD8 T cells, macrophages, and dendritic cells to reduce their activation state and limit effector functions.

1.1.2.3 Memory T cells

After resolution of an immune response, some effector CD4 and CD8 T cells transition into a long-lived memory phenotype (Ku et al. 2000; Seddon, Tomlinson, and Zamoyska 2003). These cells persist with a half-life of 1-8 years (De Boer and Perelson 2013) and recirculate through the blood and lymph. Memory cells that encounter their cognate antigen a second time do not require secondary signals like naïve cells do to become activated once again to proliferate and perform their effector functions; TCR engagement is sufficient to induce a robust response by memory T cells (Sallusto, Geginat, and Lanzavecchia 2004). This allows for the persistence of immunological memory to antigens and infectious agents across the lifetime of an organism.

1.2 Glycoproteins

Post-translational modification of proteins with carbohydrates is common and most proteins in the secretory pathway co-translationally acquire Asn-linked oligosaccharides or N-glycans in the ER and most also acquire Ser/Thr-linked oligosaccharides or O-glycans in the Golgi apparatus (Kristic and Lauc 2017). Glycosylation is dramatically varied and its presence, composition, and regulation is physiologically relevant in a wide variety of contexts (Moremen, Tiemeyer, and Nairn 2012; Spiro 2002; Lee, Qi, and Im 2015). Perturbation or dysregulation of glycan biosynthesis or processing is associated with many types of pathology, ranging from congenital

disorders of glycosylation to cancer (Ohtsubo and Marth 2006; Pinho and Reis 2015; Jaeken 2011).

1.2.1 Protein glycosylation

Protein glycosylation is critical to protein and cellular function. Glycosylation ranges from the addition of a single monosaccharide to the installation of large branching structures typified in the glycocalyx of cells that can span hundreds of nanometers and are comprised of hundreds of monosaccharide residues (Varki 2017). Over half of the proteome is glycosylated (Apweiler, Hermjakob, and Sharon 1999). Protein glycosylation is important for a wide range of physiological processes, including protein localization, cell-cell interaction or cell-matrix interaction, cell signaling, the immune response and development (Varki 1993; Merry and Merry 2005). Some highly decorated proteins can be as much as 80% carbohydrate by weight and can have oligosaccharides hundreds of nanometers long (Bansil, Stanley, and LaMont 1995). These ubiquitous modifications are installed and modified by glycosyltransferase enzymes which transfer a sugar from an activated donor unto a substrate protein, forming a glycosidic bond, and by glycosidases that hydrolyze those bonds(Lairson et al. 2008). Glycosyltransferases and glycosidases are ubiquitously expressed (Paquette, Moller, and Bak 2003).

In mammals the predominant types of glycosylation on secreted proteins are branched N-glycans, typified by N-acetylglucosamine (GlcNAc) installed on an asparagine residue, and the subset of O-glycans typified by N-acetylgalactosamine (GalNAc) installed on serine or threonine residues (Ohtsubo and Marth 2006). A wide variety of less common glycosylation types also occur on secreted proteins (Spiro 2002). These post translational modifications are relatively stable and persistent. In addition, intracellular proteins are frequently modified on serine or

threonine residues with GlcNAc, which is a much more dynamic process with swift addition and removal of glycan substituents comparable to phosphorylation (Nagel and Ball 2014).

N-glycosylation occurs in the endoplasmic reticulum (ER) co-translationally (Trombetta 2003) and is important for proper folding, for engagement with trafficking proteins, and for protein-protein interactions. N-glycosylation is initiated by addition of a preformed tetradecasaccharide from the dolichol pyrophosphate oligosaccharide donor (Weerapana and Imperiali 2006). Sites of N-glycosylation have a consensus sequence of Asn-X-Ser/Thr, where glycosylated asparagine is followed by a serine or threonine, separated by a single residue that cannot be proline (Weerapana and Imperiali 2006). Subsequent action by glycosidases and glycosyltransferases trim and extend the glycan in a cell- and substrate-specific manner. Typical N-glycans are large (2-6 kDa) and contribute an appreciable volume to any glycoprotein, as an average glycoprotein contains 3-5 N-glycans on different Asn residues(Apweiler, Hermjakob, and Sharon 1999).

There are multiple types of O-glycosylation, in which the initial sugar residue is attached to serine or threonine. By far the most common type of O-glycosylation in the secretory pathway is GalNAc added through O-glycosylation, sometimes referred to as mucin-type O-glycosylation for the high abundance of this type of glycosylation on mucin proteins (Van den Steen et al. 1998). Unless otherwise specified, the subsequent use of the term “O-glycosylation” will be assumed to refer to GalNAc type O-glycosylation.

While the pathways regulating N-glycosylation are very well described many mysteries remain about the mechanisms of O-glycosylation. In contrast the co-translational addition of N-glycans, O-glycosylation occurs post-translationally in the Golgi apparatus or late ER (Spiro 2002; Rottger et al. 1998). No consensus sequence is known, which is consistent with the

catalyzing enzymes acting on folded proteins rather than linear peptide chains (Van den Steen et al. 1998; Wandall et al. 2007). Glycoproteomic cataloguing of O-glycosites has led to the development of increasingly accurate prediction algorithms. These have revealed that O-glycosylation is exceedingly common, with approximately 80% of proteins that enter the secretory pathway containing at least one predicted O-glycan (Bennett et al. 2012; Steentoft, Vakhrushev, Vester-Christensen, Schjoldager, et al. 2011).

The process of building O-glycans is highly regulated and has several steps (Van den Steen et al. 1998). The first step is the transfer of a GalNAc to a serine or threonine by one of the polypeptide N-acetylgalactosaminyltranserases that utilize UDP-GalNAc as a donor (Berger 1999; Brockhausen 1999). This family of enzymes contains ~20 enzymes (20 for mice, 22 for human), each with their own substrate preferences and differential expression across cell types (Wandall et al. 2007). The product of their reaction is the monosaccharide GalNAc attached to serine or threonine, and this is commonly recognized as the Tn antigen. The second step, addition of beta-3-linked galactose (Gal) is catalyzed by single enzyme, the T-synthase (Ju and Cummings 2002). This results in the disaccharide known alternately as the T antigen and as Core 1, which is the basis for nearly all physiologically observed O-glycans. A variety of enzymes act subsequently to elaborate Core 1 into different target structures (Ju, Otto, and Cummings 2011). Due to the combinatorial nature of branching oligosaccharides, the number of available structures that can be formed is immense, with the specific structures generated depending on the expression and regulation of the various glycosyltransferases. (Schachter and Brockhausen 1989). However, the Tn antigen is the common precursor to all of them (Brockhausen 2000) primarily by way of Core 1, which makes T synthase critical for the formation of nearly all GalNAc-type O-glycans.

1.2.1.1 T-synthase

The T-synthase enzyme, also called Core 1 β3-galactosyltransferase, (C1GalT1) is responsible for converting Tn antigen (αGalNAca1-Ser/Thr) into the disaccharide known as T antigen or core 1 (Galβ1-3GalNAcα1-Ser/Thr) (Ju and Cummings 2002). T-synthase is ubiquitously expressed in humans and is an essential enzyme in mammals, as knockout of the T-synthase gene in mice is embryonic lethal (Xia et al. 2004). While many other glycosyltransferases share high homology and in some cases are functionally redundant, T-synthase has limited homology with other galactosyltransferases (Ju, Zheng, and Cummings 2006), and no functionally redundant enzymes or family members have been found to perform the same reaction. T-synthase function is critical for all animal cells studied, and orthologs are conserved from *C. elegans* to *Homo sapiens*. Indeed the *C. elegans* and *Homo sapiens* have greater than 40% identity despite the great evolutionary distance between them (Ju, Zheng, and Cummings 2006). A key evolutionary difference observed between T-synthase orthologs is a differential requirement for the C1GalT1 specific molecular chaperone (Cosmc). Without concomitant expression of Cosmc, T-synthase orthologs expressed in vertebrates do not fold correctly and fail to achieve an active form. The T-synthase in vertebrates is not glycosylated and requires Cosmc for folding, while T-synthase orthologs in invertebrates are glycoproteins and do not require Cosmc to achieve correct folding (Ju and Cummings 2014).

T-synthase acts in the Golgi apparatus to transfer Gal from UDP-Gal in β3 linkage onto the initial GalNAc (Ju and Cummings 2002). Active T-synthase functions as a disulfide-bonded homodimer, with substituent members connected through the C-terminal transmembrane domain. The active site of the enzyme resides in the N-terminus and utilizes the DXD motif common to many glycosyltransferases for coordination of a manganese cation which serves to

break the phosphoglycosidic bond of the donor for transfer to the accepting GalNAc (Ju and Cummings 2002; Gloster 2014).

1.2.1.2 Cosmc

Protein chaperones are responsible for helping eukaryotic proteins fold correctly into their active form. These include heat shock proteins (HSPs), lectin based chaperones, cochaperones, and protein specific chaperones (Pearl and Prodromou 2006). HSPs are the most common type of general chaperone, responsible for helping nascent proteins to fold correctly (Fink 1999). HSPs function by binding hydrophobic stretches of amino acids until they can form protected hydrophobic domains (Blond-Elguindi et al. 1993). Lectin based chaperones help to fold N-glycosylated proteins and prevent them from leaving the ER before they are fully glycosylated by binding to N-glycans on nascent glycoproteins (Trombetta 2003). Lectin chaperones retain monoglycosylated N-glycans in the ER until they are correctly folded and modified (Trombetta 2003).

Some proteins require specific chaperones in addition to HSPs and other general chaperones (Hendershot and Bulleid 2000; Braakman and Bulleid 2011). Their mechanism of action is not uniform or well understood. C1GalT1 specific molecular chaperone (Cosmc) is a specific chaperone for T-synthase (Ju and Cummings 2002), and can bind to unfolded T-synthase but not to the fully mature active enzyme. Cosmc is an example of a 'private' chaperone, dedicated to a single client, and while there are many private chaperones, there are also general or 'public' chaperones that can assist folding of many glycoproteins (Anelli and Sitia 2008).

Cosmc bears 15% homology to its client T synthase and is the result of gene duplication via retrotranslocation. The *Cosmc* gene is located on the X chromosome with a single exon (Ju and Cummings 2002). Cosmc protein localizes to the ER and, in the absence of T synthase,

exists as a disulfide-bonded oligomer and has no enzymatic activity (Ju et al. 2008). Several different naturally occurring mutations of Cosmc have been described, most from tumor cell lines but also from hematopoietic cells of patients with Tn syndrome, including truncations and substitutions (Ju et al. 2008; Yu et al. 2015; Sun, Ju, and Cummings 2018). These, together with truncation experiments with recombinant protein, have helped to identify the functional regions of Cosmc (Aryal, Ju, and Cummings 2010, 2012). Oligomerization and metal-binding occurs in the C-terminal domain, while chaperone function for T-synthase is carried out by the N terminal domain (Hanes, Moremen, and Cummings 2017). Of particular note is the E152K substitution, which has the same *in vitro* activity as WT Cosmc, but causes loss of function *in vivo* (Hanes, Moremen, and Cummings 2017). This provides a key handle toward unraveling the interactions and mechanisms of action, and the cellular regulation that occurs in this pathway *in vivo*. Multiple cell lines harbor natural or engineered mutations in Cosmc, with resultant expression of the Tn antigen. Among these, Jurkat cells and LSC cells are of particular note for their widespread utility in studying the biology of T cells and colon cancer, respectively (Ju et al. 2008; Ju and Cummings 2002). Further research into these functions should yield important insights.

1.2.2 Detecting glycoproteins

Post translational modification of proteins with complex glycan structures plays a crucial role in a wide variety of physiological processes. Central to understating the mechanisms involved is identification of which proteins are glycosylated and at which sites, as well as structure of the glycan involved. Unlike proteins and nucleic acids, glycans are neither linear or templated, so determination of glycan structures requires confirmation of linkage and

connectivity in addition to composition and sequence. Multiple tools and techniques are available for answering these questions, and recent advances have expanded abilities in this area(Cummings and Pierce 2014). However, structural elucidation of glycan substituents and their attachment to glycoconjugates remains a nontrivial problem.

Detection and characterization of glycoproteins is accomplished by several complementary techniques (Roth, Yehezkel, and Khalaila 2012). Chemical detection allows for the determination of carbohydrate presence and quantity, lectin-based detection identifies particular structural motifs based on binding modalities, and mass-spectrometry enables determination of the composition and site occupancy of specific glycans or glycoproteins.

Carbohydrate content in a sample can be confirmed via chemical detection or by chemical based stains(Mulloy et al. 2015). Because many carbohydrates contain vicinal diols they are susceptible to oxidation by periodic acid. The resultant aldehydes can react with hydrazide or amine-based probes to covalently label the target glycan, which can then be visualized. This technique can be applied to soluble samples, acrylamide gels, and *in situ* staining (Bond and Kohler 2007). Carbohydrates that have been released and have a free reducing end can be detected with chemical reaction with orcinol and sulfuric acid, which allows for estimation of carbohydrate content(Bond and Kohler 2007). Compositional analysis of a glycan, which is used to determine the relative amounts of substituent monosaccharides, is performed by mass spectrometry on the native glycan, or LC-MS of the hydrolyzed glycans and released monosaccharides, and by liquid chromatography of acid-released monosaccharides that have been labeled with a fluorophore (Pabst and Altmann 2011).

Lectins are proteins that bind to carbohydrates and are classified by structural similarities into at least 20 different families (Taylor et al. 2015). There are thousands of lectins that

collectively recognize an immense variety of structures, and over 100 different carbohydrate binding modules (CBMs) are known in proteins of all types (Consortium 2018). The comparative binding of different lectins allows for an assessment of the presence or absence of structural determinants in a sample, and lectin blots and affinity isolation can be used to great effect (Hirabayashi 2004). Due to their relative ease of isolation, many of the analytically useful lectins are plant-derived, often produced in quantity by plants as passive defenses, however, animal and fungi-derived lectins are in use as well (Li et al. 2011). While the number of well-characterized lectins is large, some of the more commonly used lectins and their recognized ligands are listed in Table (1.2). Determinants recognized by lectins are typically characterized by hapten inhibition assays which identify the soluble oligosaccharides capable of preventing ligand binding (Hirabayashi 2004).

Lectins are utilized in blotting applications, affinity chromatography, and microarray assays to isolate and identify ligands based on carbohydrate presence and structure. While the affinity of a lectin for its specific target can be high avidity especially when present in multimeric structures, lectins frequently are used in low affinity applications to recognize a class of more weakly bound structures (Hirabayashi 2004; Hirabayashi, Hashidate, and Kasai 2002). However, the fine specificity of a lectin is frequently unknown and a lectin may be misleading by binding to multiple similar structures that do not necessarily contain the same monosaccharide constituents. It is usually not possible to determine a glycan structure solely through lectin binding and chemical analysis (Pabst and Altmann 2011).

Nuclear magnetic resonance (NMR) imaging is a complementary technique that has the sensitivity to determine anomeric centers and connectivity, and is thus extremely useful in situations where it is feasible. (Lundborg and Widmalm 2011; Mulloy et al. 2015).

Unfortunately, the analysis of data from a complex oligosaccharide is nontrivial, and the stringent requirement for pure homogeneous samples is restrictive compared to other methods of analysis. Thus, NMR is not appropriate or very useful for mixtures or poorly characterized samples.

Mass spectrometry (MS) is a powerful method for elucidation of glycan structure as well as determination of site occupancy of glycoproteins, especially with recent advances in instrumentation and data analysis (Cummings and Pierce 2014; Dalpathado and Desaire 2008). Nevertheless, even with high accuracy and high sensitivity instrumentation, glycans have multiple features that present challenges for such analysis. First, many monosaccharides are isomeric forms of the same chemical structure distinguished by their stereochemistry. Thus, a determination of mass by MS can identify a monosaccharide as a hexose but not differentiate between glucose, galactose, or other isobaric isomers. This difficulty is exacerbated as the complexity of a glycan increases. Glycosidic linkages can occur at alternate branching positions of each substituent sugar, creating a host of potential isobaric species that are not immediately distinguishable. Fortunately, particular linkages produce specific diagnostic ions when analyzed by tandem MS, so careful method design and comparison to known standards allows one to extract linkage information from MS data(Hsiao et al. 2017; Veillon et al. 2017). Second, carbohydrates ionize poorly in their unmodified form. Chemical modification such as permethylation greatly increases the sensitivity of MS detection (Veillon, Zhou, and Mechref 2017). It should be noted, however, that such treatment eliminates possible modification such as sulfation that can occur on glycans, and those modifications must be tested for separately. Ionization methods suited for proteomics such as high energy collision dissociation (HCD) tend

to destroy the glycan without generating useful ions, so other methods such as electron transfer dissociation (ETD) with low energy ionization are required (Han and Costello 2013).

In addition to instrumentation and methodological challenges, the branched structures of glycans presents a much more computationally intensive problem than the linear peptides generally encountered in proteomics. Improvement of computer algorithms for searching and deconvolution of glycan MS data has enabled a much more powerful analysis of the combinatorial possibilities of glycan substituent ions. These advances have recently been reviewed (Veillon, Zhou, and Mechref 2017; Hong et al. 2017).

Despite improvements in MS based glycomic analysis, the combined analysis of glycoproteins remains a difficult problem, as ionization of proteins and their glycan modifications necessarily require different protocols and sample handling. Even identification of the sequence and composition of both the glycan and aglycan components of a modified protein does not allow unambiguous assignment of specific glycan structures to particular modification sites. This is especially problematic in mixed samples as compared to purified proteins.

Glycoproteomics has also recently advanced in identifying modification sites across the proteome. This is pioneered by the field of O-GlcNAcylation where a large number of nuclear factors and other cytosolic proteins are dynamically modified with O-GlcNAc (Nagel and Ball 2014; Khidekel et al. 2003; Khidekel et al. 2004). Affinity isolation of such proteins using lectin WGA followed by proteolysis and MS analysis has identified many sites of modification and helped generate algorithms for predictions of O-GlcNAcylation. This approach benefits from the homogenous nature of the monosaccharide O-GlcNAc in this type of modification and therefore is not directly applicable to analysis of other glycosylation. However, the utilization of *Cosmc* knockout cell lines termed Simple Cells (SC) in which the O-glycans are homogeneously

truncated to the Tn antigen by deletion has allowed for the use of *Cosmc*, has allowed for the use of an analogous approach using VVA lectins(Steentoft, Vakhrushev, Vester-christensen, and Schjoldager 2011; Steentoft, Vakhrushev, Vester-Christensen, Schjoldager, et al. 2011; Steentoft C 2013). This advance has greatly expanded our understanding of O-GalNAc glycosite determination and prediction. Unfortunately, this method requires utilization of large sample amounts and is therefore not applicable to primary tissues or patient samples.

1.2.3 Glycoprotein linked pathology

Glycosylation plays a central role in many physiological processes and its dysregulation or disfunction result in a variety of pathologies. For the purposes of this dissertation, the focus will be limited to disfunction of the O-glycosylation pathway and changes that arise in cancer. These include Tn syndrome and the emergence of Tn and sTn in tumor cells and the appearance of Tn antigen in IgA nephropathy. The role of glycans in genetic and acquired human diseases as well as the immunological targeting of glycans as antigens are reviewed elsewhere.

1.2.3.1 Tn syndrome

Tn syndrome is a rare blood disorder in which the Tn antigen is expressed on patient blood cells of all lineages. The syndrome was identified in 1959 by Dausset et al who identified a 65-year old patient whose freshly collected erythrocytes exhibited unusual agglutination properties termed polyagglutinatibility, since they seemed to be agglutatinated by all donor sera tested regardless of blood group of the donor, as well as being better agglutinated at 4°C compared to 37°C (Dausset, Moullec, and Bernard 1959). Interestingly, this patient was identified with acquired hemolytic anemia after an infection of unknown origin. The patient exhibited leukopenia, thrombocytopenia and a chronic hemolytic anemia. While the antigen was not molecularly characterized, it was termed the T nouvelle or Tn antigen because of its close

relationship but non-identity with the previously identified T antigen. The hint that Tn syndrome was associated with an unusual and novel glycan antigen was provided by Gunson et al who found that lectins binding GalNAc seemed to better agglutinate Tn erythrocytes than normal erythrocytes (Gunson, Stratton, and Mullard 1970). Dahr et al identified the Tn antigen as GalNAcα1-Ser/Thr on glycoproteins (Dahr, Uhlenbruck, and Bird 1974).

Tn syndrome itself can present pathologically, but because it is rare it is not routinely screened for and the baseline incidence is not known (Berger 1999). Our laboratory discovered that Tn syndrome is caused by somatic mutation of *Cosmc* in hematopoietic stem cells (Ju 2005). It may also possibly arise through inappropriate silencing of the *Cosmc* gene (Berger 1999; Ju and Cummings 2002; Xia et al. 2004). In any case, *Cosmc* is lost in a subset of the hematopoietic stem cell population and results in a heterologous population of cells, i.e. mosaicism, depending on the proportion of stem cells affected by the mutation. The proportion of cells bearing Tn antigen varies patient to patient, with up to 90% of erythrocytes and lower percentages of leukocytes being affected (Berger EG 1994; Brouet et al. 1983; Judson et al. 1983). In some individuals with Tn syndrome there is a mild pathology, which is interesting because there is evidence from animal models that various blood cells are adversely affected by loss of *Cosmc* (Wang et al. 2010; Wang et al. 2012). Presumably, in humans, the deleterious effects are limited to the affected cells and are not dominant phenotypically, as the patients are mosaic and have a large proportion of normal blood cells of all lineages, which allows unaffected stem cells to compensate and overcome any deficiencies. Yet the evidence that Tn syndrome can be associated with leukopenia, thrombocytopenia, or chronic hemolytic anemia (Berger 1999), suggests that the compensatory effect may be incomplete. These indications invite a closer investigation of the role of O-glycosylation in blood cell biology. The pattern of greater

incidence of Tn antigen expression on some cell types over others in Tn syndrome patients suggests that certain blood lineages are more dependent on Cosmc than others. Importantly, in such comparisons T cells are consistently lowest in incidence of Tn expression compared to other blood cells in the patients, indicating that they may be the lineage most susceptible to loss of *Cosmc* (Brouet JC 1983; Judson PA 1983; Berger EG 1994). T cells are therefore a prime model for understanding the physiological importance of O-glycosylation.

1.2.3.2 Cancer

Altered glycosylation in cancer is a burgeoning area of study because of the profound and potentially diagnostic glycolytic changes that occur with oncogenesis (Peracaula et al. 2008; Hakomori and Cummings 2012; Meany and Chan 2011). Tumors frequently have altered glycoprofiles relative to normal tissue, and glycosylation plays multiple roles in oncogenic transformation. For example, the galectin family of carbohydrate binding proteins have been implicated in a multiple specific tumor processes, occasionally in contradictory ways (Bartolazzi 2018; Chetry et al. 2018). Galectins differ in their fine specificity but the family as a whole recognize N-acetyllactosamine (LacNAc) repeats, and alterations in glycosylation result in differential binding of galectins to tumor vs normal cells (Danguy, Camby, and Kiss 2002). Multiple galectin family members promote angiogenesis in tumors, (Funasaka, Raz, and Nangia-Makker 2014; Troncoso et al. 2014; Thijssen and Griffioen 2014) or affect cell cycle progression (Fischer et al. 2005; Liu and Rabinovich 2005), but galectins 1 and 8 in particular enable disaggregation and promote metastasis, which enables tumor progression and propagation but also renders disaggregated cells more susceptible to death by anoikis (Rizqiawan et al. 2013; Sanchez-Ruderisch et al. 2011; Levy et al. 2001). In contrast, galectin 3 promotes aggregation and prevents anoikis in tumors(Zhao et al. 2010; Sanchez-Ruderisch et al. 2011). Alteration in

other types of glycosylation have similarly disparate effects on cancer development and progression. While there is no single unifying effect of abnormal glycosylation of cancerous cells, it is clear from these many examples that glycosylation is disrupted in many forms of cancer and that this disruption has profound impacts on disease progression and prognosis.

Among the abnormal glycosylation types known as tumor associated carbohydrate antigens (TACAs), Tn and Sialyl Tn (sTn) are of particular interest because of their extremely wide expression on solid tumors and their complete absence from normal tissues. Tn and sTn are expressed by more than 80% of human carcinomas, and are associated with poor prognosis in 24 different cancers (Ju et al. 2008; Springer 1984), including cervical, lung adenocarcinoma, breast and gastric and colorectal carcinomas (Numa et al. 1995; Laack et al. 2002; Fernandez et al. 2005; Kakeji et al. 1991; Konno et al. 2002). Accordingly, these antigens could be useful biomarkers for diagnosis of carcinomas, and efforts to develop them as such are currently underway. Multiple attempts have been made to utilize Tn and sTn antigens as tumor vaccines, with the aim of activating a patient's immune system against the tumor, but no reliable induction of anti-Tn and anti-sTn antibodies has yet been achieved (Jin et al. 2019; Julien, Videira, and Delannoy 2012; Heimburg-Molinaro et al. 2011). Moreover, in our hands the culture of hybridomas producing anti-Tn antibodies has been uniquely fraught, with abnormally high rates of hybridoma loss, suggesting that the act of synthesizing anti-Tn antibody may be situationally detrimental to the producing cell (unpublished observations). Taken together, these difficulties hint that these particular carbohydrate antigens may be poor targets for the immune system, perhaps explaining why tumors succeed and progress while expressing them. Beyond immune evasion, it seems probable that Tn and sTn are exerting a direct benefit to the cells that express them, possibly through altering the function of specific O-glycoproteins. Characterizing which

proteins bear Tn or sTn in a tumor context is an important avenue for investigating tumor biology, but the altered glycosylation pathways of tumor cells make the current predictive algorithms for O-glycosite modification less reliable, and direct detection is necessary. This is a difficult problem, as current techniques require large numbers of cells, and are therefore infeasible for patient samples.

In contrast to solid tumors, in which Tn and sTn are very common, these antigens have only rarely been detected in leukemias, and never in lymphomas (Ju et al. 2008; Roxby et al. 1992; Fu et al. 2016). This discrepancy is striking and suggests that blood cell biology is uniquely reliant on an intact O-glycosylation pathway. If one could discover the underlying mechanism by which Tn expressing lymphomas are disadvantaged, key aspects of lymphocyte biology might be illuminated, and routes to therapeutic treatment of leukemias might be opened.

1.2.3.3 Other pathologies

O-glycosylation has been implicated in additional human diseases including IgA nephropathy and Henoch Schonlein Purpura. IgA nephropathy is the most common type of gomerulonephritis and causes renal failure (Allen, Harper, and Feehally 1997; Allen et al. 1997). It is characterized by the deposition of immune complexes containing IgA and IgM in the kidney. The cause remains unclear, but studies have correlated IgA nephropathy with alternate O-glycosylation of the hinge region of IgA (Lehoux et al. 2014; Novak et al. 2011). It has been shown that healthy humans have distinct species or ‘glycoforms’ of IgA which bear either T or Tn on the hinge region O-glycosites (Lehoux and Ju 2017). Total IgA levels including levels of the Tn bearing IgA species are elevated in IgA nephropathy patients; this elevation likely contributes to the formation of the pathological immune complexes (Lehoux et al. 2014). It remains unclear whether the Tn-bearing IgA arises from a distinct process or from a general

deficiency of the O-glycosylation pathway antibody secreting plasma cells. Research in this area is ongoing and understanding the role of O-glycosylation in IgA nephropathy promises to lead to improved therapies for patients with this disease.

Kidney deposition of IgA immune complexes also occurs frequently in patients with Henoch-Schönlein Purpura, an autoimmune disease that target blood vessels and induces vasculitis(Saulsbury 2007). Histologically these depositions are indistinguishable from IgA nephropathy hinting as a possible shared mechanism in the two conditions (Allen et al. 1998; Saulsbury 1997; Nakazawa et al. 2019).

1.3 T cell glycosylation

Glycosylation plays an important role in at least four areas of T cell biology: differentiation, homing, antigen recognition, and cell signaling events. The first two categories involve the interaction of proteins expressed by T cells interacting with glycans on other cells, while the second two involve glycans expressed on the T cells themselves. The specific effects of glycosylation on T cell biology is an area of active inquiry and many open questions remain. Thus far it is clear that glycosylation plays an indispensable role in T cell lineage commitment, progression through the stages of thymus maturation, and functionality of mature T cells. Recent discoveries have advanced the field's understanding of this area of T cell biology, as reviewed recently by Pereira et al(Pereira et al. 2018), but a great deal remains unknown about the impact of glycosylation on T cell interactions.

1.3.1 Lineage commitment and thymic selection

Glycosylation is a crucial regulator of thymocyte differentiation and impacts development and selection at every developmental stage. Alterations of specific glycans block or modulate thymocyte development, resulting in reduced thymopoiesis or in an altered immune

repertoire. This can directly impact the capacity of the immune system to respond to disease, its ability to eliminate cancerous cells, and its propensity for autoimmunity

Seeding of the thymus by bone-marrow derived cells begins with homing of thymus-seeding progenitors (TSPs) to the thymus(Dzhagalov and Phee 2012). This process depends on the interaction of P-selectin on the thymic epithelium with its ligand P-selectin glycoprotein ligand (PSGL-1) on TSPs. (Rossi et al. 2005) PSGL-1 is a particularly well characterized example of the importance of specific glycosylation, as the presence of a core-2 sialyl-Lewis x structure at position T57 of PSGL-1, together with sulfation of three nearby tyrosine residues Y46, Y48, and Y51, is absolutely required for P-selectin binding (Leppänen et al. 2003). Thus, deficiency in the O-glycosylation pathway or in the processing and expression of PSGL-1 within TSPs impairs seeding of the thymus. (Sultana et al. 2012). However, the overall glycosylation of PSGL-1 in neither human nor murine T cells has been definitively characterized biochemically.

After the entry of TSPs to the thymus, T lineage commitment and progression through the CD4-CD8- double negative (DN) stages of thymocyte development depend on Notch signaling (Shah and Zuniga-Pflucker 2014). Signaling through Notch receptors is regulated by its glycosylation, depending on the extension of O-linked fucose with GlcNAc. (Rampal et al. 2005; Matsuura et al. 2008) Deficiencies in this pathway result in altered thymocyte population frequencies, while inappropriate expression of glycosyltransferases blocks T lineage commitment, resulting in differentiation to B cells instead (Song et al. 2016; Koch et al. 2001; Visan et al. 2006).

Early checkpoints of DN thymocyte development ensure that a functional T cell receptor (TCR) beta chain has been generated through genetic recombination, after which the cells undergo clonal expansion and enter the CD4+CD8+ double positive (DP) developmental stage

(Dzhagalov and Phee 2012). This clonal expansion depends on the intracellular O-GlcNAc glycosylation pathway, as OGT knockouts are blocked at this step (Shih, Hao, and Krangel 2011). OGT knockouts have substantially decreased numbers of thymocytes and mature T cells. Part of this is due to reduced production of IL-2 in response to TCR stimulation, but other factors are also involved (Shih, Hao, and Krangel 2011). Inhibition of the enzyme that removes O-GlcNAc, OGA, has inconsistent effects on T cell activation and development depending on the experimental conditions (Abramowitz and Hanover 2018). OGT also appears to be required for tumorgenesis in some types of T cell lymphoma (Hanover, Chen, and Bond 2018). Taken together it is clear that proper O-GlcNAcation is crucial for the proliferation and survival of T cells.

Thymocytes dynamically modulate their global surface glycosylation throughout thymic maturation. The most thoroughly investigated of these changes is the degree of sialic acid capping. Many N- and O-glycans are capped with sialic acid during synthesis in the ER, and the specific sialyl transferases expressed change with developmental stage(Marino et al. 2008). Changes in sialylation of O-glycans has only been indirectly studied through the binding of the lectin peanut agluttinin (PNA), which strongly recognizes non-sialylated Core-1 O-glycan disaccharide (Swamy et al. 1991). PNA has long been used to differentiate stages of thymocyte development and to demarcate regions of the thymus. (Reisner, Linker-Israeli, and Sharon 1976; Alvarez et al. 2006; Balcan, Gümüş, and Sahin 2008; Balcan et al. 2008). Thymocytes at the DN stage have a highly sialylated glycome and bind poorly to PNA, while those at the DP stage have greatly reduced levels of sialic acid and bind PNA well. Upon reaching the SP stage, thymocytes once again increase the incorporation of sialic acid onto their glycans and then lose binding to PNA(Sharon 1983; Wu et al. 1996). These sialylation changes are functionally relevant, though

poorly understood, as shown by the requirement of proper sialylation for thymocytes to progress through the DN stages. Deficiency in the sialyltransferase ST6Gal1 reduces the number of DN thymocytes (Marino et al. 2008; Bi and Baum 2009). Moreover, the loss of ST3Gal1 results in an altered TCR repertoire, suggesting that sialylation is also involved in TCR binding interactions and influences selection events (Moody et al. 2001). In the case of CD8+ SP thymocytes, sialylation level inversely correlates with the avidity of binding to class-I MHC, and therefore regulates negative selection of SP cells (Moody et al. 2003).

Differential glycosylation with differentiation state is also observed with N-glycan branching. The β1,6 N-acetylglucosaminltransferase-V (GnT5 or *Mgat5*) catalyzes branching of N-glycans and creates ligands for multiple galectins(Rapoport, Kurmyshkina, and Bovin 2008). This results in a regularly spaced galectin-glycoprotein lattice of immune signaling molecules, restricting the ability of those molecules to cluster together and signal, while also increasing their half-life by preventing endocytic removal from the cell surface (Demetriou et al. 2001). Thus, dysregulation of this pathway either by loss of GnT5, deficiency in galectin 3, or competitive removal of galectins with lactose, results in enhanced clustering of TCR and coreceptors, lowering the activation threshold for TCR signaling (Smith et al. 2018; Clark and Baum 2012; Dias et al. 2014). Modulation of TCR signal strength by N-glycan branching impacts positive and negative selection of thymocytes, such that loss of GnT5 resulted in fewer thymocytes passing the selection checkpoints(Liu et al. 2008), while its overexpression resulted in more cells passing and the formation of a broader TCR repertoire (Demetriou et al. 2001).

Beyond the GnT5-dependent formation of glycoprotein lattices, galectins have multiple and varied effects on T cell development and signaling. The target ligands that mediate these effects are not positively identified and differ from study to study. It is likely that there are

multiple interactions with pleiotropic effects. Galectin 1 and galectin 3 have the greatest observed effect on thymocytes and peripheral T cells, and the high expression of galectin 1 by thymic epithelial cells supports a physiological role for these interactions in normal thymocyte development (Earl, Bi, and Baum 2010).

Galectin 1 has also been reported to induce apoptosis in immature thymocytes, specifically those that have failed positive selection by the inability to generate a productive TCR gene through recombination and those that have failed negative selection due to strong TCR engagement (Pace, Hahn, and Baum 2003; Perillo et al. 1995; Perillo et al. 1997). The role of galectins in T cell function is further discussed in the next section, and has been reviewed elsewhere (Cedeno-Laurent and Dimitroff 2012).

Other dynamic changes to glycosylation have been observed through lectin staining of thymic tissues (Paessens et al. 2007; Alvarez et al. 2006), and it is probable that additional steps of thymocyte development depend upon proper and specific glycosylation beyond those already described.

