Elucidating the role of the family of GalNAc-Transferases in aberrant protein O-glycosylation in the progression of epithelial ovarian cancer

Thèse

Razan Sheta

Doctorat en Médecine Expérimentale
Philosophiae Doctor (Ph.D.)

Québec, Canada

Elucidating the role of the family of GalNAc-Transferases in aberrant protein O-glycosylation in the progression of epithelial ovarian cancer

Thèse

Razan Sheta

Sous la direction de:

Dimcho Bachvarov, Directeur de Recherche

RÉSUMÉ

Le cancer épithélial de l'ovaire (CEO) est la forme de cancer gynécologique la plus létale. Ainsi, la compréhension des changements moléculaires associés à ce cancer métastatique ovarien peut mener à l'identification de nouvelles cibles thérapeutiques essentielles. La glycosylation, une modification post-traductionnelle, joue un rôle important dans de nombreuses fonctions cellulaires. Cette glycosylation participe à des événements physiopathologiques majeurs durant la progression tumorale. De plus, il a été prouvé que l'expression aberrante des structures glycanes interfère avec des mécanismes cellulaires comme l'adhésion, la migration et la prolifération des cellules.

Dans ce contexte, notre laboratoire a récemment montré que le gène codant pour la protéine N-acétylgalactosaminyltransférase 3 (*GALNT3*), membre de la famille des GalNAc-Transférases (GalNAc-Ts), est hypométhylé et que la protéine *GALNT3* est plus fortement exprimée dans les tumeurs CEO dont la sévérité est de grade élevé ("high-grade (HG) serous"), en comparaison avec des tumeurs à potentiel malin faible ("low malignant potential (LMP) ") et des tissus ovariens normaux.

Ces observations indiquent un fort potentiel oncogénique pour le gène *GALNT3* dans les stades avancés du CEO. Ces premières constatations suggèrent également que la surexpression de *GALNT3* peut jouer un rôle important dans la tumorigenèse du CEO en augmentant sa dissémination via une O-glycosylation de type mucine aberrante. Ces glycosylations anormales peuvent donc être impliquées dans la carcinogenèse ovarienne et nécessitent une étude approfondie.

Dans ce projet de recherche, nous proposons d'approfondir les observations déjà obtenues *in vitro* en utilisant un modèle *in vivo* chez la souris, afin d'élucider le rôle fonctionnel de la *GALNT3* et d'autres membres de cette famille dans la progression du CEO. A partir d'une étude de glycoprotéomique indépendante de la masse, qui a permis d'identifier des glycopeptides intacts ou métaboliquement marqués, ce projet de recherche a rendu possible la définition précise du rôle de *GALNT3* dans la O-glycosylation des cibles de type mucine au sein des cellules CEO. Ainsi, via une recherche ciblée dans la base de données « Swiss-Prot » du protéome humain, nous avons trouvé plusieurs centaines de glycoprotéines et glycopeptides uniques, différemment exprimés dans les clones cellulaires dépourvus en

GALNT3 KD. Par la suite, nous avons identifié les gènes codant pour ces glycoprotéines et glycopeptides. Nous avons notamment trouvé, parmi la liste, un groupe de gènes impliqués dans le métabolisme cellulaire dont les modifications post-traductionnelles sont, de manière intéressante, principalement supprimées dans les clones *GALNT3* KD.

De plus, nous nous sommes intéressés aux autres membres de la famille des GalNAc-Ts dans le CEO et nous avons montré que de multiples membres et pas uniquement *GALNT3* peuvent jouer un rôle important dans la dissémination et la progression du CEO. De plus, une découverte très intéressante fut la redondance possible des rôles joués par certains membres de la famille des GalNAc-Ts dans le CEO. Ainsi, nous avons identifié *GALNT6* qui serait, à l'image de *GALNT3,* impliquée dans la dissémination et la progression du CEO. Cette implication du *GALNT6* est supportée par le fait que cette protéine a les mêmes fonctions que *GALNT3,* suggérant un effet compensatoire de *GALNT6* en absence de *GALNT3*. Pour tester cette hypothèse, nous avons abolie l'expression des deux protéines *GALNT3* et *GALNT6*, *in vivo*, et nous avons observé une effet significatif sur la formation des tumeurs et la survie des animaux. Pour la suite de ce projet, nous proposons d'analyser la structure glycane des différentes glycoprotéines identifiées dans les cellules cancéreuses, afin de déterminer les altérations des modifications O-glycanes suite à la perte d'expression de *GALNT3* et d'autres membres de la famille des GalNAc-Ts. En conclusion, notre étude contribue à comprendre la participation du glycoprotéome dans la tumorigenèse du CEO et à identifier d'autres cibles de type mucine ou des O-glycoprotéines dont l'expression aberrante serait modulée dans le CEO. Ainsi, pris dans son ensemble, ce projet de recherche montre la possibilité de discriminer entre des cellules cancéreuses et des cellules contrôles via les glycosylations de leurs protéines et permet d'entrevoir la glycobiologie comme une voie prometteuse pour l'identifier de nouveaux biomarqueurs pour le diagnostic du CEO.

ABSTRACT

Epithelial ovarian cancer (EOC) is the most lethal gynecologic malignancy, thus understanding the molecular changes associated with ovarian cancer metastasis could lead to the identification of essential therapeutic targets. Glycosylation is a post-translational modification (PTM) of proteins playing a major role in various cell properties. Glycosylation participates in major pathophysiology events during tumor progressions, and the aberrant expression of glycan structures was shown to interfere with cell properties such as cell adhesion, migration, and proliferation. The lab has previously identified the polypeptide N-acetylgalactosaminyltransferase 3 (*GALNT3*) gene, a member of the GalNAc-Transferases (GalNAc-Ts) gene family, as hypomethylated and overexpressed in high-grade (HG) serous EOC tumors, compared to low malignant potential (LMP) EOC tumors and normal ovarian tissues. Taken together, the data obtained were indicative of a strong oncogenic potential of the *GALNT3* gene in advanced EOC and suggest that *GALNT3* overexpression might contribute to EOC dissemination through aberrant mucin O-glycosylation, thus specifying some of the putative mechanisms of abnormal glycosylation implicated in ovarian carcinogenesis, which warrant further investigation. The current research project focused on expanding the *in vitro* observations obtained by using animal models to investigate *in vivo* the functional significance of *GALNT3* and other close members of the GalNAc-Ts gene family in serous EOC progression. Moreover, by applying a mass-independent chemical glycoproteomics platform to characterize intact, metabolically labeled glycopeptides, this project more profoundly characterized the role of *GALNT3* in aberrant O-glycosylation of mucin-like targets in EOC cells. Isotopically recorded ions were searched against the Swiss-Prot human proteome; and data obtained were indicative of hundreds of unique glycoproteins and glycopeptides that were differentially expressed upon *GALNT3* KD. Related gene groups were identified, and interestingly, genes implicated in mechanisms of cellular metabolic functions, and PTMs were found to be predominantly suppressed in *GALNT3* KD clones. In accordance, we also investigated the role of other members of the GalNAc-T family in EOC and we showed that multiple members and not only *GALNT3* can play an important role in EOC cancer dissemination and progression. One very interesting finding was the redundant role some

members of the GalNAc-T family members play in EOC. We investigated the compensatory functions of *GALNT3* and *GALNT6*, and we were able to demonstrate these two genes can impose that synthetic backup. Furthermore, we found that and their ablation can affect animal survival and tumor formation as observed both *in vivo* and *in vitro*. In continuation of this work, this project will focus on analyzing the glycan structures of those differentially expressed glycoproteins, to further examine the specific O-glycans alterations associated with the *GALNT3* and other members of the GalNAc-Ts upon gene knockout (KO). Fully elaborated glycopeptides can reveal structural details of the glycoproteome, thus our results could give important information on the glycome in EOC cells, and the identification of other O-glycoproteins/mucin-like targets whose aberrant expression may be modulated by these in EOC. Taken together, the ability to mark differences in the glycosylation of proteins between cancer cells and control cells can emphasize glycobiology as a promising field for potential biomarker identification.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 1: Introduction

Chapter 2: A metabolic labeling approach for glycoproteomic analysis reveals altered glycoprotein expression upon *GALNT3* knockdown in ovarian cancer cells

Chapter 3: Proteomic dataset for altered glycoprotein expression upon *GALNT3* knockdown in ovarian cancer cells

Chapter 5: General Discussion

LIST OF FIGURES

Chapter 1: Introduction

Chapter 2: A metabolic labeling approach for glycoproteomic analysis reveals altered glycoprotein expression upon GALNT3 knockdown in ovarian cancer cells

Chapter 3: Altered expression of different GalNAc-Transferases (GalNAc-Ts) is associated with disease progression and poor prognosis in women with high-grade serous ovarian cancer

Chapter 4: Elucidating the role of the polypeptide N-acetylgalactosaminyltransferase 3 (*GALNT3*) and its closest homolog-*GALNT6* in mediating aberrant O-glycosylation associated with ovarian cancer dissemination

List of supplemental figures of Chapter 4: Elucidating the role of the polypeptide N-acetylgalactosaminyltransferase 3 (*GALNT3*) and its closest homolog-*GALNT6* in mediating aberrant O-glycosylation associated with ovarian cancer dissemination

Annex

LIST OF ABBREVIATIONS

ABH	Lewis antigens and blood group
Ac4GalNAc	Tetraacetylated N-acetylgalactosamine
Ac4GalNAz	Peraceetylated N-azidoacetylgalactosamine
ADC	Antibody-drug conjugates
AGC	Automatic gain control
AKAP4	A-kinase anchor protein 4
ALSA	Antibody-lectin sandwich array
AMPK	AMP-activated protein kinase
ANXA4	Annexin A4
ASM	Acid sphingomyelinase
ATG5	Autophagy protein 5
B4GALT1	Beta-1,4-Galactosyltransferase 1
BAC	Bacterial artificial chromosome
BCAAs	Branched-chain amino acids
BCAT1	Branched chain amino-acid transaminase 1
BCAT2	Branched chain amino-acid transaminase 2
BOT	Borderline ovarian tumors
CA-125	Cancer antigen 125
Cas9	CRISPR associated protein 9
CCC	Clear cell carcinoma
CCRCC	Clear cell renal cell carcinoma
CEO	Cancer épithélial de l'ovaire
CF	Cystic fibrosis
CLL	Chronic lymphocytic leukemia

Con A	Concanavalin A
CRC	Colorectal cancer
CRISPR	Clustered regularly interspaced palindromic repeats
CT	Cytoplasmic tail
CuAAC	Copper-catalyzed azide-alkyne cycloaddition
DDOST	Dolichyl-diphosphooligosaccharide-protein glycosyltransferase
DSB	Double strand breaks
EC	Endometrioid carcinoma
ECM	Extracellular matrix
EGF	Epidermal growth factor
EGFR	Anti-epidermal growth factor receptor
ELF3	Transcription factor E74-like factor 3
EMT	Epithelial–mesenchymal transition
EOC	Epithelial ovarian cancer
ER	Endoplasmic reticulum
ETD	Electron transfer dissociation
FAK	Focal adhesion kinase
FDA	Food and Drug Administration
FFPE	Formalin-fixed paraffin-embedded
FIGO	Federation of Gynecology and Obstetrics
FL	Follicular lymphoma
FOLR	Folate receptor
FTC	Familial tumoral calcinosis
FTE	Fallopian tube epithelial cells
GALE	UDP-galactose 4′-epimerase
GalNAc	N-acetylgalactosamine
GalNAc-Ts	Galnac transferases
GalNAz	N-azidoacetylgalactosamine

GALNT3	N-acetylgalactosaminyltransferase 3
GCA	Gastric adenocarcinoma
GEMMs	Genetically engineered mouse models
GlcNAc	N-acetylglucosamine
GlcNAz	N-azidoacetylglucosamine
GO	Gene ontology
GOLPH3	Golgi phosphoprotein 3
gRNA	Guide RNA
GT1	Catalytic luminal domain
HCC	Hepatocellular carcinoma
HDL	High-density lipoprotein
HE-4	Human epididymis protein 4
HGSC	High-grade serous carcinoma
HHS	Hyperostosis syndrome
Hic-5	Hydrogen peroxide-inducible clone-5
HILIC	Hydrophilic interaction liquid chromatography
HOSE	Human ovarian surface epithelial
IHC	Immunohistochemistry
ILK	Integrin-linked kinase
IMS	Imaging mass spectrometry
IP	Intraperitonial
IPA	Ingenuity pathway analysis
IsoTaG	Isotopic targeted glycoproteomics
IV	Intravenous
KD	Knockdown
KO	Knockout
LC/MS	Liquid chromatography/mass spectrometry
LFQ	Label free quantification

LGSC	Low grade serous carcinoma
LMP	Low-malignant potential
LSCC	Laryngeal squamous cell carcinoma
LTQ	Linear trap quadrupole
LWAC	Lectin weak affinity chromatography
MALDI	Matrix-assisted laser desorption/ionization
MAP1B	Microtubule-associated protein 1B
MAPK	Mitogen-activated protein kinases
MC	Mucinous carcinoma
MCA	Multiple correspondence analyses
MD	Moderately differentiated
MET	Mesenchymal-epithelial transition
MnSOD	Manganese superoxide dismutase
MSI	Mass spectrometry imaging
MSP	Methylation-specific PCR
MTDH	Metadherin
MTHFD2	Mitochondrial folate-coupled dehydrogenase
MUC1	Mucin-1
NACT	Neoadjuvant chemotherapy
nanoLC	Nanoscale capillary liquid chromatography
NB	Neuroblastoma
NEOC	Non-epithelial ovarian cancers
NF-Kappa B	Nuclear factor kappa-light-chain-enhancer of activated B cells
NHEJ	Non-homologous end joining
NID1	Nidogen-1
NK	Natural killer
NSCLC	Non-small cell lung cancers
OC	Ovarian cancer

ORF	Open reading frame
OSCC	Oral squamous cell carcinoma
OSE	Ovarian surface epithelial cells
OST	Oligosaccharyltransferases
OST	Oligosaccharyltransferases
PAM	Protospacer adjacent motif
PARP	Poly ADP ribose polymerase
PD	Poorly differentiated
PDAC	Pancreatic ductal adenocarcinoma
PFS	Progression-free survival
PI3K	Phosphoinositide 3-kinase complex
PKA	Protein kinase A
PLA	Proximity ligation assay
PTM	Post-translational modifications
PTMs	Ost-translational modifications
RCC	Renal cell carcinoma
ROCK	The rhoa Kinase
RRMP2	Ribonucleotide reductase M2
RRSO	Risk-reducing salpingo-oophorectomy
SC	Subcutaneous
SCC	Squamous cell carcinoma
shRNA	Short hairpin RNA
Sia	Sialic acid
SILAC	Stable isotope labeling with amino acids in cell culture
siRNA	Small interfering RNA
SMPD1	Sphingomyelin phosphodiesterase 1
SNO	Superficial scrapings from normal ovaries
SOC	Serous ovarian carcinomas

SOD2	Magnesium superoxide dismutase
SPEG	Solid-phase extraction of N-linked glycopeptides
SSM	Superficial spreading melanoma
ST6GalNAc1	ST6 N-Acetylgalactosaminide Alpha-2,6-Sialyltransferase 1
STT3B	STT3B, Catalytic Subunit Of The Oligosaccharyltransferase Complex
SV40	Simian virus 40
TAA	Tumor associated antigen
TALENs	As transcription activator-like effector nuclease
TGFβ1	Transforming growth factor beta
TL	Transcription activator-like
TM	Transmembrane domain
TMA	Tissue microarrays
TMAs	Tissue microarrays
TOF	Time of flight
TUBB3	Tubulin beta 3
UDP-GalNAc	UDP-N-acetylgalactosamine
UDP-GlcNAc	UDP-N-acetylglucosamine
VNTR	Variable number of tandem repeats
VVA	Vicia villosa agglutinin
WGA	Wheat germ agglutinin
ZFN	Zinc-finger nucleases

“Take pride in how far you have come and have faith in how far you can go.”

— Pravinee Hurbungs

ACKNOWLEDGMENTS

This thesis represents not only the written work put together in this manuscript, but it is a milestone of the hard work and achievements I have accomplished in the past few years. I have been given some unique and challenging opportunities that have allowed me to create tools for performing many of the complex research presented here. This thesis presents the many lessons I have learned whether they were technical or personal. This work is also a presentation of the work of dozens of people who I wish to thank and acknowledge.

First and foremost I wish to thank my supervisor, professor Dimcho Bachvarov. He has been supportive form the beginning. Dr. Bachvarov has supported me not only by providing assistance in this research project; he has also guided me all throughout the years. Dr. Bachvarov's support and mentorship has really helped me become more confident in myself, and my scientific abilities. I also highly appreciate the opportunities given by Dr. Bachvarov to present my scientific findings at multiple Oncology and Ovarian cancer meetings. Finally, during the difficult times when writing this thesis, Dr. Bachvarov gave me the moral support and the freedom I needed to move on.

I would also like to thank my friend and colleague Adnen Faddaoui for his tremendous help in introducing me to all of the experimental techniques used in the lab, also for his intellectual scientific input in many parts of my project. I also would like to acknowledge Mrs. Magdalena Bachvarova who has assisted in many of the complex experiments presented in this thesis, and for making the workplace a comfortable place to work.

I would also thank all members of our collaboration teams, for without their help, scientific experience and knowledge this thesis project wouldn't have come together. Ms. Florence Roux-Dalavi has provided me with tremendous help and support when faced with very challenging concepts of mass-spectrometry and complex data acquisition. I would also like to specially thank Dr. Christina Woo for introducing us to the IsoTag approach and teaching us the methodology needed to initiate our project. I would like to thank her for hosting me in her laboratory at Harvard Medical School, for which she has provided me with an opportunity to examine the challenging work environment of glycoproteomics.

I would like to thank the gynecologic oncology surgeons for providing us with patient specimens as well as insight into the clinical aspects of ovarian cancer. We are also

extremely grateful to the women with ovarian cancer who generously donated their tissue samples to support our research.

I would also like to thank my thesis committee who has helped in guiding me through my pre-doc exam, which has prepared me my thesis exam. Thank you Dr. Stéphane Gobeil for being a great advisor who helped in the conception and application of several of the approaches employed in this study.

I would like to thank organizations at Laval University Cancer Research Centre for providing my with Bourse d'excellence du CRC. In addition special thanks to the Glyconet society, for providing me with the opportunity to travel to Boston and work under the supervision of Dr. Christina Woo.

Lastly I would like to thank my family starting from my biggest supporter and role model, my father Ammar Sheta. My father is my best friend, he has given me the greatest chance to travel and study abroad, which has opened many doors that got me to where I am right now. This thesis was impossible without his support and encouragement. I would also like to thank my sister Reeman Sheta for always knowing how to provide me with the source of energy to keep going. She has supported me whenever I needed it.

I would like to give special thanks to my husband Abid Oueslati; he has been one of my biggest supporters. He has helped my immensely with his scientific input and guidance. He provided me with ideas that have helped me progress well in my research work. I would also like to thank my daughter Reema Oueslati for being understanding during the long days and nights spent away from home.

PREFACE

The scientific findings and data presented in this thesis is the product of three years of work as a PhD graduate student in the laboratory of Dr. Dimcho Bachvarov. I have contributed as the first author to five manuscripts and 1 review.

Chapter 1

Consists of an introduction on EOC, and glycosylation pathways involved in normal development and cancer. This chapter covers different views on the clinical and molecular aspects of EOC, and a detail description of the glycosylation pathways specifically mucin type O-glycosylation and its role in cancer development. Glycosylation is a new examined hallmark of cancer, and this chapter discusses the strong correlation of glycosylation with tumor initiation, progression and metastasis. Thus it was essential to discuss the possible role glycosylation plays in the progression of EOC.

Chapter 2

This chapter is a presentation of two of my first author articles on which I was fortunate to work on with multiple collaborators. The manuscripts are entitled "A metabolic labeling approach for glycoproteomic analysis reveals altered glycoprotein expression upon *GALNT3* KD in ovarian cancer cells" published in the Journal of Proteomics and "Proteomic dataset for altered glycoprotein expression upon *GALNT3* knockdown in ovarian cancer cells" published in Data in Brief. This project involved the application of a highly complex and innovative approach, which I was very fortunate to work on in Dr. Bachvarov's lab. I wrote the manuscripts, which were then corrected by Dr. Bachvarov and coauthors. In collaboration with Dr. Woo and the proteomics facility at the CHUL, I was able to generate almost all the figures in the manuscripts. This work was made possible through the fruitful collaborations and contributions of the aforementioned co-authors.

Chapter 3

Presents a clinical scientific paper entitled "Altered expression of different GalNAc-Transferases (GalNAc-Ts) is associated with disease progression and poor prognosis in women with high-grade serous ovarian cancer" currently under review in the International Journal of Oncology. I have significantly contributed to this paper by preparing tissue microarrays for the analysis and scoring of the various proteins included in this study. I have generated all figures and performed all the statistical analysis. Data collection was made possible by the contribution of all the medical doctors listed as co-authors in the manuscript. We are the first lab to examine multiple GalNAc-Transferases in a subset of EOC patient samples, presenting data that indicate the implication of multiple and not one GalNAc-Transferase in the progression and poor outcome in EOC patients.

Chapter 4

Presents a scientific paper currently in preparation entitled "Elucidating the role of the polypeptide N-acetylgalactosaminyltransferase 3 (*GALNT3*) and its closest homolog-*GALNT6* in mediating aberrant O-glycosylation associated with ovarian cancer dissemination". I was also privileged to work on such a complex project examining biosynthetic backup of multiple GalNAc-Transferases, in addition to examining their functional redundancy in EOC cell lines. Our work presented in this chapter is the first to examine the effect multiple GalNAc-Transferases gene knockouts can have on the survival of animals. Indicative of the role multiple GalNAc-Transferases play in the metastatic cascade of EOC. I have prepared all figures and collected all the data; *in vivo* studies were carried out in Dr. Vanderhyden's lab. This paper will have a very high impact on the examination of multiple gene targets in EOC.

Chapter 5

This chapter briefly discusses the work gathered in this manuscript.

Annex

An ‘Annex’ section: Annex. This section includes the scientific review Dr. Bachvarov and I worked on when we first embarked on the glycosylation study entitled “Role of aberrant glycosylation in ovarian cancer dissemination” published in Biomedical reviews.

PRÉFACE

Les résultats scientifiques et les données présentées dans cette thèse sont le fruit de trois années de travail en tant que doctorante au laboratoire de Dr. Dimcho Bachvarov. J'ai contribué en tant que première auteure à cinq manuscrits et une revue scientifique.

Chapitre 1

Le chapitre 1 comprend une introduction sur le CEO et les voies de glycosylation impliquées dans le développement normal et dans le cancer. Ce chapitre couvre différents aspects cliniques et moléculaires du CEO, et une description détaillée des voies de glycosylation, spécifiquement la O-glycosylation du type mucine et son rôle dans le développement du cancer. La glycosylation est une nouvelle caractéristique du cancer, et ce chapitre traite de la forte corrélation entre la glycosylation et l'initiation, la progression et la métastase de la tumeur. Ainsi, il était essentiel de discuter du rôle possible joué par la glycosylation dans la progression du CEO.

Chapitre 2

Ce chapitre est une présentation de mes articles en tant que première auteure, dans lequels j'ai eu la chance de travailler avec plusieurs collaborateurs. Les manuscrits sont intitulés " A metabolic labeling approach for glycoproteomic analysis reveals altered glycoprotein expression upon *GALNT3* knockdown in ovarian cancer cells", publié dans le journal Journal of Proteomics. et "Proteomic dataset for altered glycoprotein expression upon *GALNT3* knockdown in ovarian cancer cells" publié dans le journale Data in Brief. Ce projet impliquait la mise en place d'une approche hautement complexe et novatrice, durant laquelle j'ai eu beaucoup de chance de travailler dans le laboratoire de Dr. Bachvarov. J'ai écrit le manuscrit, qui a ensuite été corrigé par Dr. Bachvarov et les co-auteurs. En collaboration avec Dr. Woo et la plateforme de protéomique au CHUL, j'ai pu générer la majorité des figures du manuscrit. Ce travail a été rendu possible grâce aux collaborations fructueuses et aux contributions des co-auteurs susmentionnés.