1.3.2 Signaling in mature T cells

Many of the glycan-dependent effects of thymocyte development involve TCR signal strength and cell-cell interactions at the immune synapse. Thus, it is not surprising that these same glycosylation changes regulate peripheral activation of mature T cells as well, since signaling through the TCR is central both to T cell development in the thymus and to T cell activation and effector function in the periphery. Glycosylation of TCR and its coreceptors is known to impact signal strength, duration, and activation threshold, through aggregation of

receptors on the membrane, protection of signaling components from proteases, and restricting nonspecific interactions (Rudd et al. 1999). These varied effects of glycosylation on TCR signaling can impact peripheral T cell function in a similar modality to their impact on thymocyte selection. Specific roles have been illuminated for sialylation, N-glycan branching, and the formation of galectin ligands.

Peripheral T cells exhibit differential sialylation by activation state that mirrors the changes observed between thymic maturation stages. Naïve cells have high levels of sialic acid on their glycans, but upon activation they remodel their glycans with sialidases and reduce their expression of sialyltransferases, resulting in reduced levels of sialic acid on effector T cells (Brennan 2006, Antanopolous 2012). Effector T cells that later convert to memory phenotype reacquire higher levels of sialic acid, similar to the levels seen for naïve cells (Harrington 2000). These sialylation changes have been shown to impact antigen signaling and peripheral tolerance (Pappu 2004, Brennan 2005, Lubbers 2018).

Peripheral T cells also modulate the branching of N-glycans by GnT5. Activated CD4 and CD8 T cells upregulate GnT5 and have increased branching, which increases the activation threshold of the TCR, as discussed in the previous section (Dias 2018). Reduced N-glycan branching in T cells biases towards a Th1 rather than Th2 response and is associated with autoimmunity in mice and ulcerative colitis in humans, presumably through a lowered activation threshold similar to observations in thymocytes (Dias et al. 2018; Smith et al. 2018; Morgan et al. 2004). Beyond its T cell-intrinsic effect, GnT5 also impacts TCR signaling through the interacting cell. N-glycan branching on MHC promotes binding to the TCR, and GnT5 KO in antigen presenting cells abrogates T cell stimulation (Demetriou 2001, Ryan 2012, Pereira 2018).

Galectins also function extrathymically to modulate T cells. Galectins are induced in dendritic cells and in sites of inflammation, supporting a role for galectins in the resolution of immune responses (Ilarregui 2009). Galectin 1 is consistently immunomodulatory, inducing IL-10 expression and reducing the effects of IFN-γ, either by directly acting on T cells or by inducing dendritic cells to produce signals such as IL-27 to produce the same effect. (van der Leij et al. 2004; Ilarregui et al. 2009)Galectin 1 has also been reported to induce apoptosis not only in immature thymocytes, but also activated T cells and leukemic T cells (Pace, Hahn, and Baum 2003; Perillo et al. 1995; Perillo et al. 1997). This proapoptotic effect is reported to depend upon branched N-glycans on CD45, and be inhibited by sialyl core 1 O-glycans on CD45 and CD43 (Perillo et al. 1995; Earl, Bi, and Baum 2010). Some have questioned whether galectin 1 truly induces apoptosis, or simply the extracellular exposure of phosphatidyl serine (PS), as the results differ based on the methods used to preserve galectin function (Stowell et al. 2009; Stowell et al. 2008), but PS exposure is a prophagocytic signal for macrophages (Elliott 2005) and induction of PS exposure in thymocytes and activated T cells would likely result in death and removal with or without activation of apoptosis.

Modulation of T cell responses is also mediated by the receptor PD-1 and its binding partner PD-L1 (Blank and Mackensen 2007). These inhibitory receptors that are responsible for dampening an immune response once the infection has been resolved and are currently a highly successful therapeutic target in cancer patients (Gong 2018). The interaction of PD-1 and PD-L1 requires proper glycosylation of both molecules. Inhibitory signaling is enhanced by core fucosylation of its N-glycans (Okada et al. 2017) and by extension of its N-glycans with poly-N-acetyllactosamine (Li et al. 2018) N-glycosylation of PD-L1 promotes immunomodulation as non-glycosylated PD-L1 is subject to proteolytic degradation (Li et al. 2016).

While some effects of glycosylation have been traced to specific glycoproteins, there are a host of accessory proteins that participate in the immune synapse between a T cell and an antigen presenting cell, and it is not always clear which glycoprotein is mediating a glycan-dependent effect on T cell biology (Evans et al. 2003; Dustin and Chan 2000; Haslam et al. 2008). Because they are two of the most abundantly expressed proteins on the surface of T cells and are highly decorated with O-glycans, CD45 and CD43 are prime candidates to investigate (Clark and Baum 2012; Jones et al. 1994). Both act to modulate T cell activation. CD45 participates in the immune synapse where it acts as a phosphatase to inhibit TCR signaling and, while CD43 is excluded from the immune synapse and reduces T cell activation, adhesion, and proliferation (Thurman et al. 1998; Sperling et al. 1998; Mody et al. 2007). The five CD45 isoforms are alternative splice variants that are differentially expressed on T cell subtypes (Earl and Baum 2008). The membrane proximal region contains nearly a dozen N-glycosylation sites, and each isoform presents a different glycosylation profile (Earl, Bi, and Baum 2010). In addition, the alternatively spliced portion of the protein is an extended mucin-like domain that is heavily decorated with O-glycans. CD43 also bears an extended mucin-like domain, and is the major sialic acid-rich protein on the T cell surface (Rosenstein 1999). Glycosylation of both CD43 and CD45 is modulated during T cell differentiation and activation, and by their abundance contribute to the observed differences in staining such as with PNA and SNA lectins (Clark and Baum 2012). Some studies attribute phenomena like susceptibility to Galectin 1 induced apoptosis to the O-glycosylation of CD43 and CD45: these proteins bear core 2 O-glycans on immature thymocytes, which are susceptible, but bear sialylated core 1 on mature thymocytes, which are protected (Earl and Baum 2008; Garcia and Berger 2005). Further

research is needed to establish the mechanisms behind O-glycosylation effects on T cell differentiation and activation, and the glycoproteins through which those effects are mediated.

1.3.3 T cell homing

Homing to specific effector organs is critical for immune cells and glycosylation is essential to correct homing on all cell types that have been examined. Glycosylation is a critical element of leukocyte homing in general (McEver 1997 JCI) and lymphocyte homing in specific (Rosen 2004 ARI). Homing of T lymphocytes occurs in multiple stages, termed rolling, arrest, and diapedesis. These steps occur any time a T cell extravasates into a secondary lymphoid organ or into a target tissue (von Andrian and Mempel 2003; Madri and Graesser 2000).

The first step of extravasation is rolling, in which a T cell adheres weakly to endothelial cells of blood vessels, slowing it and prolonging contact times. This process is mediated by selectins expressed by the T cell binding to carbohydrate ligands on the surface of the endothelial cells (Picker 1993, Ley 2004). Surface expression of L-selectin is induced upon T cell activation, and is also retained on memory T cells (Hengel et al. 2003; Schlub et al. 2010). L-selectin ligands include sialyl-Lewis X carbohydrate structures and their sulfated derivatives expressed on a variety of proteins on endothelial cells including GlyCAM-1, MadCAM-1, and CD34 (Rosen 2004). These ligands are induced upon exposure to inflammatory signals such as TNFα, and their expression is one of the characteristics of high endothelial venules (HEV) through which activated immune cells transmigrate to reach inflamed tissue (Girard, Moussion, and Förster 2012). Without proper recognition of glycoprotein ligands by T cells, extravasation cannot begin (Gauguet et al. 2004; Frommhold et al. 2008).

The second step of extravasation is arrest, in which a T cell initiates strong adhesion to endothelial cells, stopping its motion through the blood. Arrest is mediated by integrins

expressed on the T cell, especially LFA-1, binding to protein ligands on the endothelial cells in responses regulated by chemokine signaling (Walling and Kim 2018). While glycosylation does not play a direct role in this binding event, some integrins have been shown to bind with different affinities based on their glycosylation state, and altered glycosylation or mutation of glycosites generally results in reduced half-life and accelerated degradation of integrins.(Janik 2010, Kariya 2017, Marisco 2018, Hou 2015).

Finally, T cells transmigrate between endothelial cells to leave the blood stream and enter their target tissue. This involves the concerted action of multiple glycoprotein receptor-ligand pairs, including adherens junction protein JAM-A (Strell 2008, Kummer 2018). The N-glycosylation of JAM-A regulates its interaction with migrating leukocytes and promotes the formation of tight junctions. (Scott & Burridge, 2015). Altered JAM-A glycosylation reduces the ability of leukocytes to extravasate.

These examples illustrate the known role of glycosylation in each of the steps of lymphocyte homing. Other glycoproteins that participate in each step have yet to be investigated regarding the importance of their glycosylation state on the process.

1.3.4 T cell recognition of glycan-dependent antigens

Canonically, T cells recognize peptide antigens presented in the context of MHC, through the T cell receptor (TCR)(Williams, Peh, and Elliott 2002; Villadangos 2001). However, there are several additional modalities through which T cells also recognize glycans or glycoproteins. These specialized recognition events enable recognition of glycopeptides, glycolipids, and

zwitterionic polysaccharides (ZPS). While less common than peptide antigens, such MHC-dependent glycan and glycopeptide antigens induce robust T cell responses.

When antigen presenting cells process proteins in the lysosomal pathway for presentation on MHC, the resultant peptides are usually linear peptide chains with no post-translational modification (Goldberg 2002, Lennon 2002). However, sometimes a glycopeptide will retain part or all of a substituent glycan moiety when loaded onto MHC for presentation (Vlad et al. 2002; Deck et al. 1995; Carbone and Gleeson 1997). Because detection methods to identify MHC-loaded peptides often isolate the peptides in conditions such as low pH that are damaging to attached glycans, it is unclear how widespread this phenomenon is (Bozzacco 2013). Glycopeptides on the MHC have been observed both in cases where the glycan extrudes from the peptide-MHC surface and interacts with the TCR and in cases where the glycan is buried in the MHC pocket and participates in anchoring the glycopeptide in place (Speir et al. 1999) (Apostolopoulos et al. 2003; Avci et al. 2011). This allows for differential recognition of alternatively glycosylated proteins, such that T cells can distinguish and respond to altered glycosylation states including those that arise in cancers (Lucas, Apicella, and Taylor 2005). Moreover, even when a glycan moiety is not retained on a peptide presented to a T cell, the glycosylation of the source protein can influence the activity of proteases during antigen processing, resulting in different peptides being generated from alternate glycoforms of a given protein (Szabó et al. 2009; Bernard et al. 2010). Glycan-dependent recognition of T cell antigens and glycan-dependent processing of T cell antigens both may serve as important mechanisms for immune recognition of neoantigens in tumors. Such recognition would not rely on mutation of the underlying amino acid sequence but rather the emergent property of altered glycosylation, which is common in many cancers as discussed in section 1.2.3.2.

Glycolipids extracted from the membranes of target cells are also recognized antigenically by a subset of T cells. Glycolipids are presented on the noncanonical MHC-like molecule CD1d and recognized by a special class of T cells known as natural killer T (NKT) cells (Moody et al. 2001; Hava et al. 2005). This class of T cells bears a restricted set of TCRs with limited variability relative to those found on classical T cells (Pellicci et al. 2011). The TCR recognizes a bound lipid head group, and binding strength is modulated by the differential presentation of that head group based on the seating of the lipid moiety into the CD1d binding pocket(Girardi and Zajonc 2012). Strong binding by an NKT cell results in the induced death of the target cell by apoptosis or direct killing.

Capsular polysaccharides derived from bacteria usually do not elicit a T cell -dependent immune response. However, a subset that bear both positive and negative charges, known as zwitterionic polysaccharides (ZPS), are potent activators of T cells (Duan, Avci, and Kasper 2008). ZPS are characteristic of certain bacterial cell walls, including the human pathogens *B. fragilis, S. aureus,* and *S. pneumoniae,* and are useful immune targets (Neff et al. 2016). The uncommon zwitterionic nature of ZPS allows them to bind to alpha-helical proteins in such a way that they can bind directly into the antigen binding groove of MHC-II for presentation to T cells. ZPS are processed and loaded onto MHC-II in the lysosome, similar to peptide antigens (Duan, Avci, and Kasper 2008). Once presented on the surface of cells, the ZPS-MHC complex binds to the TCR of canonical T cells, which go on to provide activating and class-switching signals to B cells for antibody production (Kalka-Moll et al. 2002).

T cells are most commonly thought to respond solely to linear peptide antigens, but the recognition of processed glycopeptides and ZPS in the context of MHC-I and MHC-II provides direct evidence for glycan-dependent recognition of antigens by traditional T cells. In addition,

NKT cells are a specialized subset of T lymphocytes that are dedicated to the recognition of glycolipid antigens.

1.4 Summary

T cells are an essential element of adaptive immunity, and their development in the thymus is carefully regulated and monitored to ensure production of a competent but not autoreactive pool of antigen-reactive T cells. Once in the periphery, the activation, homing, persistence, and memory functions of T cells are likewise regulated to preserve a careful balance of immune function without pathological activation. Each of these steps, including both the proinflammatory and immunomodulatory signals received, as well as the intercellular interactions that promote localization of T cells to the proper site, are mediated by glycoproteins. T cell development and activation is characterized by global alterations in the glycocalyx, and multiple knockout models have demonstrated that perturbation of the N- and O-glycosylation pathways results in altered development or function of T cells. These effects range from frank loss of cellularity and blockage in cell maturation to nuanced differences in TCR pool composition and cell reactivity.

Perturbations in glycosylation result in additional pathologies beyond T cell dysfunction. In particular, O-glycan truncation due to loss of *Cosmc* is a frequent feature of human carcinomas and is an indicator of poor prognosis. Identification of the salient glycoproteins that participate in oncogenic and metastatic processes could lead to therapeutic interventions, but current techniques of glycoprotein and glycosite analysis are not feasible for small samples of complex mixtures, such as primary patient tissues.

This chapter has introduced the current knowledge of T cell biology, glycosylation, and their interaction as well as the current approaches to glycoproteomic analysis. Chapter 2 will

present advances in the understanding of the role glycosylation plays in T cell maturation and survival. Chapter 3 will present advances in methodology for sensitive identification of O-glycoproteins in samples of limited quantity including primary patient samples. Finally, Chapter 4 will discuss open questions of the field and future directions of research.

1.5 Tables and Figures

Table 1.1: Features of CD4 T cell subtypes

Cell Type	Th1	Th2	Th17	TFH	Treg
Inducing Cytokines	IFN-γ, IL-12	IL-4, IL-10	IL-1β, IL-6, IL-23	IL-6, IL-21, CXCL13	TGF-β (for induced Tregs)
Defining Transcription Factors	T-bet, STAT-1	STAT-6, GATA-3	RORγT, STAT-3	Bcl-6, STAT-3	Foxp3
Predominant Cytokines Produced	IL-2, IL-12, IFN-γ, TNF-α	IL-4, IL-5, IL-25, IL-10, 1L-13	IL-21, 22, 24, 26, IL-17A, IL-17F	IL-21, IL-6, IL-10	TGF-β, IL-10
Target Cell Type	CD8 T cells, Macrophages	Granulocytes	Neutrophils	B Cells	T cells, Dendritic Cells, Macrophages
Immune Context	Viral, Bacterial infection	Extracellular parasites	Inflammation, autoimmunity	Antibody class switching and refinement	Attenuation of Immune Response

Table 1.2: A selection of frequently used lectins for structural determination of glycans.

Lectin Name and abbreviation	**Typical ligand**
Wheat germ agglutinin (WGA)	O-GlcNAc
Peanut agglutinin (PNA)	Core 1
Concanavalin A (ConA)	High Mannose, hybrid type, complex N-glycans
Sambucus nigra agglutinin (SNA)	2-6 linked sialic acid
Helix pomatia agglutinin (HPA)	α linked GalNAc
Vicia villosa agglutinin (VVA)	α linked GalNAc
Artocarpus integrifolia agglutinin (Jacalin)	Galactose
Maackia amurensis leukoagglutinin (MAL-I)	2-3 linked sialic acid
Maackia amurensis erythroagglutinin (MAL-II)	2-3 linked sialic acid on O-glycans
Ricinus comnnunis agglutinin (RCA-1)	Galactose
Ulex europaeus agglutinin (UEA)	Fucose
Galectins	LacNAc
Phaseolus vulgaris erythroagglutinin (E-PHA)	Bisecting GlcNAc on N-Glycans
Phaseolus vulgaris leukoagglutinin (L-PHA)	β6 branched N-glycans
Datura stramonium agglutinin (DSA)	2,4 branched N-glycans

Chapter 2: Cosmc is required for T cell persistence in the periphery

A portion of this chapter has been accepted by *Glycobiology*:

Cosmc is required for T cell persistence in the periphery

Christopher E. Cutler[1,2], Mark B. Jones[1,4], Alicia A. Cutler[3], Amanda Mener[2], Connie M. Arthur[2], Sean R. Stowell[2], Richard D. Cummings[1,4]

[1]Beth Israel Deaconess Medical Center, Department of Surgery, Boston, MA, USA

[2]Emory University School of Medicine, Atlanta, GA, USA

[3]University of Colorado, Boulder, CO, USA

[4]Harvard Medical School Center for Glycoscience, Harvard Medical School, Boston, MA, USA

2.1 Abstract

T lymphocytes, a key arm of adaptive immunity, are known to dynamically regulate O-glycosylation during T cell maturation and when responding to stimuli; however, the direct role of O-glycans in T cell maturation remains largely unknown. Using a conditional knockout of the gene (*C1GalT1C1* or *Cosmc*) encoding the specific chaperone Cosmc, we generated mice whose T cells lack extended O-glycans (T cell Conditional *Cosmc* Knock Out or TCKO mice) and homogeneously express the truncated Tn antigen. Loss of *Cosmc* is highly deleterious to T cell persistence, with near-complete elimination of *Cosmc*-null T cells from spleen and lymph nodes. Total T cell counts are 20% of wild type, among which only 5% express the truncated glycans, with the remaining 95% consisting of escapers from *Cre*-mediated recombination. TCKO thymocytes were able to complete thymic maturation but failed to populate the secondary lymphoid organs both natively and upon adoptive transfer to wild type recipients. Our results demonstrate that extended O-glycosylation is required for the establishment and maintenance of the peripheral T cell population. [170 words]

2.2 Introduction

T lymphocytes are a crucial element of the adaptive immune system that recognize potentially harmful antigens through the T cell receptor (TCR) and coordinate the immune response (Koch and Radtke 2011). T cell dysfunction can lead to life threatening infections or auto-immune disorders. T cells must home to secondary lymph organs (SLO), including the spleen and lymph nodes, to survey antigens and protect the body from pathogen invasion (Nolz, Starbeck-Miller, and Harty 2011; Kumar, Connors, and Farber 2018; Freitas and Rocha 1999). In order for T cells to mature, home to their target locations, and appropriately bind foreign targets, the TCR and associated glycoproteins must be correctly expressed, folded, and modified (Nolz and Harty 2014; Kawashima and Fukuda 2012; Dustin and Chan 2000).

As T cells develop from bone-marrow derived precursors in the thymus (Trnková, Pastoreková, and Petrik 2012), one essential step in their maturation is generation of a functional TCR by genomic rearrangement. This unique TCR is subsequently tested for functionality and self-reactivity (Heusmann 1991), where T cells that fail these tests are induced to undergo apoptosis and never reach circulation or the SLO (Dustin and Chan 2000). The signals for regulating and testing TCR generation are mediated by glycoproteins, but the role that glycosylation plays directly in these events is largely unknown. Similarly, glycoproteins mediate signaling for other major events in T cell maturation, activation, and homing (Pereira et al. 2018; Love and Bhandoola 2011; Daniels, Hogquist, and Jameson 2002; Bi and Baum 2009; Jones 2018). The glycosylation state of these glycoproteins is temporally regulated during maturation and activation indicating a substantial regulatory role for glycosylation in these processes (Balcan et al. 2008; Balcan, Gümüş, and Sahin 2008; Paessens et al. 2007).

More than 80% of proteins with a secretory signal sequence are predicted to be modified with O-glycosylation (Steentoft, Vakhrushev, Vester-Christensen, Schjoldager, et al. 2011; Apweiler, Hermjakob, and Sharon 1999), and O-glycans influence a protein's structure, stability, and binding interactions (Seguchi et al. 1991; Wang et al. 2012; Daniels et al. 2001; Cummings and Pierce 2014). O-glycosylation of proteins begins co-translationally in the Golgi apparatus with the addition of an α-linked N-acetylgalactosamine (GalNAc) to serine or threonine, catalyzed by any of more than 20 polypeptide:UDP-GalNAc transferase enzymes (Tenno et al. 2007). This initiating structure, GalNAc linked to Ser/Thr, the Tn antigen, is almost invariably modified by the addition of β3-linked galactose to form the Core-1 disaccharide Galβ3GalNAcα1-Ser/Thr, which forms the precursor for extended O-glycan structures. This crucial reaction is catalyzed by a single enzyme, C1GalT1 or T-synthase. Production of functional T-synthase enzyme requires the chaperone Cosmc encoded on the X chromosome (*C1GalT1C1* or *Cosmc*), without which nascent T-synthase aggregates and is degraded (Ju and Cummings 2002). Loss of Cosmc or T-synthase results in the truncation of all Core-1 type O-glycans to the Tn antigen. The requirement for Cosmc chaperone is complete and quantitative, such that in cells lacking Cosmc there is no residual T-synthase activity **(Ju, Otto, and Cummings 2011)**. Reinforcing the importance of the enzymatic reaction from this single gene, global Cosmc KO animals result in embryonic lethality (Wang et al. 2010).

T cell O-glycans are dynamically regulated both during maturation and immune responses (Starr et al. 2003; Hernandez et al. 2007). There is some indication that specific O-glycans participate in particular stages of T cell development and activity (Moore et al. 2008; Pereira et al. 2018), but the overall importance of O-glycans in T cell maturation remains largely unknown. Thus, we investigated the general role of O-glycans in T cell biology by taking advantage of

Cosmc deletion to abrogate T-synthase activity. Cosmc$^{-/-}$ cells have truncated O-glycans presenting as Tn antigen, providing an ideal model to evaluate the importance of complex O-glycan structures.

To evaluate the role that O-glycans play after T lineage commitment, we generated a T cell specific *Cosmc* KO (TCKO) model and examined the effect on T cells in the thymus, circulation, and SLO. We determined that loss of Cosmc and consequential truncation of O-glycans is highly deleterious to T cell persistence, with near-complete elimination of peripheral Tn antigen expressing T cells. TCKO thymocytes complete thymic maturation but fail to populate the SLO. Co-transfer of WT and TCKO thymocytes shows that TCKO cells have reduced ability to home to SLO and are not maintained in circulation. We demonstrate that Cosmc, and therefore extended O-glycosylation, is a critical element in the establishment and maintenance of the peripheral T cell population.

2.3 Results

2.3.1 Lck-Cre drives T cell specific knockout of Cosmc

In order to specifically examine the role of O-glycosylation in T cells, we created a T cell Conditional *Cosmc* Knock Out (TCKO) mouse. The mouse was produced by crossing the *Cosmcfl/fl* mouse (Wang et al. 2012) with a mouse expressing Cre recombinase from the T cell specific *Lck* proximal promoter (Hennet et al. 1995) (**Fig 2.1A**). The proximal promoter was selected because it is most active during thymocyte maturation, beginning at the early CD8^{-}CD4^{-} double negative (DN) stage, while *Lck* expression from the distal promoter is more active during activation of mature T cells (Shi and Petrie 2012). Mouse genotypes were determined by PCR (**Fig 2.1B**). All TCKO mice used in subsequent experiments constitutively express Cre in a T

cell specific manner and are homozygous *Cosmc*$^{fl/fl}$ females or hemizygous *Cosmc*$^{fl/y}$ males, as *Cosmc* is located on the X chromosome. To determine whether T-synthase activity is reduced in TCKO cells, we performed a fluorescence-based T-synthase enzyme activity assay on lysate from bulk thymocytes of TCKO and Endothelial and Hematopoietic Conditional *Cosmc* Knock Out (EHCKO) mice (**Fig 2.1C**). EHCKO mice express Cre recombinase from the hematopoietic specific promoter *Tie2* (Wang et al. 2012). EHCKO thymocyte derived lysate have appreciable specific activity of T-synthase, though reduced relative to an immortalized human neutrophil cell line (HL-60). This indicates very low level of *Cosmc* deletion in the T cell compartment, consistent with prior characterization of the EHCKO model (Wang et al. 2012). In contrast, TCKO thymocytes exhibit a reduced specific activity relative to both HL-60 and EHCKO lysates. However, T-synthase activity in TCKO thymocyte lysate is still substantially higher than the activity of the negative control, an immortalized human T cell line with a known T-synthase deficiency (Jurkat) (Ju et al. 2008). This indicates that the T cell population in TCKO mice is somewhat heterogeneous, containing some cells that have not undergone Cre-mediated *Cosmc* deletion. Notwithstanding the incomplete deletion of *Cosmc*, T-synthase activity is significantly reduced in TCKO thymocytes, making it an effective model to examine the role of T-synthase in T cell maturation and function. We next characterized the hematopoietic compartment of TCKO mice by measuring the abundance of major blood cell types in TCKO and wild type (WT) littermates. Blood counts of TCKO mice revealed significant lymphopenia relative to littermate controls (**Fig 2.1D**). The levels of granulocytes were unchanged and there was a small but significant decrease in the number of monocytes. Together these data indicate that *Cosmc* can be specifically deleted in T cells using this TCKO model, resulting in decreased T-synthase activity and reduced levels of circulating, peripheral T cells.

2.3.2 Cosmc KO results in dramatically reduced T cell numbers

To determine if all circulating T cell subtypes were affected equally, we examined the peripheral T cell populations in the TCKO mouse. $CD4^+$ and $CD8^+$ T cells were identified and analyzed by flow cytometry (**Fig 2.2A**). While circulating levels of B cells were unchanged in TCKO mice compared to WT, as expected, the median cell numbers of both the $CD4^+$ and the $CD8^+$ T cells were only 20% that of WT littermates (**Fig 2.2B, Table 2.I**). This decrease in T cell numbers highlights the effect of *Cosmc* deletion. A significantly higher percentage of both $CD4^+$ and $CD8^+$ cells express Tn antigen relative to cells from WT mice, where Tn antigen is undetectable. However, surprisingly >90% of circulating T cells in TCKO mice do not express Tn antigen. The median percentage of Tn positive cells was 5% for $CD4^+$ T cells and 1% for $CD8^+$ T cells (**Fig 2.2B, Table 2.II**). This very small percentage of cells expressing the Tn antigen led us to hypothesize that Tn antigen positive cells are depleted from circulation in one of possibly three ways: 1) Tn antigen positive T cells may be sequestered in SLO; 2) Tn antigen positive T cells may be prevented from entering circulation; or 3) Tn antigen positive T cells may be actively removed from circulation.

We then interrogated SLOs for T cell population amounts, as compared to circulation. Splenic T cells mirrored the circulating T cells in both overall T cell levels and percentage of Tn antigen expression (**Fig 2.2C**). The populations of both $CD4^+$ and $CD8^+$ T cells were significantly reduced by three- to five-fold, with a median count at 22% and 36% of WT levels, respectively. Splenic TCKO T cells expressed the Tn antigen at the same low percentages seen in blood (**Fig 2.2C**). In the case of the peripheral lymph nodes (PLN, pooled lymphocytes from inguinal, axillary, and brachial nodes) and the mesenteric lymph node (MLN), $CD4^+$ T cell

counts were significantly reduced in TCKOs and $CD8^+$ T cell counts also trended lower. The median percentage of Tn positive T cells in both PLN and MLN was even lower than observed in blood and spleen (**Fig 2.2D** and **E, Table 2.I**). As the reduced T cell counts and Tn+ status in blood and spleen are reflected in the lymph nodes, we conclude that Tn+ T cells are not being sequestered in SLO.

Although the percentage of Tn expressing cells is consistently lower in $CD8^+$ T cells than in $CD4^+$ T cells of TCKO mice, within the total T cell population (including Tn- cells) the relative ratio of $CD8^+$ to $CD4^+$ cells in TCKO do not differ significantly from WT (**Fig 2.2F**). This indicates that the persistence of both $CD4^+$ and $CD8^+$ T cells is negatively impacted by *Cosmc* deletion, with $CD8^+$ T cells having enhanced sensitivity to *Cosmc* deficiency, but that proportional numbers of both populations escaped *Cosmc* deletion. While loss of *Cosmc* may have a more severe impact on persistence of $CD8^+$ T cells than $CD4^+$ T cells, the qualitative effect on both populations appears to be similar.

Having observed that total numbers of circulating T cells are five-fold lower in TCKO mice than in WT and the median percentage of T cells expressing Tn antigen does not exceed 5% in any peripheral lymphoid compartment, we analyzed whether *Cosmc*-deleted cells are escaping detection. We measured the percentage of T cells expressing the Tn antigen by co-staining with the Tn-specific antibody Bags6 and the GalNAc-binding lectins VVA and HPA, both of which were consistent with the antibody-based detection (data not shown). These data suggest that the majority of peripheral T cells represent those that escaped Cre recombination and loss of *Cosmc*. Thus, these data indicate that T cells lacking *Cosmc* are either unable to reach or are deficient in their ability to populate the periphery, and that cells escaping *Cosmc* deletion have a survival advantage in establishing the peripheral T cell pool.

2.3.3 Cosmc KO does not block thymocyte maturation

We next examined whether T cells from TCKOs survive thymic maturation and selection. Although the *Lck* Cre transgene is activated very early at the $CD4^{-}CD8^{-}$ double negative (DN) stage, thymocyte populations (**Fig 2.3A**) at all maturation stages were not significantly different in numbers between TCKO mice and their WT littermates. These results demonstrate that in contrast to the results regarding peripheral T cells, loss of *Cosmc,* as driven by *Lck* Cre, does not result in dramatic loss of TCKO thymocyte populations (**Fig 2.3B**). At all maturation stages the percentage of T cells expressing Tn antigen was significantly higher in TCKO than WT. The percentage of T cells expressing the Tn antigen is significantly and substantially higher in maturing T cells in the thymus than in T cells isolated from the blood or peripheral lymph organs (57% at the highest in the thymus compared to 5% at the highest in the periphery) (**Fig 2.3C, Table 2.II**), which indicates that the strong deleterious effect of *Cosmc* loss on T cells does not occur during thymocyte maturation.

Due to the large differences in percent of Tn positive T cells at each stage of T cell maturation, we examined the maturation profile of TCKO and WT thymocytes. As T cells mature and prepare to egress from the thymus, they upregulate levels of CD62L (L-selectin). We observed that 69% of WT single positive (SP) thymocytes express high levels of CD62L, indicating that they are preparing to exit the thymus (**Fig 2.3D**). However, in TCKO SP thymocytes there was a striking difference in CD62L expression levels between Tn+ and Tn- cells (**Fig 2.3E**). Overall, 62% of TCKO SP cells were $CD62L^{hi}$, but the $CD62L^{lo}$ SP thymocytes were only 13% Tn+, while the $CD62L^{hi}$ SP thymocytes were 50% Tn+ (**Fig 2.3F**). This result indicates that T cells with *Cosmc* deletion are successfully maturing, and Tn+ cells are enriched

in the most mature T cell populations. We hypothesize that this may be due to inappropriate retention of Tn+ cells or accelerated passage through checkpoints intended to ensure selection of immune competent, non-autoreactive cells. Accelerated passage is better supported by the data than retention because CD4 SP thymocytes from TCKO contain a reduced percentage of $CD62L^{hi}$ cells than WT, indicating that mature pre-egress thymocytes are not being enriched, though no significant difference was detected in the proportion of cells expressing high levels of other maturation markers CD24 and Qa-2 (**Fig 2.3G**). These data imply that the Tn+ cells are not depleted, even at the latest maturation stages of thymic development.

The high percentage of Tn+ T cells in the thymus is in sharp contrast to the trends seen for peripheral T cells. Whereas the mice with the highest penetrance of Tn antigen expression and therefore *Cosmc* deletion, had the most pronounced T cell loss and those with the highest T cell numbers had the least penetrance of Tn antigen expression, no such trend exists within the thymus (**Fig 2.4A**). Plotting the loss in cell numbers by compartment and population (**Fig 2.4B**) clearly demonstrates that the powerful effect of *Cosmc* deletion occurs after the events of thymic maturation. Comparing $CD4^+$ and $CD8^+$ T cell populations for Tn antigen expression shows that $CD8^+$ T cells are more strongly affected by the selective pressure against *Cosmc* deletion, and that this difference begins to appear even at the SP stage of thymic maturation (**Fig 2.4C**). Together, these data demonstrate that the selective pressure decreasing TCKO T cell persistence exists in the thymus, but predominantly acts after thymic maturation. TCKO T cells are able to mature through thymic checkpoints and are not strongly selected against within the thymus. Thus, we hypothesize that the T cell numbers in general and Tn antigen positive T cells specifically are so drastically reduced in the periphery due to one or both of two likely

explanations: 1) T cells are unable to leave the thymus, 2) T cells are actively selected against in the periphery.

2.3.4 TCKO T cells are deficient in homing to SLO

To determine whether Tn+ T cells encountered negative survival selection after leaving the thymus, we adoptively transferred thymocytes into WT animals (**Fig 2.5A**). Thymocytes isolated from the WT and TCKO mice were differentially fluorescently labeled, mixed in a 1:1 ratio, and injected into a WT host. Animals were allowed to rest for 16 hours post injection (the time required for T cells to home to SLO and recirculate (Oostendorp, Ghaffari, and Eaves 2000; Ganusov and Auerbach 2014; Srour et al. 2001)), at which point the recipient tissues were collected and analyzed. Donor splenocytes correctly homed to all peripheral lymphoid organs indicating that cells from both the TCKO and WT mice survived isolation and the adoptive transfer procedure. As expected, WT SP4 thymocytes homed at lower efficiency than WT $CD4^+$ splenocytes but were still able to home to all SLO (**Fig 2.5B**); however, TCKO thymocytes homed with significantly lower efficiency than WT for each SLO measured. Strikingly, TCKO SP4 thymocytes engrafted into the peripheral lymph nodes at less than 20% the efficiency as WT SP4 thymocytes. The results from this adoptive transfer experiment dramatically illustrate that *Cosmc* knockout is detrimental to the ability of T cells to home to SLO, and that negative selective pressure from the loss of elaborated O-glycans acts on post-thymic T cells.

2.4 Discussion

We discovered that loss of *Cosmc* causes a profound loss of T cells from the peripheral lymph system while T cell maturation was undisturbed. The dramatic global reduction of peripheral T cell numbers in TCKO animals and the further disproportionate loss of T cells expressing Tn antigen on their surface, demonstrates that appropriate elaboration of extended O-glycans is critical to persistence of T cells in the blood or SLO. Nevertheless, the presence of some Tn+ T cells in the periphery, however few, indicates that extended O-glycans are not an absolute requirement to maturation and survival. These findings are consistent with previous studies (Kao and Sandau 2006; Vainchenker et al. 1985; Berger 1999) and further indicate that T-synthase activity is especially important in T cells.