Chapitre 3

Ce chapitre présente un article scientifique clinique intitulé " Altered expression of different GalNAc-Transferases (GalNAc-Ts) is associated with disease progression and poor prognosis in womem with high-grade serous ovarian cancer ". Ce papier est actuellement en cours d'évaluation dans le journal : Journal international d'oncologie. J'ai contribué de manière significative à cet article en préparant des microarrays de tissus pour l'analyse et l'appreciation des diverses protéines incluses dans cette étude. J'ai généré toutes les données et effectué toutes les analyses statistiques. La collecte de données a été rendue possible par la contribution de tous les médecins énumérés comme co-auteurs dans le manuscrit. Nous sommes le premier laboratoire à examiner plusieurs GalNAc-Transférases dans un sous-ensemble d'échantillons de patients CEO, présentant des données qui indiquent l'implication de multiples GalNAc-Transférase dans la progression du CEO.

Chapitre 4

Ce chapitre présente un article scientifique actuellement en préparation intitulé " Elucidating the role of the polypeptide N-acetylgalactosaminyltransferase 3 (*GALNT3)* and its closest homolog-*GALNT6* in mediating aberrant O-glycosylation associated with ovarian cancer dissemination". J'ai également eu le privilège de travailler sur un projet aussi complexe qui examine la redondance fonctionnelle de plusieurs GalNAc-Transférases dans les lignées cellulaires de CEO. Notre travail présenté dans ce chapitre est le premier à examiner l'effet d'inactivation de multiples gènes de GalNAc-Transférases sur la survie des animaux, soulignant le rôle que jouent les GalNAc-transférases dans la cascade métastatique de CEO. J'ai préparé tous les résultas et recueilli toutes les données. Les études *in vivo* ont été réalisées dans le laboratoire du Dr. Vanderhyden. Cet article aura un impact très important sur l'examen de multiples cibles de gènes dans le CEO.

Chapitre 5

Ce chapitre expose brièvement le travail recueilli dans ce manuscrit.

Annexe

Une section «Annexe»: Annexe. Cette section inclut la revue scientifique préparée par Dr. Bachvarov et moi même intitulée " Role of aberrant glycosylation in ovarian cancer dissemination" publiée dans le journal Biomedical reviews.

Chapter 1: Introduction

1.1 Ovarian cancer

1.1.1 Overview of ovarian cancer

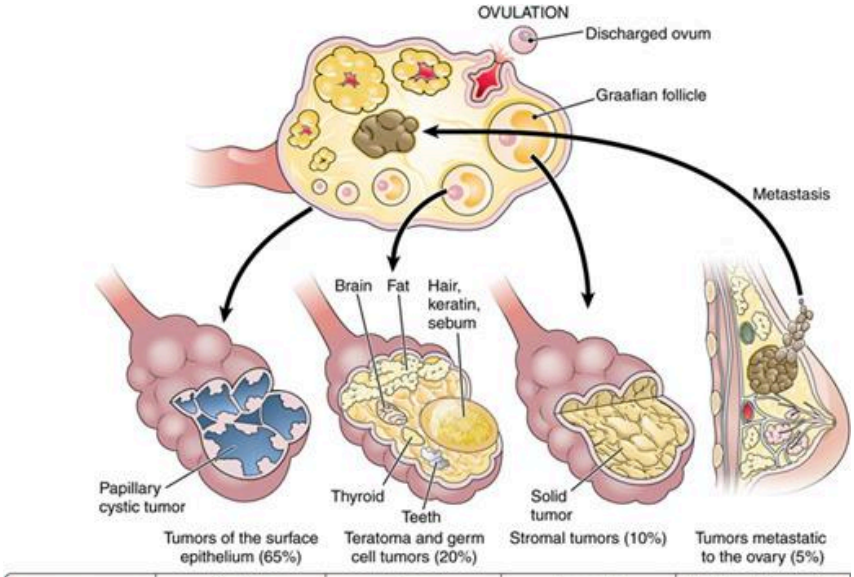
Ovarian cancer is one of the deadliest gynaecologic malignancies amongst women in the Western world, killing over 14,000 women each year (1). Ovarian cancer is the fifth leading cause of cancer deaths amongst women (2). Additionally, ovarian cancer 5-year survival rate has been reported to be less than 46 percent (3). Although ovarian cancer is not very common, it is the deadliest it is the deadliest cancer among all gynecologic malignancies, that is why it has been referred to as the "silent killer" (4). This is indeed due to the fact that ovarian cancer patients do not present with defined symptoms, especially not for the early stages of the disease (4). Ovarian cancer is one of the fewest cancers that cannot be referred to a single cancer type. Ovarian cancers refer to a subset of distinct types of cancers that not only originate from the ovary, but also it refers to those cancers that involve the ovary. Recent studies have suggested that some ovarian cancers arise from the fallopian tubes, and subsequently metastasize to the ovary, while other studies suggest that ovarian cancers arise from cells that are not considered intrinsic to the ovary such as the endometrium (5). Research is now more focused on properly defining ovarian cancer, since this disease exists with different origins, risk factors, genetic mutations, and prognoses. Till today, there is still difficulty in understanding where the different ovarian cancers arise and how they behave, which indicates the importance for improvements in the diagnosis, detection and treatment of these cancers.

1.1.2 The different types of ovarian cancer

There are three main categories of cells that make up the ovaries (epithelial cells, germ cells and stromal cells) (6). Each of these cell types can develop into a distinct type of ovarian tumor (6, 7) (Figure 1.1).

Epithelial ovarian cancer (EOC), which originates from cells covering the outer surface of the ovaries (Figure 1.1), account for 90% of all malignant ovarian cancers, and the American Cancer Society reported 21,550 cases of EOC in 2009 and these numbers were predicted to increase (*8, 9*), and commonly occurs in postmenopausal women. EOC is

classified into five different subtypes (8), which will be discussed in more detail in the following sections. Accordingly, non-epithelial ovarian cancers (NEOC), are not very common, accounting for 5-10% of all ovarian cancers (*10*). NEOC include: germ cell tumors (GCTs) and sex cord or stromal cell tumors (SCSTs) (*10*) (Figure 1.1). GCTs make up around five percent of ovarian cancers, and this type of tumor originates from cells that produced the eggs (ova). Patients presenting with this type of cancer belong to women in their early 20s (*10*). GCTs are also divided into different subtypes, and the most common types include teratomas and dysgerminomas, in addition to other not very commonly observed subtypes (*10*) (Figure 1.1). SCSTs account for around 3-5% of malignant ovarian tumors, these tumors originate from the connective tissue cells that cover and hold the ovaries together and produce the two female hormones estrogen and progesterone (*10*). SCSTs occur in women of all ages, and the main subtypes include: fibroma, granulosa cell tumors and sertoli-leydig cell tumors (*10*) (Figure 1.1). In addition to tumors that originate from ovarian cells, ovarian tumors may also originate from cells of other primary malignancies (*11*). Tumors metastasizing to the ovaries account for 5% of all ovarian tumors, and this rate is not considered to be very low since the ovaries are known to be common metastatic sites. The most common primary malignancies that metastasize to the ovaries include breast, colorectal, lung and gastric cancers (*11, 12*) (Figure 1.1).



Cell of origin:	Surface epithelial cells	Germ cells	Sex cord, stromal cells	Metastasis to ovaries (Krukenberg)
Proportion (%) of ovarian tumors:	65-70%	15-20%	5-10%	5%
Proportion (%) of malignant ovarian tumors:	90%	3-5%	2-3%	5%
Age group affected:	>20 years	>25 years	All ages	Adults
Types:	• Serous tumor • Mucinous tumor • Endometrioid tumor • Uncommon types • Cystadenocarcinoma	• Teratoma • Dysgerminoma • Uncommon types	• Fibroma • Granulosa-theca cell tumor • Sertoli-Leydig cell tumor	Metastases from: • Breast • Lung • Colon • Other sites

Figure 1.1. The origin and type of EOC tumors
The table summarizes details on the frequency, age range, and common subtypes of ovarian tumors (*13*).

1.1.3 The five histopathological subtypes of EOC

EOCs have been divided into 5 major subtypes which include: 1) high-grade serous carcinoma (HGSC) 2) endometrioid carcinoma (EC) 3) clear cell carcinoma (CCC) 4) low grade serous carcinoma (LGSC) and 5) mucinous carcinoma (MC) (*14*) (Figure 1.2). The naming of these EOC subtypes is dependent on how closely the tumor cells found within the tissues resemble the normal cells lining of the different organs found in the female genitourinary tract (*15*).

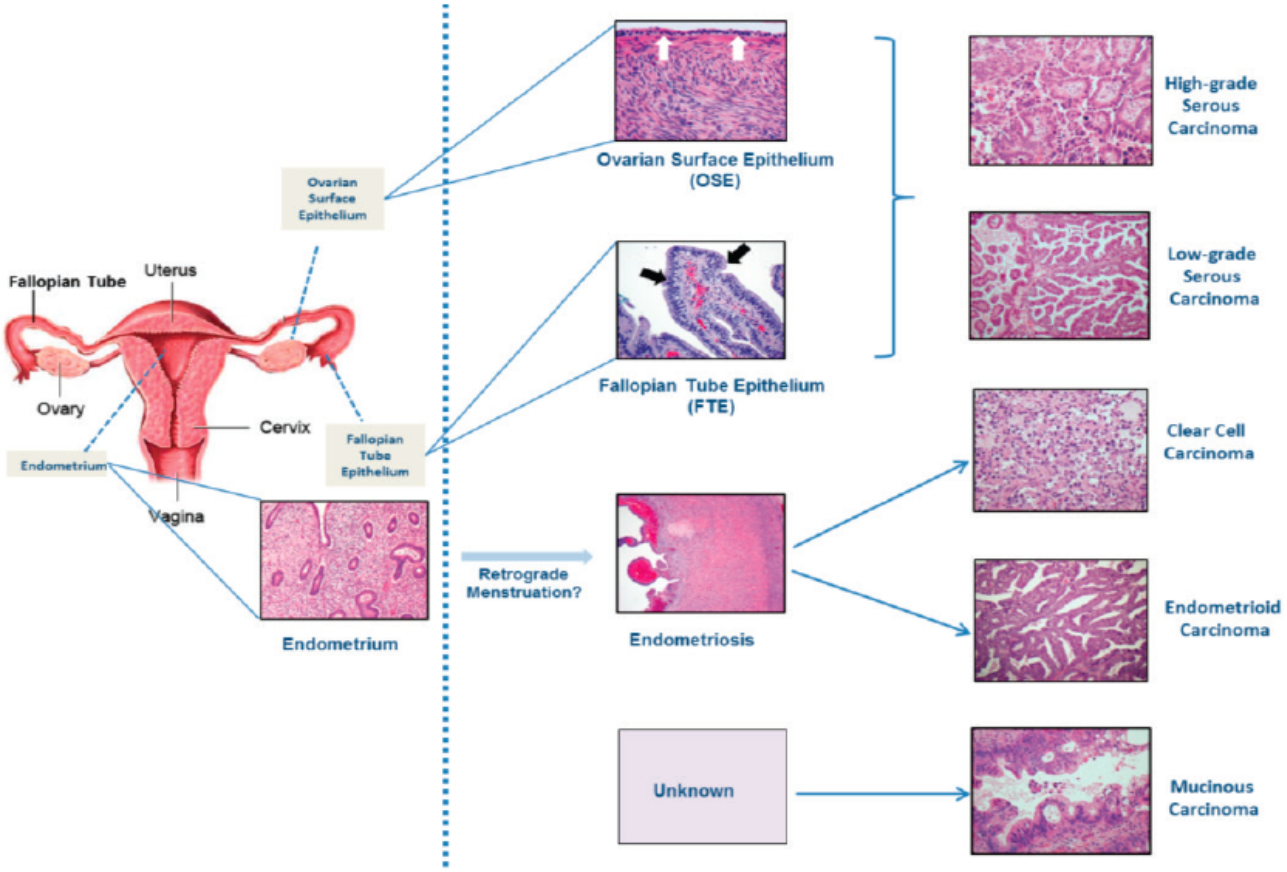


Figure 1.2. The different cellular origins of EOCs
The figure summarizes literature data suggesting that many ovarian carcinomas originate from outside the ovaries. The figure includes a detailed classification of the different ovarian carcinomas and the fact that they arise from different tissue types (*16*).

1.1.3.1 High-Grade Serous Carcinoma (HGSC)

HGSC is the most common type of EOC, and it accounts for more than 74% of all EOCs (*17, 18*). HGSC appears in patients at the later stages of the disease (16). This type of tumor is characterized with a solid growth found within the tumor cells, usually defined with abnormal nuclei of various shapes and sizes (16). Another known feature of these cells is their level of cellular proliferation (*17, 18*). Patients with HGSC, present with large masses on their ovaries in addition to metastatic spreading into the omentum, and other abdominal regions (16). Genetic characterization of HGSC show that this type of cancer is genetically unstable, with mutations in the tumor suppressor genes: *TP53, BRCA1* and *BRCA2* (*19*).