We discovered that, following *Cosmc* deletion in TCKO mice, there was an overall reduction in T cell levels relative to WT, and in that reduced population a surprisingly low percentage (<5%) expressed the Tn antigen. While the number of peripheral T cells expressing the Tn antigen was vanishingly small, an appreciable number of circulating T cells that escaped *Cosmc* deletion remained. Reduced total T cell numbers indicates that 'escaper' T cells, i.e. those in which *Cosmc* was not deleted efficiently by Cre activity, fail to expand and fill the available niche. This is an important result, as compensatory expansion of a minor subset of escaper T cells dividing to achieve comparable cell counts to WT animals is common in cases where a majority of T cells have a survival defect and the lack of such expansion in the TCKO mice was surprising (Martin et al. 2013). The reduction in total T cell numbers observed in TCKO mice indicates that another mechanism is preventing the usual compensatory expansion. One possibility is that T cells in TCKO mice are activated in the periphery, inducing *Lck* expression and inappropriate production of Cre. *Lck* expression upon activation is predominantly driven

from the distal promoter (Chiang and Hodes 2016); however, if expression of the proximal promoter were also induced, Cre would be expressed resulting in post-thymic *Cosmc* deletion. These T-synthase-deficient cells would experience the same selective pressure that eliminated the earlier Tn expressing thymic emigrants, maintaining the lymphopenia. However, because extrathymic expression of *Lck* is controlled by the distal rather than the proximal promoter, it is unlikely for recombination to occur in peripheral T cells that had already escaped recombination. Therefore, the Tn positive T cells in the periphery are most likely to be persistent Tn positive thymic emigrants, rather than mature Tn negative cells converted to Tn positive by post-egress deletion of *Cosmc*.

The dramatically low number of Tn antigen-expressing T cells we detected in the TCKO model parallels findings in human patients with Tn syndrome. Tn syndrome is a relatively rare disorder arising from spontaneous somatic mutation of *Cosmc* in hematopoietic stem cells, and is characterized by Tn antigen expression on all hematopoietic subpopulations (Ju 2005; Vainchenker et al. 1985; Berger 1999). Examination of peripheral blood cells from these patients revealed that while Tn antigen occurs on a moderate to high proportion of most blood cell types (70-95% of erythrocytes, 50-80% of megakaryocytes, 25-90% of granulocytes, and 15-65% of B cells) the occurrence on T cells was far lower, ranging only between 2-5% of total T cells (Brouet JC 1983; Judson PA 1983; Berger EG 1994). This finding matches closely to our TCKO results, in which fewer than 5% of surviving T cells expressed the Tn antigen. Thus, T cells are particularly susceptible to perturbation of O-glycosylation and the TCKO model is an ideal method for further study on the unique requirements of T cells for extended O-glycosylation and modeling Tn syndrome.

We observed that <60% of CD4 single positive thymocytes expressed the Tn antigen, the marker of Cre mediated *Cosmc* deletion, and that the percentage in other thymic subpopulations was even lower. Previous studies that also used the *Lck* Cre construct for gene deletion documented greater than 80% recombination efficiency in T cells (Shi and Petrie 2012). This divergence in recombination efficiency suggests that the loss of *Cosmc* is impacting not only the survival and homing of peripheral T cells, but possibly also the development of thymocytes and/or their ability to successfully pass positive and negative selection checkpoints. We observed a steady increase in the proportion of each subpopulation that expressed the Tn antigen, with more mature stages having a higher percentage of Tn positive cells.

This expected increase in percentage of Tn positive cells with maturation does not indicate a survival advantage for Tn positive cells during the selection process, because in cultured cell lines there is a seven-day delay between loss of Cosmc and the appearance of Tn antigen on the cell surface (Hofmann et al. 2015). This delay between *Cosmc* deletion and Tn antigen expression corresponds to the expected time to turn over functional Cosmc and T-synthase. Thymic maturation from $CD4^-CD8^-$ double negative (DN) T cell progenitor to mature $CD4^+$ or $CD8^+$ single positive (SP) cell takes approximately two weeks, a gradual increase of Tn positive cells in maturing populations is expected. While we detected increases in the percentage of Tn positive T cells, there was high variability between biological replicates in expression levels (SD =25%). This high degree of variability between individual mice in the eventual proportion of Tn+ cells is surprising and suggests that the effect of *Cosmc* deletion is subject to stochastic processes. One such stochastic process is TCR recombination, which is the target of the thymic selection events. Therefore, signaling through the TCR and associated coreceptors is a promising potential mechanism by which O-glycan truncation impacts thymocyte fitness.

However, it is clear from the proportion of Tn positive cells that pass selection and reach the mature SP stage that loss of extended O-glycans does not preclude productive TCR engagement and associated signaling.

Of additional interest is our finding that the most mature cells, i.e. those expressing high levels of L-selectin (CD62L), are highly enriched for Tn antigen expression (58% Tn positive among $CD62L^{hi}$ cells, compared to 14% Tn positive among $CD62L^{lo}$ cells). This enrichment of Tn positive mature thymocytes might indicate an enhancement of maturation signals in Tn positive cells, such that they are driven towards a mature phenotype earlier than their Tn negative counterparts, or it might indicate that Tn positive cells are slower to egress from the thymus. Because normal thymocyte development results in upwards of 90% of cells undergoing apoptosis due to failure to pass positive or negative selection checkpoints, and such apoptotic cells are quickly cleared by resident phagocytes (Surh and Sprent 1994), it is not possible to easily determine the relative survival rate of Tn positive and Tn negative thymocytes in TCKO animals. Determining the effect of Cosmc loss on thymocyte survival will require further new experimental models and further study, but could further illuminate the role of O-glycosylation in thymocyte development, selection, and survival.

Because T cells are required for B cell development, complete loss of peripheral T cells such as occurs upon deficiency of the common gamma chain consequently results in an accompanying absence of B cells, which is termed Severe Combined Immunodeficiency Syndrome (Buckley 2004). However, very small numbers of T cells are apparently sufficient to provide the necessary factors for B cell development, as patients with DiGeorge syndrome have severe T lymphopenia but do not lack T cells entirely, yet have no reduction in B cell numbers (Derfalvi et al. 2016). Therefore it is not unexpected that the B cell numbers in our TCKO model would not differ from WT, despite the extensive crosstalk between the two

cell types. Once the T cell compartment of the TCKO model is more fully characterized, it would be interesting to evaluate the ways in which B cells are affected by the altered and reduced T cell population.

The TCKO model we describe here exhibits reduced overall T cell levels, appreciable Tn expression, and a low percentage of Tn antigen positive T cells in circulation and the peripheral lymphoid organs. In contrast, the EHCKO mouse, in which *Cosmc* deletion is driven by the earlier *Tie2* promoter, had no detectable Tn expression in the thymus (Wang et al. 2010). This indicates that in addition to a survival and homing defect in thymic and peripheral T lymphocytes, *Cosmc* deletion also affects the differentiation or homing of early thymic progenitors. This deficiency could present as impaired development in the bone marrow, egress from the bone marrow, homing to the thymus, a combination of those three effects, or others. This *Cosmc* requirement in early T cell maturation combined with our present findings indicate that proper elaboration of O-glycans is essential to all stages of T cell development, from early differentiation in the bone marrow to maturation in the thymus and to appropriate homing to immune organs. Investigating the mechanism by which extended O-glycosylation modulates each of these stages of T cell development will illuminate the role of carbohydrates in: T cell survival signaling events; the quality control systems ensuring only correctly mature T cells are released into the body; and in the signaling pathways influencing T cell homing.

We established that murine L-selectin (CD62L) is upregulated disproportionately on mature Tn+ thymocytes (**Fig 2.3**). In addition to serving as a marker of thymocyte maturation, CD62L is an adhesion molecule necessary for the homing of T cells to SLO. CD62L binds to O-glycosylated ligands on endothelial cells, but it is unknown whether O-glycans themselves are present on CD62L, and whether they play any role in protein function. Prediction algorithms indicate that murine, but not human, L-selectin might bear an O-glycosylation site on the

protein's stem region (Steentoft C 2013). Thus, it is possible that the defect in T cell homing in the TCKO mouse model is partially mediated by dysfunction of L-selectin due to improper glycosylation. Investigating the impact of post translational modification on selectin function could lead to better understanding of leukocyte tethering and homing and provide avenues to manipulate the trafficking of immune cells.

We employed the well-established adoptive transfer method to analyze homing and migration characteristics of T cells (Matheu MP 2011). We found that TCKO T cells engrafted at less than 30% the efficiency of WT T cells. While adoptive transfer is widely accepted in the field, it is primarily used for the transfer of cells isolated from spleen or other SLO. A few others have undertaken adoptive transfer of mature thymocytes to determine the homing and survival of recent thymic emigrants (Houston, Higdon, and Fink 2011; Kim et al. 2016). In agreement with their findings, we found that mature $CD4^+$ SP thymocytes are less competent at entering SLO than splenic $CD4^+$ T cells on a short time scale (16h) corresponding to tissue homing (**Fig 2.5**). We also discovered that TCKO cells are further deficient relative to WT counterparts, only entering the SLO at approximately one fourth the efficiency of WT cells. An important caveat to interpreting the findings that TCKO cells engrafted to SLO at lower rates than WT T cells, is that the TCKO cells are also less numerous in the blood, meaning that the effect is not simply due to a defect in homing. If the only effect of *Cosmc* deletion in TCKO cells was reduced ability to enter SLO, there would be an enrichment of TCKO cells in the blood. Thus, the data support our conclusion that T cells expressing the Tn antigen are actively removed from circulation. Indeed, a series of early transplant studies support targeted removal of partly de-glycosylated T cells from circulation. In these studies, splenocytes and thymocytes were desialylated prior to transfusion into recipient rats. Desialylated splenocytes successfully arrived to recipient SLO at

approximately half the frequency of mock treated cells, and a corresponding increase was seen in the number of cells sequestered in the liver (Gesner and Ginsburg 1964; Woodruff and Gesner 1969). In the case of thymocytes, the effect of desialylation was even greater with near complete abrogation of splenic localization of the transfused cells, with the bulk of transfused cells located in the liver (Berney and Gesner 1970). While these classic experiments did not distinguish between lymphocyte subsets, the results are consistent with specific capture of T cells in the liver. This effect would be unique to T cells with perturbed glycosylation, as no loss or hepatic capture of circulating B cells, granulocytes, erythrocytes, or platelets has been observed in studies with genetic or enzymatic alteration of glycosylation in those cells (Wang et al. 2012; Wang et al. 2010). Thus, if there is a specific and novel glycan-receptor in the liver for partly-de-glycosylated T cells or TCKO cells, it cannot simply recognize the Tn antigen. The targeted removal of Tn antigen expressing T cells that we observe indicates that the effect is T cell specific and not a general response to cells expressing Tn antigen (**Fig 2.5C**).

We suspect that Tn+ T cells are being captured and removed from circulation in the liver. The liver is a presumed site of homeostatic T cell death (Crispe and Huang 1994; Mehal, Azzaroli, and Crispe 2001) and the majority of thymocytes transplanted by Berney et al localized to the liver (Berney and Gesner 1970). However, despite these suggestive findings, the swift disappearance of dying cells in our adoptive transfer experiments makes it difficult for us to conclusively determine and how the Tn+ T cells would be specifically targeted. The canonical mediator of glycan-based removal of circulating factors is the Ashwell-Morell receptor (AMR), which is expressed primarily in the liver and is responsible for extracting prothrombotic factors from the blood, including von-Willebrand factor (Grewal 2010). However, the pathway for AMR-mediated removal has never been shown to function on cell-sized targets, and the known

ligands for AMR are moieties of N-linked glycans, so it is unlikely that AMR is mediating the sequestration of TCKO T cells in the liver.

It is possible that altered glycosylation of cell surface proteins makes TCKO mutant T cells targets for receptor-mediated recognition and destruction. Alternatively, the truncated glycans could influence the innate hardiness and durability of the circulating cells, resulting in increased fragility to stressors of circulation. Whether the liver mediates removal of Tn positive T cells, and how Tn expressing T cells are selectively targeted are exciting mechanistic questions to pursue.

Our discovery that truncated O-glycosylation adversely affects both $CD4^+$ and $CD8^+$ T cells contrasts prior work on two other enzymes in the pathway. Both ST3Gal-I and C2GnT act immediately after T-synthase, catalyzing the addition of a sialic acid (Sia) or an N-acetylgalactosamine (GalNAc) residue, respectively. When ST3Gal-I was deleted, preventing the sialylation of Core-1 O-glycans, $CD8^+$ T cells were almost completely absent from blood and SLO, while peripheral $CD4^+$ T cells were unaffected; neither CD4 nor CD8 SP thymocytes were reduced in number (Priatel et al. 2000). The authors ruled out a direct effect on CD8 itself or the associated signaling through the TCR, so the reason for the dichotomy between $CD4^+$ and $CD8^+$ cells remains unclear (Kao and Sandau 2006). The ST3Gal-I KO partially phenocopies our TCKO model, and the similar mechanisms may be responsible for the loss of $CD8^+$ T cells in both models, especially since we observed enhanced clearance of Tn expressing $CD8^+$ T cells relative to CD4 T cells in TCKO animals. In contrast, ectopic expression of C2GnT, forcing the branching of Core 1 glycans into Core 2 based structures, resulted in no defect of T cell development but did cause reduced immune function in affected T cells (Tsuboi and Fukuda 1997). That result demonstrates that the composition and structure of O-glycans on T cells

impacts and modulates the cell-cell interactions necessary for effector function. However, our work shows that beyond modulation of signals the presence of those glycans is also a requirement for survival.

Taken together our results demonstrate that T cells require extended O-glycosylation for proper physiology in the thymus and SLO. One of the most commonly studied cultured T cell lines, the Jurkat cells, express the Tn antigen (Piller, Piller, and Fukuda 1990), due to dramatically reduced expression of T-synthase resulting from a mutation in the *Cosmc* gene (Ju and Cummings 2002), resulting in near-complete loss of enzyme activity and global expression of the Tn antigen. Much of the foundational work describing T cell receptor signal ligand effector function has been carried out in this T-synthase deficient cell line and likely differs substantially from T-synthase-sufficient cells and from physiological T cell activity. This highlights a potentially major caveat of the many studies regarding T cell signaling and response carried out in Jurkat cell lines and the need to examine T cell function utilizing primary T cells *ex vivo* or *in vivo* under physiological conditions.

2.5 Materials and Methods

Animals: Mice were kept in a specific pathogen-free barrier facility. Animal studies were performed according to the Institutional Animal Care and Use Committee protocol approved by Emory University and Beth Israel Deaconess Medical Center and in compliance with guidelines from the National Institutes of Health. T cell specific Cosmc knock out mice were generated by crossing *Cosmc*$^{flox/flox}$ animals (Wang et al. 2012) with *Lck-Cre* transgenic mice (JAX Lab

#003802, (Hennet et al. 1995)). Mouse genotypes were determined by PCR of DNA from tail snips, using the primer sequences: CosmcFor: GCAACACAAAGAAACCCTGGG, CosmcRev: TCGTCTTTGTTAGGGGCTTGC, LckCreFor: TGTGAACTTGGTGCTTGAGG, LckCreRev: CAGGTTCTTGCGAACCTCAT.

T cell isolation: Blood was collected by cheek bleed. For organ collection, mice were euthanized by CO_2 asphyxiation and organs of interest were collected, including the thymus, spleen, mesenteric, axial, inguinal, and brachial lymph nodes. Organs were pressed through a 70 µm cell strainer in cold phosphate buffered saline, supplemented with 2% FBS. Splenocytes and whole blood were treated with red blood cell lysis buffer. Once isolated, cells were immunolabeled with a panel of antibodies against cell surface markers CD4 (RM4-5, Biolegend 100553), CD8 (5H10, Invitrogen MCD0804), CD90.2 (30-H12, Invitrogen 12-0903-82), CD62L (MEL-14, Biolegend 104405), CD24 (30-F1, Invitrogen 12-0241-82), Qa2 (695H1-9-9, Biolegend 121703), Tn antigen (Bags6 (Cao et al. 1995)), VVA lectin (Vector Labs B-1235). Secondary staining with anti-IgM (Invitrogen A 21042), anti-IgG (Invitrogen A-11029), and streptavidin (Biolegend 405206) was used when necessary. Live/Dead staining was performed with propidium iodide (Invitrogen BMS500PI). Whole blood was analyzed on a Drew Scientific Hemavet to obtain CBC counts.

Flow cytometry: Cells were suspended in PBS containing 2% FBS at 10^7/ml and immunolabeled on ice with appropriate panels of antibodies at 1-5 ug/ml, washed, and analyzed on BD LSR II or Beckman Coulter CytoFLEX LX flow cytometer. Data were collected with FACSDiVa (BD Biosciences) and analyzed using Flowjo (v 10) (BD Biosciences) software.

Adoptive transfer: Splenic T cells were isolated from WT mice by magnetic depletion (CD4 T cell isolation kit, Miltenyi 130-104-454). Thymocytes isolated from WT and TCKO mice and were enriched for CD4 single positive cells by magnetic depletion against CD8 (CD8 microbeads, Miltenyi 130-117-044). WT and TCKO cells were then fluorescently labeled with 1:1,000 and 1:10,000 dilution of CellTrace Violet (Biolegend 100553) for TCKO and WT cells, respectively. TCKO and WT cells were mixed 1:1. A total of 10^8 cells were injected into WT hosts through the tail vein. Hosts were sacrificed after 16 hours and blood and lymphoid organs collected and analyzed by flow cytometry as described above.

Glycosyltransferase Assays: T-synthase activity of purified T-synthase and cell extracts were measured using the acceptor GalNAcα-phenyl as previously described (Ju, 2013). Briefly, cell extract or purified enzyme was incubated with 100 mM MES, pH 6.8, 0.2% Triton X-100, 20 mM MnCl2, 1 mM GalNAcα1-*O*-phenyl, 0.4 4-methylumbelliferyl-UDP-Gal, 2 mM ATP, and excess O-glycosidase, at 37 °C for 60 min and stopped by raising the pH to 9.6. Fluorescence of released 4MU was measured on a PerkinElmer Victor3 plate reader (Ex: 355 nm; Em: 460 nm.)

Statistical Analysis: Data were analyzed for statistical significance by Mann-Whitney, t-test, and ANOVA as indicated in the individual figure legends. Tests were performed using GraphPad Prism version 8 (Graphpad Software). A p-value of less than 0.05 was considered statistically significant. Regression was performed in Microsoft Excel, using the power fit and R-squared values are reported for each regression.

Acknowledgements

The authors gratefully acknowledge the help and support of Jamie Heimburg-Molinaro in preparing this manuscript, and Raj Aryal and Melinda Hanes for use of recombinant T-synthase. BaGs6 antibody was a kind gift from the late Georg Springer. Support for this work is from the NIH (Grants P41GM103694, R01AI101982, P01HL085607, and U01CA168930 to RDC).

2.6 Tables and Figures

Table 2.1: Total cell counts

		WT			TCKO		
Organ	*Cell Type*	*Median Cell Count*	*n*	*SD*	*Median Cell Count*	*n*	*SD*
Thymus	DN	4.60E+06	19	6.23E+06	5.11E+06	20	1.63E+07
Thymus	DP	7.67E+07	19	3.11E+07	5.18E+07	20	2.32E+07
Thymus	SP4	7.03E+06	19	3.77E+06	4.22E+06	20	1.26E+07
Thymus	SP8	1.90E+06	19	2.89E+06	1.07E+06	20	1.40E+07
Blood	B	7.29E+03	53	2.28E+03	7.21E+03	27	2.83E+03
Blood	CD4	1.03E+03	53	3.61E+02	1.21E+02	27	2.54E+02
Blood	CD8	6.83E+02	53	2.20E+02	1.45E+02	27	2.65E+02
Spleen	B	5.08E+07	18	2.71E+07	5.10E+07	15	1.98E+07
Spleen	CD4	1.62E+07	18	6.97E+06	3.50E+06	15	5.07E+06
Spleen	CD8	1.08E+07	18	1.16E+07	3.88E+06	15	4.24E+06
PLN	B	1.86E+06	5	9.51E+05	1.41E+06	4	1.54E+06
PLN	CD4	1.93E+06	5	9.41E+05	3.93E+05	4	6.47E+05
PLN	CD8	1.98E+06	5	7.57E+05	4.53E+05	4	4.56E+05
MLN	B	1.62E+06	4	9.08E+05	2.62E+06	5	2.30E+06
MLN	CD4	1.59E+06	4	9.00E+05	5.94E+05	5	5.29E+05
MLN	CD8	1.01E+06	4	9.65E+05	5.33E+05	5	4.33E+05

Notes: DN= double negative, DP= double positive, SP= single positive, SD= standard deviation

Table 2.2: Percent Tn positive cells

		WT			**TCKO**		
Organ	*Cell Type*	*Median Percent Tn positive*	*n*	*SD*	*Median Percent Tn positive*	*n*	*SD*
Thymus	DN	0.335	20	0.5734	2.810	23	5.105
Thymus	DP	0.055	20	0.8024	14.500	23	16.76
Thymus	SP4	0.335	20	0.5415	52.600	23	24.56
Thymus	SP8	0.105	20	0.7404	23.400	23	22.29
Blood	CD4	0.000	88	2.669	4.700	63	23.52
Blood	CD8	0.000	98	2.253	1.000	64	11.53
Spleen	CD4	0.044	19	0.3522	4.980	17	12.72
Spleen	CD8	0.024	19	0.2593	1.000	17	6.81
PLN	CD4	0.035	9	0.02917	2.450	9	2.427
PLN	CD8	0.008	9	0.01833	0.110	9	0.2215
MLN	CD4	0.062	9	3.761	3.825	10	23.88
MLN	CD8	0.088	9	1.48	0.770	10	4.283

Notes: DN= double negative, DP= double positive, SP= single positive, SD= standard deviation

Figure Legends

Figure 2.1 Generation of a T cell specific *Cosmc* knockout mouse. A Genetic approach to generate a T cell specific *Cosmc* knock out (TCKO) by crossing a mouse carrying *Cosmc* flanked by *loxP* sites to a mouse driving Cre expression from the proximal *Lck* promoter. **B** Representative PCR genotyping results showing the presence of the WT and/or floxed *Cosmc* allele (top) and the presence or absence of the transgenic *Lck Cre* (bottom). **C** T-synthase enzymatic activity of purified T-synthase and lysates from immortalized human neutrophil cell line (HL-60), immortalized human T cell cell line with deficient T-synthase activity (Jurkat), T cells isolated from low penetrance EHCKO mice, and T cells isolated from the TCKO mice. Analysis by one-way ANOVA, n=2. **D** Populations of lymphocytes, granulocytes, and monocytes in blood of WT and TCKO mice measured by Complete Blood Count (CBC) analysis. Analysis by t-test with Bonferroni correction for multiple comparisons. WT n=15, TCKO n=8, * indicates $p< 0.05$, ns=not significant.

Figure 2.2 Analysis of T cells from TCKO blood, spleen, and lymph nodes. A The gating strategy and representative populations for flow cytometric analysis of lymphocyte populations. **B-E** (left panels) The cell counts of non-T cells, $CD4^+$ T cells, and $CD8^+$ T cells was measured for WT and TCKO mice in blood (**B**), spleen (**C**), peripheral lymph nodes (inguinal, axial and brachial lymph nodes (PLN)) (**D**), and mesenteric lymph nodes (MLN) (**E**). **B-E** (right panels) The percent of $CD4^+$ and $CD8^+$ T cells that express detectable Tn antigen in each organ as noted, for WT and TCKO mice**. F** The ratio of $CD4^+$ T cells to $CD8^+$ T cells in secondary lymphoid organs as noted for WT and TCKO mice was calculated from the flow cytometry data. Blood: WT n=53, TCKO n=27, Spleen: WT n=18, TCKO n=15, peripheral LN: WT= 5, KO=4, MLN:

WT n=4, TCKO n=5. All comparisons by Mann Whitney test with Bonferroni correction for multiple comparisons. ns = not significant.

Figure 2.3: Effect of TCKO on thymocyte populations. A Gating strategy for thymocyte analysis. **B** Absolute counts of thymocytes from each stage of T cell maturation (DN, DP, SP) recovered from the thymus of WT and TCKO quantified by flow cytometry. Median and quartiles are indicated by dashed lines. **C** Percent of thymocytes that were positive for Tn antigen are presented for each stage of T cell maturation. Median and quartiles are indicated by dashed lines. **D-E** Representative flow plots of CD4 SP thymocytes from WT (**D**) and TCKO (**E)** for Tn antigen and CD62L expression. **F** Quantification of CD4 SP thymocytes analyzed for simultaneous expression of L-selectin (CD62L) and Tn antigen (n=5). **G** Percent of CD4 SP cells expressing high levels of maturation markers (CD62L, CD24, Qa-2) in WT or TCKO mice were analyzed. Error bars represent standard deviation. For **B** and **C**, WT n=19 TCKO n= 20. For **F** and **G**, n=5 for both WT and TCKO. Comparisons for **B, C,** and **F** were made by Mann Whitney test with Bonferroni correction for multiple comparisons, and for **G** a 2-way ANOVA was used; * indicates $p<0.05$, ns = not significant.

Figure 2.4: Difference in Tn expressing cells across T cell maturation. A Total number of CD4 T cells in the blood, spleen, and thymus of individual TCKO mice, plotted against the percent of those cells expressing the Tn antigen. Trendline is a power curve fit, with R^2 value displayed. **B** Visual representation of the difference in total cell count between TCKO and WT mice, calculated by dividing the median WT count by the median TCKO count for B cells, $CD4^+$ T cells, and $CD8^+$ T cells. Bars below the axis indicate a reduced number of cells in TCKO. **C** Percentage of T cells expressing Tn antigen at specified development stages and compartments. Mean values and standard error are plotted for $CD4^+$ and $CD8^+$ populations. Statistical

differences between CD4 and CD8 cells are indicated, as calculated by Student's t test with Bonferroni correction for multiple comparisons.

Figure 2.5 Adoptive transfer of WT and TCKO Thymocytes. A Thymocytes from WT and TCKO mice were isolated, fluorescently labeled, then mixed at a 1:1 ratio before transfer into WT hosts by tail vein injection. Hosts were sacrificed 16 hours post injection, and their tissues analyzed by flow cytometry for the presence of transferred cells. **B** Quantification of distribution of adoptively transferred WT and TCKO CD4 SP thymocytes and WT CD4 splenocytes, calculated as a percent of injected ratio relative to host CD4 T cells, normalized to transferred splenocytes in blood. Two outliers were excluded by Grubbs test, and statistical significance calculated by Student's t test with Bonferroni correction for multiple comparisons. * indicates $p<0.05$, ns = not significant. **C** Proposed model of Tn antigen positive T cell clearance from circulation.

Figure 2.1

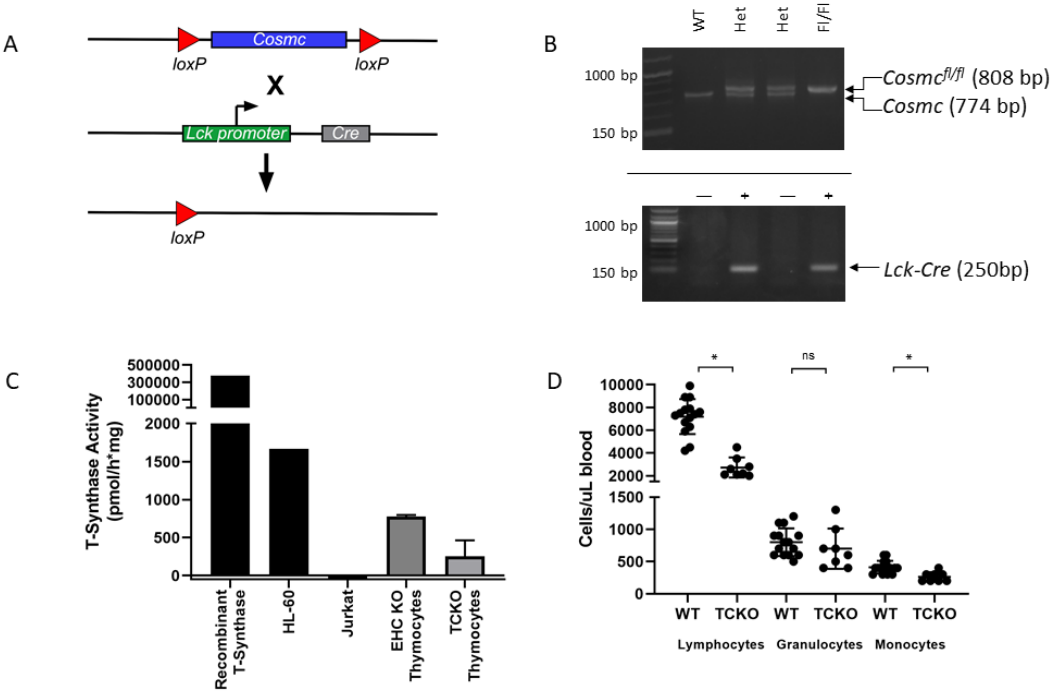

A
Cosmc
loxP
loxP
X
Lck promoter
Cre
loxP
B
WT
Het
Het
Fl/Fl
1000 bp
150 bp
Cosmc fl/fl (808 bp)
Cosmc (774 bp)
— + — +
1000 bp
150 bp
Lck-Cre (250bp)
C
T-Synthase Activity (pmol/h*mg)
500000
300000
100000
2000
1500
1000
500
0
Recombinant T-Synthase
HL-60
Jurkat
EHC KO Thymocytes
TCKO Thymocytes
D
Cells/uL blood
10000
8000
6000
4000
2000
1500
1000
500
0
*
ns
*
WT
TCKO
WT
TCKO
WT
TCKO
Lymphocytes
Granulocytes
Monocytes

Figure 2.2

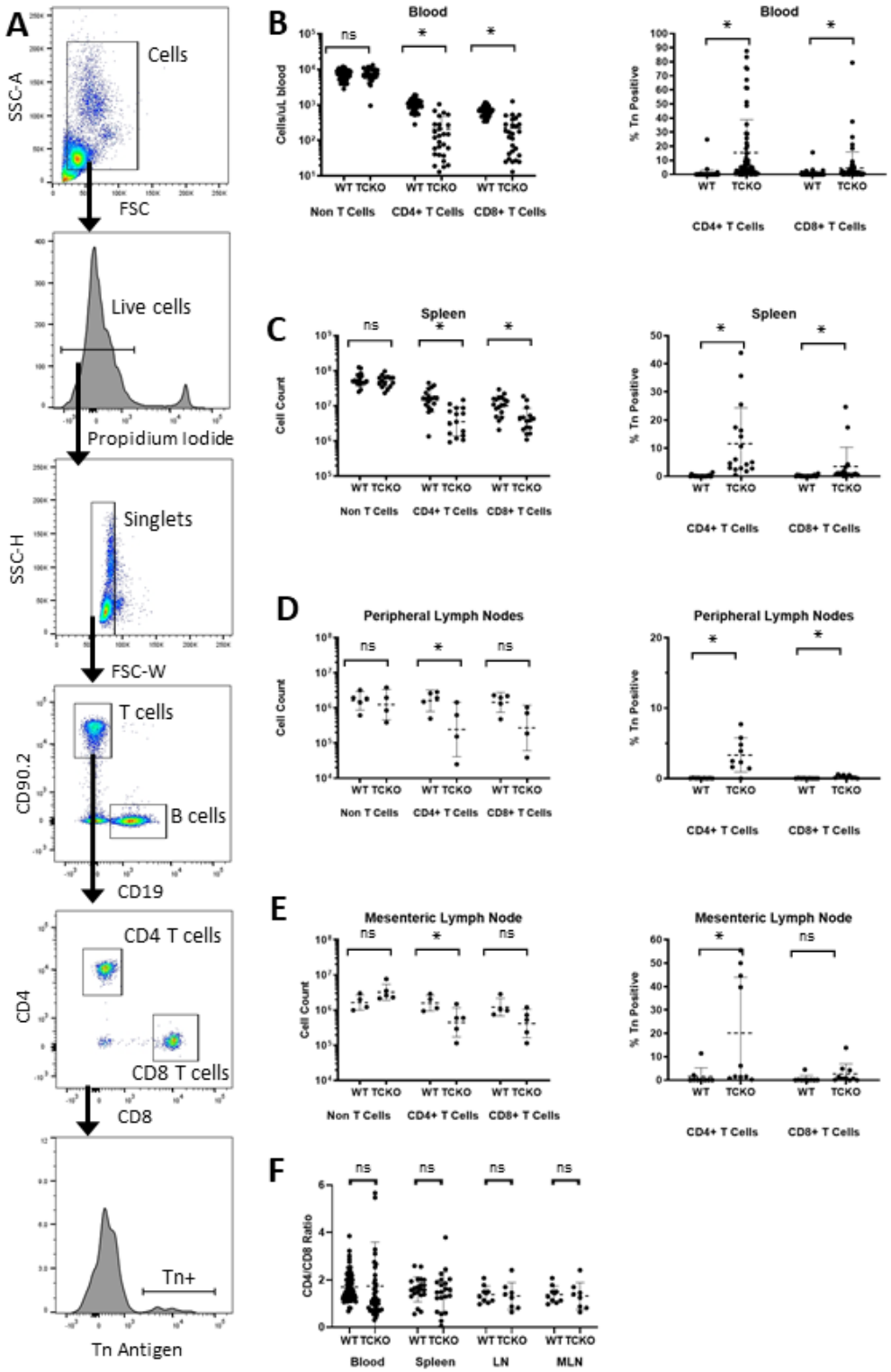

A
SSC-A
Cells
FSC
Live cells
Propidium Iodide
SSC-H
Singlets
FSC-W
CD90.2
T cells
B cells
CD19
CD4
CD4 T cells
CD8 T cells
CD8
Tn+
Tn Antigen
B
Blood
Cells/uL blood
ns
*
WT TCKO
Non T Cells
CD4+ T Cells
CD8+ T Cells
% Tn Positive
C
Spleen
Cell Count
D
Peripheral Lymph Nodes
E
Mesenteric Lymph Node
F
CD4/CD8 Ratio
Blood
Spleen
LN
MLN