1.1.3.2 Endometrioid carcinoma (EC)

EC ovarian cancer has been suggested to be 50% present at Stage I, and usually involves the two ovaries (*20*). This is the more favorable subtype of EOC, and women treated for this cancer subtype have better survival and overall outcomes (*21*). However, the remaining 50% of EC patients diagnosed at later stages show poor survival rates (*21*). As seen from figure 1.2, EC is an EOC carcinoma characterized with a defined link to endometrial tissue (*22-24*). One of the major challenges associated with this disease is that around 5% of ECs are usually associated with synchronous uterine EC, when diagnosed (*25, 26*). These cases have shown to be complicated, since pathologists have difficulties determining if the endometrial and ovarian tumors represent two independent carcinomas or a single cancer type arising in a single organ and metastasizing to another (*25, 26*). Genetic mutations that define this EOC subtype include: *CTNNB1, PIK3CA, KRAS, ARID1A, PTEN*, and *PPP2R1A* (*27*). Furthermore, *CTNNB1* mutations are common genetic alterations in ECs, and interestingly they show to be unique to this type of cancer since they are rarely observed in the other EOC subtypes (*27*).

1.1.3.3 Clear cell carcinoma (CCC)

The nomenclature of CCC was given to this subtype since the CCC tumor cells display large profusions of clear cytoplasm, and this is in part due to the accumulation of intracytoplasmic glycogen in these cells (*28*) (Figure 1.2). Diagnosis of this subtype is mainly classified at Stage I, but nonetheless, prognosis of patients with advanced stages of CCC is unfavorable (*29-31*). Similar to EC, CCC is associated with endometriosis (*32*), and it was shown that CCCs is more commonly associated with cysts or benign tumors when defined at the early stages (*32*). CCCs frequently contain mutations in *ARID1A*, and *PIK3CA* (*33, 34*). Other CCC mutations, such as *PPP2R1A*, *PTEN, KRAS*, and *TP53*, have also been identified, but not as frequently as those described earlier (*35, 36*).

1.1.3.4 Low-grade serous carcinoma (LGSC)

LGSCs are considered to be the less common type of ovarian tumors, as lower than 20% of LGSCs are diagnosed at stage I (*37-39*). Interestingly, LGSCs have been suggested to be more chemoresistant compared to HGSCs (*40*), and survival rates of women diagnosed at the later stages of this disease have not improved in the last decade (*41*). Pathological examinations of this disease showed that LGSCs often arise from serous borderline tumors (SBTs) (*41*). Histological classification of LGSCs shows that the tumor cells have small nuclei characterized with low mitotic activity (Figure 1.2). There are some discrepant reports on the relationship between LGSCs and HGSCs, since it is suggested that LGSC do not progress into HGSC (*19*), while other suggest that LGSC may progress into the so called intermediate stage before evolving into HGSCs (*42-44*). Genetic mutations characteristic of LGSCs include mutations of *KRAS* or *BRAF* and *ERBB2* (*19, 45*).

1.1.3.5 Mucinous carcinoma (MC)

MCs are the least common subtype of EOC (*46*). Mucinous tumors do not have a clear defined origin, and it has been suggested that MCs may originate from other cancer types such as colorectal cancer (*47*). Colorectal cancers are primary tumors that have been shown to frequently metastasize to the ovaries, thus mimicing primary MC tumors (*47*). Patients diagnosed with MC at the advanced stages of the disease show very poor survival outcome, while those patients diagnosed at the earlier stages show very good prognosis (*48, 49*). Histological classification of MCs show characteristics of columnar cells with basal nuclei, and the cytoplasm of these cells show very low mucin staining (Figure 1.2) (*50*). Genetic mutations of MCs include mutations in the *KRAS* and *TP53* genes (*51*). Additionally, *ERBB2* amplification has been reported in more than 10% of MCs (*52*), suggesting that both the *ERBB2* amplification and the *KRAS* mutation might be associated with improved survival in MC patients (*52*).

1.1.4 EOC tumor staging and grading

Complicated diagnosis of EOC has hindered its early detection, since most of the time, EOC is diagnosed at advanced stages, when the cancer has spread outside its tissue of origin and metastasized throughout the abdominal area (*53*). Early diagnosis of EOC has

shown to be difficult due to the fact that localized ovarian tumors are usually asymptomatic, while advanced disease present with symptoms such as abdominal pain, pelvic masses and vaginal bleeding (*53*). Surgeons and pathologists have developed classification systems to better categorize the grading and stages of this cancer type. Surgical grouping has allowed for a better classification of the extent of the disease ordered by examination of the pelvic mass, amount of ascites fluid present, level of capsule rupture, diaphragm inspection, region of metastasis, and small and large bowel examination (*54*). Table 1.1 summarizes the Federation of Gynecology and Obstetrics (FIGO) staging of EOC criteria (*54*).

Table 1.1 FIGO stage grouping for primary carcinoma of the ovary

FIGO Stage Grouping for Primary Carcinoma of the Ovary

CURRENT FIGO STAGE GROUPING	
Stage I	Growth limited to the ovaries
Stage IA	Growth limited to one ovary; no ascites
Stage IAi	No tumor on the external surface, and no capsule rupture
Stage IAii	Tumor present on the external surface, or capsule ruptured, or both
Stage IB	Growth limited to both ovaries; no ascites
Stage IBi	No tumor on the external surface; capsule intact
Stage IBii	Tumor present on the external surface, or capsule ruptured, or both
Stage IC	Tumor either stage IA or IB but with ascites present or with positive peritoneal washings
Stage II	Growth involving one or both ovaries with pelvic extension
Stage IIA	Extension of metastases to the uterus or tubes
Stage IIB	Extension to other pelvic tissues
Stage IIC	Tumor either stage IIA or IIB but with ascites present or with positive peritoneal washings

Stage III	Growth involving one or both ovaries with intraperitoneal metastases outside the pelvis or positive retroperitoneal nodes. Tumor limited to the true pelvis with histologically proved malignant extension to small bowel or omentum.
Stage IV	Growth involving one or both ovaries with distant metastases. If pleural effusion is present, there must be positive cytology to allot a case to stage IV. Parenchymal liver metastases indicate stage IV.

NEWLY ADOPTED FIGO STAGE GROUPING*

Stage I	Growth limited to the ovaries
Stage IA	Growth limited to one ovary; no ascites. No tumor on the external surface; capsule intact.
Stage IB	Growth limited to both ovaries; no ascites. No tumor on the external surfaces; capsules intact.
Stage IC†	Tumor either Stage IA or IB but with tumor on the surface of one or both ovaries, or with capsule ruptured, or with ascites present containing malignant cells or with positive peritoneal washings.
Stage II	Growth involving one or both ovaries with pelvic extension.
Stage IIA	Extension and/or metastases to the uterus and/or tubes.
Stage IIB	Extension to other pelvic tissues.
Stage IIC†	Tumor either Stage IIA or IIB but with tumor on the surface of one or both ovaries, or with capsule(s) ruptured, or with ascites present containing malignant cells or with positive peritoneal washings.
Stage III	Tumor involving one or both ovaries with peritoneal implants outside the pelvis and/or positive retroperitoneal or inguinal nodes. Superficial liver metastasis equals Stage III. Tumor is limited to the true pelvis but with histologically verified malignant extension to small bowel or omentum.
Stage IIIA	Tumor grossly limited to the true pelvis with negative nodes but with histologically confirmed microscopic seeding of abdominal peritoneal surfaces.
Stage IIIB	Tumor of one or both ovaries with histologically confirmed implants of abdominal peritoneal surfaces, none exceeding 2 cm in diameter. Nodes negative.
Stage IIIC	Abdominal implants >2 cm in diameter and/or positive retroperitoneal or inguinal nodes.
Stage IV	Growth involving one or both ovaries with distant metastasis. If pleural effusion is present there must be positive cytologic test results to allot a case to Stage IV. Parenchymal liver metastasis equals Stage IV.

The table includes information on stages I to IV ovarian cancers with a detailed description of each stage characteristic that is defined histologically following the FIGO criteria (*54*).

In addition to classifying ovarian tumors according to their FIGO stage, doctors have developed a classification system divding EOC into two types (Type I and Type II) based on the shared features of the different cancer types (*55*) (Figure 1.3). Type I carcinomas are low-grade cancers that are less aggressive, and are considered to be genetically stable, and surgically classified to arise form endometriosis or benign tumors. Type II cancers are high-grade EOCs classified to be biologically highly aggressive tumors, with a high tendency to metastasize from their primary lesions (*56*). Moreover, Type II carcinomas of the ovary have been genetically characterized to carry frequent mutations in *TP53*, *BRCA1* and *BRCA2* (*56*).

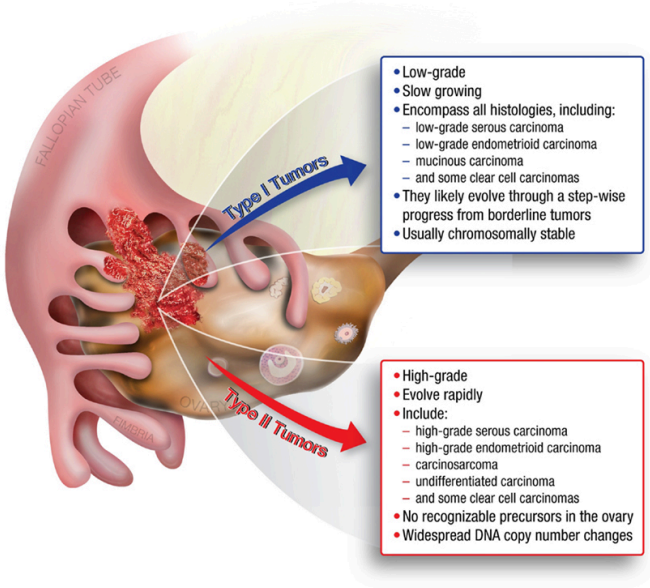


Figure 1.3. The two major histologic subtypes of EOC and their sub classifications
Type I tumors include low-grade, slow growing carcinomas known to develop from borderline tumors that develop from the ovarian surface epithelium, inclusion cysts, or endometriosis. Type II tumors include high-grade and rapidly growing carcinomas known to spread beyond the ovaries (*57*).

Additionally, pathologists classify EOC by grading the cancer, which allows examining the degree of similarity between EOC to healthy cells (*58*). When examined, the grading of the cancerous tissue helps pathologists make predictions on features of the cancer type such as its spreading potential, which in turn helps in improving prognosis (*58*). Healthy tissues consist of several cell types that are usually found clustered together, and when examining

tumor tissue, pathologists compare their cell characteristics to the normal healthy tissue and have developed a method to classify them accordingly: **a)** if the cancer cells display similar cellular architecture to normal cells, but contain a different type of cell grouping then it is referred to as a low-grade tumor or differentiated tumor (*15*); **b)** if the cancer cells look very different from healthy normal cells, they are then referred to as high-grade tumors or poorly differentiated tumors (*15*). There are not only two classifications for these cancer types, but some ovarian tumors are known as borderline cancers or borderline ovarian tumors (BOT) and are characterized by the absence of invasive stoma (*50*). BOTs have also been referred to as low-malignant potential tumors (LMP), they are also called borderline since these tumors present with cellular characteristics found in benign tissue or invasive cancers (*50*).

Importantly, epithelial EOC are mostly of the serous type and are graded as either:

1) LGSC -- low-grade serous carcinoma
2) HGSC -- high-grade serous carcinoma

While the other subtypes have been proposed to be graded, as follows:

Grade 1: The cancerous tissue is well differentiated, for which is appears to contain many normal-looking cells.

Grade 2: The tissue is moderately differentiated, and these cells appear to look more abnormal than normal.