Figure 2.3

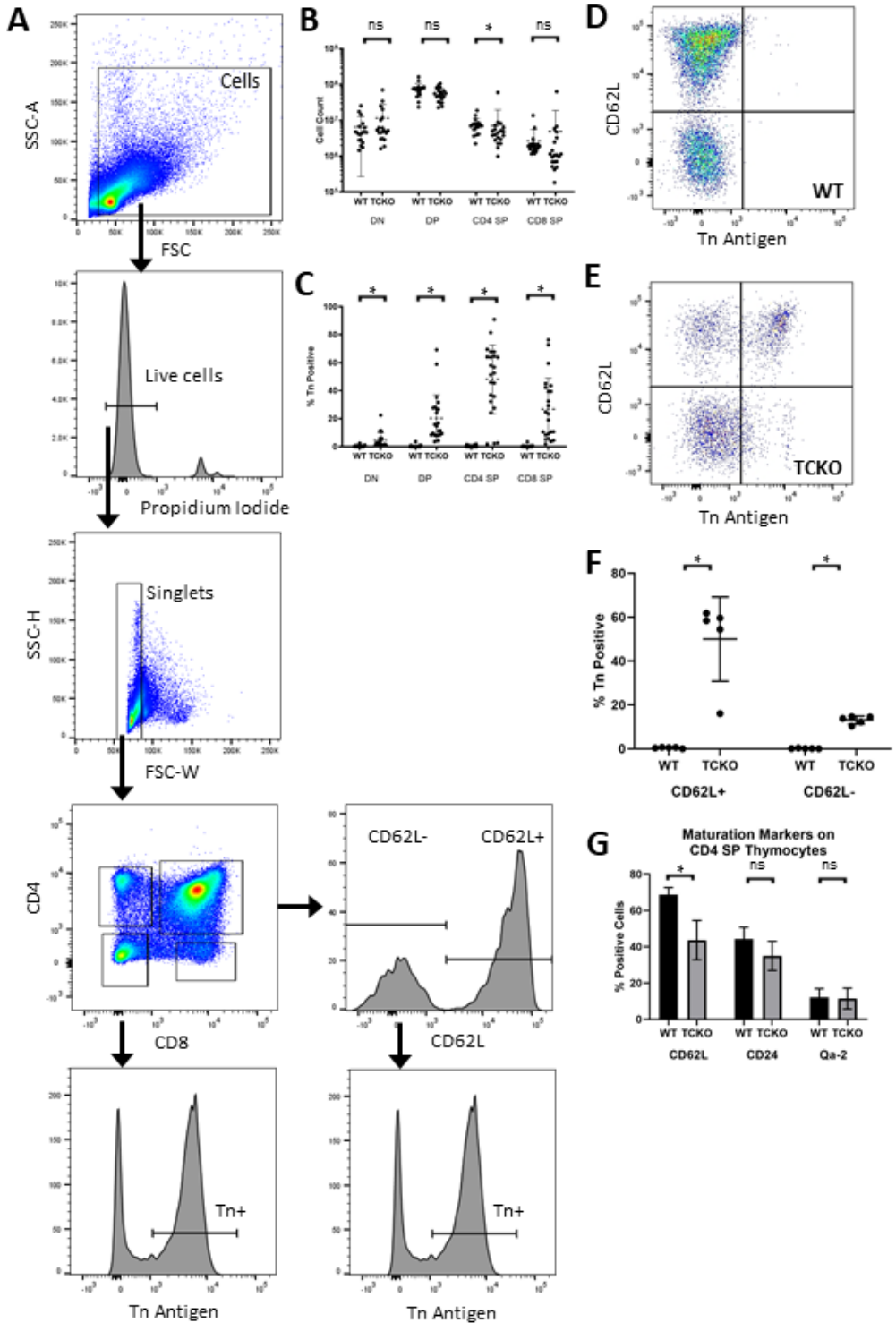

A
SSC-A
Cells
FSC
Live cells
Propidium Iodide
SSC-H
Singlets
FSC-W
CD4
CD8
Tn+
Tn Antigen
CD62L-
CD62L+
CD62L
Tn+
Tn Antigen
B
ns
ns
*
ns
Cell Count
WT TCKO
DN
DP
CD4 SP
CD8 SP
C
*
% Tn Positive
D
CD62L
WT
Tn Antigen
E
CD62L
TCKO
Tn Antigen
F
% Tn Positive
WT
TCKO
CD62L+
CD62L-
G
Maturation Markers on CD4 SP Thymocytes
% Positive Cells
CD62L
CD24
Qa-2

Figure 2.4

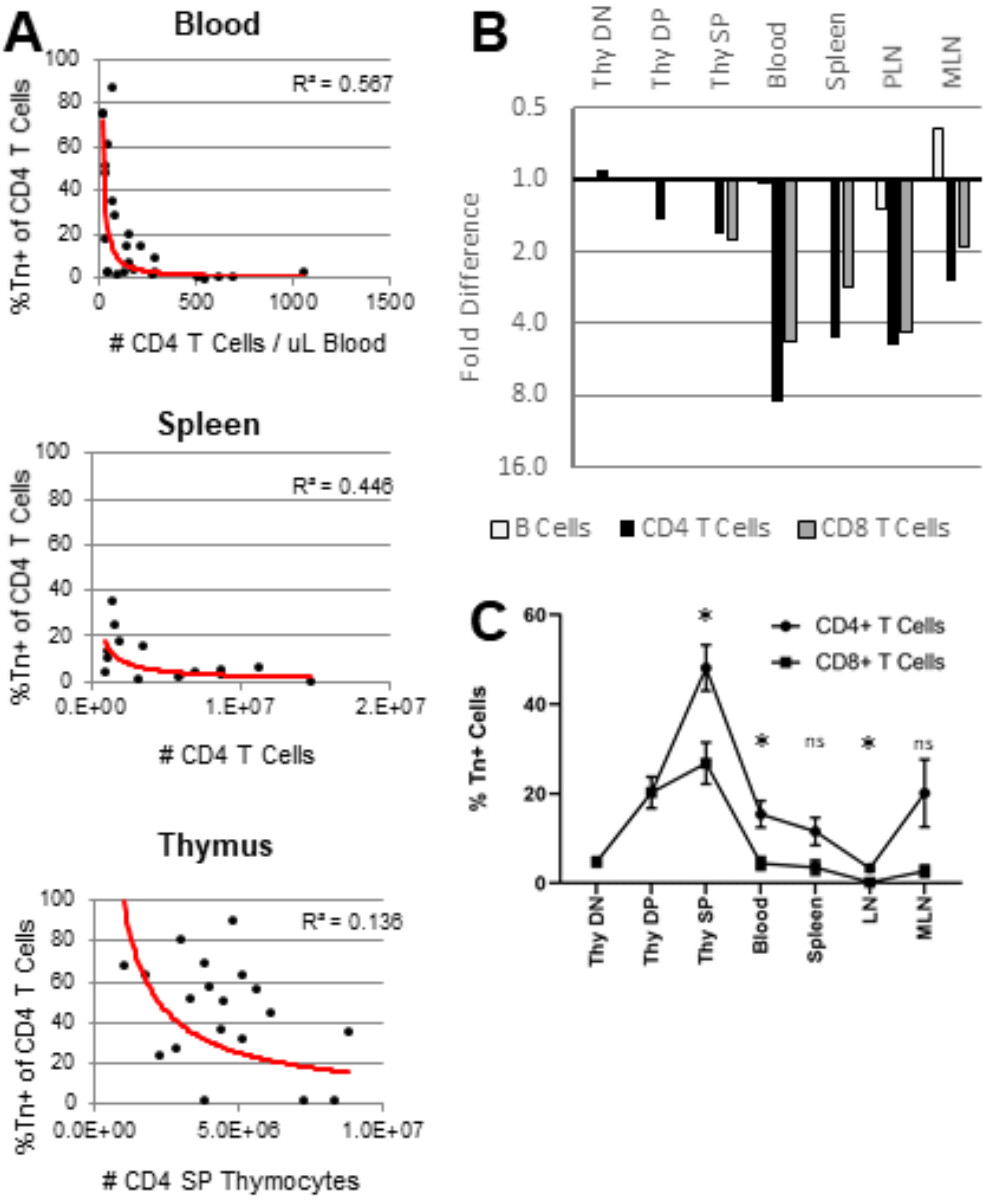
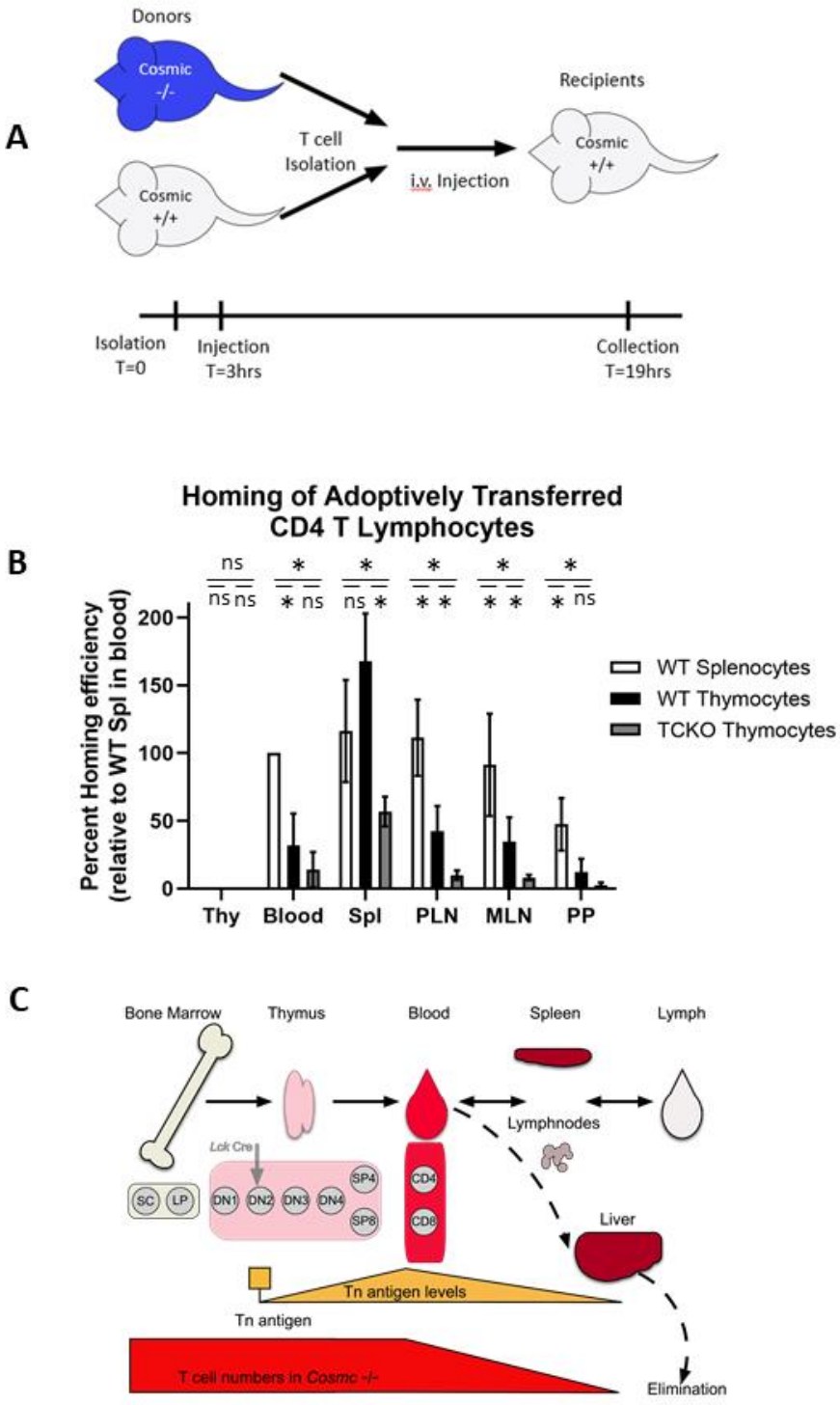

A
Blood
R² = 0.567
%Tn+ of CD4 T Cells
CD4 T Cells / uL Blood
Spleen
R² = 0.446
%Tn+ of CD4 T Cells
CD4 T Cells
Thymus
R² = 0.136
%Tn+ of CD4 T Cells
CD4 SP Thymocytes
B
Thy DN
Thy DP
Thy SP
Blood
Spleen
PLN
MLN
Fold Difference
B Cells
CD4 T Cells
CD8 T Cells
C
CD4+ T Cells
CD8+ T Cells
% Tn+ Cells
ns
Thy DN
Thy DP
Thy SP
Blood
Spleen
LN
MLN

Figure 2.5


A
Donors
Cosmic -/-
Cosmic +/+
T cell Isolation
Recipients
Cosmic +/+
i.v. Injection
Isolation T=0
Injection T=3hrs
Collection T=19hrs
B
Homing of Adoptively Transferred CD4 T Lymphocytes
Percent Homing efficiency (relative to WT Spl in blood)
200
150
100
50
0
ns ns ns
* ns
ns *
* *
* *
* ns
Thy
Blood
Spl
PLN
MLN
PP
WT Splenocytes
WT Thymocytes
TCKO Thymocytes
C
Bone Marrow
Thymus
Blood
Spleen
Lymph
Lymphnodes
Lck Cre
SC
LP
DN1
DN2
DN3
DN4
SP4
SP8
CD4
CD8
Liver
Tn antigen levels
Tn antigen
T cell numbers in Cosmc -/-
Elimination

Chapter 3: A Sensitive Glycoproteomic Method for Identifying O-glycosylated Proteins Using Low Sample Quantities

3.1 Abstract

Abnormal glycosylation occurs in a majority of human carcinomas and is associated with adverse outcomes. One essential step in exploring the glycobiology of cancers is identification of glycoproteins and the specific sites of modification. Unlike N-glycans, O-glycans are not added to conserved sequons and prediction of sites of O-glycosylation remains imprecise despite recent improvements in prediction algorithms. Current state-of-the-art techniques for O-glycoproteomic identification rely on protein enrichment via weak lectin affinity chromatography, which requires large sample amounts of genetically modified cell lines. To be clinically relevant, identification of aberrant glycosylation patterns must be carried out on small patient biopsies. We developed a chemoenzymatic labeling approach that would allow for higher affinity isolation and enhanced sensitivity, permitting reduced sample quantities. We demonstrate that recombinant human ST6GalNAc1 can utilize biotin-linked CMP-sialic acid donor sugar to label target substrates. We also discover that ST6GalNAc1 contains a heavily glycosylated mucin-like domain and autocatalyzes its own sialylation. A truncated form of the enzyme containing only the glycosyltransferase domain functions well for labeling glycopeptides, glycoproteins, whole cells, and cell lysates.

3.2 Introduction

Post-translational glycosylation influences a host of glycoprotein properties including folding, conformation, oligomerization, binding, and proteolytic resistance. Because over half of the proteome is glycosylated (Apweiler, Hermjakob, and Sharon 1999), perturbed or dysregulated glycosylation has immense potential to adversely affect nearly all cellular and physiological processes, and indeed has been linked to a wide range of pathologies (see Chapter 1.2.3). In particular, cancer is frequently associated with major changes in glycosylation (Peracaula et al. 2008; Hakomori and Cummings 2012).

One exceedingly common abnormal glycan that arises in cancer is the Tn antigen, composed of α-linked N-acetylgalactosamine (GalNAc) on serine or threonine, and its α2,6-sialylated derivative sialyl-Tn (sTn) (Hakomori 2001). Tn antigen is a biosynthetic precursor to all GalNAc-type O-glycans, which are predicted to occur on 80% of proteins entering the secretory pathway, but in healthy tissue Tn antigen is invariably extended with other saccharide residues and is not exposed on the surface of healthy tissue(Steentoft C 2013). In contrast, Tn and sTn appear on more than 80% of human carcinomas (Springer 1984; Ju et al. 2008), at rates varying by tumor type, and are associated with poor prognosis in many classes of cancer. (Numa et al. 1995; Laack et al. 2002; Fernandez et al. 2005; Kakeji et al. 1991; Konno et al. 2002). While a mechanism by which these abnormal glycans contribute to tumor progression is not yet known, there are indications from animal models that defects in the O-glycosylation pathway that lead to the formation of Tn can directly promote oncogenesis in some tissue types (Hofmann et al. 2015; Fu et al. 2016).

The cause of Tn antigen appearance in tumors is most commonly traced to loss of expression of the chaperone Cosmc, either by direct mutation of the *Cosmc* (*C1GalT1C1*) gene,

mutation of the promoter region, or epigenetic silencing (Ju et al. 2008; Yu et al. 2015). Cosmc is required for biosynthesis of active T-synthase enzyme, which is responsible for extending Tn antigen into the T antigen, also known as the core 1 structure, by addition of a β1,3 linked galactose (Gal)(Ju and Cummings 2002). Core 1 is the biosynthetic foundation of nearly all GalNAc type O-glycans, and loss of functional T-synthase due to lack of Cosmc chaperone results in truncation of O-glycans to the Tn monosaccharide. In some cases, Cosmc-competent tumor cells also express some Tn antigen, which has been attributed to mislocalization of glycosylation enzymes in the secretory pathway (Sun, Ju, and Cummings 2018; Christiansen et al. 2014; Gill et al. 2013).

The sTn antigen is created by addition of sialic acid to Tn antigen by the activity of the enzyme ST6GalNAc1, and sTn expression in tumors is associated with high expression of ST6GalNAc1 (Tsuji and Takashima 2014; Sewell et al. 2006; Patani, Jiang, and Mokbel 2008). When analyzed in vitro, ST6GalNAc1 is also capable of acting on core 1 to form sialyl T (sT) antigen, but its *in vivo* activity is restricted to converting Tn to sTn (Marcos et al. 2004). The related enzyme ST6GalNAc2 is also capable of performing this reaction to sialylate sTn and sT *in vitro*, but in counterpart to ST6GalNAc1 its observed *in vivo* activity is limited to conversion of T to sT (Samyn-Petit et al. 2000; Marcos et al. 2004).

Knowing which proteins bear which glycan modification is essential to understanding their function and the ways in which those functions can be perturbed by alterations in glycosylation. Perturbed glycosylation can lead not only to altered glycan composition, but also to changes in glycosite occupancy across a protein. Thus, it is necessary to directly analyze primary samples in order to identify glycosylation differences between patient tumors and cultured cell lines. Recent advances in proteomic identification of O-glycoproteins and their sites

of glycosylation using weak lectin affinity chromatography with Cosmc knockout “simple cells” (SC) has allowed for the refinement of O-glycosite prediction algorithms(Steentoft, Vakhrushev, Vester-Christensen, Schjoldager, et al. 2011). However, these techniques require large amounts of sample that are infeasible for primary tissues and cells. Here we demonstrate a sensitive technique for proteomic identification of O-glycoproteins that could be utilized reliably with minute sample quantities.

3.3 Results

3.3.1 Recombinant ST6GalNAc1 is active, accepts biotin-linked CMP-Neu5Ac, and acts to autosialylate itself.

In order to identify sites of O-glycosylation, we devised a system to tag all available O-glycosylation sites with biotin. ST6GalNAc1 preferentially sialylates Tn antigen monosaccharide and to a lesser extent T antigen disaccharide if enzyme concentrations are high enough (cite). To take advantage of this enzyme activity in an *in vitro* assay, we designed a His tagged secreted recombinant human ST6GalNAc1 and expressed it in HEK293 Freestyle (239F) cells, then purified from conditioned media (Fig 3.1A). To ascertain whether the purified ST6GalNAc1 was functional, the enzyme was combined with a biotinylated donor sugar (5’biotinyl-CMP-Neu5Ac, structure in Fig 3.1B) and HEK SimpleCell (SC) lysate *in vitro.* SC lines are glycoengineered *Cosmc* gene knockouts that homogeneously express truncated O-GalNac glycoproteins due to lack of T-synthase activity (Steentoft, Nat met, 2011), and therefore contain a wide variety of suitable substrates for ST6GalNAc1 to modify. The enzyme activity assay revealed that the purified ST6GalNAc1 was functional and catalyzed addition of biotinylated sugars onto proteins (Fig 3.1B). However, unexpectedly, rather than detecting multiple bands corresponding to the abundant substrates in HEK SC lysate, only a major band

appeared on Western blots of the reaction mixture. The lone band corresponded to the size of purified ST6GalNAc1 (Fig 3.1B) and was present in reactions run in the absence of HEK SC lysate. Because the band appeared in conditions with only enzyme and biotinylated sugar donor, we concluded that ST6GalNAc1 was acting to autosialylate itself, rather than incorporate the biotin tag onto other glycoproteins.

Supporting the inference of autosialylation, the band of both the freshly purified ST6GalNAc and the primary band after the enzyme activity assay migrate at 100kDa, based on a protein ladder, substantially larger than the predicted 67 kDa for the peptide sequence alone, indicating that the protein bears substantial post-translational modification, likely to be glycosylation. We therefore performed a gel-shift assay to see if the migration pattern changed after treatment with glycosidases. PNGaseF treatment to remove all N-glycans resulted in an appreciable gel shift, while O-glycanase did not, indicating the presence of N-glycans, but not providing evidence for the O-glycans that could be acceptors of its enzymatic activity. O-glycanase is only active toward the core 1 disaccharide Galβ1-3GalNAcα1-Ser/Thr and cannot release the Tn antigen alone, thus interpretations have to be limited(Fig 3.1C) We performed lectin blots of recombinant ST6GalNAc1 and found it was recognized by both VVA, which recognizes terminal GalNAc including the Tn antigen, and PNA, which recognizes the T antigen (Fig 3.1D). We then directly analyzed the mass of the purified recombinant protein using MALDI mass spectrometry (MS), which revealed an observed mass of 82 kDa (Fig 3.1E). These results indicate that approximately ~15 kDa of carbohydrates are present on ST6GalNAc1, as the recombinant amino acid sequence predicts a protein size of ~67 kDa. The protein contains four predicted sites of N-glycosylation, which typically have a mass of 2-3 kDa each, but can be

larger (cite). Thus O-glycosylation is present at some level detectable by lectin blot and could account for as much as 7 kDa of the observed mass.

To directly evaluate the O-glycosylation of ST6GalNAc1 we analyzed tryptic peptides by liquid-chromatography-linked mass spectrometry (LC-MS). Out of 83 serine and threonine residues in the sequence, 49 were predicted to bear O-glycans (cite NetOGlyc 4.0, which is an extraordinarily number for a non-mucin glycoprotein). Our LC-MS identified a total of 42 occupied sites, 9 of which did not overlap with the software prediction (Table 3.1, Figure 3.2A). These sites were occupied by a variety of glycans including both T and Tn structures and their sialylated forms, with some microheterogeneity at individual sites (Table 3.1, Fig 3.2B&C). Note that the data did not provide linkage or stereochemistry information, so the structures are presumed based on mass. In particular, the sialyl T structures denoted in Table 3.1 and Figure 3.2 could be either form of sT, e.g., the more common standard form with sialic acid linked α2,3 to the terminal Gal, or the rarer alternate form with sialic acid linked α2,6 to the initiating GalNAc. Because ST6GalNAc1 is capable of generating the latter linkage (Marcos et al. 2004), and has been shown to autosialylate (Fig 1), it is likely that some or all of the sT trisaccharides observed here take the alternate form. Taken together, these findings reveal a heavily O-glycosylated domain on ST6GalNAc1 that bears a large number of valid targets for its own activity to autogenerate the sTn antigen, supporting the unexpected autocatalytic activity observed.

The unexpected presence of a highly decorated mucin-like domain on ST6GalNAc1 prompted the question of whether this was a unique feature of this enzyme, or if it might be shared by other sialyltransferases. We compared predicted O-glycosylation results from the NetOGlyc 4.0 prediction algorithm for the human sialyltransferases, a large number of enzymes (Table 3.2). We found that among human sialyl transferases ST6GalNAc1 is unique in its high

degree of O-glycosylation. The 49 O-glycosylation sites predicted for ST6GalNAc1 is almost three times more than the sialyltransferase with the second most predicted sites (ST6Gal2; 18 sites) and more than eight times the median across all human sialyltransferases (6 sites). We also compared predictions for ST6GalNAc1 orthologs in other species (Table 3.3). We found that ST6GalNAc1 orthologs uniformly have high numbers of predicted O-glycosylation sites, with the predicted number of sites ranging from 12 in zebrafish to 69 in bats, and a median of 45 sites among the 12 species analyzed. That this high degree of post-translational modification differs from related enzymes and is highly conserved across orthologs of this glycosyltransferase in many animal species strongly suggests a conserved biological role for the high degree of O-glycosylation on ST6GalNAc1.

3.3.2 Truncation of ST6GalNAc1

Because the observed and predicted sites of O-glycosylation on ST6GalNAc1 clustered into the mucin-like stem region, away from the glycosyltransferase domain, we hypothesized that a truncated form of the enzyme could be utilized for our chemoenzymatic labeling approach (Fig 3.3A). With the goal of avoiding the intense autoglycosylation that eclipsed any other labeling in the previous assay, we designed a truncated construct that only contained the glycosyltransferase domain with a 6-His epitope tag, hereafter ST6GalNAc1-trunc. This truncated protein was expressed in 293F cells and purified from conditioned media (Fig 3.3B). We characterized the full length and truncated constructs using multiple techniques to check for oligomerization. Protein crosslinking showed that ST6GalNAc1-trunc but not ST6GalNAc1 formed a cross-linkable oligomeric form (Fig 3.3C). On blue native PAGE, both ST6GalNAc1 and ST6GalNAc1-trunc exhibited at least one dimer band, with ST6GalNAc1 also appearing as multiple higher-order species (Fig3.3D) These results suggest that the mucin-like stem region

has an anti-adhesive effect, and that without it the truncated glycosyltransferase tends to cluster into homotypic dimers or oligomers. However, evaluation with an analytical size exclusion chromatography column with multi-angle light scattering (SEC-MALS) gave different results, indicating monomeric protein (Fig 3.3D). Quantification of the SEC-MALS data established the protein mass, glycan mass, and hydrodynamic radius for each protein (Table 3.4). The SEC-MALS indicated monodisperse proteins with monomeric molar masses, suggesting that while ST6GalNAc1-trunc has an apparent tendency to oligomerize, this interaction is weak and in the diluting conditions of size exclusion chromatography the binding is reversible and the protein acts as a monomer. The protein and glycan masses calculated from the SEC-MALS data indicate that the ST6GalNAC1-trunc mutant contains 6 kDa less sugar than ST6GalNAc1, which corresponds to our earlier estimates of the contribution of O-glycosylation to the glycoprotein mass.

To determine what cofactors and substrates could participate in the enzymatic reaction of ST6GalNAc1, we also performed thermal shift assays utilizing SYPRO Orange and measured the melting temperature of recombinant ST6GalNAc1 in the presence of various additives. The results are presented in Table 3.5. Neither addition of metals nor their chelation with EDTA altered the Tm of ST6GalNAc1, indicating that metal binding is not a feature of enzyme stability or catalysis for ST6GalNAc1. The exception is zinc, which had a high error range in the Tm reading and appeared to destabilize the protein. In comparison, both free monosaccharide GalNAc and GalNAc alpha-linked to 4MU stabilized the protein indicating binding to ST6GalNAc1, as expected for a viable enzymatic substrate. The sugar donor CMP-Neu5Ac did not alter the protein's Tm. This is not unexpected, as the binding of the high-energy sugar donor

is expected to be less energetically favorable than binding to the post-catalysis products including free CMP and a sialylated substrate.

3.3.3 Enzymatic activity of ST6GalNAc1-trunc on glycopeptides

To measure the activity of ST6GalNAc1-trunc we utilized an HPLC shift assay with a model Tn glycopeptide 2GP-1, which contains a sequence corresponding to human PSGL-1. Enzymatic addition of sialic acid by ST6GalNAc1-trunc resulted in a readily observable earlier elution time of the modified glycopeptide. Our reaction buffer for glycosyltransferase reactions includes metals, which are required by some glycosyltransferases for enzymatic activity, but chelation of metal ions by EDTA did not diminish the activity of ST6GalNAc1-trunc (Fig 3.4A). By sampling the reaction mixture at multiple timepoints, the clear conversion from the Tn glycopeptide to one bearing sialyl Tn was observed, with complete saturation of available substrate reached within seven hours of reaction (Fig 3.4B and 3.4C).

Expanding from experiments with a single glycopeptide, the activity of ST6GalNAc1-trunc was further characterized on a microarray composed of multiple model Tn-glycopeptides. These peptides include regions of Muc1, Muc2, and the IgA stem region, and are synthesized with one or more Tn antigen installed at a variety of sites in order to compare fine specificities of carbohydrate binding proteins and enzymes. Arrays were treated separately with ST6GalNAc1-trunc or T-synthase and probed to determine the glycopeptides labeled by each enzyme. ST6GalNAc1-trunc modified nearly all Tn containing glycopeptides, with a few exceptions marked by red arrows in the figure (Fig 3.4D). Those exceptions include GalNAc attached to serine as a single amino acid, and three different Tn glycosites on a synthesized peptide of the IgA stem region, a structure that we previously characterized as unusually inaccessible to glycosyltransferases(Lehoux et al. 2014). The glycopeptides resistant to labeling by

ST6GalNAc1-trunc were also resistant to modification by T-synthase. In addition, while T synthase acted on many of the model glycopeptides, there were several substrates of ST6GalNAc1-trunc that were not modified by T synthase, indicating that ST6GalNAc1-trunc may be more permissive in its activity than T synthase. This is promising for our proposed method, as a greater breadth of action allows for a wider range of glycoproteins to be labeled for proteomic identification.

3.3.4 Activity of ST6GalNAc-1-trunc on glycopeptides

To confirm that ST6GalNAc1-trunc was only acting to label expected substrates (Tn and possibly core 1 O-glycans), we treated glycoprotein arrays with St6GalNAc1-trunc, with and without glycanase pretreatment (Fig 3.5A). Lectin staining of glycanase treated arrays confirmed the removal of sialic acids by neuraminidase (Fig 3.5B) and the removal of N-glycans by PNGaseF (Fig 3.5C). In contrast, O-glycosidase showed poor activity toward the glycopeptide array, with only a modest reduction in PNA lectin binding (Fig 3.5D). Subsequent action of ST6GalNAc1-trunc on glycanase-treated arrays confirmed that the sialylation activity by the recombinant enzyme was concentrated on the O-glycans (Fig 3.5E). Staining of untreated arrays with three Tn-reactive probes, VVA and HPA lectins, and Rebags6 antibody, partially overlapped staining profile of ST6GalNAc1-trunc treatment (Fig 3.F). Specifically, bovine submaxillary mucin (BSM) was recognized by both ST6GalNAc1-trunc and the Tn-reactive probes. BSM is known to contain high levels of Tn antigen and is an expected substrate of ST6GalNAc1-trunc (Yamada et al. 2010). In addition, IgA, which is a combination of both IgA_1 that contains multiple O-glycans in its hinge region including the Tn antigen, and IgA_2 which lacks O-glycans,In addition, IgA was labeled by ST6GalNAc1-trunc consistent with reported O-glycosylation of the hinge region. Normal IgA_1 is known to contain a mixture of Tn and core 1

glycans both of which could potentially serve as substrates for ST6GalNAc1-trunc. ST6GalNAc1-trunc also labeled casein and neural cell adhesion molecule (NCAM), both known O-glycosylated proteins (Pisano et al. 1994; Suzuki et al. 2003) (Fig 3.5E). The reduction in labeling activity on IgA, NCAM, and casein following O-glycanase treatment indicates that some of the labeling was due to ST6GalNAc1-trunc acting on core 1, as O-glycosidase removes core 1 but not Tn antigen. The incomplete action of O-glycosidase (Fig 3.5D) indicates that we cannot draw firm conclusions as yet about whether residual labeling of these proteins reflects labeling of Tn antigen, as is likely for IgA, or otherwise glycanase-resistant core 1-based structures, as is likely for casein and NCAM. Nevertheless, it does suggest that pretreatment with O-glycanase enhances the specificity of detection toward labeling of Tn antigen, which can be exploited in a glycoproteomic workflow in the event that labeling of core 1 and other extended structures is not desired, such as if one only wishes to identify proteins bearing truncated Tn antigen. In contrast, N-glycanase had no effect on ST6GalNAc1-trunc labeling confirming that there were no off-target sialylation of N-glycans (Fig 3.5 E).

3.3.5 ST6GalNAc1-trunc acts effectively on cells and cell lysates

To assess the activity of the truncated enzyme, we treated WT HEK cells (which present the T antigen) and SC HEK cells (which present the Tn antigen) with ST6GalNAc1-trunc and biotinylated sugar donor, then visualized the incorporation with fluorescent streptavidin binding via flow cytometry (Fig 3.6A). Both cell lines were labeled by the enzyme, consistent with ST6GalNAc1-trunc being capable of sialylating both T and Tn antigens substrates. When the cells were pretreated with O-glycanase, which removes T antigen disaccharide but does not act on Tn antigen monosaccharide, WT cells were no longer labeled with biotin while SC cells were labeled equivalently to those not pretreated with O-glycanase. Titration of the biotinylated sugar

donor revealed that 10µM of CMP-(9'biotinyl)Neu5Ac was sufficient to saturate all modifiable sites on the HEK SC surface (Fig 3.6B), and that concentration was used for subsequent experiments.

To further characterize the labeling activity of ST6GalNAc1-trunc and to evaluate labeling across multiple cell types, we treated lysates of six WT and SC cell line pairs (Fig 3.6C). The degree of Tn antigen in the lysates was assessed by western blot using the antibody ReMab6, which binds specifically to mucin-type di-GalNAc repeats (Matsumoto, in prep) (Fig 3.6C) and VVA lectin, which recognizes terminal α-GalNAc (Fig 3.6D). This analysis confirmed the presence of Tn antigen in SC but not WT cell lines. When applied to lysates, ST6GalNac1-trunc robustly labeled SC lines but had only limited activity on WT cell lysates (Fig 3.6E). The majority of WT cell lines exhibited little to no labeling with the exception of the MKN45 and Colo205 cell lines. While both cell lines had appreciable label incorporation, the labeling was much higher on the corresponding SC lines. Furthermore, glycoproteins in a wide range of molecular weights from 20 to 200 kDa are labeled, suggesting that ST6GalNAc1-trunc recognizes a broad variety of substrates. Together these results show that ST6GalNAc1-trunc robustly incorporates biotinylated sugars into a wide range of substrates in many different cell lines. Additionally, while the enzyme can act on both T and Tn antigens, it has the greatest activity on Tn antigens.

3.4 Discussion

O-glycosylation is important to the function of many proteins and cellular processes, but the lack of a conserved sequon for O-glycan addition make prediction of O-glycosites difficult. Direct confirmation of site of O-glycosylation through the use of weak lectin affinity chromatography linked mass spectrometry has greatly improved the library of O-glycosites and

consequently the accuracy of prediction algorithms(Steentoft C 2013; Steentoft, Vakhrushev, Vester-christensen, and Schjoldager 2011). However, the technique requires large numbers of cultured cell lines and is not applicable to patient or other primary samples. We sought to create a sensitive approach for detection of O-glycoproteins in such samples by the chemoenzymatic labeling. ST6GalNAc1 is able to act on the universal O-glycan precursor Tn antigen and is therefore a promising approach for global O-glycan detection. The recombinant protein constructs we generated were capable of accepting biotin linking sialic acid substrates and attaching them to target glycoproteins. This biotin labeling will enable improved isolation and detection in future proteomic studies.

We discovered that ST6GalNAc1 is unique among human sialyltransferases (members of the GT29 family) in having a heavy decorated mucin-like domain with a predicted 49 glycosites and 42 sites observed in this study. This represents a strong confirmation of the highly glycosylated nature of this enzyme. Only 15 sites have been observed in past data (Steentoft C 2013). Orthologs of ST6GalNAc1 share this highly glycosylated mucin-like structure across vertebrate species as distant from humans as zebrafish. This indicates a physiological purpose to the conserved glycosylation and suggests a possible mode of regulation for this enzyme. Because we saw ST6GalNAc1-trunc was more likely to self oligimerize than the full length construct, we hypothesize that one function of the mucin-like stem is anti-adhesive and prevents inappropriate clustering of ST6GalNAc1 with itself or with other enzymes.

ST6GalNAc1 is expressed in very low levels in most cells which corresponds to an absences of sTn antigen in ordinary tissues (Clark et al. 2003; Chik et al. 2014; Sewell et al. 2006). The emergence of sTn in tumors is correlated with high expression of ST6GalNAc1 (Sewell et al. 2006; Ozaki et al. 2012; Christiansen et al. 2014). This correlation has been

hypothesized to reflect the nature of sTn as product of ST6GalNAc1. However, our discovery that ST6GalNAc1 catalyzes the autosialylation of its own glycans clustered on the mucin-like stem region suggests an alternate possibility in which the predominate carrier of sTn in tumor cells is ST6GalNAc1 itself. Further investigation into whether ST6GalNAc1 bears Tn or sTn in normal conditions will help to determine whether the autosialylation of ST6GalNAc1occurs in healthy cells. A careful analysis of the glycoproteins bearing sTn in tumor samples is also warranted because the glycobiology of these cells will also differ based on which glycoproteins are being acted on by ST6GalNAc1. Knowledge of which proteins, ST6GalNAc1 or others, are physiological presenters of sTn in tumors will aid in future diagnostics and investigation of tumor biology.

The heterogeneity of glycan structures on neighboring sites observed on ST6GalNAc1 raises exciting questions about the specificities of individual glycotransferases. We observed Tn, sTn, T, and sT antigens in close proximity to one another but only detected microheterogeneity at a minority of sites. This is indicative of restricted specificities of T synthase, ST6GalNAc1, and potentially ST3Gal4 for glycans in different glycopepide contexts. Further investigation of the consistency of these differences and their extension to other glycoproteins is a promising avenue for research of O-glycoprotein biosynthesis.

Pilot use of ST6GalNAc1-trunc for labeling of Tn positive cell lysates with biotinylated sialylic acid confirmed that this technique allows for high affinity isolation of O-glycoproteins. Future application of the technique to glycoproteomics holds significant promise and investigations in this direction are planned pending synthesis of additional biotinylated substrate.

3.5 Materials and Methods

Lectins, antibodies, and reagents

Biotinylated lectins, VVA, PNA, SNA, ConA and Streptavidin-HRP conjugate were purchased from Vector labs. Cyanine 5-labeled streptavidin, streptavidin agarose, and BCA protein assay kit were purchased from Thermo Scientific. Neuraminidase A, PNGaseF, and O-glycosidase were from NEB. Roche cOmplete™ protease inhibitor was from Sigma. The beta-actin antibody was from NovusBio. Peroxidase- and Alexa Fluor 488-labeled secondary antibodies (goat anti-mouse IgG and IgM) were from SeraCare and Thermo Scientific, respectively. ReBags6 antibody was produced in our lab.

The HCF-1 protein for microarray experiments was kindly provided by Suzanne Walker. Biotinylated-CMP-Neu5Ac was a generous gift from Gert-Jan Boons. Sequencing grade trypsin was from Promega. The nitrocellulose microarray slides were purchased from Grace Bio-Labs.

Cloning, protein expression and purification of recombinant ST6GalNAcI

Human ST6GalNAc1 was inserted behind the transferrin signal sequence followed by an N-terminal 8x Histidine tag in pcDNA3.1(+). The stem and catalytic domain of ST6GalNAc1 consisted of residues 38-600, and the truncated construct of the catalytic domain included residues 248-600. The plasmid was transfected with polyethyleneimine into HEK293 Freestyle (Thermo Fisher) cells and the culture supernatant was collected 5-7 days later (Subedi et al. 2015). ST6GalNAcI was purified from the culture supernatant using NiNTA nickel affinity chromatography, dialyzed into 20 mM Tris pH 7.5, 250 mM NaCl and stored at -80°C until use.

Cell lines.

Simple Cell (SC) lines were gifts from Henrick Clausen (University of Copenhagen) and all other cell lines were obtained from ATCC. HEK293S and MDA-MB-231 cells were maintained in Dulbecco's Modified Eagle Media (DMEM-1956) supplemented with 10% Fetal Bovine Serum (FBS). Jurkat cells, Colo205, and MKN45 cells were maintained in RPMI supplemented with 10% FBS. A mixture of 45% Ham's F12, 45% DMEM, and 10% FBS was used for LS174T cells. MCF7 and MCF7-SC were grown in DMEM (1885) supplemented with 0.01% insulin and 10% FBS. All cells were maintained at 37°C with 5% CO_2.

Oligomerization studies

For chemical crosslinking, Cosmc, IgG (Equitech-Bio, Inc), and BSA (Sigma) proteins were dialyzed into 20 mM sodium phosphate pH 7.5, 250 mM NaCl. Crosslinking was conducted by addition of 0.5 mM disuccinimidyl glutarate (DSG, ThermoFisher) for 30 min at room temperature at a protein concentration of 0.3 mg/ml. Reactions were quenched with 1/10 volume of 100 mM Tris pH 7.5 for 30 min at room temperature and analyzed by SDS-PAGE and Coomassie G-250 staining. For blue native PAGE, proteins were adjusted to 0.3 mg/ml with 10 mM Tris pH 7.5, 50 mM NaCl, and sample buffer including Ponceau S and glycerol was added. BN-PAGE (Invitrogen) was run at 4°C with Coomassie G-250 dye in the Cathode buffer and stained with Coomassie G-250.

SEC-MALS.

Proteins were preparatively purified using a Superdex 200 10/300 (GE Healthcare) with phosphate buffered saline as running buffer. Analytical size exclusion chromatography with a Superdex 200 10/300 with multi-angle light scattering (Wyatt Dawn Heleos II) and dynamic light scattering (Wyatt) was performed with an Agilent 1620 chromatography. BSA

(ThermoFisher) was used as a protein standard and to align peaks. Data was analyzed with the Astra 7 software and molecular masses were calculated with a glycoconjugate analysis. The dn/dc values for protein and glycan were 0.185 and 0.15, respectively.

Enzyme reaction conditions

Buffers for glycosidase and glycosyltransferase reactions were 50 mM sodium acetate pH 5.5, 5 mM calcium chloride for Neuraminidase, 50 mM sodium phosphate pH 7.5 for PNGaseF and O-glycanase, and 50 mM cacodylate, pH 6.5, with 10 mM magnesium and manganese chloride, for ST6GalNAc1. All incubations were at 37°C, for 1 hr, 8 hr, or overnight, depending on the treated material. ST6GalNAc1 reactions were performed with 50 ug/ml of enzymes and included no donor sugar, 10 uM CMP-Neu5Ac, or 10 uM 5'-biotinyl-CMP-Neu5Ac, unless otherwise indicated.

HPLC shift assay with model Tn-glycopeptide.

Synthetic glycopeptide 2GP-1 was prepared and used as in (Heimburg-Molinaro et al. 2009). Glycopeptide at 10uM concentration were treated with ST6GalNAc1 with or without CMP-Neu5Ac. Reactions were quenched by the addition of 0.1% trifluoro-acetic acid, and analyzed by HPLC using a Grace Vydac column with a gradient of acetonitrile from 5-50% over 60 minutes.

Microarray experiments

Glycoprotein microarray experiments were performed as described (McKittrick and Noll, in preparation). All glycoproteins were printed on nitrocellulose slides in phosphate buffered saline from a solution of 100ug/ml protein concentration. Microarray slides were blocked with a solution of SuperG Blocking buffer (GraceBio-Labs) and stored dry at -20°C. Immediately

before use slides were hydrated and washed. Treatment with Neuraminidase A (New England Biolabs) was performed in 50 mM sodium acetate pH 5.5, 5 mM calcium chloride for 10 hrs at 37°C. Treatment with PNGaseF (New England Biolabs) and O-glycosidase (New England Biolabs) was performed in 50 mM sodium phosphate pH 7.5 overnight at 37°C. ST6GalNAc1 treatment was carried out in 50 mM cacodylate pH 6.5, 150 mM sodium chloride, and 0.1% Tween-20, and 0.5 mM 5'biotinyl-CMP-Neu5Ac for 4 hrs at 37°C. Microarrays were probed with biotinylated lectins (0.1 ug/ml SNA and ConA, 1ug/ml PNA) and incorporated biotinyl-glycoproteins were detected with Streptavidin-cyanine5 (0.5 ug/ml).

Cell labeling and flow cytometry.

HEK293 wild type and SC cells were lifted with EDTA and washed in Hank's balanced salt solution (HBSS). Cells were labeled or mock treated with ST6GalNAc1 (50 ug/ml), 5'-biotinyl-CMP-Neu5Ac (0.5 mM), in Tris-buffered saline at 37°C for 1 hr. Cells were analyzed with streptavidin, ReBags6, and VVA lectin by flow cytometry on a FACS Calibur cytometer (BD Biosciences).

Cell lysate labeling and western blots.

Cells were detached with 1mM ethylenediaminetetraacetic acid (EDTA) and collected by centrifugation and stored at -20°C until use. Cell were lysed in 20 mM sodium acetate pH 6.0 with 150 mM NaCl, 4mM $CaCl_2$ with 1% TritonX-100 with Complete protease inhibitor (Roche), and centrifuged 10 minutes at 20k x *g*. Total protein content of cell lysates was assessed with BCA assay (Thermo Fisher). Fifty microliters of cell lysate at a concentration of 2 mg/ml were treated with Neuraminidase A (New England Biolabs) for 6 hours at 37°C before adding recombinant 50 μg of ST6GalNAcI and biotinylated CMP-Sialic acid to 500 μM and

incubated overnight at 37°C. Asialo-BSM was included as a positive control. Labeled samples were analyzed by western blotting with Streptavidin-HRP, and unlabeled samples were analyzed by western blotting with ReMab6 antibody(Matsumoto et al. 2016) and PNA and VVA lectins (Vector labs), and loading controls were performed with an anti-β-actin antibody (Novus Bio).

Metal binding assays.

Protein was dialyzed into 20 mM Tris pH 7.4, 250 mM NaCl. For treatment with a cation chelator, 1mM EDTA was added and incubated at 4°C for 1 hr before dialysis into 20 mM Tris pH 7.4, 250 mM NaCl. Thermal shift assays were performed with 3 uM ST6GalNAc1 protein and SYPRO Orange (ThermoFisher) at a final concentration of 5x with 1mM divalent cation present. Fluorescence intensity measurements were recorded from 21°C to 95°C in 1.0°C steps, with a CFX96 Real-Time System with a C1000 Thermal cycler (Bio-Rad). Data were fit to a Boltzman curve with sloping baselines with OriginPro (OriginLab) to locate the inflection point (T_m). Assays were performed in triplicate and mean and standard deviation are reported.

In vitro biotinylation with ST6GalNAc1 and enrichment

Cells were lifted with EDTA, resuspended in buffer (20 mM Tris pH 7.4, 150 mM sodium chloride, 1% TritonX-100, and cOmplete™ protease inhibitor) and lysed with sonication. Lysate was cleared by centrifugation and treated with O-glycosidase before in vitro biotinylation. Enzymatic labeling was performed by adding 5'-biotinyl-CMP-Neu5Ac and ST6GalNAc to 100 **µ**M and 50 ug/ml, respectively, with 50 mM cacodylate, pH 6.5, and 10 mM magnesium and manganese chloride, and incubation at 37°C for 8 hr to overnight. Unreacted biotinylated sugar donor was removed by passing samples through a desalting column and collecting the void volume. Enrichment was performed by capture on streptavidin resin, washing with 1% SDS, 6M

urea, and PBS. After reduction and alkylation, bound glycoproteins were released by overnight trypsin digestion. The supernatant was collected, desalted with a Sep-Pak C18 column, dried and resuspended in 0.1% formic acid with 5% acetonitrile.

Mass spectrometry and data analysis

Mass spectrometry data was collected on an Orbitrap Fusion Lumos Tribrid spectrometer (Thermo Scientific) downstream of a Dionex UltiMate 3000 UHPLC system (Thermo Scientific). Technical replicates were performed and each sample was analyzed in triplicate. Samples were separated on a self-packed C18 fused-silica column. MS^1 data was collected in the Orbitrap at 120,000 resolution (400-1600 *m/z* range) AGC was set to 400,000 and the maximum ion time was 50 ms. HCD MS^2 spectra were collected and data was processed with Proteome Discoverer 2.1 (Thermo Scientific)

Acknowledgements

We appreciate members of the Cummings laboratory for critical discussions. This work was supported, in whole or in part, by National Institutes of Health Grants GM068559 (to R. D. C.).

3.6 Tables and Figures

Table 3.1. Predicted and Observed O-glycosylation on Recombinant ST6GalNAc1

Each serine or threonine in ST6GalNAc1 is listed. Predictions were made with NetOGlyc 4.0 server, observed glycosylation sites and glycan compositions were determined with ProteomeDiscover software. In the absence of specific monosaccharide identification, compositions of HexNAc, HexNAcNeuAc, HexHexNAc, and HexHexNAcNeuAc were presumed to be Tn, sTn, T, and sT, respectively. An asterisk (*) in the observed column indicates a site previously observed as glycosylated, as reported in the online Glycodomain Viewer (https://glycodomain.glycomics.ku.dk/uniprot/Q9NSC7 Steentoft 2013)

Residue	Predicted	Observed	Glycans Observed
T 35	**yes**	**yes**	sTn
S 38	**yes**	no	-
T 43	**yes**	**yes**	T
S 50	**yes**	**yes**	Tn or sTn
S 53	**yes**	**yes**	Tn or sTn
S 59	**yes**	no	-
T 63	**yes**	no	-
T 68	**yes**	**yes**	Tn
T 69	**yes**	**yes**	Tn
T 83	**yes**	**yes**	Tn
T 85	**yes**	**yes**	T
T 91	**yes**	no	-
T 92	**yes**	no	-
T 113	**yes**	no *****	-
S 121	**yes**	**yes**	Tn
T 127	**yes**	**yes ***	Tn, T, sT
T 131	**yes**	**yes ***	sTn
S 133	**yes**	**yes ***	Tn, sTn, sT
S 143	**yes**	**yes**	Tn, sTn, sT
T 146	**yes**	**yes ***	sT
S 150	**yes**	**yes**	Tn
S 153	**yes**	**yes**	sTn
T 156	**yes**	no	-

T	158	**yes**	**yes**	Tn
T	159	**yes**	**yes**	Tn, sTn
T	166	**yes**	**yes**	sTn, sT
T	170	**yes**	**yes**	Tn
S	172	**yes**	no	-
T	174	**yes**	no	-
S	176	**yes**	**yes**	T
T	185	**yes**	no	-
T	186	**yes**	**yes**	T
T	189	**yes**	no	-
S	194	**yes**	**yes**	sTn
T	202	**yes**	**yes** *	T
S	206	**yes**	no *	-
T	207	**yes**	**yes** *	Tn
T	209	**yes**	**yes**	Tn
T	215	**yes**	**yes** *	Tn
T	216	**yes**	**yes** *	T
T	229	**yes**	**yes** *	Tn, sTn
S	237	**yes**	**yes** *	T
T	239	**yes**	**yes** *	Tn
T	240	**yes**	**yes** *	Tn, sTn
S	253	**yes**	**yes**	sTn
S	264	no	**yes**	sTn
T	272	no	**yes**	Tn
T	273	no	**yes**	Tn
S	277	**yes**	no	-
S	283	**yes**	no	-
S	285	no	no	-
T	297	no	no	-
S	302	no	no	-
S	308	no	no	-
S	328	no	no	-
T	335	**yes**	no	-
S	349	**yes**	no	-
S	354	no	no	-
T	359	no	**yes**	sTn
S	372	no	no	-
S	380	no	no	-
S	388	no	**yes**	Tn
T	401	no	no	-
T	403	no	no	-

S	404	no	no	-
T	409	no	no	-
S	412	no	no	-
T	414	no	no	-
S	416	no	no	-
T	443	no	**yes**	Tn, sTn
T	457	no	**yes**	sT
S	460	no	**yes**	T
S	501	no	no	-
T	503	no	no	-
T	515	no	no	-
T	516	no	no	-
T	523	no	no	-
S	532	no	no	-
T	538	no	no	-
S	545	no	**yes**	sTn
T	551	no	no	-
S	552	no	no	-
T	590	no	no *****	-

Table 3.2: Number of predicted O-GalNAc glycosites for human sialyltransferases

Predictions performed with NetOGlyc 4.0 server.

Human STs	
ST3GAL1	3
ST3GAL2	5
ST3GAL3	7
ST3GAL4	2
ST3GAL5	2
ST3GAL6	1
ST6GAL1	12
ST6GAL2	18
ST6GALNAC1	**49**
ST6GALNAC2	3
ST6GALNAC3	2
ST6GALNAC4	8
ST6GALNAC5	7
ST6GALNAC6	5
ST8SIA1	2
ST8SIA2	8
ST8SIA3	7
ST8SIA4	8
ST8SIA5	1
ST8SIA6	12

Table 3.3: Number of predicted O-GalNAc glycosites for ST6GalNAc1 orthologs

Predictions performed with NetOGlyc 4.0 server for selected species. Table is organized by increasing phylogenetic distance from human.

ST6GalNAc1 orthologs	
Human	**49**
Gorilla	45
Mouse	27
Bat	69
Dog	52
Cat	45
Cow	48
Donkey	54
Chicken	41
Pigeon	31
Frog	21
Zebrafish	12

Table 3.4: Quantification of SEC-MALS Data for ST6GalNAc1 and ST6GalNAc1-trunc

	Molar Mass (kDa) **CALCULATED**	**(PREDICTED)**	**Hydrodynamic Radius**
ST6GalNAc1	91.2 ± 0.6%		6.7 nm ± 5%
Protein	78.9 ± 0.6%	(66.8)	
Glycan	12.3	(15)	
ST6GalNAc1 Trunc	56.3 ± 0.7%		3.9 nm ± 5%
Protein	46.7 ± 0.7%	(42.7)	
Glycan	9.6	(9)	
BSA	63.6 ± 0.6%	(66.5)	3.5 nm ± 4%

Table 3.5: Thermal shift of ST6GalNAc1 with chemical additives

Additive	**ST6GalNac1 Tm (°C)**
untreated	53.3 ± 0.2
EDTA treated	52.1 ± 0.1
$CaCl_2$	52.6 ± 0.6
$MnCl_2$	52.5 ± 0.2
$MgCl_2$	52.3 ± 0.1
$ZnCl_2$	53.6 ± 7
$FeCl_2$	52.1 ± 0.1
CMP-Neu5Ac	52.9 ± 0.1
GalNAc	51.5 ± 0.1
4MU-αGalNAc	51.6 ± 0.5

FIGURE LEGENDS

Figure 3.1: Recombinant ST6GalNAc1 catalyzes its own sialylation and bears a mixture of T and Tn O glycans. A. Recombinant ST6GnT1 enzyme was expressed in HEK293T cells and purified over a nickel column. Fractions were separated by PAGE and visualized with Coomassie blue staining. **B** Sialylation reactions were carried out by combining purified recombinant ST6GnT1 enzyme, biotinylated CMP-sialic acid donor, and simple cell (HEKSC) lysate and detected with streptavidin-HRP by western blot. **C** Recombinant ST6GalNAc1 was treated with the indicated glycosidase enzymes and visualized by PAGE and Coomassie blue staining. **D** Recombinant ST6GnT1 enzyme was analyzed by lectin blot using VVA and PNA lectins. **E** Whole recombinant ST6GnT1 enzyme was analyzed by MALDI.

Figure 3.2 ST6GalNAc1 contains many occupied O-glycosites. **A** ST6GalNAc1 was digested by trypsin and analyzed by LC-MS. Identified glycosylation sites were compared to those predicted by the NetOGlyc 4.0 server, and represented as a Venn diagram. See also Table 3.1. **B** Observed glycans were mapped to their positions on ST6GalNAc1 and plotted based on composition. Stereochemistry of monosaccharides is presumed, and representation of sTn is not intended to communicate linkage, only the presence of sialic acid. **C** Sequence of ST6GalNAc1. Potential sites of O-glycosylation are bold, predicted sites are grey, observed sites are yellow.

Figure 3.3 Truncated ST6GalNAc1 has reduced carbohydrate content. A Domain map of the full length and truncated enzyme constructs. **B** Truncated enzyme was expressed in HEK293T cells and purified over nickel column, with fractions visualized by Coomassie blue staining. **C** Recombinant ST6GalNAc1 constructs and control proteins were crosslinked with DSG and visualized by gel electrophoresis for oligomeric laddering. **D** Recombinant ST6GalNAc1 and control proteins were analyzed by Blue Native PAGE. **E** Purified full-length

and truncated ST6GalNAc1 were analyzed by SEC-MALS. Quantification of protein and glycan mass as well as hydrodynamic radius were calculated and are presented in Table 3.4.

Figure 3.4 ST6GalNAc1-trunc acts on a range of Tn-bearing glycopeptides A Model glycopeptide 2GP-1 was incubated with ST6GalNAc1-trunc and CMP-Neu5Ac with or without added EDTA for metal chelation. Reaction products were visualized using HPLC. **B** 2GP-1 was incubated with ST6GalNAc1-trunc and CMP-Neu5Ac and the reaction progress was monitored by analyzing the reaction mixture with HPLC. **C**. Quantification of appearance of the sTn-2GP-1 product from B. **D**. Glycopeptide microarrays were either incubated with ST6GalNAc1-trunc and biotin-linked CMP-Neu5Ac and stained with streptavidin, or incubated with recombinant T-synthase and UDP-Gal and stained with PNA lectin. Fluorescence of labeled glycopeptides was measured as a readout of enzyme activity.

Figure 3.5 ST6GalNAc1-trunc labels O-glycoproteins on a microarray A Schematic of experiment: Glycoprotein microarrays were pre-treated neuraminidase, followed by N-glycanase or O-glycosidase. Treated microarrays were then sialylated with truncated ST6GalNAc1 and biotinylated sugar donor and probed with streptavidin or lectins. **B** Glycosidase activity was confirmed by staining with SNA lectin after neuraminidase treatment, **C** ConA lectin after PNGaseF treatment, and **D** PNA lectin after O-glycosidase treatment. **E** Comparison of biotin incorporation with or without glycosidase pretreatment. **F** Untreated microarrays were probed with three different reagents to detect Tn antigen: VVA and HPA lectins and ReBags6 monoclonal antibody.

Figure 3.6 ST6GalNAc1-trunc selectively labels Tn-positive cells and successfully incorporates a biotin on Tn-glycoproteins A HEK WT and HEKSC cells were treated with truncated ST6GnT1 enzyme and biotinylated donor sugar, with or without O-glycanase

pretreatment. Cells were then stained with streptavidin and assessed by flow cytometry. **B** HEKSC cells were treated and analyzed as in A, with varying concentrations of biotinylated donor sugar. **C** and **D** Cell lysates from WT (W) and SC (S) cell line pairs were probed by western blot with ReMab6, or with VVA lectin, with anti-β-actin used as a loading control. **E** Lysates of WT and HEKSC lines were treated with truncated ST6GalNAc1, with or without biotinylated donor sugar. Treated lysates were probed with streptavidin by western blot.

Figure 3.1

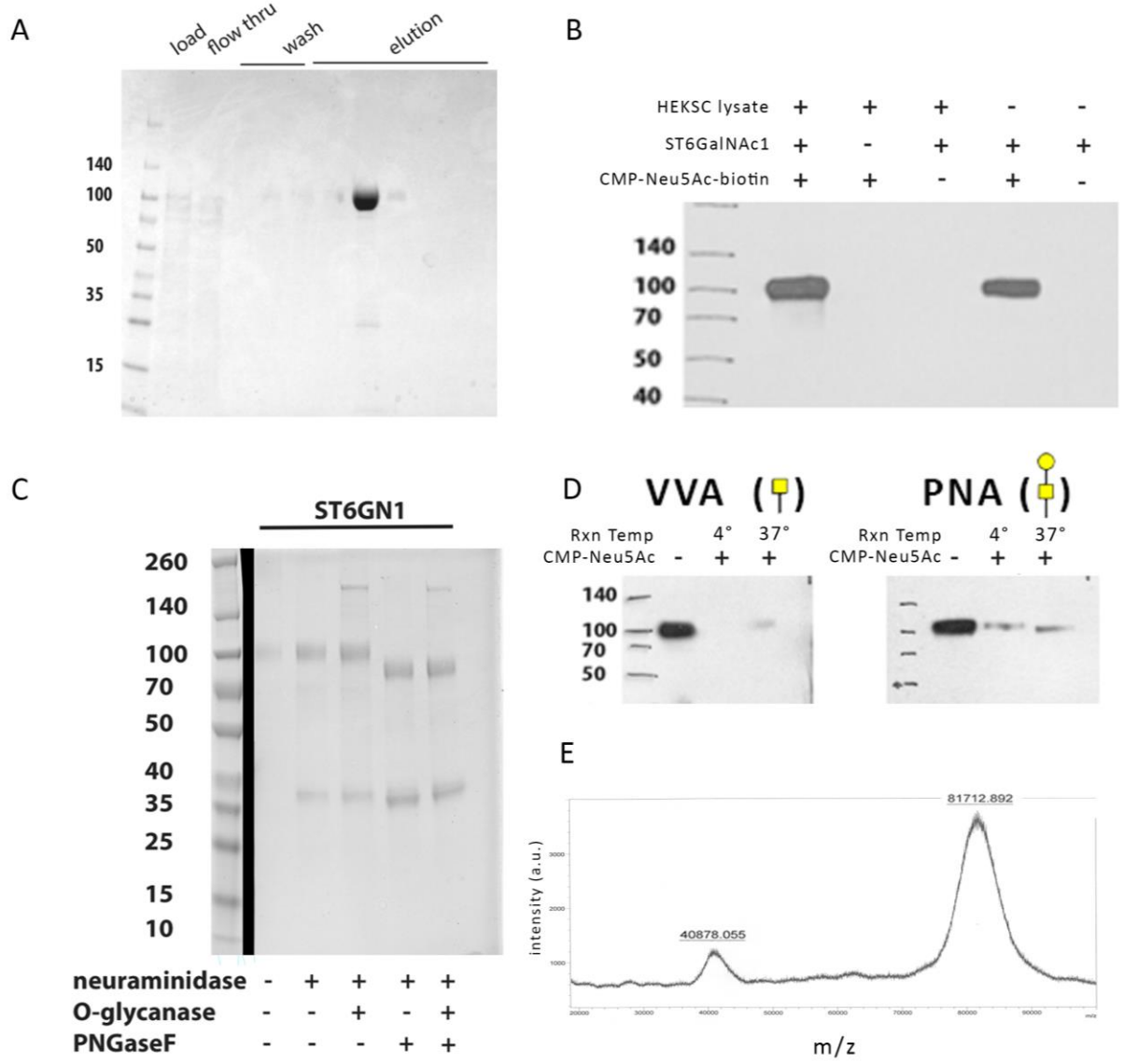
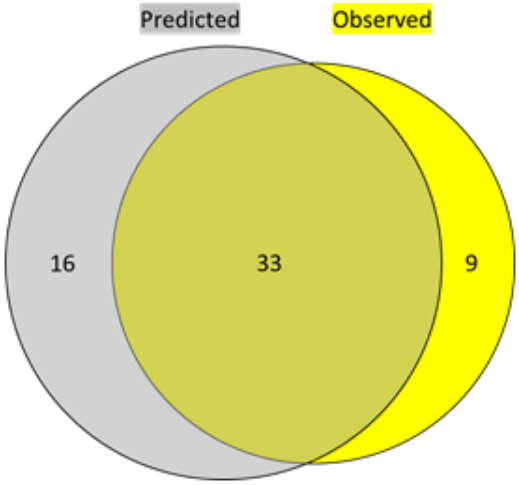
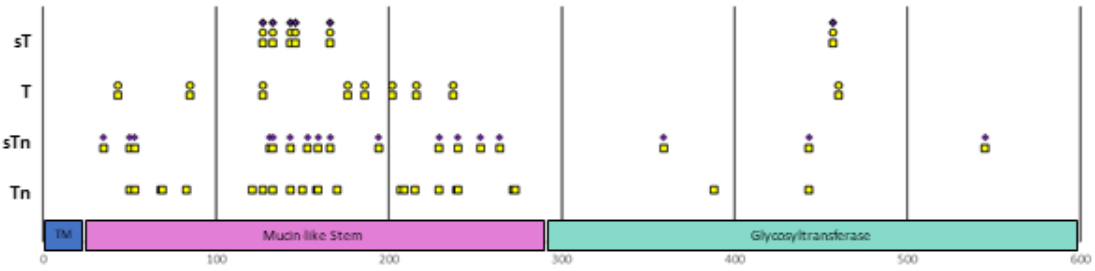

A
load
flow thru
wash
elution
140
100
50
35
15
B
HEKSC lysate + + + - -
ST6GalNAc1 + - + + +
CMP-Neu5Ac-biotin + + - + -
140
100
70
50
40
C
ST6GN1
260
140
100
70
50
40
35
25
15
10
neuraminidase - + + + +
O-glycanase - - + - +
PNGaseF - - - + +
D
VVA
PNA
Rxn Temp 4° 37°
CMP-Neu5Ac - + +
140
100
70
50
E
81712.892
40878.055
intensity (a.u.)
m/z

Figure 3.2

A



B



C

```
001 MRLAVGALLV CAVLGLCLAE DQVDPRLIDG KDPQTKPSRH QRTENIKERS
051 LQSLAKPKSQ APTRARRTTI YAEPVPENNA LNTQTQPKAH TTGDRGKEAN
101 QAPPEEQDKV PHTAQRAAWK SPEKEKTMVN TLSPRGQDAG MASGRTEAQS
151 WKSQDTKTTQ GNGGQTRKLT ASRTVSEKHQ GKAATTAKTL IPKSQHRMLA
201 PTGAVSTRTR QKGVTTAVIP PKEKKPQATP PPAPFQSPTT QRNQRLKAAN
251 FKSEPRWDFE EKYSFEIGGL QTTCPDSVKI KASKSLWLQK LFLPNLTLFL
301 DSRHFNQSEW DRLEHFAPPF GFMELNYSLV QKVVTRFPPV PQQQLLLASL
351 PAGSLRCITC AVVGNGGILN NSHMGQEIDS HDYVFRLSGA LIKGYEQDVG
401 TRTSFYGFTA FSLTQSLLIL GNRGFKNVPL GKDVRYLHFL EGTRDYEWLE
451 ALLMNQTVMS KNLFWFRHRP QEAFREALHM DRYLLLHPDF LRYMKNRFLR
501 SKTLDGAHWR IYRPTTGALL LLTALQLCDQ VSAYGFITEG HERFSDHYYD
551 TSWKRLIFYI NHDFKLEREV WKRLHDEGII RLYQRPGPGT AKAKN
```

Figure 3.3

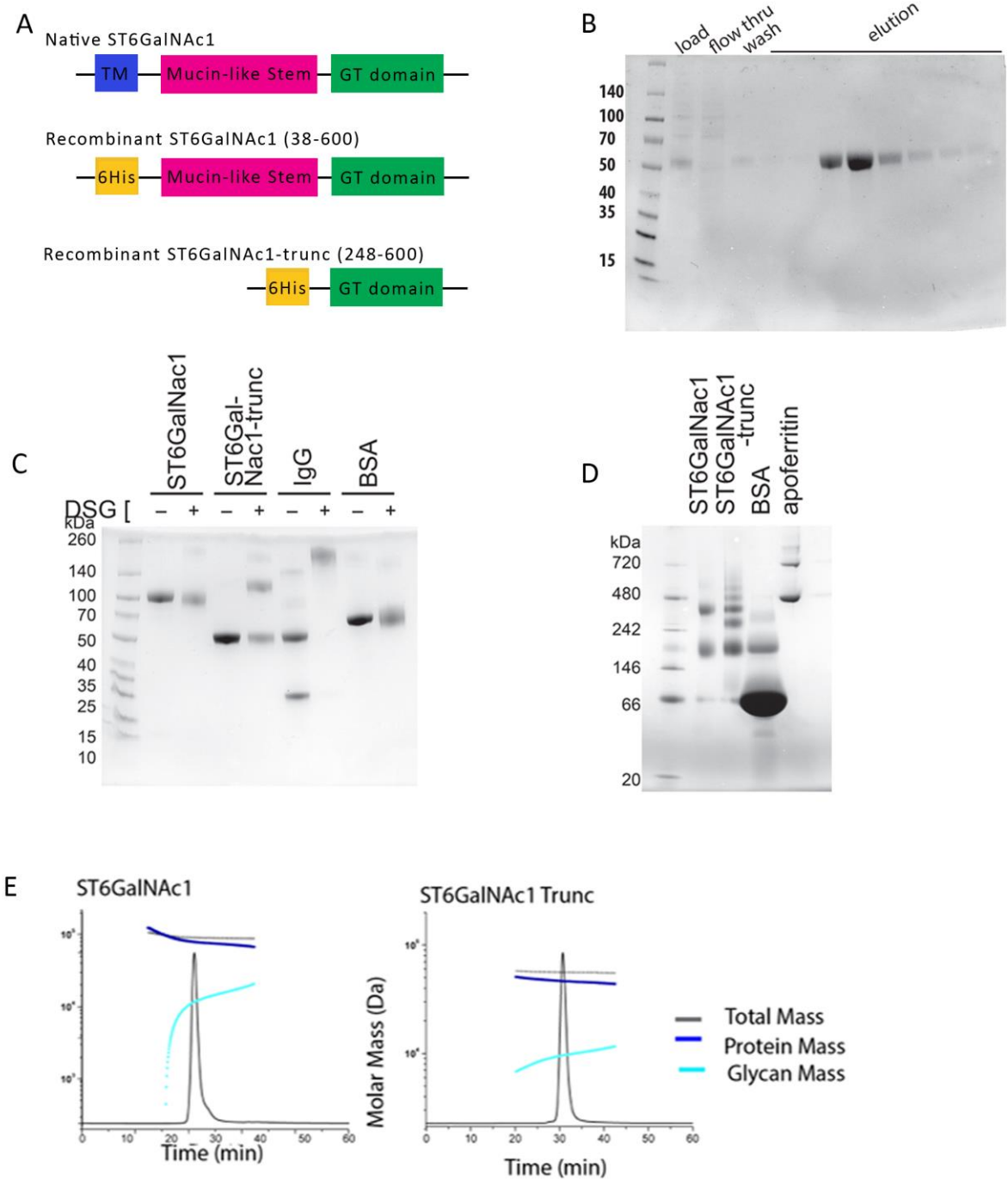

A
Native ST6GalNAc1
TM
Mucin-like Stem
GT domain
Recombinant ST6GalNAc1 (38-600)
6His
Mucin-like Stem
GT domain
Recombinant ST6GalNAc1-trunc (248-600)
6His
GT domain
B
load
flow thru
wash
elution
140
100
70
50
40
35
15
C
ST6GalNac1
ST6Gal-Nac1-trunc
IgG
BSA
DSG [
– + – + – + – +
kDa
260
140
100
70
50
40
35
25
15
10
D
ST6GalNac1
ST6GalNAc1 -trunc
BSA
apoferritin
kDa
720
480
242
146
66
20
E
ST6GalNAc1
ST6GalNAc1 Trunc
Molar Mass (Da)
Time (min)
Time (min)
Total Mass
Protein Mass
Glycan Mass