Grade 3: The tissue is very poorly differentiated, and most of the cells in these tissues appear highly abnormal. (*58*)

1.1.5 Carcinogenesis and genetics of EOC

Like many other cancer types, EOC develops as a result of genetic mutations, and as acknowledged, these mutations play a major role in the acquisition of differentiated cellular functions that pertain and are not limited to enhanced growth and proliferation, increased cellular invasion and migration, angiogenesis, increased drug resistance and evasion of the immune system (*59*). There are two possible ways for the acquisition of genetic abnormalities, either through **a)** germline or inherited mutations, or **b)** acquired somatic mutations (*60*). These mutations are then known to result in either activating or inactivating tumor suppressors and oncogenes (*60*). One common example of inherited genetic

mutations in EOC is the inactivation of the tumor suppressor genes *BRCA1* and *BRCA2* (*61*). Population based studies on EOC patients indicate that between 10 and 15% of these patients have mutations in the *BRCA1* or the *BRCA2* genes (*62, 63*). Moreover, other genetic mutations have been described in EOC, as these gene mutations have been detected through the application of genome-wide association studies (*64*). These studies indicate inherited mutations of the genes *MLH1, MSH2, MSH6, PMS2, EPCAM,* and *TP53* (*64*). Similarly, mutations acquired in EOC tumors are mostly found in HGSCs, with a common mutation in the *TP53* gene (*65*). Furthermore, somatic mutations in the *BRCA1* and *BRCA2* genes also appear to occur in more than 30% of HGSCs (*66*). Likewise, rare somatic mutations have also been found in HGSCs, including activation and inactivation of the following genes: *BRIP1, CHEK2, NF1, FAT3, CCNE1, PIK3CA, KRAS, MYC, PTEN, RB1, NF1, CDKN2A* and *RAD51C* (*56, 67*). Table 1.2 summarizes the list of identified mutations in EOC and their suggested functions in the different EOC subtypes (*50*).

Table 1.2. Characteristic mutations found in EOCs

Carcinoma Subtype	Gene Activation	Gene Inactivation
HGSC	• No frequent mutations	• *BRCA1*, *BRCA2*: tumor suppressor genes that help repair DNA damage or destroy cells if DNA cannot be repaired • *TP53*: a tumor suppressor commonly mutated in cancers and crucial for genomic stability and DNA repair
LGSC	• *BRAF*: an oncogene involved with intracellular signaling involved with directing cell growth • *KRAS*: an oncogene that recruits and activates proteins necessary for tumor growth	• Unknown
MC	• *BRAF* • *ERBB2*: a receptor tyrosine kinase that can interact with signaling molecules to promote tumor growth and block cell death. It is commonly amplified or overexpressed in cancers • *KRAS*	• *CDKN2A*: a tumor suppressor that induces cell cycle arrest • *RNF43*: a tumor suppressor that inhibits WNT signaling pathway • *TP53*
EC	• *CTNNB1*: operates as a signal transducer to regulate gene transcription and cell–cell adhesion • *KRAS* • *PIK3CA*: promotes cell survival, growth, and migration • *PPP2R1A*: regulates signaling pathways that inhibit cell growth and division	• *ARID1A*: regulates gene by altering the accessibility of transcription factors to DNA • *BRCA1*, *BRCA2* • *PTEN*: a tumor suppressor that inhibits cellular growth. This gene is often mutated in cancer
CCC	• *KRAS* • *PIK3CA* • *PPP2R1A*	• *ARID1A* • *PTEN* • *TP53*

The table includes a list of the defined genetic alterations found in the different ovarian cancer subtypes. The gene list includes genes found either activated or inactivated in ovarian cancer (*50*).

In addition to genetic modifications found in EOC, epigenetic modifications have also showed to play additional roles in the EOC etiology (*68*). These types of modifications alter DNA accessibility to major transcription factors that in turn modify gene expression by either promoting or inhibiting their expression (*68*). The most common epigenetic modifications in EOC are DNA methylation, histone modifications, and more recently the suugested role of noncoding RNAs (ncRNAs) in the metastasis of EOC (*68, 69*). Multiple epigenomic studies have identified hyper and hypomethlyated genes and examined their role in the progression and carcinogenesis of EOC; most of them associated with poor

prognosis of the disease (*70*). One such example is the identified promoter methylation of the *BRCA1* gene, found in more than 10% of HGSC patients (*71*). Studies from our lab also identified a role of aberrant DNA methylation in EOC, where data demonstrated that DNA hypermethylation occurs in low-malignant potential (borderline) tumors and these hypermethylated genes predominantly included key developmental/homeobox genes (*72*). Interestingly, contrary to DNA hypermethylation, significant DNA hypomethylation was observed almost exclusively in high-grade (grade 3) serous EOC tumors (*72*). These reports suggest that more focused research in the area of epigenetics may aid in the development of new EOC gene targets and more specific EOC therapies (*68*). Furthermore, studies in the field of proteomics and metabolomics have also helped in providing new understanding in EOC development, and specifying new detection and monitoring methods for this disease (*73, 74*). Proteomics studies have allowed for a close examination of important signaling pathways in EOC (*74, 75*). Thus, if research can have a better understanding of multiple signaling pathways suggested to play a role in EOC, then more sensitive and targeted therapies can be better explored and applied in the field.

1.1.6 EOC metastasis

EOC is one of the deadliest cancer types, especially since it presents at later or advanced stage of the disease (*76*). EOC patients show non-specific symptoms, making it harder to diagnose, in addition EOC is characterized with a highly aggressive metastatic profile, and is thus known to exhibit a high metastatic potential (*76*). Metastasis of EOC, like many other cancer types is known to progress in a manner reflected by documented changes in the cell's biochemistry, morphology, migratory and invasive patterns (*76*). Nonetheless, the biological behavior of EOC is unique compared to other cancer types, since the process of EOC metastasis seems to occur more rapidly (*77*). It has been suggested that EOC from primary tumor sites can metastasize in a passive mechanism, which is made possible by the characterized physiological movement of peritoneal fluid to the peritoneum and omentum (*77*). The microenvironment where EOC cells to spread provides a unique placement for attachment and metastasis of these cells (*76*). For instance, the pelvic area provides EOC cells with high potential for metastasizing, since the organs in the pelvic region lack barriers that would usually hinder the spreading or metastasis of these tumor cells (*76*).

Additionally, the epithelial–mesenchymal transition (EMT) pathway is also another factor that promotes the invasive and migratory capacity of EOC cells (*78*). EMT contributes to the generation of circulating tumor cells from epithelial cancers such as EOCs, by promoting tumor cell intravasation to form mico-metastasis allowing for clonal outgrowth (*79*). Multiple studies have examined different gene pathways and protein types that regulate extracellular matrix degradation and coordinate cell junction assembly in EOC. Examinations of newly identified protein and gene receptors and their role in EMT induced EOC metastasis, include: the A-kinase anchor protein 4 (*AKAP4*) (*80*), Annexin A2 (*81*), Nidogen-1 (*NID1*) (*82*), the Golgi phosphoprotein 3 (*GOLPH3*) (*83*), the transcription factor E74-like factor 3 (*ELF3*) (*84*), in addition to two genes studied by our group: **a)** *Hic-5* where we examine its role in EOC tumorigenesis through the regulation of the EMT pathway in a TGF-β-1 indpendent manner (TGF-β-1 is a transforming growth factor known to induce EMT, which is an essential process for cancer cell progression) (*85, 86*), **b)** the mannose receptor gene *LY75*, where our group examined its role in modulating EOC dissemination by regulating both the EMT and mesenchymal-epithelial transition (MET) pathways (*87*). Additionally, microRNAs are now more highly examined for their role in controlling signaling pathways that mediate EOC metastasis, studies have also examined several regulatory microRNAs that are shown to be involved in the formation of cancer stem cells, EMT, increasing migratory capacity of cancer cells, the enhanced formation of spheroids, apoptosis, autophagy, angiogenesis, formation and development of ascites. These EOC-related microRNAs include miR-145, mir-21, miR-31, miR-506, miR-101, miR-200, miR-214, and miR-25, each of them playing specific roles in the different EOC histotypes (*76*). Moreover, EOCs are also known for their peritoneal recurrence, and this occurs due to the availability of a protective niche for EOC cells, enhancing their ectopic survival by reprogramming genetic pathways that induces their capacity to become more aggressive when metastasized (*88-91*).

1.1.7 Risk factors of EOC

Research in EOC has led to the identification of several risk factors associated with EOC development, in addition to factors that have showed a protective role for the developing of the disease (*92*). Table 1.3 list the most described types of risks factors associated with

EOC (*92*). Although this list has been properly identified by researchers, doctors and pathologists, some shortcomings are linked to these risk factors in that they are most usually associated with the least least aggressive types of EOC, and thus are not directly linked with the more lethal types of EOC such as HGSC (*92*).

Table 1.3. Risk factors for EOC

Increased Risk	Decreased Risk
Age	Oral contraceptive (OC) use
Family history of cancer	Oophorectomy
Hereditary cancer syndromes	Hysterectomy
Obesity	Tubal ligation
Nulliparity	Lactation
Hormone replacement therapy	Salpingectomy
Increased numbers of lifetime ovulatory cycles	Bilateral salpingo-oophorectomy (BSO)

The list of factors that either increase or decrease the risk of developing ovarian cancer (*92*).

As with many cancer types, EOC risk increases with the patient's age; thus, the age-adjusted incidence rate for women ages 65 and older is currently suggested to be more than five times higher than women younger than 65 (1). This incidence rate is due in part to the long-term exposure to environmental factors that help induce genetic alterations in the patients system (1). In addition to age being a leading risk factor in EOC, the hereditary aspect plays a big role in EOC development, since as discussed earlier; there are multiple germline mutations that can induce the risk of EOC development (*93-96*). Hereditary type genetic mutations account for around 15% of all EOCs (*97*), as previously discussed, *BRCA1* and *BRCA2* are the most described genes known for their role in increasing the risk of EOC development (*98-100*). Less examined hereditary EOC gene mutations include mutations in the genes *CHEK2, PALB2, RAD51* family, *BRIP1*, and *BARD1* (*101*). Moreover, reproductive factors such as parity, infertility and breastfeeding have been also suggested as factors that can either increase or decrease the risk of developing EOC (*102*). Parity for example has been suggested to play a protective role, as data shows that pregnancies decrease the risk of EOC development by as high as 40% (*103*). Also, data

suggests that breastfeeding reduces the risk by around 8% (*104*). Moreover, studies have linked infertility with the possibility of increasing the risk of EOC (*105*). Similarly, the use of fertility drugs has also been suggested to increase the risk of EOC development (*105*). Hormone use has also been linked with EOC, for instance steroid hormones have been suggested to enhance EOC risk (*106*). Studies have also examined the role of hormones in EOC development when studying the link between postmenopausal hormone replacements and the possible increase in EOC risk (*107*). This data suggested that hormone therapy is highly associated with EOC development, with a calculated 20% increase in the chances of developing the disease (*107*). Other factors include lifestyle aspects such as obesity, diet, physical activity and smoking, which all were suggested to associate with EOC risks (*108*). Diet and EOC risk, have been examined by the Women's Health Initiative, and data show that there is a 40% lower risk for EOC development in women following low-fat diets (*109, 110*). Interestingly, studies also link the consumption of green tea in decreasing the chance of developing EOC (*111, 112*). Finally, an association between smoking and EOC risk has been suggested, but this relationship is shown to significantly vary by ovarian cancer subtype (*113*). One such study presented analysis that show a 7% increase in the risk of developing EOC in women smokers compared to nonsmokers, but as stated, they found that this direct link varies depending on the EOC histological subtype (*113*).

1.1.8 Prevention and detection of EOC

Preventative approaches, focused on reducing EOC recurrence, are dependent on ways to modulate female hormone cycles, while more invasive approaches include surgery that is predominantly focused on the removal or modification of gynecological tract components, such as the fallopian tubes (salpingectomy), ovaries (oophorectomy), and uterus (hysterectomy) (*114*). One such example of preventative surgery is known as risk-reducing salpingo-oophorectomy (RRSO), which is an approach mainly used for women that are *BRCA1/2* mutation carriers (*115*). Moreover and as mentioned, hormonal therapy is the less invasive preventative method, but hormonal intervention has shown to produce conflicting results on EOC risk (*116, 117*). Thus, more research studies need to be conducted to better stratify the type of hormones to use and the period in which they can be applied. As previously mentioned, EOCs have relatively high mortality rates in women in the Western

world due to the lack of presenting symptoms, hence research is currently more focused on finding alternatives for early detection of the disease. One such approach is the use of biomarkers as assay methods to better detect the cancer and its stage. Several biomarkers have been described to be useful in the prognosis of patients at the early stages, such markers include the cancer antigen 125 (CA-125) and human epididymis protein 4 (HE-4) (*118*). Nonetheless, these tumor markers have not been the best strategy for early detection in all EOC subtypes due to their heterogeneity associated with this disease. CA-125 is the most promising tumor marker used in EOC detection; it became recognized when an antibody was developed against this antigen in the 80s (*119*). Since then, studies revealed that more than 80% of women with advanced EOC display significantly elevated serum levels of CA-125 (*120*). Interestingly, it was found that CA-125 levels in patients serum directly correlated with the stage of EOC in addition to being a predictive tool for chemotherapy response, further supporting that CA-125 can be only applied as a useful marker of disease progression (*121, 122*). Although showing great promise, CA-125 has also been found to be relatively unspecific and/or insensitive for the early detection of EOC, since CA-125 can also be elevated in cases of benign EOC tumors, in cases with patients presenting with inflammation, and even in non-EOC tumors, and interestingly, in patients with EOC, CA-125 can sometimes be undetected (*123, 124*). Similarly, HE-4 is also a recently discovered biomarker for the detection of EOC, and has been shown to be most effective in differentiating between normal ovarian and EOC tissue (*125, 126*). Similar to CA-125, HE-4 is not a unique biomarker for those women with EOCs, but also displays elevated levels in patients with other types of gynecologic tumors (*127, 128*).