Figure 3.4

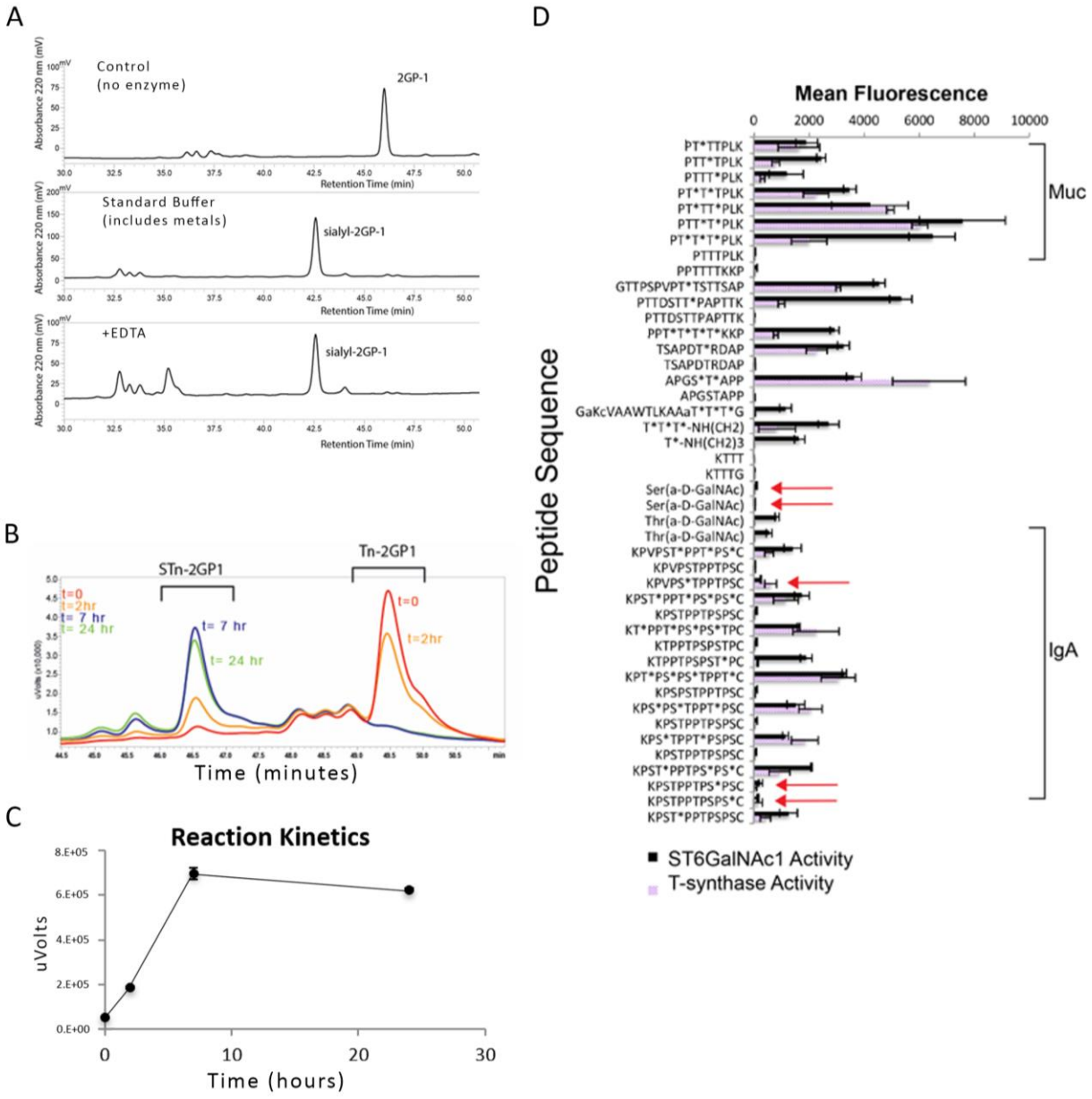

A
Control
(no enzyme)
2GP-1
Standard Buffer
(includes metals)
sialyl-2GP-1
+EDTA
sialyl-2GP-1
Absorbance 220 nm (mV)
Retention Time (min)
B
STn-2GP1
Tn-2GP1
t=0
t=2hr
t= 7 hr
t= 24 hr
Time (minutes)
C
Reaction Kinetics
uVolts
Time (hours)
D
Mean Fluorescence
Peptide Sequence
PT*TTPLK
PTT*TPLK
PTTT*PLK
PT*T*TPLK
PT*TT*PLK
PTT*T*PLK
PT*T*T*PLK
PTTTPLK
PPTTTTKKP
GTTPSPVPT*TSTTSAP
PTTDSTT*PAPTTK
PTTDSTTPAPTTK
PPT*T*T*T*KKP
TSAPDT*RDAP
TSAPDTRDAP
APGS*T*APP
APGSTAPP
GaKcVAAWTLKAAaT*T*T*G
T*T*T*-NH(CH2)
T*-NH(CH2)3
KTTT
KTTTG
Ser(a-D-GalNAc)
Ser(a-D-GalNAc)
Thr(a-D-GalNAc)
Thr(a-D-GalNAc)
KPVPST*PPT*PS*C
KPVPSTPPTPSC
KPVPS*TPPTPSC
KPST*PPT*PS*PS*C
KPSTPPTPSPSC
KT*PPT*PS*PS*TPC
KTPPTPSPSTPC
KTPPTPSPST*PC
KPT*PS*PS*TPPT*C
KPSPSTPPTPSC
KPS*PS*TPPT*PSC
KPSTPPTPSPSC
KPS*TPPT*PSPSC
KPSTPPTPSPSC
KPST*PPTPS*PS*C
KPSTPPTPS*PSC
KPSTPPTPSPS*C
KPST*PPTPSPSC
Muc
IgA
ST6GalNAc1 Activity
T-synthase Activity

Figure 3.5

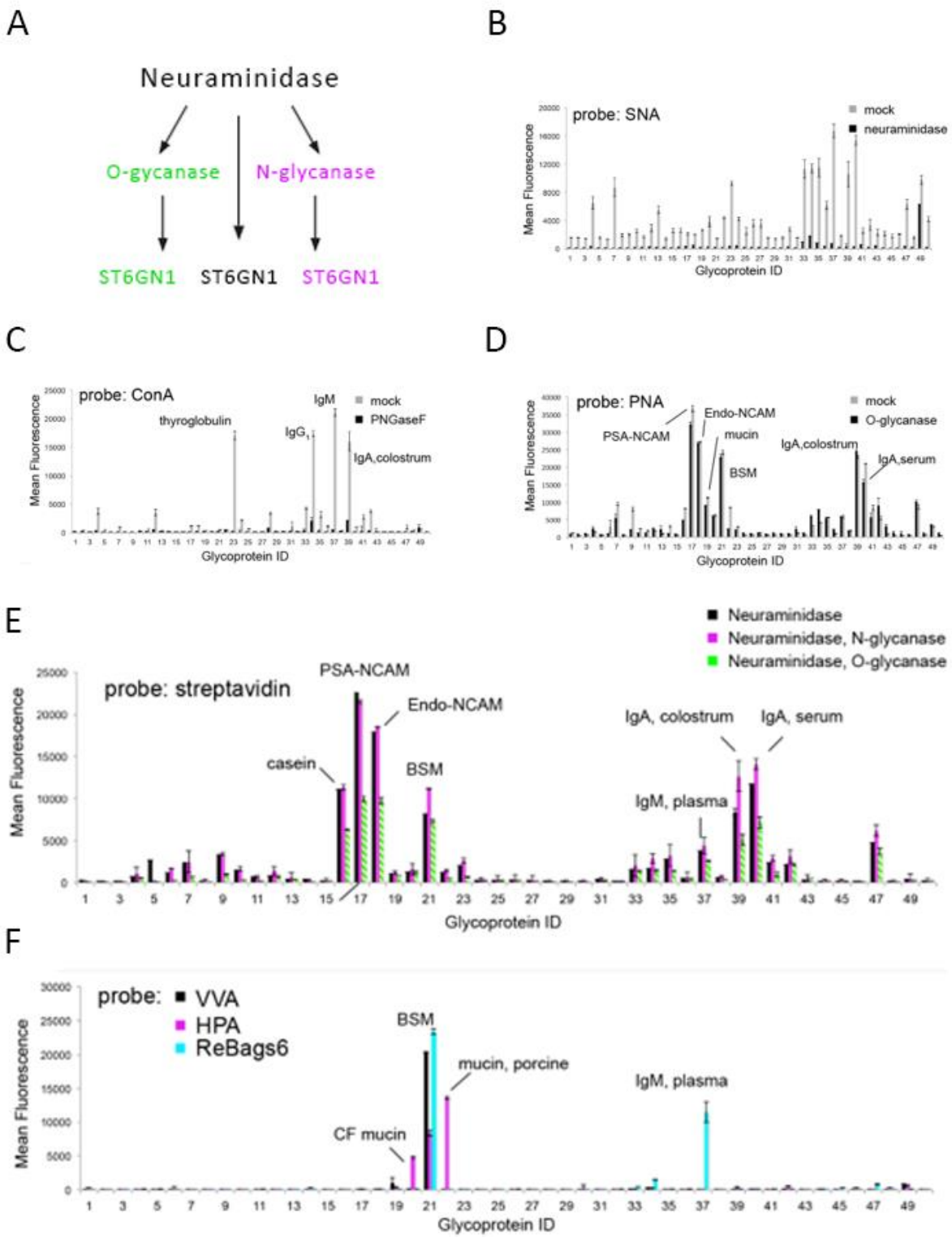
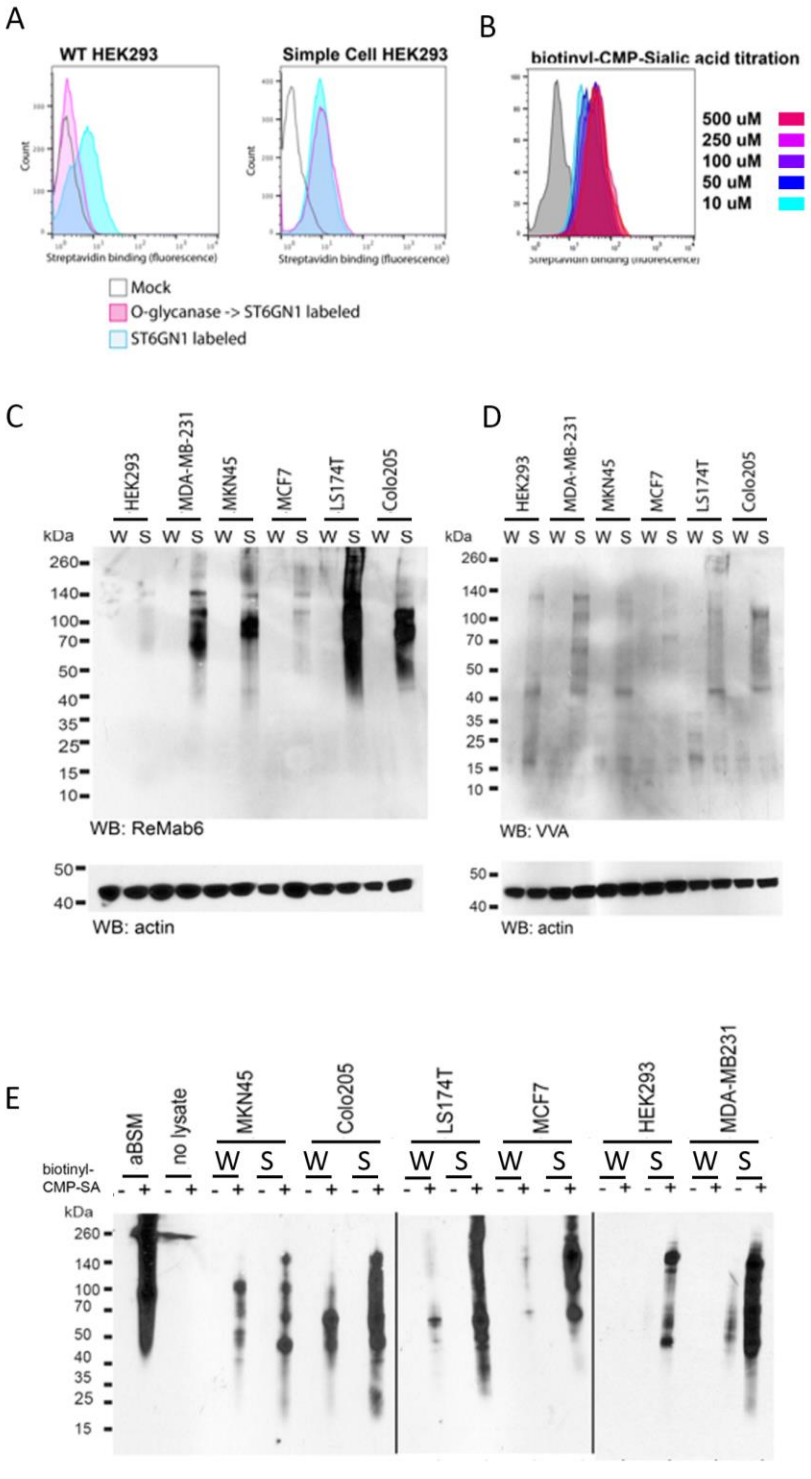

A
Neuraminidase
O-gycanase
N-glycanase
ST6GN1
ST6GN1
ST6GN1
B
probe: SNA
mock
neuraminidase
Mean Fluorescence
Glycoprotein ID
C
probe: ConA
thyroglobulin
IgM
IgG1
IgA,colostrum
mock
PNGaseF
Mean Fluorescence
Glycoprotein ID
D
probe: PNA
PSA-NCAM
Endo-NCAM
mucin
BSM
IgA,colostrum
IgA,serum
mock
O-glycanase
Mean Fluorescence
Glycoprotein ID
E
Neuraminidase
Neuraminidase, N-glycanase
Neuraminidase, O-glycanase
probe: streptavidin
PSA-NCAM
Endo-NCAM
casein
BSM
IgA, colostrum
IgA, serum
IgM, plasma
Mean Fluorescence
Glycoprotein ID
F
probe:
VVA
HPA
ReBags6
BSM
mucin, porcine
IgM, plasma
CF mucin
Mean Fluorescence
Glycoprotein ID

Figure 3.6


A
WT HEK293
Simple Cell HEK293
Count
Streptavidin binding (fluorescence)
Mock
O-glycanase -> ST6GN1 labeled
ST6GN1 labeled
B
biotinyl-CMP-Sialic acid titration
500 uM
250 uM
100 uM
50 uM
10 uM
C
HEK293
MDA-MB-231
MKN45
MCF7
LS174T
Colo205
kDa
W S
260
140
100
70
50
40
35
25
15
10
WB: ReMab6
WB: actin
D
WB: VVA
E
aBSM
no lysate
biotinyl-CMP-SA

Chapter 4: Discussion and future directions

4.1 Introduction

T cell dependent immunity is crucial for adaptive defenses against infectious and oncogenic challenges and T cell glycosylation is essential for proper T cell function. The processes of T cell development and immune activation are strongly tied to changes in glycosylation, as outlined in section 1.3. Altered glycosylation is also a prominent feature of other pathologies including cancer, as described in section 1.2.3. Therefore, understanding the role of glycosylation in protein and cellular functions promises to provide insight into normal and pathological processes, and to open multiple avenues for therapeutic intervention in a wide variety of diseases. As presented in Chapter 2, extended O-glycosylation is vital to T cell survival in the periphery and appropriate T cell homing to SLO. Chapter 3 outlined our initial forays into identifying and characterizing proteins that carry the cancer associated Tn and sTn antigens. The discoveries made in this work further the understanding of T cell requirements for glycosylation and indicate promising avenues of future research into immunological and oncological phenomena. This chapter discusses the import of these discoveries and the direction of prospective research with respect to T cell glycosylation, lymphocyte homing, cancer glycobiology, and glycoproteomics.

4.2 Unique T cell requirement of O-glycans

The most novel finding of the work described in this dissertation is the discovery that T cells have a critical post-thymic requirement for *Cosmc*. While glycosylation as a whole and O-glycosylation in particular are generally important for many cell and tissue types, the concordance of observations between Tn syndrome patients, lectin staining of cancer cells, and the TCKO model discussed in section 2.4 suggests that the survival defect we observed for T cells lacking *Cosmc* is, in fact, an effect unique to T cells. To understand the mechanism by

which this effect is exerted, it will be important to compare multiple cell types to discern what similarities and differences exist in the role of glycosylation in different blood cell lineages. Indeed, others in our laboratory have investigated a B cell specific Cosmc KO that has provided evidence for a general role of lymphocyte glycans in extravasation, but suggests that hepatic removal of cells expressing the Tn antigen is not a shared feature between T cells and other cell types (manuscript in preparation).

Additionally, the fate of defective or apoptotic cells in the body has not been thoroughly investigated and the site of T cell removal is largely unknown (Crispe and Huang 1994). There have been indications in past work that the liver is a site for the removal of moribund T cells and that the concentration of T cells in the liver may be a specific effect of glycosylation state, as neuraminidase treatment results in hepatic capture of a subset of adoptively transferred lymphocytes (Iwakoshi et al. 1998; Gesner and Ginsburg 1964; Woodruff and Gesner 1969; Berney and Gesner 1970; Crispe et al. 2000). Our own unpublished observations suggest that this is a feature restricted to T cell biology and is not shared by B cells or other leukocytes. The mechanism of T cell specific glycan dependent recognition in the liver is an important question for future study. The liver itself is a uniquely immunomodulatory organ with high levels of inhibitory signals that prevent activation of immune cells both in infection and in allogenic transplant (Robinson, Harmon, and O'Farrelly 2016). It is unknown whether this feature is related to the liver's apparent role in disposal of T cells at the end of their life cycle. The question of whether such functions are coupled could have implications for organ transplantation and T cell-based therapies.

4.3 Lymphocyte homing

The currently accepted model of lymphocyte homing involves a multistep process of cellular activation and receptor ligand interactions, described in greater depth in section 1.3.3. These steps include key lectin-glycan binding interactions, but until now the data has only indicated a role for glycans expressed on endothelial cells binding with ligands on lymphocytes. The findings in chapter 2 provide evidence that perturbation of O-glycans expressed by T cells can profoundly reduce the ability of T cells to home to the secondary lymphoid organs (SLO) such as spleen and lymph nodes. The indication that the O-glycans on T cells directly impact their ability to enter SLO is an intriguing addition to the existing T cell homing model. Whether this suggests an additional requirement for T cell glycosylation in order to complete one of the dogmatically accepted steps of rolling, arrest, and extravasation, or whether it could be evidence for an additional step in the process remains unclear.

My finding that O-glycosyaltion on T cells is required for T cell homing highlights the possibility of further features of lymphocyte homing that remain to be discovered including but not limited to interactions with glycans expressed on the surface lymphocytes. Such discoveries would open new avenues for therapeutic intervention in multiple autoimmune disorders and other pathologies similar to the current therapies targeting leukocyte extravasation that are in use for conditions such as multiple sclerosis, and ulcerative colitis (Ley et al. 2016; Raab-Westphal, Marshall, and Goodman 2017). Treatment of these diseases is most potently affected by preventing the inflammatory cells from reaching the sites of immune response.

The T cell specific Cosmc knock out (TCKO) model utilized in this work results in a dearth of investigable cells and is not ideal for pursuing the problem of O-glycosylation's role in T cell homing, but an inducible knockout of Cosmc in T cells or other blood cells is an exciting

possibility. The ideal model would allow for induced deletion of Cosmc in T cells upon administration of tamoxifen or similar agent, so that T cells of different maturation states and differentiation subtypes could be generated and analyzed separately. With such cells available, models of homing and extravasation that utilize *ex vivo* conditions or adoptively transferred cells could be used to dissect the process and identify the stages at which WT and Cosmc$^{-/-}$ cells differ, and the mechanisms involved.

4.4 T cell signaling and survival

The signaling pathways of T cells have been extensively investigated using a variety of primary cells and cultured cell lines, with a great deal of detail being established about the function of specific molecules and larger complexes of the immune synapse that forms between a T cell and a cell presenting an antigen to it (Gaud, Lesourne, and Love 2018; Rossjohn et al. 2015). The many avenues of analysis ensure a degree of robustness to the findings upon which current models are based, but it is important to note that many of the studies have been carried out utilizing the Jurkat cell line which carries a Cosmc mutation resulting in global expression of the Tn antigen (Ju and Cummings 2002). As observed section 2.3, Cosmc is essential for the physiological persistence of T cells, but Jurkat cells provide clear evidence that Cosmc is not required for *in vitro* survival or for the induction of signals through the TCR. Nevertheless, signal strength is critically important in the selection and activation of T cells, and Cosmc may well exert an effect on the signal strength due to the elaboration of glycans on the TCR complex or other glycoproteins participating in the immune synapse. The majority of those proteins are predicted to bear O-glycans, and CD45 and CD43 in particular are heavily O-glycosylated proteins expressed in high abundance on the surface of T cells that are known to modulate T cell signaling based on their glycosylation state and interactions with glycan binding proteins (van

Vliet et al. 2006; Mazurov et al. 2012). It would be prudent to carefully examine the source of data describing T cell signaling events and look for possible discrepancies between observations utilizing Jurkat cells and cells competent for T-synthase function. Understanding any differences illuminated by this comparison could lead to further comprehension of the modulation of immune signaling by host or microbial cells, and to potential avenues for therapeutic intervention to T cell responses. Additionally, Jurkat cells are a lab-adapted cell line and are likely to contain multiple compensatory adaptations for any growth and survival deficiencies introduced by loss of Cosmc. Therefore, comparing primary WT and Cosmc$^{-/-}$ cell types at a signaling and genetic level will allow for the identification of critical pathways that require elaborated O-glycosylation. If these pathways are also pharmacologically targetable, then it could reveal additional therapeutic avenues.

4.4 Cancer glycobiology

It has long been known that the glycosylation of tumors frequently differs from that of normal tissues, and multiple tumor-associated carbohydrate antigens (TACAs) have been described (Pinho and Reis 2015). Two of these are the Tn antigen and its sialylated derivative sialyl-Tn (sTn), and extensive work has been done to try to understand the physiological significance and underlying mechanism of appearance for these TACAs (Julien, Videira, and Delannoy 2012; Munkley and Elliott 2016; Fu et al. 2016; Ju and Cummings 2014). Tn and sTn appear on a high percentage of solid tumors (70-90%, depending on tumor type) and are an indicator of poor prognosis (Ju, Otto, and Cummings 2011), as discussed in section 1.2.3.2.

Multiple mechanisms have been found to underlie Tn or sTn expression, including direct mutation of Cosmc or T-synthase genes, mutations in their promoter regions, epigenetic silencing of those genes, and mislocalization of the enzymes(Ju et al. 2011; Munkley and Elliott

2016). Curiously, the detection of Tn or sTn antigens has not always correlated with either the gene expression levels of Cosmc and T-synthase or the detected levels of enzyme activity (Sun, Ju, and Cummings 2018; Dong et al. 2018) (Ogawa 2017). ST6GalNAc1, the enzyme utilized in this work to proteomically identify sites of O-glycosylation, is generally expressed at exceedingly low levels, often below the limit of detection for enzymatic activity (Uhlen 2015). Thus, its overexpression in tumors with detectable Tn and sTn antigen has led to hypotheses that it can outcompete T-synthase and induce the appearance of sTn directly (Ferreira et al. 2013; Dall'Olio et al. 2014). And, while ectopic expression of ST6GalNAc1 in cultured cells does result in surface expression of sTn, it does not abrogate the appearance of extended Core 1 based O-glycans as detected by lectin staining(Tsuji and Takashima 2014). This heterogeneity indicates that any dominance of ST6GalNAc1 over T synthase is incomplete and likely dependent on the glycoprotein substrate.

However, our discoveries regarding the glycosylation of ST6GalNAc1 enzyme in section 3.1 throw these suppositions into question. We found that ST6GalNAc1 itself contains many O-glycosylation sites and that the recombinant protein bears both Tn and sTn, along with more elaborated structures, suggests that the appearance of sTn in tumors expressing elevated ST6GalNAc1 may not indicate a widespread alteration of the tumor's glycosylation, but rather the direct detection of antigenic features of the enzyme itself. If this is the case, then approaches for biomarker based diagnosis and investigations of the link between sTn appearance and poor prognosis will need to focus more directly on ST6GalNAc1 itself. It also raises the question of whether ST6GalNAc1 is mislocalized to the cell surface and what affects such mislocalization may have, or whether histologic detection of sTn largely reflects intracellular localization of the

antigen. The answers to these questions will impact the direction of new antigen targeting and the modes of recognition available for immune surveillance of such tumors.

While previous publications have speculated on the competition between ST6GalNAc1 and other enzymes, both in normal physiology and in cancer biology, until now the glycan-dependent regulation of ST6GalNAc1 had not been considered. However, we found that the enzyme is heavily decorated with targets of its own activity, and that this property is conserved for orthologs across vertebrates, while at the same time being unique among other members of the same enzyme family. This fundamental discovery, discussed in section 3.2, strongly indicates that the presence of such targets is an important feature of the biology of ST6GalNAc1, for its functional regulation or intracellular localization. Investigating chimeric constructs or enzymatically dead versions of ST6GalNAc1 to discover the nature of this regulation will help elaborate our understanding of the pathway of glycosylation and how it is susceptible to dysfunction in cancer.

4.5 Glycoproteomics

Characterization of glycoproteins involves the dissecting glycan structure, identifying glycosite occupancy, and assigning structures to specific sites of modification. Each of these steps can be technically challenging, and current techniques work best on homogeneous samples and/or purified proteins. In the case of O-glycosylation, homogenous samples have been generated by deletion of Cosmc, but the enrichment of the resulting Tn-antigen-bearing glycoproteins currently relies on low affinity methods for isolation, which results in a need for a large amount of material and large sample size that is only applicable to cultured cell lines (Steentoft 2013). In order to apply current technology to primary samples or to smaller amounts of material requires an innovative approach with higher affinity isolation. Our choice of

ST6GalNAc1 in the method developed in chapter 3 is ideal in many respects for chemoenzymatic labeling for subsequent high affinity purification of glycoproteins, leveraging as it does the well-established high-affinity interaction of biotin and streptavidin, while relying on the broad activity of ST6GalNAc1 on nearly all Tn-bearing glycoproteins. Our approach in particular is suited for tissues expressing truncated O-glycans such as tumors bearing Tn or sTn. Nevertheless, other approaches utilizing other enzymes or conjugation methods are also viable and any expansion of the available techniques for glycoproteomic isolation will be welcome.

Analysis of normal tissues should be possible by similar approaches using other enzymes in the pathway of N- and O-glycosylation. For example, a wide variety of glycans are capped with sialic acid by terminal sialidases such as ST3Gal1 or ST6Gal1. Using a method that utilized these enzymes with biotinylated CMP-Neu5Ac would permit labeling and isolation of nearly all the glycoproteins in a given sample. Downstream separation based on lectin affinity or other methods of fractionation to reduce sample complexity could allow sensitive detection of many glycoforms of different proteins. Intermediate enzymes that act further downstream in the glycosylation pathway than ST6GalNAc1, but necessarily on the termini, are likely to label a smaller subset of possible glycoproteins. Chemoenzymatic addition of biotinylated donor sugars could leverage this feature to identify substrates of these enzymes and allow further dissection of the substrate specificities inherent in the pathway.

Sialyltransferases as a class are good candidates for developing these approaches as several of them have been demonstrated to tolerate bulky substituents such as biotin on the donor sugar (Capicciotti 2017, Kajihara 2007). If small sample glycoproteomics becomes widely applicable, comparison of tumor and normal tissue will be a priority as altered glycosylation

including both differences in glycan structure as well as glycosite occupancy are likely to influence tumor progression and poor outcomes associated altered glycophenotype.

On a more fundamental level, comparison of alternate tissue types should yield insight into the differential regulation of glycoproteins via differences in the glycosylation machinery. This type of tissue-specific regulation could highlight altered functions and different roles for similar proteins in different contexts.

4.6 Conclusion

Glycosylation is an essential modification of proteins that influences structure, conformation, localization, and interactions. Chapter 2 presented discoveries of the crucial role that O-glycosylation plays in T cell biology, including evidence that it is required for persistence of mature peripheral T cells. These discoveries suggest further directions of research that will further understanding of T cell life cycle, homing, and signaling. Each of these areas contains the potential for basic and translational discoveries to further understanding of human biology and to uncover therapeutic avenues for treatment of infections or autoimmune disorders.

Chapter 3 presented discoveries that ST6GalNAc1 acts to glycosylate itself, opening the way to further dissection of the regulation of this pathway which is consistently dysregulated in cancer. Targeting the antigenic carbohydrates directly has so far been unproductive, but enhanced understanding of the processes involved in this dysregulation will help to identify additional targets in the pathway that can be potentially more therapeutically tractable. Additionally, chapter 3 presented a powerful method for identifying glycoproteins in a limited sample that will lead to important insights in tumor biology. This method can be extended and adapted using other glycosyltransferases, allowing for elucidation of substrate specificities and dissection of other glycosylation pathways.

The fundamental discoveries discussed in this dissertation are novel and directly advance the field's understanding of T cell glycobiology and of cancer associated carbohydrate antigens. Moreover, these discoveries pave the way for multiple future directions of research that have great potential for increasing basic and translational knowledge in multiple disciplines.

References

Abbas, A. K., K. M. Murphy, and A. Sher. 1996. 'Functional diversity of helper T lymphocytes', *Nature*, 383: 787-93.

Abramowitz, L. K., and J. A. Hanover. 2018. 'T cell development and the physiological role of O-GlcNAc', *FEBS Lett*, 592: 3943-49.

Allen, A. C., P. S. Topham, S. J. Harper, and J. Feehally. 1997. 'Leucocyte beta 1,3 galactosyltransferase activity in IgA nephropathy', *Nephrol Dial Transplant*, 12: 701-6.

Allen, A. C., F. R. Willis, T. J. Beattie, and J. Feehally. 1998. 'Abnormal IgA glycosylation in Henoch-Schonlein purpura restricted to patients with clinical nephritis', *Nephrol Dial Transplant*, 13: 930-4.

Allen, A., S. Harper, and J. Feehally. 1997. 'Origin and structure of pathogenic IgA in IgA nephropathy', *Biochem Soc Trans*, 25: 486-90.

Alvarez, Georgina, Ricardo Lascurain, Pedro Hernández-Cruz, Daniel Tetaert, Pierre Degand, Patricia Gorocica, Blanca Espinosa, Edgar Zenteno, and Raúl Chávez. 2006. 'Differential O-glycosylation in cortical and medullary thymocytes.', *Biochimica et biophysica acta*, 1760: 1235-40.

Anelli, T., and R. Sitia. 2008. 'Protein quality control in the early secretory pathway', *EMBO J*, 27: 315-27.

Apostolopoulos, Vasso, Elizabeth Yuriev, Paul a Ramsland, Jodie Halton, Carla Osinski, Wenjun Li, Magdalena Plebanski, Hans Paulsen, and Ian F C McKenzie. 2003. 'A glycopeptide in complex with MHC class I uses the GalNAc residue as an anchor.', *Proceedings of the National Academy of Sciences of the United States of America*, 100: 15029-34.

Apweiler, R., H. Hermjakob, and N. Sharon. 1999. 'On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database', *Biochim Biophys Acta*, 1473: 4-8.

Arnold, B. 2002. 'Levels of peripheral T cell tolerance', *Transpl Immunol*, 10: 109-14.

Aryal, R. P., T. Ju, and R. D. Cummings. 2010. 'The endoplasmic reticulum chaperone Cosmc directly promotes in vitro folding of T-synthase', *J Biol Chem*, 285: 2456-62.

———. 2012. 'Tight complex formation between Cosmc chaperone and its specific client non-native T-synthase leads to enzyme activity and client-driven dissociation', *J Biol Chem*, 287: 15317-29.

Ashton-Rickardt, P. G., A. Bandeira, J. R. Delaney, L. Van Kaer, H. P. Pircher, R. M. Zinkernagel, and S. Tonegawa. 1994. 'Evidence for a differential avidity model of T cell selection in the thymus', *Cell*, 76: 651-63.

Avci, Fikri Y, Xiangming Li, Moriya Tsuji, and Dennis L Kasper. 2011. 'A mechanism for glycoconjugate vaccine activation of the adaptive immune system and its implications for vaccine design.', *Nature medicine*, 17: 1602-9.

Balcan, Erdal, Ayça Gümüş, and Mesut Sahin. 2008. 'The glycosylation status of murine [corrected] postnatal thymus: a study by histochemistry and lectin blotting.', *Journal of molecular histology*, 39: 417-26.

Balcan, Erdal, Ibrahim Tuğlu, Mesut Sahin, and Pervin Toparlak. 2008. 'Cell surface glycosylation diversity of embryonic thymic tissues.', *Acta histochemica*, 110: 14-25.

Bansil, R., E. Stanley, and J. T. LaMont. 1995. 'Mucin biophysics', *Annu Rev Physiol*, 57: 635-57.

Barry, M., and R. C. Bleackley. 2002. 'Cytotoxic T lymphocytes: all roads lead to death', *Nat Rev Immunol*, 2: 401-9.

Bartolazzi, A. 2018. 'Galectins in Cancer and Translational Medicine: From Bench to Bedside', *Int J Mol Sci*, 19.

Bennett, E. P., U. Mandel, H. Clausen, T. A. Gerken, T. A. Fritz, and L. A. Tabak. 2012. 'Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family', *Glycobiology*, 22: 736-56.

Berger, E G. 1999. 'Tn-syndrome.', *Biochimica et biophysica acta*, 1455: 255-68.

Berger EG, Thurnher M, Dinter A. 1994. 'The cell-specific expression of glycan antigens exemplified by the deficiency of beta 1,3 galactosyltransferase in the Tn-syndrome', *Transfus Clin Biol*, 1: 103-08.

Bernard, Dannie, Michael S Ventresca, Laura a Marshall, Carole Evelegh, Yonghong Wan, and Jonathan L Bramson. 2010. 'Processing of tumor antigen differentially impacts the development of helper and effector CD4+ T-cell responses.', *Molecular therapy : the journal of the American Society of Gene Therapy*, 18: 1224-32.

Berney, S. N., and B. M. Gesner. 1970. 'The circulatory behaviour of normal and enzyme altered thymocytes in rats', *Immunology*, 18: 681-91.

Bi, Shuguang, and Linda G Baum. 2009. 'Sialic acids in T cell development and function.', *Biochimica et biophysica acta*, 1790: 1599-610.

Blank, Christian, and Andreas Mackensen. 2007. 'Contribution of the PD-L1/PD-1 pathway to T-cell exhaustion: an update on implications for chronic infections and tumor evasion', *Cancer Immunology, Immunotherapy*, 56: 739--45.

Bleul, C. C., and T. Boehm. 2000. 'Chemokines define distinct microenvironments in the developing thymus', *Eur J Immunol*, 30: 3371-9.

Blond-Elguindi, S., S. E. Cwirla, W. J. Dower, R. J. Lipshutz, S. R. Sprang, J. F. Sambrook, and M. J. Gething. 1993. 'Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BiP', *Cell*, 75: 717-28.

Boehmer, Hv. 1986. 'The selection of the alpha,beta heterodimeric T-cell receptor for antigen', *Immunol Today*, 7: 333-6.

Bond, M. R., and J. J. Kohler. 2007. 'Chemical methods for glycoprotein discovery', *Curr Opin Chem Biol*, 11: 52-8.

Borowski, C., C. Martin, F. Gounari, L. Haughn, I. Aifantis, F. Grassi, and H. von Boehmer. 2002. 'On the brink of becoming a T cell', *Curr Opin Immunol*, 14: 200-6.

Bour-Jordan, H., and J. A. Blueston. 2002. 'CD28 function: a balance of costimulatory and regulatory signals', *J Clin Immunol*, 22: 1-7.

Braakman, I., and N. J. Bulleid. 2011. 'Protein folding and modification in the mammalian endoplasmic reticulum', *Annu Rev Biochem*, 80: 71-99.

Brockhausen, I. 1999. 'Pathways of O-glycan biosynthesis in cancer cells', *Biochim Biophys Acta*, 1473: 67-95.

———. 2000. 'O-linked chain glycosyltransferases', *Methods Mol Biol*, 125: 273-93.

Bromley, S. K., W. R. Burack, K. G. Johnson, K. Somersalo, T. N. Sims, C. Sumen, M. M. Davis, A. S. Shaw, P. M. Allen, and M. L. Dustin. 2001. 'The immunological synapse', *Annu Rev Immunol*, 19: 375-96.

Brouet JC, Vainchenker W, Blanchard D, Testa U, Cartron JP. 1983. 'The origin of human B and T cells from multipotent stem cells: a study of the Tn syndrome', *European Journal of Immunology*, 13: 350-52.

Brouet, Jean-Claude, William Vainchenker, Dominique Blanchard, Ugo Testa, and Jean-pierre Cartron. 1983. 'The origin of human B and T cells from multipotent stem cells: a study of the Tn syndrome', *European journal of immunology*, 13: 350-52.

Buckley, R. H. 2004. 'Molecular defects in human severe combined immunodeficiency and approaches to immune reconstitution', *Annu Rev Immunol*, 22: 625-55.

Cao, Y., U. R. Karsten, W. Liebrich, W. Haensch, G. F. Springer, and P. M. Schlag. 1995. 'Expression of Thomsen-Friedenreich-related antigens in primary and metastatic colorectal carcinomas. A reevaluation', *Cancer*, 76: 1700-8.

Carbone, F R, and P a Gleeson. 1997. 'Carbohydrates and antigen recognition by T cells.', *Glycobiology*, 7: 725-30.

Carlyle, J. R., and J. C. Zuniga-Pflucker. 1998. 'Requirement for the thymus in alphabeta T lymphocyte lineage commitment', *Immunity*, 9: 187-97.

Cedeno-Laurent, Filiberto, and Charles J Dimitroff. 2012. 'Galectin-1 research in T cell immunity: Past, present, and future', 142: 107-16.

Chetry, M., S. Thapa, X. Hu, Y. Song, J. Zhang, H. Zhu, and X. Zhu. 2018. 'The Role of Galectins in Tumor Progression, Treatment and Prognosis of Gynecological Cancers', *J Cancer*, 9: 4742-55.

Chiang, Y. J., and R. J. Hodes. 2016. 'T-cell development is regulated by the coordinated function of proximal and distal Lck promoters active at different developmental stages', *Eur J Immunol*, 46: 2401-08.

Chik, J. H., J. Zhou, E. S. Moh, R. Christopherson, S. J. Clarke, M. P. Molloy, and N. H. Packer. 2014. 'Comprehensive glycomics comparison between colon cancer cell cultures and tumours: implications for biomarker studies', *J Proteomics*, 108: 146-62.

Christiansen, M. N., J. Chik, L. Lee, M. Anugraham, J. L. Abrahams, and N. H. Packer. 2014. 'Cell surface protein glycosylation in cancer', *Proteomics*, 14: 525-46.

Chu, H. H., S. W. Chan, J. P. Gosling, N. Blanchard, A. Tsitsiklis, G. Lythe, N. Shastri, C. Molina-Paris, and E. A. Robey. 2016. 'Continuous Effector CD8(+) T Cell Production in a Controlled Persistent Infection Is Sustained by a Proliferative Intermediate Population', *Immunity*, 45: 159-71.

Clark, H. F., A. L. Gurney, E. Abaya, K. Baker, D. Baldwin, J. Brush, J. Chen, B. Chow, C. Chui, C. Crowley, B. Currell, B. Deuel, P. Dowd, D. Eaton, J. Foster, C. Grimaldi, Q. Gu, P. E. Hass, S. Heldens, A. Huang, H. S. Kim, L. Klimowski, Y. Jin, S. Johnson, J. Lee, L. Lewis, D. Liao, M. Mark, E. Robbie, C. Sanchez, J. Schoenfeld, S. Seshagiri, L. Simmons, J. Singh, V. Smith, J. Stinson, A. Vagts, R. Vandlen, C. Watanabe, D. Wieand, K. Woods, M. H. Xie, D. Yansura, S. Yi, G. Yu, J. Yuan, M. Zhang, Z. Zhang, A. Goddard, W. I. Wood, P. Godowski, and A. Gray. 2003. 'The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment', *Genome Res*, 13: 2265-70.

Clark, Mary C, and Linda G Baum. 2012. 'T cells modulate glycans on CD43 and CD45 during development and activation, signal regulation, and survival.', *Annals of the New York Academy of Sciences*, 1253: 58-67.

Consortium, C. AZypedia. 2018. 'Ten years of CAZypedia: a living encyclopedia of carbohydrate-active enzymes', *Glycobiology*, 28: 3-8.

Cosgrove, D., S. H. Chan, C. Waltzinger, C. Benoist, and D. Mathis. 1992. 'The thymic compartment responsible for positive selection of CD4+ T cells', *Int Immunol*, 4: 707-10.

Crispe, I. N., T. Dao, K. Klugewitz, W. Z. Mehal, and D. P. Metz. 2000. 'The liver as a site of T-cell apoptosis: graveyard, or killing field?', *Immunol Rev*, 174: 47-62.

Crispe, I. N., and L. Huang. 1994. 'Neonatal, moribund and undead T cells: the role of the liver in T cell development', *Semin Immunol*, 6: 39-41.

Cummings, R. D., and J. M. Pierce. 2014. 'The challenge and promise of glycomics', *Chem Biol*, 21: 1-15.

Cyster, J. G. 1999. 'Chemokines and cell migration in secondary lymphoid organs', *Science*, 286: 2098-102.

———. 2000. 'Leukocyte migration: scent of the T zone', *Curr Biol*, 10: R30-3.

Dahr, W., G. Uhlenbruck, and G. W. Bird. 1974. 'Cryptic A-like receptor sites in human erythrocyte glycoproteins: proposed nature of Tn-antigen', *Vox Sang*, 27: 29-42.

Dall'Olio, F., N. Malagolini, M. Trinchera, and M. Chiricolo. 2014. 'Sialosignaling: sialyltransferases as engines of self-fueling loops in cancer progression', *Biochim Biophys Acta*, 1840: 2752-64.

Dalpathado, D. S., and H. Desaire. 2008. 'Glycopeptide analysis by mass spectrometry', *Analyst*, 133: 731-8.

Danguy, A., I. Camby, and R. Kiss. 2002. 'Galectins and cancer', *Biochim Biophys Acta*, 1572: 285-93.

Daniels, M. A., L. Devine, J. D. Miller, J. M. Moser, A. E. Lukacher, J. D. Altman, P. Kavathas, K. A. Hogquist, and S. C. Jameson. 2001. 'CD8 binding to MHC class I molecules is influenced by T cell maturation and glycosylation.', *Immunity*, 15: 1051--61.

Daniels, Mark A., Kristin A. Hogquist, and Stephen C. Jameson. 2002. 'Sweet 'n' sour: the impact of differential glycosylation on T cell responses', *Nature immunology*, 3: 903--10.

Dausset, J., J. Moullec, and J. Bernard. 1959. 'Acquired hemolytic anemia with polyagglutinability of red blood cells due to a new factor present in normal human serum (Anti-Tn)', *Blood*, 14: 1079-93.

De Boer, R. J., and A. S. Perelson. 2013. 'Quantifying T lymphocyte turnover', *J Theor Biol*, 327: 45-87.

Deck, B, M Elofsson, J Kihlberg, and E R Unanue. 1995. 'Specificity of glycopeptide-specific T cells.', *Journal of immunology (Baltimore, Md. : 1950)*, 155: 1074-8.

Demetriou, Michael, Maria Granovsky, Sue Quaggin, and James W. Dennis. 2001. 'Negative regulation of T-cell activation and autoimmunity by Mgat5 N-glycosylation', *Nature*, 409: 733--39.

Derfalvi, B., K. Maurer, D. M. McDonald McGinn, E. Zackai, W. Meng, E. T. Luning Prak, and K. E. Sullivan. 2016. 'B cell development in chromosome 22q11.2 deletion syndrome', *Clin Immunol*, 163: 1-9.

Dias, Ana M., Alexandra Correia, Mrcia S. Pereira, Catarina R. Almeida, Ins Alves, Vanda Pinto, Telmo A. Catarino, Nuno Mendes, Magdalena Leander, M. Teresa Oliva-Teles, Lus Maia, Cristina Delerue-Matos, Naoyuki Taniguchi, Margarida Lima, Isabel Pedroto, Ricardo Marcos-Pinto, Paula Lago, Celso A. Reis, Manuel Vilanova, and Salom S. Pinho. 2018. 'Metabolic control of T cell immune response through glycans in inflammatory bowel disease', *Proceedings of the National Academy of Sciences*, 115: E4651--E60.

Dias, Ana M., Joana Dourado, Paula Lago, Joana Cabral, Ricardo Marcos-Pinto, Paulo Salgueiro, Catarina R. Almeida, Sandra Carvalho, Snia Fonseca, Margarida Lima, Manuel Vilanova, Mrio Dinis-Ribeiro, Celso A. Reis, and Salom S. Pinho. 2014. 'Dysregulation of T cell receptor N-glycosylation: a molecular mechanism involved in ulcerative colitis', *Human Molecular Genetics*, 23: 2416--27.

Dong, X., Y. Jiang, J. Liu, Z. Liu, T. Gao, G. An, and T. Wen. 2018. 'T-Synthase Deficiency Enhances Oncogenic Features in Human Colorectal Cancer Cells via Activation of Epithelial-Mesenchymal Transition', *Biomed Res Int*, 2018: 9532389.

Duan, Jinyou, Fikri Y Avci, and Dennis L Kasper. 2008. 'Microbial carbohydrate depolymerization by antigen-presenting cells: deamination prior to presentation by the MHCII pathway.', *Proceedings of the National Academy of Sciences of the United States of America*, 105: 5183-8.

Dustin, Michael L, and Andrew C Chan. 2000. 'Signaling Takes Shape in the Immune System', *Cell*, 103: 283-94.

Dzhagalov, Ivan, and Hyewon Phee. 2012. 'How to find your way through the thymus: a practical guide for aspiring T cells.', *Cellular and molecular life sciences : CMLS*, 69: 663-82.

Earl, L A, and L G Baum. 2008. 'CD45 glycosylation controls T-cell life and death.', *Immunology and cell biology*, 86: 608-15.

Earl, Lesley a, Shuguang Bi, and Linda G Baum. 2010. 'N- and O-glycans modulate galectin-1 binding, CD45 signaling, and T cell death.', *The Journal of biological chemistry*, 285: 2232-44.

Evans, Edward J, Lawrence Hene, Lisa M Sparks, Tao Dong, Christelle Retiere, Janet A Fennelly, Raquel Manso-Sancho, Jill Powell, Veronique M Braud, Sarah L Rowland-Jones, Andrew J McMichael, and Simon J Davis. 2003. 'The T cell surface--how well do we know it?', *Immunity*, 19: 213-23.

Fantini, M. C., C. Becker, G. Monteleone, F. Pallone, P. R. Galle, and M. F. Neurath. 2004. 'Cutting edge: TGF-beta induces a regulatory phenotype in CD4+CD25- T cells through Foxp3 induction and down-regulation of Smad7', *J Immunol*, 172: 5149-53.

Farber, D. L., N. A. Yudanin, and N. P. Restifo. 2014. 'Human memory T cells: generation, compartmentalization and homeostasis', *Nat Rev Immunol*, 14: 24-35.

Fernandez, M., J. Keyrilainen, R. Serimaa, M. Torkkeli, M. L. Karjalainen-Lindsberg, M. Leidenius, K. von Smitten, M. Tenhunen, S. Fiedler, A. Bravin, T. M. Weiss, and P. Suortti. 2005. 'Human breast cancer in vitro: matching histo-pathology with small-angle x-ray scattering and diffraction enhanced x-ray imaging', *Phys Med Biol*, 50: 2991-3006.

Ferreira, J. A., P. A. Videira, L. Lima, S. Pereira, M. Silva, M. Carrascal, P. F. Severino, E. Fernandes, A. Almeida, C. Costa, R. Vitorino, T. Amaro, M. J. Oliveira, C. A. Reis, F. Dall'Olio, F. Amado, and L. L. Santos. 2013. 'Overexpression of tumour-associated carbohydrate antigen sialyl-Tn in advanced bladder tumours', *Mol Oncol*, 7: 719-31.

Fink, A. L. 1999. 'Chaperone-mediated protein folding', *Physiol Rev*, 79: 425-49.

Fischer, C., H. Sanchez-Ruderisch, M. Welzel, B. Wiedenmann, T. Sakai, S. Andre, H. J. Gabius, L. Khachigian, K. M. Detjen, and S. Rosewicz. 2005. 'Galectin-1 interacts with the {alpha}5{beta}1 fibronectin receptor to restrict carcinoma cell growth via induction of p21 and p27', *J Biol Chem*, 280: 37266-77.

Freitas, A A, and B Rocha. 1999. 'Peripheral T cell survival.', *Current opinion in immunology*, 11: 152-6.

Frommhold, David, Andreas Ludwig, M Gabriele Bixel, Alexander Zarbock, Inna Babushkina, Melitta Weissinger, Sandra Cauwenberghs, Lesley G Ellies, Jamey D Marth, Annette G Beck-Sickinger, Michael Sixt, Bärbel Lange-Sperandio, Alma Zernecke, Ernst Brandt, Christian Weber, Dietmar Vestweber, Klaus Ley, and Markus Sperandio. 2008. 'Sialyltransferase ST3Gal-IV controls CXCR2-mediated firm leukocyte arrest during inflammation.', *The Journal of experimental medicine*, 205: 1435-46.

Fu, C., H. Zhao, Y. Wang, H. Cai, Y. Xiao, Y. Zeng, and H. Chen. 2016. 'Tumor-associated antigens: Tn antigen, sTn antigen, and T antigen', *HLA*, 88: 275-86.

Funasaka, T., A. Raz, and P. Nangia-Makker. 2014. 'Galectin-3 in angiogenesis and metastasis', *Glycobiology*, 24: 886-91.

Ganusov, V. V., and J. Auerbach. 2014. 'Mathematical modeling reveals kinetics of lymphocyte recirculation in the whole organism', *PLoS Comput Biol*, 10: e1003586.

Garcia, Gonzalo G., and Scott B. and Berger. 2005. 'Age-associated changes in glycosylation of CD43 and CD45 on mouse CD4 T cells', *European journal of immunology*, 35: 622--31.

Gatti, E., and P. Pierre. 2003. 'Understanding the cell biology of antigen presentation: the dendritic cell contribution', *Curr Opin Cell Biol*, 15: 468-73.

Gaud, G., R. Lesourne, and P. E. Love. 2018. 'Regulatory mechanisms in T cell receptor signalling', *Nat Rev Immunol*, 18: 485-97.

Gauguet, J. M., S. D. Rosen, J. D. Marth, and U. H. von Andrian. 2004. 'Core 2 branching beta1,6-N-acetylglucosaminyltransferase and high endothelial cell N-acetylglucosamine-6-sulfotransferase exert differential control over B- and T-lymphocyte homing to peripheral lymph nodes', *Blood*, 104: 4104-12.

Gesner, B. M., and V. Ginsburg. 1964. 'Effect of Glycosidases on the Fate of Transfused Lymphocytes', *Proc Natl Acad Sci U S A*, 52: 750-5.

Gill, David J, Keit Min Tham, Joanne Chia, Shyi Chyi Wang, Catharina Steentoft, Henrik Clausen, Emilie A Bard-Chapeau, and Frederic A Bard. 2013. 'Initiation of GalNAc-type O-glycosylation in the endoplasmic reticulum promotes cancer cell invasiveness.', *Proceedings of the National Academy of Sciences of the United States of America*, 110: E3152-61.

Girard, Jean-Philippe, Christine Moussion, and Reinhold Förster. 2012. 'HEVs, lymphatics and homeostatic immune cell trafficking in lymph nodes.', *Nature reviews. Immunology*, 12: 762-73.

Girardi, E., and D. M. Zajonc. 2012. 'Molecular basis of lipid antigen presentation by CD1d and recognition by natural killer T cells', *Immunol Rev*, 250: 167-79.

Gloster, T. M. 2014. 'Advances in understanding glycosyltransferases from a structural perspective', *Curr Opin Struct Biol*, 28: 131-41.

Grewal, P. K. 2010. 'The Ashwell-Morell receptor', *Methods Enzymol*, 479: 223-41.
Gunson, H. H., F. Stratton, and G. W. Mullard. 1970. 'An example of polyagglutinability due to the Tn antigen', *Br J Haematol*, 18: 309-16.
Hakomori, S. 2001. 'Tumor-associated carbohydrate antigens defining tumor malignancy: basis for development of anti-cancer vaccines.', *Advances in experimental medicine and biology*, 491: 369-402.
Hakomori, S. I., and R. D. Cummings. 2012. 'Glycosylation effects on cancer development', *Glycoconj J*, 29: 565-6.
Han, L., and C. E. Costello. 2013. 'Mass spectrometry of glycans', *Biochemistry (Mosc)*, 78: 710-20.
Hanes, M. S., K. W. Moremen, and R. D. Cummings. 2017. 'Biochemical characterization of functional domains of the chaperone Cosmc', *PloS one*, 12: e0180242.
Hanover, J. A., W. Chen, and M. R. Bond. 2018. 'O-GlcNAc in cancer: An Oncometabolism-fueled vicious cycle', *J Bioenerg Biomembr*, 50: 155-73.
Harty, J. T., A. R. Tvinnereim, and D. W. White. 2000. 'CD8+ T cell effector mechanisms in resistance to infection', *Annu Rev Immunol*, 18: 275-308.
Haslam, Stuart M, Sylvain Julien, Joy M Burchell, Clare R Monk, Alessio Ceroni, Oliver a Garden, and Anne Dell. 2008. 'Characterizing the glycome of the mammalian immune system.', *Immunology and cell biology*, 86: 564-73.
Hava, D. L., M. Brigl, P. van den Elzen, D. M. Zajonc, I. A. Wilson, and M. B. Brenner. 2005. 'CD1 assembly and the formation of CD1-antigen complexes', *Curr Opin Immunol*, 17: 88-94.
Heimburg-Molinaro, J., A. Almogren, S. Morey, O. V. Glinskii, R. Roy, G. E. Wilding, R. P. Cheng, V. V. Glinsky, and K. Rittenhouse-Olson. 2009. 'Development, characterization, and immunotherapeutic use of peptide mimics of the Thomsen-Friedenreich carbohydrate antigen', *Neoplasia*, 11: 780-92.
Heimburg-Molinaro, J., M. Lum, G. Vijay, M. Jain, A. Almogren, and K. Rittenhouse-Olson. 2011. 'Cancer vaccines and carbohydrate epitopes', *Vaccine*, 29: 8802-26.
Hendershot, L. M., and N. J. Bulleid. 2000. 'Protein-specific chaperones: the role of hsp47 begins to gel', *Curr Biol*, 10: R912-5.
Hengel, R. L., V. Thaker, M. V. Pavlick, J. A. Metcalf, G. Dennis, Jr., J. Yang, R. A. Lempicki, I. Sereti, and H. C. Lane. 2003. 'Cutting edge: L-selectin (CD62L) expression distinguishes small resting memory CD4+ T cells that preferentially respond to recall antigen', *J Immunol*, 170: 28-32.
Hennet, T, F K Hagen, L a Tabak, and J D Marth. 1995. 'T-cell-specific deletion of a polypeptide N-acetylgalactosaminyl-transferase gene by site-directed recombination.', *Proceedings of the National Academy of Sciences of the United States of America*, 92: 12070-4.
Hernandez, Joseph D, Jeffrey Klein, Stephen J Van Dyken, Jamey D Marth, and Linda G Baum. 2007. 'T-cell activation results in microheterogeneous changes in glycosylation of CD45.', *International immunology*, 19: 847-56.
Hirabayashi, J. 2004. 'Lectin-based structural glycomics: glycoproteomics and glycan profiling', *Glycoconj J*, 21: 35-40.
Hirabayashi, Jun, Tomomi Hashidate, and Ken-ichi Kasai. 2002. 'Glyco-catch method: A lectin affinity technique for glycoproteomics.', *Journal of biomolecular techniques : JBT*, 13: 205-18.
Hofmann, B. T., L. Schluter, P. Lange, B. Mercanoglu, F. Ewald, A. Folster, A. S. Picksak, S. Harder, A. T. El Gammal, K. Grupp, C. Gungor, A. Drenckhan, H. Schluter, C. Wagener, J. R. Izbicki, M. Jucker, M. Bockhorn, and G. Wolters-Eisfeld. 2015. 'COSMC knockdown mediated aberrant O-glycosylation promotes oncogenic properties in pancreatic cancer', *Mol Cancer*, 14: 109.
Hogquist, Kristin A, Yan Xing, Fan-Chi Hsu, and Virginia Smith Shapiro. 2015. 'T Cell Adolescence: Maturation Events Beyond Positive Selection.', *Journal of immunology (Baltimore, Md. : 1950)*, 195: 1351-7.

Hong, P., H. Sun, L. Sha, Y. Pu, K. Khatri, X. Yu, Y. Tang, and C. Lin. 2017. 'GlycoDeNovo - an Efficient Algorithm for Accurate de novo Glycan Topology Reconstruction from Tandem Mass Spectra', *J Am Soc Mass Spectrom*, 28: 2288-301.

Houston, E. G., Jr., L. E. Higdon, and P. J. Fink. 2011. 'Recent thymic emigrants are preferentially incorporated only into the depleted T-cell pool', *Proc Natl Acad Sci U S A*, 108: 5366-71.

Hsiao, C. T., P. W. Wang, H. C. Chang, Y. Y. Chen, S. H. Wang, Y. Chern, and K. H. Khoo. 2017. 'Advancing a High Throughput Glycotope-centric Glycomics Workflow Based on nanoLC-MS(2)-product Dependent-MS(3) Analysis of Permethylated Glycans', *Mol Cell Proteomics*, 16: 2268-80.

Huesmann, M, B Scott, P Kisielow, and H von Boehmer. 1991. 'Kinetics and efficacy of positive selection in the thymus of normal and T cell receptor transgenic mice.', *Cell*, 66: 533-40.

Ilarregui, Juan M., Diego O. Croci, Germn A. Bianco, Marta A. Toscano, Mariana Salatino, Mnica E. Vermeulen, Jorge R. Geffner, and Gabriel A. Rabinovich. 2009. 'Tolerogenic signals delivered by dendritic cells to T cells through a galectin-1-driven immunoregulatory circuit involving interleukin 27 and interleukin 10', *Nature immunology*, 10: 981--91.

Itano, A. A., and M. K. Jenkins. 2003. 'Antigen presentation to naive CD4 T cells in the lymph node', *Nat Immunol*, 4: 733-9.

Iwakoshi, N. N., I. Goldschneider, F. Tausche, J. P. Mordes, A. A. Rossini, and D. L. Greiner. 1998. 'High frequency apoptosis of recent thymic emigrants in the liver of lymphopenic diabetes-prone BioBreeding rats', *J Immunol*, 160: 5838-50.

Jaeken, J. 2011. 'Congenital disorders of glycosylation (CDG): it's (nearly) all in it!', *J Inherit Metab Dis*, 34: 853-58.

Jankovic, V., K. Remus, A. Molano, and J. Nikolich-Zugich. 2002. 'T cell recognition of an engineered MHC class I molecule: implications for peptide-independent alloreactivity', *J Immunol*, 169: 1887-92.

Jin, K. T., H. R. Lan, X. Y. Chen, S. B. Wang, X. J. Ying, Y. Lin, and X. Z. Mou. 2019. 'Recent advances in carbohydrate-based cancer vaccines', *Biotechnol Lett*, 41: 641-50.

Jones, Amanda Tomlinson, Bartolome Federsppiel, Lesley C Ellies, Michael J Williams, Roger Burgener, Vincent Duronio, C Anne Smith, Fumio Takei, and Hermann J Ziltener. 1994. 'Characterization of the Activation-Associated Isoform of CD43 on Murine T Lymphocytes', *Journal of immunology (Baltimore, Md. : 1950)*.

Jones, M.B. 2018. 'IgG and leukocytes: Targets of immunomodulatory α2,6 sialic acids', *Cell Immunol*, 333: 58-64.

Ju, T., Q. Zheng, and R. D. Cummings. 2006. 'Identification of core 1 O-glycan T-synthase from Caenorhabditis elegans', *Glycobiology*, 16: 947-58.

Ju, Tongzhong, and Richard D Cummings. 2002. 'A unique molecular chaperone Cosmc required for activity of the mammalian core 1 beta 3-galactosyltransferase.', *Proceedings of the National Academy of Sciences of the United States of America*, 99: 16613-8.

Ju, Tongzhong, and Richard D. Cummings. 2014. 'Core 1 β3Galactosyltransferase (C1GalT1, T-Synthase) and Its Specific Molecular Chaperone Cosmc (C1GalT1C1).' in Naoyuki Taniguchi, Koichi Honke, Minoru Fukuda, Hisashi Narimatsu, Yoshiki Yamaguchi and Takashi Angata (eds.), *Handbook of Glycosyltransferases and Related Genes* (Springer Japan: Tokyo).

Ju, Tongzhong, Grainger S Lanneau, Tripti Gautam, Yingchun Wang, Baoyun Xia, Sean R Stowell, Margaret T Willard, Wenyi Wang, Jonathan Y Xia, Rosemary E Zuna, Zoltan Laszik, Doris M Benbrook, Marie H Hanigan, and Richard D Cummings. 2008. 'Human tumor antigens Tn and sialyl Tn arise from mutations in Cosmc.', *Cancer research*, 68: 1636-46.

Ju, Tongzhong, Vivianne I Otto, and Richard D Cummings. 2011. 'The Tn antigen-structural simplicity and biological complexity.', *Angewandte Chemie (International ed. in English)*, 50: 1770-91.

Ju, Tongzhong, Baoyun Xia, Rajindra P Aryal, Wenyi Wang, Yingchun Wang, Xiaokun Ding, Rongjuan Mi, Miao He, and Richard D Cummings. 2011. 'A novel fluorescent assay for T-synthase activity.', *Glycobiology*, 21: 352-62.

Ju, Tongzhong; Cummings, Richard D. . 2005. 'Protein glycosylation: chaperone mutation in Tn syndrome', *Nature*, 437: 1252.

Judson, P A, F A Spring, M A Taylor, and D J Anstee. 1983. 'Evidence for carbohydrate-deficient forms of the major sialoglycoproteins of human platelets, granulocytes and T lymphocytes in individuals with Tn syndrome.', *Immunology*, 50: 415-22.

Judson PA, Spring FA, Taylor MA, Anstee DJ. 1983. 'Evidence for carbohydrate-deficient forms of the major sialoglycoproteins of human platelets, granulocytes and T lymphocytes in individuals with Tn syndrome', *Immunology*, 50: 415-22.

Julien, Sylvain, Paula A Videira, and Philippe Delannoy. 2012. 'Sialyl-tn in cancer: (how) did we miss the target?', *Biomolecules*, 2: 435-66.

Kakeji, Y., S. Tsujitani, M. Mori, Y. Maehara, and K. Sugimachi. 1991. 'Helix pomatia agglutinin binding activity is a predictor of survival time for patients with gastric carcinoma', *Cancer*, 68: 2438-42.

Kalka-Moll, W. M., A. O. Tzianabos, P. W. Bryant, M. Niemeyer, H. L. Ploegh, and D. L. Kasper. 2002. 'Zwitterionic polysaccharides stimulate T cells by MHC class II-dependent interactions', *J Immunol*, 169: 6149-53.

Kao, Charlly, and MM Sandau. 2006. 'The sialyltransferase ST3Gal-I is not required for regulation of CD8-class I MHC binding during T cell development', *The Journal of …*.

Kapsenberg, M. L. 2003. 'Dendritic-cell control of pathogen-driven T-cell polarization', *Nat Rev Immunol*, 3: 984-93.

Kawashima, Hiroto, and Minoru Fukuda. 2012. 'Sulfated glycans control lymphocyte homing.', *Annals of the New York Academy of Sciences*, 1253: 112-21.

Khidekel, N., S. Arndt, N. Lamarre-Vincent, A. Lippert, K. G. Poulin-Kerstien, B. Ramakrishnan, P. K. Qasba, and L. C. Hsieh-Wilson. 2003. 'A chemoenzymatic approach toward the rapid and sensitive detection of O-GlcNAc posttranslational modifications', *J Am Chem Soc*, 125: 16162-3.

Khidekel, N., S. B. Ficarro, E. C. Peters, and L. C. Hsieh-Wilson. 2004. 'Exploring the O-GlcNAc proteome: direct identification of O-GlcNAc-modified proteins from the brain', *Proc Natl Acad Sci U S A*, 101: 13132-7.

Kim, H. K., A. T. Waickman, E. Castro, F. A. Flomerfelt, N. V. Hawk, V. Kapoor, W. G. Telford, and R. E. Gress. 2016. 'Distinct IL-7 signaling in recent thymic emigrants versus mature naive T cells controls T-cell homeostasis', *Eur J Immunol*, 46: 1669-80.

Kishimoto, H., and J. Sprent. 1997. 'Negative selection in the thymus includes semimature T cells', *J Exp Med*, 185: 263-71.

Koch, U., T. A. Lacombe, D. Holland, J. L. Bowman, B. L. Cohen, S. E. Egan, and C. J. Guidos. 2001. 'Subversion of the T/B lineage decision in the thymus by lunatic fringe-mediated inhibition of Notch-1.', *Immunity*, 15: 225--36.

Koch, Ute, and Freddy Radtke. 2011. 'Mechanisms of T Cell Development and Transformation', *Annual Review of Cell and Developmental Biology*, 27: 539--62.

Konno, A., Y. Hoshino, S. Terashima, R. Motoki, and T. Kawaguchi. 2002. 'Carbohydrate expression profile of colorectal cancer cells is relevant to metastatic pattern and prognosis', *Clin Exp Metastasis*, 19: 61-70.

Kristic, J., and G. Lauc. 2017. 'Ubiquitous Importance of Protein Glycosylation', *Methods Mol Biol*, 1503: 1-12.

Ku, C. C., M. Murakami, A. Sakamoto, J. Kappler, and P. Marrack. 2000. 'Control of homeostasis of CD8+ memory T cells by opposing cytokines', *Science*, 288: 675-8.

Kumar, Brahma V., Thomas J. Connors, and Donna L. Farber. 2018. 'Human T Cell Development, Localization, and Function throughout Life', *Immunity*, 48: 202--13.

Laack, E., H. Nikbakht, A. Peters, C. Kugler, Y. Jasiewicz, L. Edler, D. K. Hossfeld, and U. Schumacher. 2002. 'Lectin histochemistry of resected adenocarcinoma of the lung: helix pomatia agglutinin binding is an independent prognostic factor', *Am J Pathol*, 160: 1001-8.

Lairson, L. L., B. Henrissat, G. J. Davies, and S. G. Withers. 2008. 'Glycosyltransferases: structures, functions, and mechanisms', *Annu Rev Biochem*, 77: 521-55.

Lee, H. S., Y. Qi, and W. Im. 2015. 'Effects of N-glycosylation on protein conformation and dynamics: Protein Data Bank analysis and molecular dynamics simulation study', *Sci Rep*, 5: 8926.

Lehoux, S., and T. Ju. 2017. 'Separation of Two Distinct O-Glycoforms of Human IgA1 by Serial Lectin Chromatography Followed by Mass Spectrometry O-Glycan Analysis', *Methods Enzymol*, 585: 61-75.

Lehoux, S., R. Mi, R. P. Aryal, Y. Wang, K. T. Schjoldager, H. Clausen, I. van Die, Y. Han, A. B. Chapman, R. D. Cummings, and T. Ju. 2014. 'Identification of distinct glycoforms of IgA1 in plasma from patients with immunoglobulin A (IgA) nephropathy and healthy individuals', *Mol Cell Proteomics*, 13: 3097-113.

Leppänen, Anne, Tadayuki Yago, Vivianne I Otto, Rodger P McEver, and Richard D Cummings. 2003. 'Model glycosulfopeptides from P-selectin glycoprotein ligand-1 require tyrosine sulfation and a core 2-branched O-glycan to bind to L-selectin.', *The Journal of biological chemistry*, 278: 26391-400.

Levy, Y., R. Arbel-Goren, Y. R. Hadari, S. Eshhar, D. Ronen, E. Elhanany, B. Geiger, and Y. Zick. 2001. 'Galectin-8 functions as a matricellular modulator of cell adhesion', *J Biol Chem*, 276: 31285-95.

Ley, K., J. Rivera-Nieves, W. J. Sandborn, and S. Shattil. 2016. 'Integrin-based therapeutics: biological basis, clinical use and new drugs', *Nat Rev Drug Discov*, 15: 173-83.

Li, Chia-Wei, Seung-Oe Lim, Ezra M. Chung, Yong-Soo Kim, Andrew H. Park, Jun Yao, Jong-Ho Cha, Weiya Xia, Li-Chuan Chan, Taewan Kim, Shih-Shin Chang, Heng-Huan Lee, Chao-Kai Chou, Yen-Liang Liu, Hsin-Chih Yeh, Evan P. Perillo, Andrew K. Dunn, Chu-Wei Kuo, Kay-Hooi Khoo, Jennifer L. Hsu, Yun Wu, Jung-Mao Hsu, Hirohito Yamaguchi, Tzu-Hsuan Huang, Aysegul A. Sahin, Gabriel N. Hortobagyi, Stephen S. Yoo, and Mien-Chie Hung. 2018. 'Eradication of Triple-Negative Breast Cancer Cells by Targeting Glycosylated PD-L1', *Cancer Cell*, 33: 187--201.e10.

Li, Chia-Wei, Seung-Oe Lim, Weiya Xia, Heng-Huan Lee, Li-Chuan Chan, Chu-Wei Kuo, Kay-Hooi Khoo, Shih-Shin Chang, Jong-Ho Cha, Taewan Kim, Jennifer L. Hsu, Yun Wu, Jung-Mao Hsu, Hirohito Yamaguchi, Qingqing Ding, Yan Wang, Jun Yao, Cheng-Chung Lee, Hsing-Ju Wu, Aysegul A. Sahin, James P. Allison, Dihua Yu, Gabriel N. Hortobagyi, and Mien-Chie Hung. 2016. 'Glycosylation and stabilization of programmed death ligand-1 suppresses T-cell activity', *Nature communications*, 7: 12632.

Li, Y., S. C. Tao, G. S. Bova, A. Y. Liu, D. W. Chan, H. Zhu, and H. Zhang. 2011. 'Detection and verification of glycosylation patterns of glycoproteins from clinical specimens using lectin microarrays and lectin-based immunosorbent assays', *Anal Chem*, 83: 8509-16.

Liu, F. T., and G. A. Rabinovich. 2005. 'Galectins as modulators of tumour progression', *Nat Rev Cancer*, 5: 29-41.

Liu, S. D., C. C. Whiting, T. Tomassian, M. Pang, S. J. Bissel, L. G. Baum, V. V. Mossine, F. Poirier, M. E. Huflejt, and M. C. Miceli. 2008. 'Endogenous galectin-1 enforces class I-restricted TCR functional fate decisions in thymocytes', *Blood*, 112: 120-30.

Love, Paul E, and Avinash Bhandoola. 2011. 'Signal integration and crosstalk during thymocyte migration and emigration.', *Nature reviews. Immunology*, 11: 469-77.

Lucas, Alexander H, Michael a Apicella, and Christopher E Taylor. 2005. 'Carbohydrate moieties as vaccine candidates.', *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 41: 705-12.

Lundborg, M., and G. Widmalm. 2011. 'Structural analysis of glycans by NMR chemical shift prediction', *Anal Chem*, 83: 1514-7.

Madri, J. A., and D. Graesser. 2000. 'Cell migration in the immune system: the evolving inter-related roles of adhesion molecules and proteinases', *Dev Immunol*, 7: 103-16.

Marcos, Nuno T, Sandra Pinho, Catarina Grandela, Andrea Cruz, Bénédicte Samyn-Petit, Anne Harduin-Lepers, Raquel Almeida, Filipe Silva, Vanessa Morais, Julia Costa, Jan Kihlberg, Henrik Clausen, and Celso A Reis. 2004. 'Role of the human ST6GalNAc-I and ST6GalNAc-II in the synthesis of the cancer-associated sialyl-Tn antigen.', *Cancer research*, 64: 7050-7.