1.1.9 Treatment methods of EOC

When women are diagnosed with EOC, they usually receive combination treatments of surgery and chemotherapy (*129-131*). Surgical management is highly dependent on the diagnosis of the patient and helps in determine the staging of the disease (*132*). For women with early stage EOC surgery has also been proven to be best suggested course of treatment since it helps in indicating subsequent treatment options (*133*). Surgery options include: hysterectomy (removal of the uterus), and bilateral salpingo-oophorectomy (removal of the fallopian tubes and ovaries) (*16*). Some patients cannot undergo surgical treatment due to

the presence of comorbidities that impede surgical intervention, and these patients are usually subjected to neoadjuvant chemotherapy (NACT) (*134*). In EOC, NACT is used to reduce EOC tumor burden in order to better enable the required subsequent surgical treatment (*135, 136*). The first type of chemotherapy treatment in EOC was dependent on the intravenous (IV) delivery of an anthracyclin known as doxorubicin, which was also administered in combination with a known DNA cross-linking agent named cyclophosphamide (*135*). Furthermore, in the 80s, cisplatin, a platinum-based alkylating agent, which works by inducing breaks in DNA, showed good response rates in cancer treatment (*137, 138*). Currently, combination of cisplatin or carboplatin and a taxane such as taxol or docetaxel, are becoming the standard chemotherapy treatment of advanced EOC (*139*). Taxane derivatives such as taxol and docetaxel work by disrupting microtubule dynamics by permanently binding to β- tubulin subunits in dividing cells, which have been shown to lead to mitotic arrest and cell death (*140*).

Moreover, targeted therapies are now becoming more interesting options for cancer treatments; examples of targeted therapies in EOC include HER2/Neu protein amplification therapy, and therapeutics for anti-epidermal growth factor receptor (*EGFR*) used as anti-*EGFR* for targeting protein amplification (*141, 142*). Other targeted therapies include the use of poly ADP ribose polymerase (PARP) inhibitors, as the Food and Drug Administration (FDA) approved the PARP inhibitor olaparib for use as a maintenance regimen for patients with advanced and/or relapsed EOCs (*143, 144*). More trials are currently investigating the use of PARP inhibitors alone or in combination with chemotherapy, and initial results show promising improvements in the progression-free survival (PFS) of patients (*143, 144*). Olaparib has been approved for EOC patients with a *BRCA* mutation who have received three chemotherapy regimens prior to the use of olaparib (*143, 144*). Additional trials are now in progress to investigate the application of other PARP inhibitors as treatment and maintenance option for EOC patients. With all the proposed methods of EOC treatment, unfortunately most women diagnosed with advanced EOC show relapse or recurrence of the disease (*145*). Patient relapse has been shown to be dependent on the staging, the histologic subtype, and residual presence of the cancer (*145*). Most of the women displaying recurrence are diagnosed through the detection of observed increases in CA-125 in their serum in addition to the application of current sophisticated

imaging technologies for tumor mass detection (*146, 147*).

1.1.10 Experimental models of EOC

Scientific research is focused on modeling EOC biology, and several experimental models have evolved in the past decade. These experimental models have opened a door for advancing EOC translational research. A list of the current experimental models is presented in Figure 1.4 (*57*). Research efforts have shown great promise in developing *in vitro* and *in vivo* models that can better describe the possible treatment methods for EOCs (*57*).

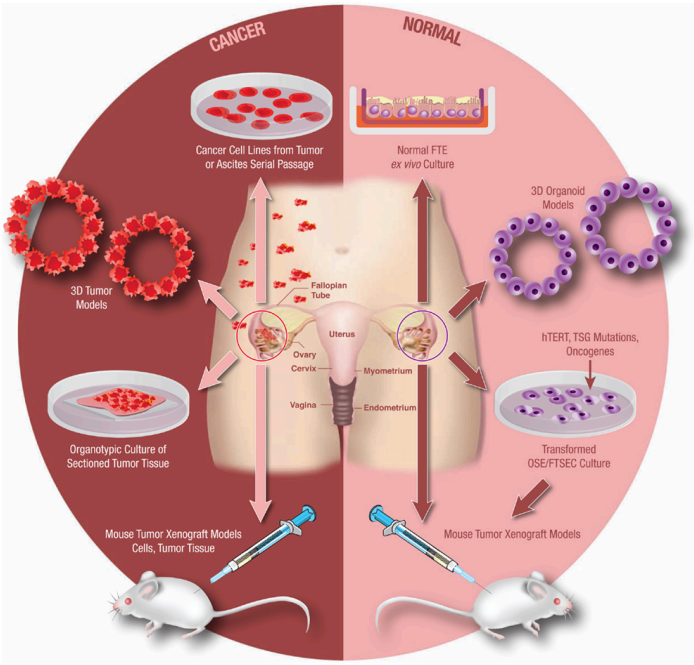


Figure 1.4. Model systems from primary human tissues
A set of the experimental model systems that include both *in vitro* and *in vivo* approaches have been developed over the past few years to study pathways of EOC. The models include approaches that aim at understanding the biology of EOC cancer cells, as well as the study of benign ovarian tissue. Advances have also been developed to expand on the 2D cell culture models into 3D and organoid cell cultures from both tissues and ascites fluid, these approaches have showed great promise and provided important insights into the biology of EOC. The development of unique xenograft animal models has also shown great promise in the field of studying EOC. Collectively these approaches have allowed for the evolvement of our understanding of major molecular pathways of EOC (*57*).

1.1.10.1 EOC cell lines

EOC cell lines represent the most common model in EOC research today, and working with *in vitro* cancer cell cultivation has allowed for an improved understanding of the biology of the cancer and the cancer cell molecular and biochemical features. Scientists were able to develop cell lines representative of the different types of EOC, but no single cell line can be used as a global representation of all EOCs. Table 1.4 summarizes a list of these ovarian cell lines and their proposed tissue of origin.

Table 1.4. List of EOC cell lines

Histologic Subtype	Molecular Features	Cell Lines
HGSC	Near universal somatic *TP53* mutations, high-frequency *BRCA1* and *BRCA2* alterations, extensive genomic instability, common *MYC* and *CCNE1* amplifications	CAOV3*, CAOV4*, COV318*, COV362*, COV504*, JHOS2, JHOS4, Kuramochi*, OAW28*, OVCAR3*, OVCAR4*, OVCAR5, OVKATE, OVSAHO, PEA1, PEA2, PEO1, PEO4, PEO14, PEO23, SNU119
EC	Few *TP53*, *BRCA1*, *BRCA2* mutations, frequent *ARID1A* mutations, common *PTEN*, *PIK3CA*, and *CTNNB1* mutations or loss	A2780*, SKOV3*, TOV112D*
CCC	Few *TP53*, *BRCA1*, *BRCA2* mutations, frequent *ARID1A*, *PIK3CA*, and *PTEN* mutations or loss, high expression of *HNF1B*	2008, JHOC-5, JHOC-7, JHOC-9, OVMANA*, OVTOKO*, RMG-2
MC	Some *TP53* mutations, few *BRCA1*, *BRCA2* mutations, frequent *KRAS* mutations, some *ERBB2* amplification	COV644, MCAS*
Other	Mixed features of more than one subtype, features precluding classification, or conflicting classification by referenced sources	ES2*, IGROV1*, OV90*, OVCAR8*, TOV21G*

The table lists the different cell lines originating from the five subtypes of ovarian cancer, including a description of their distinctive molecular features (*16*).

These EOC cells are grown in culture dishes in a single monolayer, and these cells have shown to spread well across the cell culture plate (Figure 1.4). Recently, other methods of culturing cancer cells have been developed, and one such system is the use of three-dimensional (3D) cell culture systems, which seem to better reflect the tumor microenvironment (*148*). The interesting property of growing EOC cell in 3D cultures is

that growth in spheroids has been shown better mimic the histological morphology of tumors from which the cells were originally derived (*148*). Additionally, 3D culture accurately mimics cellular characteristics such as cellular proliferation, drug response, cell invasion, and phenotypic heterogeneity found within the EOC tumor (*148*) (Figure 1.4).

1.1.10.2 Genetically engineered animal models

Genetically engineered mouse models (GEMMs) represent another model system extensively used in EOC research, and these models have been suggestive to better replicate the morphological, and the biological features of the different EOC histotypes (*57*). GEMMs have not only been used for understanding tumor biology in a research laboratory setting, but have also been shown to be useful for preclinical testing and development of EOC treatment strategies (*57*). In order to develop a certain GEMM that shows characteristics of a specified histotype of EOC, it is important to modify the list of genes that have been characterised to be specifically altered in a particular EOC subtype (*57*). Modulation of the genes listed in Table 1.2 is usually the common approache to best replicate the distinctive features of the different EOC subtypes (*57*). The process of activating or inactivating these selected genes of interest can be accomplished by gene modifying techniques such as the use of viral proteins such as short hairpin RNA (shRNA), in addition to the more recently developed genome targeted approaches that target DNA binding proteins such as transcription activator-like effector nuclease (TALENs), and the recent development of the clustered regularly interspaced short palindromic repeats (CRISPR)/Cas genome editing systems (*149*). TALENs work by using specific restriction enzymes that target DNA sequences through the binding of transcription activator-like (TAL) effector through their central domain responsible for DNA binding targeting to specified DNA sequences (*149*). Once bound, artificial nucleases bind to the DNA target site and induce double strand breaks (DSB) followed by the activation of non-homologous end joining (NHEJ) in the cell (*149*). Similar to the TALEN system, the clustered regularly interspaced palindromic repeats (CRISPR)/associated protein 9 (Cas9) gene-editing system was developed and is proven to be less expensive than use of TALENs for gene editing, but is also preferred for its capability to perform site-specific gene mutations with minimal off target effects (149). One of the more reliable strategies in which the CRISPR system works

is dependent on the Cre-lox technology, which works by the use of two major biological molecules - Cas9 and guide RNA (gRNA) in addition to the Cre-Lox system (*150*). In this system, a single gRNA recognizes a 20 nucleotide target sequence that is found adjacent to a 5′ NGG 3′ protospacer adjacent motif (PAM), then the Cas9 then introduces a DSB in the specified target sequence (*151*). A cassette that includes the Cre recombinase is then inserted into the gene of interest, the Cre-Lox system is a site-specific recombination system known to excise bacterial artificial chromosome (BAC) (*152*). The Cre recombinase enzyme, originally derived from the P1 bacteriophage recognizes specific DNA sequences called the Lox sites where the DNA found between the two Lox sites can then be excised (*153*). The complex formation of these macromolecules allows for the highly selective targeting of specific DNA sequences (149), which will subsequently cause DNA damage eliciting the activity of NHEJ, resulting in the excision of DNA inactivating the desired gene target (149). These genetic engineering technologies utilized in the production of GEMMs make use of two types of cell lines either the ovarian surface epithelial cells (OSE), and the use of fallopian tube epithelial cells (FTE) since as previously mentioned FTEs epithelial cells are now being suggested to be the origin of HGSCs (*16*).

1.1.10.3 Xenograft models of EOC

GEMMs have provided good systems to examine gene functions and disease progression of a particular subtype of EOC; indeed, mouse xenograft models are also one of the most utilized experimental systems used in the field of cancer research (*154*). Xenograft models were shown to be effective when using immunocompromised or immunodeficient mice, which thus far has shown to provide an extensive research tool in the field of EOC research (*154*). Research in EOC makes good use of xenografts since it appeared to be an excellent approach to characterize HGSC events especially when paired with *in vitro* studies that aim at characterizing metastatic transformations upon genetically modifying different genes playing essential role in EOC progression and metastasis (*155*). The biggest challenge facing the application of xenograft models in EOC is the proper examination of the tumor formation and how it best reflects the histological type of tumor formation, which has appeared to be largely dependent on cancer cell injection site (*156*). Therefore many groups have focused on the application of proper models to best reflect the different types of tumor

development of EOC including **1)** subcutaneous (SC) **2)** intraperitoneal (IP) and **3)** orthotopic models (*156*). The most described types of EOC xenograft models involve tumor cells injected either subcutaneously or intraperitoneally (*157-159*).

There are differnet models of ovarian carcinoma metastasis occurring by either direct extension from the ovarian or fallopian tumor to neighboring organs or by direct detachment of cancer cells from the primary tumor site (*77*). Interestingly, the most highly reported method of implantation of ovarian cancer has been examined in EOC, where EOC spreads by IP sloughing, lymphatic invasion, and hematogenous dissemination (*160*). IP dissemination is the most common type of dissemination in EOC; since after malignant cells have evaded from the ovarian capsule, they are shed from the tumor surface into the peritoneal cavity where they follow normal routes of peritoneal fluid (*161*). Therefore it has been suggested that IP injection of cancer cells in animal models can accurately model advanced disease, as EOC implantation frequently appear disseminated throughout the peritoneum (*162, 163*).

1.2 Glycosylation and cancer

1.2.1 Role of glycosyaltion in cancer development

Cancer cells have been characterized to undergo activation, rapid growth, adherence and invasion of neighboring cells/tissues, and these processes are also linked to changes in the cell's glycosylation profiles (*164*). Studies have provided strong evidence for the existence of important links between abnormal glycosylation and tumorigenesis. Glycosylation changes characterized in cancer cells follow a variety of forms, as glycan alterations can be associated with loss or gain of expression, depending on the cell type and the specific glycan's structure. Aberrant glycosylation in cancer cells could affect certain ligand-receptor interactions and more importantly, could favor cancer cell proliferation, migration and invasion/metastasis. Furthermore, aberrant glycosylation of glycan structures is now a new hallmark common of neoplastic transformation. Determining glycan structures is an essential step towards establishing a more comprehensive understanding of their roles in normal and pathological conditions, including cancer.

1.2.2 The chemistry and biology of protein glycosylation

Glycosylation represents the most complex form of all proteins' post-translational modifications (PTMs), and is defined by the regulated process of adding lipids and carbohydrates to proteins, as this regulation depends on various building enzymes and proteases that allow for the diversification of protein function (*165*). Many studies have advanced our understanding of the major roles of glycans, but no evidence has yet comprehensively explained their role or function under different normal and pathological conditions, which is mainly due to their enormous diversity. Glycans have been shown to play crucial roles in the development, the growth and the functionality of organisms (*166*). Cell surface glycans are mainly implicated in recognizing molecules that aid in the process of communication and adhesion (*167*). Analysis of the SWISS-PROT database showed that the majority of proteins of the sequon (which is a sequence of successive amino acids that act as sites of attachment site for polysaccharides) are glycosylated, as more than 50% of all proteins are characterized as glycoproteins (*168*). Glycosylation of proteins is relatively rare in prokaryotes and quite common in eukaryotes. Three types of glycosylation of proteins are currently known, including C-linked glycosylation, N-linked glycosylation, and O-linked glycosylation. However, C-glycosylation, representing mannosylation of C2 of the indole ring of tryptophan residues is quite rare (*169*), as the major glycosylation types represent N-linked and O-linked (mucin) glycosylation (*165*).

1.2.2.1 N-linked glycosylation

N-linked-(or asparagine-type) glycosylation represents a fundamental and extensive PTM that results in the covalent attachment of an oligosaccharide onto asparagine residues of the polypeptide chains (*170*). The N-linked glycosylation of glycoproteins takes place in both prokaryotic and eukaryotic cells (*170*). Through the attachment of glycans at asparagine residues, N-linked glycosylation promotes protein folding by enhancing solubility and mediating interactions between nascent proteins and cellular proteins, such as chaperones (*170*). Three key processes in N-linked glycosylation have been identified: a) the lipid-mediated assembly of monosaccharides into glycans, which is performed by various enzymes in the endoplasmic reticulum (ER) - Golgi complex; b) the acceptance of a glycan

by the consensus sequence Asn-X-Ser/Thr (X for any amino acid except proline), and c) the oligosaccharyltransferase-catalyzed attachment of the glycan to the side chain of the asparagine residue (*170*). The initiation step of N-glycan synthesis starts with the transfer of GlcNAc-P from UDP-GlcNAc to membrane-bound Dol-P to form GlcNAc-P-P-Dol, and this process is catalyzed by the enzyme GlcNAc-1-phosphotransferase (*171*). See Figure 1.5 for an illustration of N-type glycosylated proteins.

1.2.2.2 O-linked (mucin) glycosylation

O-glycosylation of proteins represents the most diverse PTM form (*167*). There are several O-glycosylation types, and they have been characterized and differentiated based on their biosynthesis and linkage processes. These include: O-GalNAc or mucin-type, O-mannose, O-xylose, O-fucose, O-glucose, O-galactose and O-GlcNAc (*167*). The initiation process of O-glycosylation, except for O-GalNAc, O-Xyl and O-GlcNAc, occurs in the ER (*167*). O-GalNAc initiation of proteins takes place in the Golgi (*167*) and the O-GlcNAc glycosylation takes place in the cytosol and the nucleus (*172*). O-linked glycosylation mostly happens on serine and threonine residues (*173*), although as recently demonstrated, O-glycosylation can also be found on tyrosine residues (*174*). The initiation step of O-glycosylation is more complex than N-glycosylation and involves the transfer of different monosaccharides to each of the six O-linked glycans (*167*). O-GalNAc glycosylation is more distinct as it involves two steps including the initiation step, which is directed by 20 GalNAc-transferases, a family of enzymes known as the UDP N-acetylgalactosamine: polypeptide N-acetyl galactosaminyl transferases (GalNAc-Ts) (*175*). Thus, the O-GalNAc initiation step involves the transfer of the monosaccharide GalNAc from UDP-GalNAc to the hydroxyl group of the serine, threonine or tyrosine residues found in the target protein substrate, which is followed by a processing step, where 30 or more glycosyltransferases create distinct glycan structures (*175, 176*). See Figure 1.5 for an illustration of O-type glycoproteins. As described above, cellular glycosylation mechanisms and their biosynthetic pathways are very complex and have shown to be fundamental to the changes in glycan processing and divergence. Numerous studies have shown that alterations in surface glycans are linked to disease progression and most importantly in cancer, as these alterations play pivotal roles in cancer initiation and progression (*177, 178*).

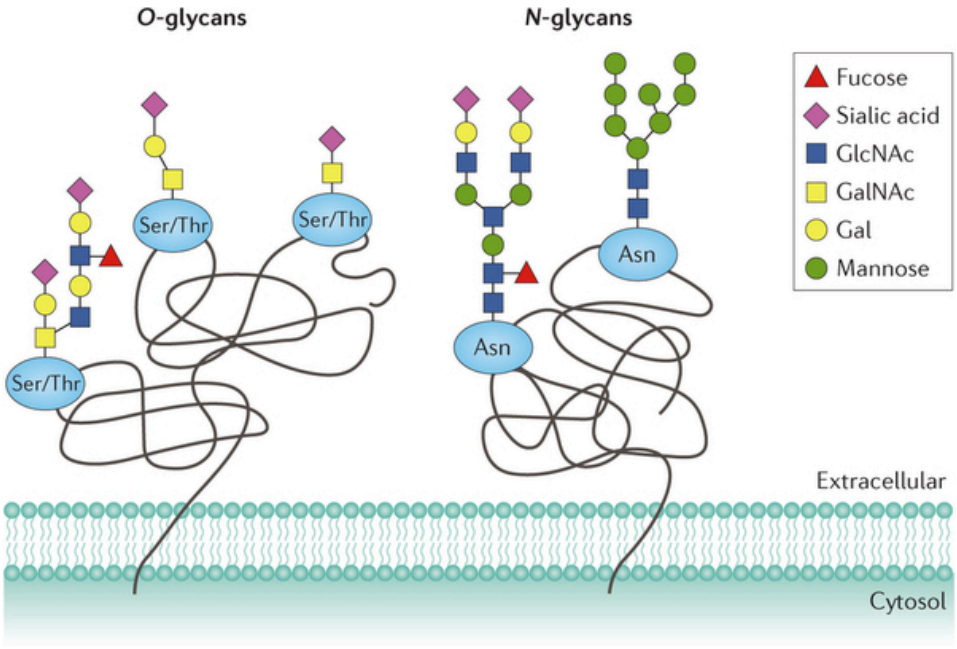


Figure 1.5. N- and O-glycosylated proteins
Proteins that are N-glycosylated are glycosylated by N-linkage to asparagine residues, while O-glycosylated proteins are glycosylated by O-linkage to serine and/or threonine residues. O-glycans can be extended to produce different core and terminal structures that can be sialylated and/or fucosylated. N-glycans contain a common pentasaccharide core region consisting of three mannose and two N-acetyl-glucosamine (GlcNAc) subunits attached to an asparagine residue that can be further modified by the addition of terminal GlcNAc, galactose and sialic acid moieties (*179*).

1.2.2.3 Synthesis of O-glycan structures

There exists hundreds of different O-glycans structures in mucin type glycans, and these O-glycans are characterized by their structural variation, which is in part due to the fact that in mammalian mucins, there exists a minimum of eight different O-glycan core structures (Table 1.5) (*180*). The simplest form of O-glycans is characterized by a single N-acetylgalactosamine (O-GalNAc) residue that is found linked to a serine or threonine (*181*). This core structure is referred to as the Tn antigen, and is the precursor structure to all O-glycans (*181*). The Tn antigen acts as the building block for all other core structures by the help of the two genes C1GALT1 and C1GALT1C1 coding for the transferase β3-Gal-transferase responsible for the conversion of the Tn antigen to the core 1 structure (*180,*

182) (Figure 1.6) and (Table 1.5). The most common type of O-glycan is the core 1 O-GalNAc glycan also known as the T antigen which gets elongated to form more complex core structures by the action of the enzymes GCNT1, -3 or -4 and ST3GAL1 (*182*) (*180*). The two antigenic cores (Tn antigen and T antigen) get modified with sialic acid allowing for the formation of the known structures sialylated Tn or T antigens (*180*). The second core structure is the core 2 structure and contains an N-acetylglucosamine branching attached to the core 1 structure, and the transferase responsible for this conversion is the β6-GlcNAc-transferase (Table 1.5) and (Figure 1.6) (*180*). Core 2 structures are not only common to mucin type glycoproteins but are also found in other types of glycoproteins (*180*). Most importantly the other core structures (core 3 and 4) are only found in secreted type mucin glycoproteins (*180*). The core 3 structure is synthesized by core 3 β3-GlcNAc-transferase (B3GNT6), and its activity is not highly detectable in several cell types (*180*), while the core 4 structure is synthesized by the core 4 β6-GlcNAc-transferase, similar to that responsible for the core 2 structure synthesis (*183*) (Figure 1.6) and (Table 1.5). Core structures 5-8 (Table 1.5), are considered to be the less common glycan type, and are confined to certain tissue types, thus limited information on the synthesis of O-glycan core structures 5-8 (*181*). Essentially, the resulting core structure depends on the relative activity of the glycosyltransferase responsible for their core elongation.

Table 1.5. Structures of O-glycan cores and antigenic epitopes found in mucins.
Adapted with permission from (*184*).