Marino, J. H., C. Tan, B. Davis, E. S. Han, M. Hickey, R. Naukam, A. Taylor, and K. S. and Miller. 2008. 'Disruption of thymopoiesis in ST6Gal I-deficient mice', *Glycobiology*, 18: 719--26.

Marrack, P., and J. Kappler. 1997. 'Positive selection of thymocytes bearing alpha beta T cell receptors', *Curr Opin Immunol*, 9: 250-5.

Martin, Christopher E, Kwesi Frimpong-Boateng, Darina S Spasova, John C Stone, and Charles D Surh. 2013. 'Homeostatic proliferation of mature T cells.', *Methods in molecular biology (Clifton, N.J.)*, 979: 81-106.

Matheu MP, Cahalan MD, Parker I. 2011. 'General approach to adoptive transfer and cell labeling for immunoimaging', *Cold Spring Harbor Protocols*, 2011.

Matsumoto, Yasuyuki, Matthew R Kudelka, M. S. Hanes, S. Lehoux, J. Heimburg-Molinaro, T. Ju, and R D Cummings. 2016. 'Development of Defined Human Chimeric anti-Tn Monoclonal Antibody', *Glycobiology*, 26: 1437.

Matsuura, Aiko, Makiko Ito, Yuta Sakaidani, Tatsuhiko Kondo, Kosuke Murakami, Koichi Furukawa, Daita Nadano, Tsukasa Matsuda, and Tetsuya Okajima. 2008. '\textless i \textgreater O \textless /i \textgreater -Linked \textless i \textgreater N \textless /i \textgreater -Acetylglucosamine Is Present on the Extracellular Domain of Notch Receptors', *Journal of Biological Chemistry*, 283: 35486--95.

Mazurov, D., A. Ilinskaya, G. Heidecker, and A. Filatov. 2012. 'Role of O-glycosylation and expression of CD43 and CD45 on the surfaces of effector T cells in human T cell leukemia virus type 1 cell-to-cell infection', *J Virol*, 86: 2447-58.

Meany, D. L., and D. W. Chan. 2011. 'Aberrant glycosylation associated with enzymes as cancer biomarkers', *Clin Proteomics*, 8: 7.

Mehal, W. Z., F. Azzaroli, and I. N. Crispe. 2001. 'Immunology of the healthy liver: old questions and new insights', *Gastroenterology*, 120: 250-60.

Merkenschlager, M., D. Graf, M. Lovatt, U. Bommhardt, R. Zamoyska, and A. G. Fisher. 1997. 'How many thymocytes audition for selection?', *J Exp Med*, 186: 1149-58.

Merry, A. H., and C. L. Merry. 2005. 'Glycoscience finally comes of age', *EMBO Rep*, 6: 900-3.

Mody, P. D., J. L. Cannon, H. S. Bandukwala, K. M. Blaine, A. B. Schilling, K. Swier, and A. I. Sperling. 2007. 'Signaling through CD43 regulates CD4 T-cell trafficking', *Blood*, 110: 2974-82.

Moody, Anne Marie, Daniel Chui, Pedro A Reche, John J Priatel, Jamey D Marth, and Ellis L Reinherz. 2001. 'Developmentally Regulated Glycosylation of the CD8αβ Coreceptor Stalk Modulates Ligand Binding', *Cell*, 107: 501-12.

Moody, Anne Marie, Simon J North, Bruce Reinhold, Steven J Van Dyken, Mark E Rogers, Maria Panico, Anne Dell, Howard R Morris, Jamey D Marth, and Ellis L Reinherz. 2003. 'Sialic acid capping of CD8beta core 1-O-glycans controls thymocyte-major histocompatibility complex class I interaction.', *The Journal of biological chemistry*, 278: 7240-6.

Moore, Gregory T C, Steven J Brown, Adam C Winterhalter, Mark Lust, Evelyn J Salvaris, Carly Selan, Harshal H Nandurkar, Paul V Desmond, Peter J Cowan, and Anthony J F d'Apice. 2008. 'Glycosylation changes in hFUT1 transgenic mice increase TCR signaling and apoptosis resulting in thymocyte maturation arrest.', *Molecular immunology*, 45: 2401-10.

Moremen, K. W., M. Tiemeyer, and A. V. Nairn. 2012. 'Vertebrate protein glycosylation: diversity, synthesis and function', *Nat Rev Mol Cell Biol*, 13: 448-62.

Morgan, R., G. Gao, J. Pawling, J. W. Dennis, M. Demetriou, and B. Li. 2004. 'N-acetylglucosaminyltransferase V (Mgat5)-mediated N-glycosylation negatively regulates Th1 cytokine production by T cells', *J Immunol*, 173: 7200-8.

Mulloy, B., A. Dell, P. Stanley, and H. Prestegard J. 2015. 'Structural Analysis of Glycans.' in rd, A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, A. G. Darvill, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar and P. H. Seeberger (eds.), *Essentials of Glycobiology* (Cold Spring Harbor (NY)).

Munkley, Jennifer, and David J. Elliott. 2016. 'Hallmarks of glycosylation in cancer', *Oncotarget*.

Murphy, K. M., and S. L. Reiner. 2002. 'The lineage decisions of helper T cells', *Nat Rev Immunol*, 2: 933-44.

Nagel, A. K., and L. E. Ball. 2014. 'O-GlcNAc transferase and O-GlcNAcase: achieving target substrate specificity', *Amino Acids*, 46: 2305-16.

Nakazawa, S., R. Imamura, M. Kawamura, T. Kato, T. Abe, H. Iwatani, K. Yamanaka, M. Uemura, H. Kishikawa, K. Nishimura, M. Tajiri, Y. Wada, and N. Nonomura. 2019. 'Evaluation of IgA1 O-glycosylation in Henoch-Schonlein Purpura Nephritis Using Mass Spectrometry', *Transplant Proc*, 51: 1481-87.

Neff, C. P., M. E. Rhodes, K. L. Arnolds, C. B. Collins, J. Donnelly, N. Nusbacher, P. Jedlicka, J. M. Schneider, M. D. McCarter, M. Shaffer, S. K. Mazmanian, B. E. Palmer, and C. A. Lozupone. 2016. 'Diverse Intestinal Bacteria Contain Putative Zwitterionic Capsular Polysaccharides with Anti-inflammatory Properties', *Cell Host Microbe*, 20: 535-47.

Nolz, Jeffrey C, and John T Harty. 2014. 'IL-15 regulates memory CD8+ T cell O-glycan synthesis and affects trafficking.', *The Journal of clinical investigation*, 124: 1013-26.

Nolz, Jeffrey C, Gabriel R Starbeck-Miller, and John T Harty. 2011. 'Naive, effector and memory CD8 T-cell trafficking: parallels and distinctions.', *Immunotherapy*, 3: 1223-33.

Novak, J., Z. Moldoveanu, B. A. Julian, M. Raska, R. J. Wyatt, Y. Suzuki, Y. Tomino, A. G. Gharavi, J. Mestecky, and H. Suzuki. 2011. 'Aberrant glycosylation of IgA1 and anti-glycan antibodies in IgA nephropathy: role of mucosal immune system', *Adv Otorhinolaryngol*, 72: 60-3.

Numa, F., N. Tsunaga, T. Michioka, S. Nawata, H. Ogata, and H. Kato. 1995. 'Tissue expression of Sialyl Tn antigen in gynecologic tumors', *J Obstet Gynaecol (Tokyo 1995)*, 21: 385-9.

Ohtsubo, Kazuaki, and Jamey D. Marth. 2006. 'Glycosylation in Cellular Mechanisms of Health and Disease', *Cell*, 126: 855--67.

Okada, Masahiro, Shunsuke Chikuma, Taisuke Kondo, Sana Hibino, Hiroaki Machiyama, Tadashi Yokosuka, Miyako Nakano, and Akihiko Yoshimura. 2017. 'Blockage of Core Fucosylation Reduces Cell-Surface Expression of PD-1 and Promotes Anti-tumor Immune Responses of T Cells', *Cell Reports*, 20: 1017--28.

Oostendorp, R. A., S. Ghaffari, and C. J. Eaves. 2000. 'Kinetics of in vivo homing and recruitment into cycle of hematopoietic cells are organ-specific but CD44-independent', *Bone Marrow Transplant*, 26: 559-66.

Ozaki, H., H. Matsuzaki, H. Ando, H. Kaji, H. Nakanishi, Y. Ikehara, and H. Narimatsu. 2012. 'Enhancement of metastatic ability by ectopic expression of ST6GalNAcI on a gastric cancer cell line in a mouse model', *Clin Exp Metastasis*, 29: 229-38.

Pabst, M., and F. Altmann. 2011. 'Glycan analysis by modern instrumental methods', *Proteomics*, 11: 631-43.

Pace, K. E., H. P. Hahn, and L. G. Baum. 2003. 'Preparation of recombinant human galectin-1 and use in T-cell death assays', *Methods Enzymol*, 363: 499-518.

Paessens, Lutz C, Juan J García-Vallejo, Rosette J Fernandes, and Yvette van Kooyk. 2007. 'The glycosylation of thymic microenvironments. A microscopic study using plant lectins.', *Immunology letters*, 110: 65-73.

Paquette, S., B. L. Moller, and S. Bak. 2003. 'On the origin of family 1 plant glycosyltransferases', *Phytochemistry*, 62: 399-413.

Patani, N., W. Jiang, and K. Mokbel. 2008. 'Prognostic utility of glycosyltransferase expression in breast cancer', *Cancer Genomics Proteomics*, 5: 333-40.

Pearl, L. H., and C. Prodromou. 2006. 'Structure and mechanism of the Hsp90 molecular chaperone machinery', *Annu Rev Biochem*, 75: 271-94.

Pellicci, Daniel G, Andrew J Clarke, Onisha Patel, Thierry Mallevaey, Travis Beddoe, Jérôme Le Nours, Adam P Uldrich, James McCluskey, Gurdyal S Besra, Steven a Porcelli, Laurent Gapin, Dale I Godfrey, and Jamie Rossjohn. 2011. 'Recognition of β-linked self glycolipids mediated by natural killer T cell antigen receptors.', *Nature immunology*, 12: 827-33.

Peracaula, R., S. Barrabes, A. Sarrats, P. M. Rudd, and R. de Llorens. 2008. 'Altered glycosylation in tumours focused to cancer diagnosis', *Dis Markers*, 25: 207-18.

Pereira, M. S., I. Alves, M. Vicente, A. Campar, M. C. Silva, N. A. Padrao, V. Pinto, A. Fernandes, A. M. Dias, and S. S. Pinho. 2018. 'Glycans as Key Checkpoints of T Cell Activity and Function', *Front Immunol*, 9: 2754.

Perillo, N. L., K. E. Pace, J. J. Seilhamer, and L. G. Baum. 1995. 'Apoptosis of T cells mediated by galectin-1', *Nature*, 378: 736-9.

Perillo, N. L., C. H. Uittenbogaart, J. T. Nguyen, and L. G. Baum. 1997. 'Galectin-1, an endogenous lectin produced by thymic epithelial cells, induces apoptosis of human thymocytes', *J Exp Med*, 185: 1851-8.

Piller, V., F. Piller, and M. Fukuda. 1990. 'Biosynthesis of truncated O-glycans in the T cell line Jurkat. Localization of O-glycan initiation', *J Biol Chem*, 265: 9264-71.

Pinho, Salom S., and Celso A. Reis. 2015. 'Glycosylation in cancer: mechanisms and clinical implications', *Nature Reviews Cancer*, 15: 540--55.

Pisano, A., N. H. Packer, J. W. Redmond, K. L. Williams, and A. A. Gooley. 1994. 'Characterization of O-linked glycosylation motifs in the glycopeptide domain of bovine kappa-casein', *Glycobiology*, 4: 837-44.

Priatel, J J, D Chui, N Hiraoka, C J Simmons, K B Richardson, D M Page, M Fukuda, N M Varki, and J D Marth. 2000. 'The ST3Gal-I sialyltransferase controls CD8+ T lymphocyte homeostasis by modulating O-glycan biosynthesis.', *Immunity*, 12: 273-83.

Raab-Westphal, S., J. F. Marshall, and S. L. Goodman. 2017. 'Integrins as Therapeutic Targets: Successes and Cancers', *Cancers (Basel)*, 9.

Rampal, Raajit, Annie S. Y. Li, Daniel J. Moloney, Stephanie A. Georgiou, Kelvin B. Luther, Aleksandra Nita-Lazar, and Robert S. Haltiwanger. 2005. 'Lunatic Fringe, Manic Fringe, and Radical Fringe Recognize Similar Specificity Determinants in \textless i \textgreater O- \textless /i \textgreater Fucosylated Epidermal Growth Factor-like Repeats', *Journal of Biological Chemistry*, 280: 42454--63.

Rapoport, E. M., O. V. Kurmyshkina, and N. V. Bovin. 2008. 'Mammalian galectins: structure, carbohydrate specificity, and functions', *Biochemistry (Mosc)*, 73: 393-405.

Reisner, Yair, Mariana Linker-Israeli, and Nathan Sharon. 1976. 'Separation of mouse thymocytes into two subpopulations by the use of peanut agglutinin', *Cellular Immunology*, 25: 129-34.

Rizqiawan, A., K. Tobiume, G. Okui, K. Yamamoto, H. Shigeishi, S. Ono, H. Shimasue, M. Takechi, K. Higashikawa, and N. Kamata. 2013. 'Autocrine galectin-1 promotes collective cell migration of squamous cell carcinoma cells through up-regulation of distinct integrins', *Biochem Biophys Res Commun*, 441: 904-10.

Robinson, M. W., C. Harmon, and C. O'Farrelly. 2016. 'Liver immunology and its role in inflammation and homeostasis', *Cell Mol Immunol*, 13: 267-76.

Roncarolo, M. G., R. Bacchetta, C. Bordignon, S. Narula, and M. K. Levings. 2001. 'Type 1 T regulatory cells', *Immunol Rev*, 182: 68-79.

Roncarolo, M. G., M. K. Levings, and C. Traversari. 2001. 'Differentiation of T regulatory cells by immature dendritic cells', *J Exp Med*, 193: F5-9.

Rosen, S. D. 2004. 'Ligands for L-selectin: homing, inflammation, and beyond', *Annu Rev Immunol*, 22: 129-56.

Rossi, Fabio M V, Stephane Y Corbel, Jasmeen S Merzaban, Douglas a Carlow, Klaus Gossens, Jeffrey Duenas, Leslie So, Lin Yi, and Hermann J Ziltener. 2005. 'Recruitment of adult thymic progenitors is regulated by P-selectin and its ligand PSGL-1.', *Nature immunology*, 6: 626-34.

Rossjohn, J., S. Gras, J. J. Miles, S. J. Turner, D. I. Godfrey, and J. McCluskey. 2015. 'T cell antigen receptor recognition of antigen-presenting molecules', *Annu Rev Immunol*, 33: 169-200.

Roth, Ziv, Galit Yehezkel, and Isam Khalaila. 2012. 'Identification and quantification of protein glycosylation', *International Journal of Carbohydrate Chemistry*, 2012.

Rottger, S., J. White, H. H. Wandall, J. C. Olivo, A. Stark, E. P. Bennett, C. Whitehouse, E. G. Berger, H. Clausen, and T. Nilsson. 1998. 'Localization of three human polypeptide GalNAc-transferases in HeLa cells suggests initiation of O-linked glycosylation throughout the Golgi apparatus', *J Cell Sci*, 111 (Pt 1): 45-60.

Roxby, D. J., M. B. Pfeiffer, A. A. Morley, and M. A. Kirkland. 1992. 'Expression of the Tn antigen in myelodysplasia, lymphoma, and leukemia', *Transfusion*, 32: 834-8.

Rudd, Pauline M., Mark R. Wormald, Robyn L. Stanfield, Mingdong Huang, Niklas Mattsson, Jeffrey A. Speir, Jeannine A. DiGennaro, Jacquelyn S. Fetrow, Raymond A. Dwek, and Ian A. Wilson. 1999. 'Roles for glycosylation of cell surface receptors involved in cellular immune recognition', *Journal of Molecular Biology*, 293: 351--66.

Russell, J. H., and T. J. Ley. 2002. 'Lymphocyte-mediated cytotoxicity', *Annu Rev Immunol*, 20: 323-70.

Sakaguchi, S., and N. Sakaguchi. 2005. 'Regulatory T cells in immunologic self-tolerance and autoimmune disease', *Int Rev Immunol*, 24: 211-26.

Sallusto, F., J. Geginat, and A. Lanzavecchia. 2004. 'Central memory and effector memory T cell subsets: function, generation, and maintenance', *Annu Rev Immunol*, 22: 745-63.

Samyn-Petit, B., M. A. Krzewinski-Recchi, W. F. Steelant, P. Delannoy, and A. Harduin-Lepers. 2000. 'Molecular cloning and functional expression of human ST6GalNAc II. Molecular expression in various human cultured cells', *Biochim Biophys Acta*, 1474: 201-11.

Sanchez-Ruderisch, H., K. M. Detjen, M. Welzel, S. Andre, C. Fischer, H. J. Gabius, and S. Rosewicz. 2011. 'Galectin-1 sensitizes carcinoma cells to anoikis via the fibronectin receptor alpha5beta1-integrin', *Cell Death Differ*, 18: 806-16.

Santana, M. A., and Y. Rosenstein. 2003. 'What it takes to become an effector T cell: the process, the cells involved, and the mechanisms', *J Cell Physiol*, 195: 392-401.

Saulsbury, F. T. 1997. 'Alterations in the O-linked glycosylation of IgA1 in children with Henoch-Schonlein purpura', *J Rheumatol*, 24: 2246-9.

———. 2007. 'Clinical update: Henoch-Schonlein purpura', *Lancet*, 369: 976-8.

Schachter, H., and I. Brockhausen. 1989. 'The biosynthesis of branched O-glycans', *Symp Soc Exp Biol*, 43: 1-26.

Schlub, T. E., V. P. Badovinac, J. T. Sabel, J. T. Harty, and M. P. Davenport. 2010. 'Predicting CD62L expression during the CD8+ T-cell response in vivo', *Immunol Cell Biol*, 88: 157-64.

Seddon, B., P. Tomlinson, and R. Zamoyska. 2003. 'Interleukin 7 and T cell receptor signals regulate homeostasis of CD4 memory cells', *Nat Immunol*, 4: 680-6.

Seguchi, T, R K Merkle, M Ono, M Kuwano, and R D Cummings. 1991. 'The dysfunctional LDL receptor in a monensin-resistant mutant of Chinese hamster ovary cells lacks selected O-linked oligosaccharides.', *Archives of biochemistry and biophysics*, 284: 245-56.

Sewell, Robert, Malin Bäckström, Martin Dalziel, Steven Gschmeissner, Hasse Karlsson, Thomas Noll, Jochem Gätgens, Henrik Clausen, Gunnar C Hansson, Joy Burchell, and Joyce Taylor-Papadimitriou. 2006. 'The ST6GalNAc-I sialyltransferase localizes throughout the Golgi and is responsible for the synthesis of the tumor-associated sialyl-Tn O-glycan in human breast cancer.', *The Journal of biological chemistry*, 281: 3586-94.

Shah, D. K., and J. C. Zuniga-Pflucker. 2014. 'An Overview of the Intrathymic Intricacies of T Cell Development', *The Journal of Immunology*, 192: 4017--23.

Sharon, N. 1983. 'Lectin receptors as lymphocyte surface markers', *Adv Immunol*, 34: 213-98.

Shi, Jianjun, and Howard T Petrie. 2012. 'Activation kinetics and off-target effects of thymus-initiated cre transgenes.', *PloS one*, 7: e46590.

Shih, Han-Yu, Bingtao Hao, and Michael S. Krangel. 2011. 'Orchestrating T-cell receptor α gene assembly through changes in chromatin structure and organization', *Immunologic Research*, 49: 192--201.

Shortman, K., M. Egerton, G. J. Spangrude, and R. Scollay. 1990. 'The generation and fate of thymocytes', *Semin Immunol*, 2: 3-12.

Smith, Logan K., Giselle M. Boukhaled, Stephanie A. Condotta, Sabrina Mazouz, Jenna J. Guthmiller, Rahul Vijay, Noah S. Butler, Julie Bruneau, Naglaa H. Shoukry, Connie M. Krawczyk, and Martin J. Richer. 2018. 'Interleukin-10 Directly Inhibits CD8+ T Cell Function by Enhancing N-Glycan Branching to Decrease Antigen Sensitivity', *Immunity*, 48: 299--312.e5.

Smyth, M. J., J. M. Kelly, V. R. Sutton, J. E. Davis, K. A. Browne, T. J. Sayers, and J. A. Trapani. 2001. 'Unlocking the secrets of cytotoxic granule proteins', *J Leukoc Biol*, 70: 18-29.

Song, Yinghui, Vivek Kumar, Hua-Xing Wei, Ju Qiu, and Pamela Stanley. 2016. 'Lunatic, Manic, and Radical Fringe Each Promote T and B Cell Development', *The Journal of Immunology*, 196: 232--43.

Speir, J a, U M Abdel-Motal, M Jondal, and I a Wilson. 1999. 'Crystal structure of an MHC class I presented glycopeptide that generates carbohydrate-specific CTL.', *Immunity*, 10: 51-61.

Sperling, A. I., J. R. Sedy, N. Manjunath, A. Kupfer, B. Ardman, and J. K. Burkhardt. 1998. 'TCR signaling induces selective exclusion of CD43 from the T cell-antigen-presenting cell contact site', *J Immunol*, 161: 6459-62.

Spiro, R. G. 2002. 'Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds', *Glycobiology*, 12: 43R-56R.

Sprent, J. 1995. 'Central tolerance of T cells', *Int Rev Immunol*, 13: 95-105.

Springer, G F. 1984. 'T and Tn, general carcinoma autoantigens.', *Science (New York, N.Y.)*, 224: 1198-206.

Srour, E. F., A. Jetmore, F. M. Wolber, P. A. Plett, R. Abonour, M. C. Yoder, and C. M. Orschell-Traycoff. 2001. 'Homing, cell cycle kinetics and fate of transplanted hematopoietic stem cells', *Leukemia*, 15: 1681-4.

Starr, Timothy K, Mark A Daniels, Michelle M Lucido, Stephen C Jameson, and Kristin A Hogquist. 2003. 'Thymocyte sensitivity and supramolecular activation cluster formation are developmentally regulated: a partial role for sialylation.', *Journal of immunology (Baltimore, Md. : 1950)*, 171: 4512-20.

Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT, Lavrsen K, Dabelsteen S, Pedersen NB, Marcos-Silva L, Gupta R, Bennett EP, Mandel U, Brunak S, Wandall HH, Levery SB, Clausen H. 2013. 'Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology', *EMBO J*, 32: 1478-88.

Steentoft, Catharina, Sergey Vakhrushev, Malene B Vester-christensen, and Katrine T-b G Schjoldager. 2011. 'A novel method to mine the elusive O-glycoproteome (GalNAc-type) using Zinc finger nuclease glycoengineered SimpleCells', *Nature methods*.

Steentoft, Catharina, Sergey Y Vakhrushev, Malene B Vester-Christensen, Katrine T-B G Schjoldager, Yun Kong, Eric Paul Bennett, Ulla Mandel, Hans Wandall, Steven B Levery, and Henrik Clausen. 2011. 'Mining the O-glycoproteome using zinc-finger nuclease-glycoengineered SimpleCell lines.', *Nature methods*, 8: 977-82.

Stowell, S. R., S. Karmakar, C. M. Arthur, T. Ju, L. C. Rodrigues, T. B. Riul, M. Dias-Baruffi, J. Miner, R. P. McEver, and R. D. Cummings. 2009. 'Galectin-1 induces reversible phosphatidylserine exposure at the plasma membrane', *Mol Biol Cell*, 20: 1408-18.

Stowell, S. R., Y. Qian, S. Karmakar, N. S. Koyama, M. Dias-Baruffi, H. Leffler, R. P. McEver, and R. D. Cummings. 2008. 'Differential roles of galectin-1 and galectin-3 in regulating leukocyte viability and cytokine secretion', *J Immunol*, 180: 3091-102.

Subedi, G. P., R. W. Johnson, H. A. Moniz, K. W. Moremen, and A. Barb. 2015. 'High Yield Expression of Recombinant Human Proteins with the Transient Transfection of HEK293 Cells in Suspension', *J Vis Exp*: e53568.

Sultana, D. A., S. L. Zhang, S. P. Todd, and A. Bhandoola. 2012. 'Expression of Functional P-Selectin Glycoprotein Ligand 1 on Hematopoietic Progenitors Is Developmentally Regulated', *The Journal of Immunology*, 188: 4385--93.

Sun, X., T. Ju, and R. D. Cummings. 2018. 'Differential expression of Cosmc, T-synthase and mucins in Tn-positive colorectal cancers', *BMC Cancer*, 18: 827.

Surh, C D, and J Sprent. 1994. 'T-cell apoptosis detected in situ during positive and negative selection in the thymus.', *Nature*, 372: 100-3.

Suzuki, M., K. Angata, J. Nakayama, and M. Fukuda. 2003. 'Polysialic acid and mucin type o-glycans on the neural cell adhesion molecule differentially regulate myoblast fusion', *J Biol Chem*, 278: 49459-68.

Swamy, M. J., D. Gupta, S. K. Mahanta, and A. Surolia. 1991. 'Further characterization of the saccharide specificity of peanut (Arachis hypogaea) agglutinin', *Carbohydr Res*, 213: 59-67.

Szabó, Tamás G, Robin Palotai, Péter Antal, Itay Tokatly, László Tóthfalusi, Ole Lund, György Nagy, András Falus, and Edit I Buzás. 2009. 'Critical role of glycosylation in determining the length and structure of T cell epitopes.', *Immunome research*, 5: 4.

Taylor, M. E., K. Drickamer, R. L. Schnaar, M. E. Etzler, and A. Varki. 2015. 'Discovery and Classification of Glycan-Binding Proteins.' in rd, A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, A. G. Darvill, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar and P. H. Seeberger (eds.), *Essentials of Glycobiology* (Cold Spring Harbor (NY)).

Tenno, Mari, Kazuaki Ohtsubo, Fred K Hagen, David Ditto, Alexander Zarbock, Patrick Schaerli, Ulrich H von Andrian, Klaus Ley, Dzung Le, Lawrence a Tabak, and Jamey D Marth. 2007. 'Initiation of protein O glycosylation by the polypeptide GalNAcT-1 in vascular biology and humoral immunity.', *Molecular and cellular biology*, 27: 8783-96.

Thery, C., and S. Amigorena. 2001. 'The cell biology of antigen presentation in dendritic cells', *Curr Opin Immunol*, 13: 45-51.

Thijssen, V. L., and A. W. Griffioen. 2014. 'Galectin-1 and -9 in angiogenesis: a sweet couple', *Glycobiology*, 24: 915-20.

Thurman, E. C., J. Walker, S. Jayaraman, N. Manjunath, B. Ardman, and J. M. Green. 1998. 'Regulation of in vitro and in vivo T cell activation by CD43', *Int Immunol*, 10: 691-701.

Trnková, K, S Pastoreková, and J Petrik. 2012. 'Novel approaches to antiviral and anticancer immunotherapy.', *Acta virologica*, 56: 271-82.

Trombetta, E. S. 2003. 'The contribution of N-glycans and their processing in the endoplasmic reticulum to glycoprotein biosynthesis', *Glycobiology*, 13: 77R-91R.

Troncoso, M. F., F. Ferragut, M. L. Bacigalupo, V. M. Cardenas Delgado, L. G. Nugnes, L. Gentilini, D. Laderach, C. Wolfenstein-Todel, D. Compagno, G. A. Rabinovich, and M. T. Elola. 2014. 'Galectin-8: a matricellular lectin with key roles in angiogenesis', *Glycobiology*, 24: 907-14.

Tsuboi, S, and M Fukuda. 1997. 'Branched O-linked oligosaccharides ectopically expressed in transgenic mice reduce primary T-cell immune responses.', *The EMBO journal*, 16: 6364-73.

Tsuji, Shuichi, and Shou Takashima. 2014. 'ST6 N-Acetylgalactosaminide Alpha-2, 6-Sialyltransferase 1 (ST6GALNAC1)', *Handbook of Glycosyltransferases and Related Genes*: 715-25.

Ueno, T., F. Saito, D. H. Gray, S. Kuse, K. Hieshima, H. Nakano, T. Kakiuchi, M. Lipp, R. L. Boyd, and Y. Takahama. 2004. 'CCR7 signals are essential for cortex-medulla migration of developing thymocytes', *J Exp Med*, 200: 493-505.

Vainchenker, William, Giovanna Vinci, Ugo Testa, Annie Henri, Antonio Tabilio, Marie-pierre Fache, Henri Rochant, and Jean-pierre Cartron. 1985. 'Presence of the Tn Antigen on Hematopoietic Progenitors from Patients with the Tn Syndrome', *Journal of Clinical Investigation*, 75: 541-46.

Van den Steen, P., P. M. Rudd, R. A. Dwek, and G. Opdenakker. 1998. 'Concepts and principles of O-linked glycosylation', *Crit Rev Biochem Mol Biol*, 33: 151-208.

van der Leij, J., A. van den Berg, T. Blokzijl, G. Harms, H. van Goor, P. Zwiers, R. van Weeghel, S. Poppema, and L. Visser. 2004. 'Dimeric galectin-1 induces IL-10 production in T-lymphocytes: an important tool in the regulation of the immune response', *J Pathol*, 204: 511-8.

van Vliet, Sandra J., Sonja I. Gringhuis, Teunis B. H. Geijtenbeek, and Yvette van Kooyk. 2006. 'Regulation of effector T cells by antigen-presenting cells via interaction of the C-type lectin MGL with CD45', *Nature immunology*, 7: 1200--08.

Varki, A. 1993. 'Biological roles of oligosaccharides: all of the theories are correct', *Glycobiology*, 3: 97-130.

———. 2017. 'Biological roles of glycans', *Glycobiology*, 27: 3-49.

Veillon, L., Y. Huang, W. Peng, X. Dong, B. G. Cho, and Y. Mechref. 2017. 'Characterization of isomeric glycan structures by LC-MS/MS', *Electrophoresis*, 38: 2100-14.

Veillon, L., S. Zhou, and Y. Mechref. 2017. 'Quantitative Glycomics: A Combined Analytical and Bioinformatics Approach', *Methods Enzymol*, 585: 431-77.

Villadangos, J. A. 2001. 'Presentation of antigens by MHC class II molecules: getting the most out of them', *Mol Immunol*, 38: 329-46.

Visan, Ioana, Julie S. Yuan, Joanne B. Tan, Kira Cretegny, and Cynthia J. Guidos. 2006. 'Regulation of intrathymic T-cell development by Lunatic Fringe? Notch1 interactions', *Immunological reviews*, 209: 76--94.

Vlad, a. M., S. Muller, M. Cudic, H. Paulsen, L. Otvos, F.-G. Hanisch, and O. J. Finn. 2002. 'Complex Carbohydrates Are Not Removed During Processing of Glycoproteins by Dendritic Cells: Processing of Tumor Antigen MUC1 Glycopeptides for Presentation to Major Histocompatibility Complex Class II-restricted T Cells', *Journal of Experimental Medicine*, 196: 1435-46.

von Andrian, U. H., and T. R. Mempel. 2003. 'Homing and cellular traffic in lymph nodes', *Nat Rev Immunol*, 3: 867-78.

Walling, B. L., and M. Kim. 2018. 'LFA-1 in T Cell Migration and Differentiation', *Front Immunol*, 9: 952.

Wandall, H. H., F. Irazoqui, M. A. Tarp, E. P. Bennett, U. Mandel, H. Takeuchi, K. Kato, T. Irimura, G. Suryanarayanan, M. A. Hollingsworth, and H. Clausen. 2007. 'The lectin domains of polypeptide

GalNAc-transferases exhibit carbohydrate-binding specificity for GalNAc: lectin binding to GalNAc-glycopeptide substrates is required for high density GalNAc-O-glycosylation', *Glycobiology*, 17: 374-87.

Wang, Yingchun, Shawn M Jobe, Xiaokun Ding, Hyojung Choo, David R Archer, Rongjuan Mi, Tongzhong Ju, and Richard D Cummings. 2012. 'Platelet biogenesis and functions require correct protein O-glycosylation.', *Proceedings of the National Academy of Sciences of the United States of America*, 109: 16143-8.

Wang, Yingchun, Tongzhong Ju, Xiaokun Ding, Baoyun Xia, Wenyi Wang, Lijun Xia, Miao He, and Richard D Cummings. 2010. 'Cosmc is an essential chaperone for correct protein O-glycosylation', *Proceedings of the National Academy of Sciences*, 107: 9228-33.

Webb, S. R., and J. Sprent. 1990. 'Tolerogenicity of thymic epithelium', *Eur J Immunol*, 20: 2525-8.

Weerapana, E., and B. Imperiali. 2006. 'Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems', *Glycobiology*, 16: 91R-101R.

Weninger, W., N. Manjunath, and U. H. von Andrian. 2002. 'Migration and differentiation of CD8+ T cells', *Immunol Rev*, 186: 221-33.

Williams, A., C. A. Peh, and T. Elliott. 2002. 'The cell biology of MHC class I antigen presentation', *Tissue Antigens*, 59: 3-17.

Woodruff, J. J., and B. M. Gesner. 1969. 'The effect of neuraminidase on the fate of transfused lymphocytes', *J Exp Med*, 129: 551-67.

Wu, Wenyu, Patricia H Harley, Jennifer A Punt, Susan O Sharrow, and Kelly P Kearse. 1996. 'Identification of CD8 as a Peanug Agglutinin (PNA) Receptor Molecule on Immature Thymocytes', *Journal of Experimental Medicine*, 184.

Xia, L., T. Ju, A. Westmuckett, G. An, L. Ivanciu, J. M. McDaniel, F. Lupu, R. D. Cummings, and R. P. McEver. 2004. 'Defective angiogenesis and fatal embryonic hemorrhage in mice lacking core 1-derived O-glycans', *J Cell Biol*, 164: 451-9.

Yamada, K., S. Hyodo, M. Kinoshita, T. Hayakawa, and K. Kakehi. 2010. 'Hyphenated technique for releasing and MALDI MS analysis of O-glycans in mucin-type glycoprotein samples', *Anal Chem*, 82: 7436-43.

Yu, X., Z. Du, X. Sun, C. Shi, H. Zhang, and T. Hu. 2015. 'Aberrant Cosmc genes result in Tn antigen expression in human colorectal carcinoma cell line HT-29', *Int J Clin Exp Pathol*, 8: 2590-602.

Zerrahn, J, W Held, and D H Raulet. 1997. 'The MHC reactivity of the T cell repertoire prior to positive and negative selection.', *Cell*, 88: 627-36.

Zhao, Q., M. Barclay, J. Hilkens, X. Guo, H. Barrow, J. M. Rhodes, and L. G. Yu. 2010. 'Interaction between circulating galectin-3 and cancer-associated MUC1 enhances tumour cell homotypic aggregation and prevents anoikis', *Mol Cancer*, 9: 154.

Zinkernagel, R. M., G. N. Callahan, J. Klein, and G. Dennert. 1978. 'Cytotoxic T cells learn specificity for self H-2 during differentiation in the thymus', *Nature*, 271: 251-3.

Zlotoff, Daniel A, and Avinash Bhandoola. 2011. 'Hematopoietic progenitor migration to the adult thymus.', *Annals of the New York Academy of Sciences*, 1217: 122-38.

Zuniga-Pflucker, J. C., and M. J. Lenardo. 1996. 'Regulation of thymocyte development from immature progenitors', *Curr Opin Immunol*, 8: 215-24.