O-Glycan	**Structure**
Core	
Tn antigen	GalNAcαSer/Thr
Sialyl-Tn antigen	Siaα2-6GalNAcαSer/Thr
Core 1 or T antigen	Galβ1-3GalNAcαSer/Thr
Core 2	GlcNAcβ1-6(Galβ1-3) GalNAcαSer/Thr

O-Glycan	**Structure**
Core 3	GlcNAcβ1-3GalNAcαSer/Thr
Core 4	GlcNAcβ1-6(GlcNAcβ1-3) GalNAcαSer/Thr
Core 5	GalNAcα1-3GalNAcαSer/Thr
Core 6	GlcNAcβ1-6GalNAcαSer/Thr
Core 7	GalNAcα1-6GalNAcαSer/Thr
Core 8	Galα1-3GalNAcαSer/Thr

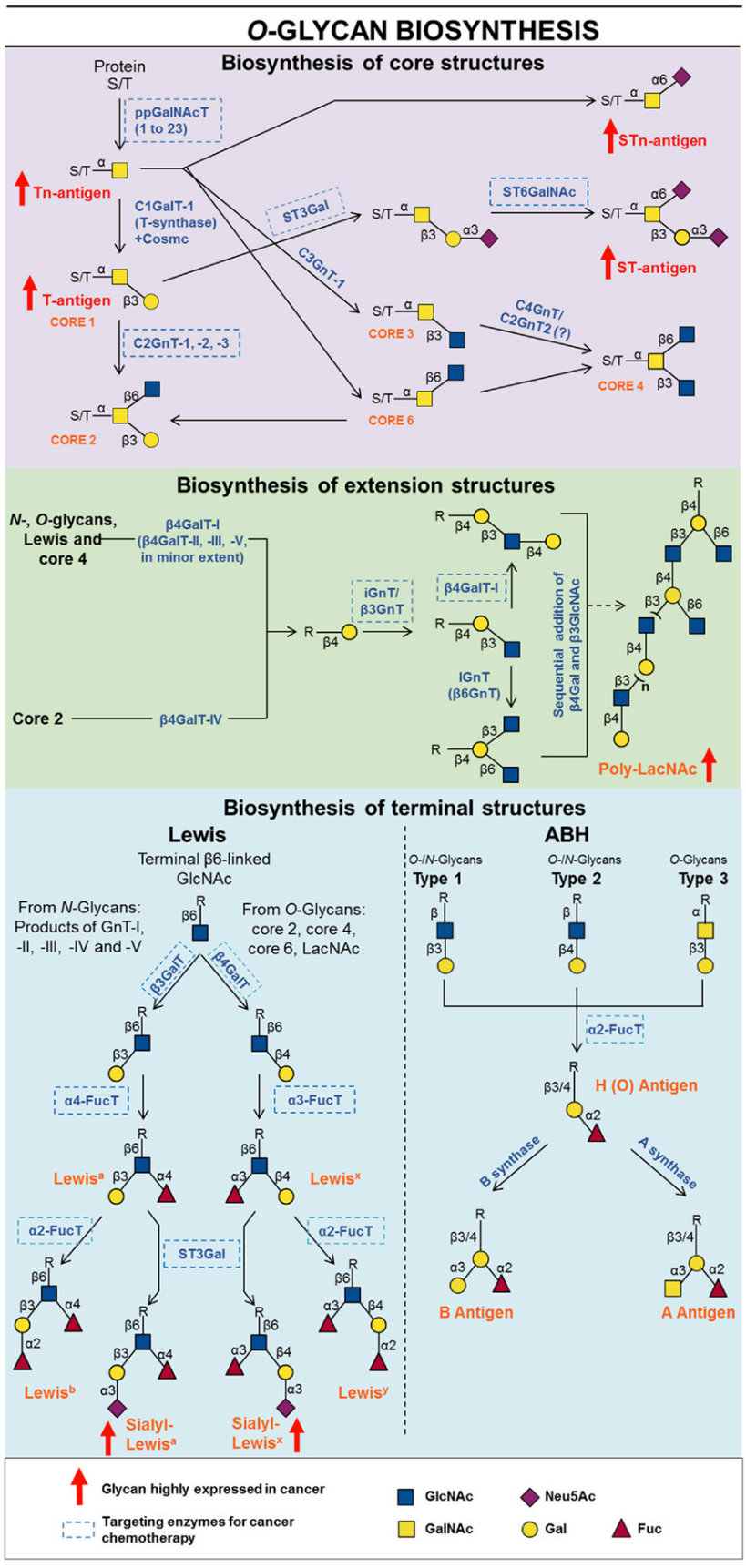

O-GLYCAN BIOSYNTHESIS
Biosynthesis of core structures
Protein
S/T
ppGalNAcT
(1 to 23)
Tn-antigen
C1GalT-1
(T-synthase)
+Cosmc
T-antigen
CORE 1
C2GnT-1, -2, -3
CORE 2
ST3Gal
C3GnT-1
CORE 3
CORE 6
ST6GalNAc
STn-antigen
ST-antigen
C4GnT/
C2GnT2 (?)
CORE 4
Biosynthesis of extension structures
N-, O-glycans,
Lewis and
core 4
β4GalT-I
(β4GalT-II, -III, -V,
in minor extent)
Core 2
β4GalT-IV
iGnT/
β3GnT
β4GalT-I
IGnT
(β6GnT)
Sequential addition of
β4Gal and β3GlcNAc
Poly-LacNAc
Biosynthesis of terminal structures
Lewis
Terminal β6-linked
GlcNAc
From N-Glycans:
Products of GnT-I,
-II, -III, -IV and -V
From O-Glycans:
core 2, core 4,
core 6, LacNAcs
β3GalT
β4GalT
α4-FucT
α3-FucT
Lewisa
Lewisx
α2-FucT
ST3Gal
Lewisb
Lewisy
Sialyl-
Lewisa
Sialyl-
Lewisx
ABH
O-/N-Glycans
Type 1
O-/N-Glycans
Type 2
O-Glycans
Type 3
α2-FucT
H (O) Antigen
B synthase
A synthase
B Antigen
A Antigen
Glycan highly expressed in cancer
Targeting enzymes for cancer
chemotherapy
GlcNAc
GalNAc
Neu5Ac
Gal
Fuc

Figure 1.6. Schematic representation of the biosynthesis and processing of O-linked glycans
The figure above describes the biosynthetic pathways of O-glycosyaltion. The top panel gives a detailed description of the biosynthesis of the main core structures (Tn-antigen, T-antigen, sTn antigen, and sT-antigen). The middle panel contains a description of the bioynsthesis of the extension strcutures for all of N-, and O-glycans, in addition to Lewis and Core 4 structures. Finally the bottom panel describes the pathway of the biosynathesis of terminal structures for all of of N-, and O-glycans as well as the ABH blood group antigens (*185*).

1.2.3 Terminal glycan linkages/modifications and their role in cancer development

The elongation of the major core structures described thus fur are what characterize them as separate glycans found on the peptide backbone, but additionally these core structures can be modified by the addition of terminal sugar groups that can dictate the biological role and overall properties of these glycans and essentially the glycoproteins with these modifications. The modifications that these glycans are known for include: **1)** sialylation, **2)** fucosylation, **3)** blood group and **4)** Lewis antigens (*186*). Interestingly, glycan alterations have also been suggested to play important role in cancer progression through aberrantly regulating tumor cell proliferation, migration, invasion, cellular attachment, and angiogenesis (*187*). Figure 1.5 demonstrates the different pathways involved in the biosynthesis and processing of O-linked oligosaccharides.

1.2.3.1 Lewis antigens and blood group antigens

Lewis antigens and blood group (ABH) antigens are fucosylated terminal saccharide structures found attached to both O- and N-linked glycans (Figure 1.6) (*188*). They are referred to as blood group antigens since they were initially observed on red blood cells, but have now also been discovered to be present on other types of cells (*189*). The three known types of Lewis and ABH antigens are: a) type 1, Galβl-3GlcNAcβ; b) type 2, Galβl-4GlcNAcβ; and c) type 3, Galβl-3GalNAcα (*190*). Out of the three types of Lewis structures, type 3 represent core 1 structure of O-glycans and has been found to be exclusively attached to O-glycans (*190*). Changes such as loss or gain of expression in any

of these antigens have shown negative prognosis in several cancer types such as lung cancer, adneocarcinomas, and prostate cancer (*191-193*). Moreover, a study examining Lewis structures in EOC showed that serous and mucinous benign adenomas and LMPs have elevated levels of blood groups and Lewis antigens compared to HGSCs (*194*). Lewis antigens are derivatives of the type 1 and 2 precursors mentioned earlier. Lewis antigens are synthesized by the addition of αFuc to GlcNAc and are thus known as Le^a or Le^x antigens (*195*), while αFuc structures found attached to H types 1 and 2 antigens give rise to the Le^b and Le^y antigens (*195*). The most common observed Lewis structure found to be upregulated in cancers such as stomach, breast, lung, colon and EOC is the Le^Y antigen (*196-198*). This antigen has been suggested to be a good potential target in all these cancer types for it is found to be highly elevated compared to normal cells or tissues (*198*). Interestingly, cancer vaccines and immune-conjugated chemotherapy against the Le^Y show great promise in clinical studies (*199-201*). Lewis antigens can also be modified by the addition of a sialic acid to the Le^a and Le^x , referred to as sialyl-Le^a (SLe^a) and sialyl-Le^x (SLe^x) antigen (*202*). Similar to the Le^Y antigen, elevated expression of the SLe^a antigen seems to be a good marker for examining the metastatic potential of several cancer types such as colorectal, and pancreatic cancers (*203, 204*). In addition, studies have also analyzed the overexpression of the SLe^x antigen and its correlation with the aggressiveness and poor prognosis in prostate and lung cancers (*205, 206*). Several cancers have also shown to exhibit an overexpression of both types of antigens SLe^a and SLe^x , and their expression was correlated with the level and degree of malignancy of that cancer type (*207*). It has been suggested that the observed overexpression of these Lewis antigens maybe in part due to the upregulation of the genes coding for these enzymes (*208*).

1.2.3.2 Sialylation

All of the core structures listed in Table 1.5 can be sialylated, as sialylation occurs through the transfer of a sialic acid by a specific subset of sialyltransferases, and these transferases are cell type specific and also dependent on the cellular differentiation state (*209*). The addition of sialic acid to glycans adds a negative charged residue to these terminal structures, which effectively helps altering their roles in various cell surface properties (*210*). The first core structure (Tn antigen) can be sialylated by a α2, 6-sialyl transferase

(ST6GalNAc1), which transforms it to a Sialyl-Tn (sTn) (sTn-Neu5Acα2- 6GalNAcα-O-Ser/Thr) disaccharide (*181*) (see Table 1.5 and Figure 1.5). Core 1 or the T antigen can also be sialylated to form a sialyl-T (sTn) trisaccharide antigen, as the sialic acid gets linked to the core 1 structure by the action of the ST3Gal1 transferase to form an alpha-2,3 linkage NeuAcα2-6 (Galβ1-3) GalNAc-O-Ser/Thr to the core structure (*210, 211*) (Figure 1.6). While transferases responsible for the sialylation of the other core structure have not been as extensively examined, STn is the simplest but most examined type of sialylated mucin-type O-glycan (*210*). Sialyltransferases can compete with other elongating glycosyltransferases, which in turn prevents proper elongation of glycan structures on glycoproteins, and thus they act as terminal modifications for these glycans (*212*). With this type of competition, it is often evident that there are hypoglycosylated structures giving rise to more STn and Tn structures, which are usually observed in cancer cells of different origin and absent in normal cells or tissues as hypoglycosylated forms (*212*). These short and sialylated glycan structures were found to be expressed in more than 80% of human epithelial carcinomas, while absent or not as highly expressed in other cancer types (*182, 212*). Their expression has been linked with worse outcome and lower overall survival of epithelial cancer patients, which is why they were later on referred to as pan-carcinoma antigens (*182, 212*). Moreover, the presence of STn modifications on protein glycans was frequently found in patients' serum and tumor samples, and their overexpression has been linked with poor survival of women with ovarian and breast cancers (*213-217*). Elevated expression of sialylated antigens plays a role in tumorigenesis by inhibiting cellular adhesion of cancer cells through the electrostatic repulsion of negative charges (*218*), and thus promoting cellular detachment from the tumor microenvironment (*217*). Recent reports have also demonstrated that the glycol mucins MUC16 and MUC1 carry sialyl Lewis^{a} (SLe^{a}), and sialyl Lewis^{x} (SLe^{x}) (*219*) in both borderline and HGSCs, suggesting that these modified Lewis structures can be used as biomarkers for EOC (*219*). It is believed that the high expression of the sTn antigens in cancer is in part due to an increase in the expression of the sialyl transferase ST6GalNAc1, allowing for more terminal sialic acids vs core structure synthesis (*68, 210*). Another hypothesis is that this phenomenon is not only due to the high activity of the ST6GalNAc1 transferase, but may additionally be

due to the loss of activity of the main core transferase C1GALT1 (Core 1 transferase), or a lower activity of the B3GNT6 (Core 3 transferase).

1.2.3.3 Fucosylation

Glycan structures can be additionally modified by a fucose saccharide, and there exists two types of fucosylated modifications (*220*). The first is described as the addition of a fucose to the terminal glycan and is thus referred to as terminal fucosylation giving rise to the blood group antigens described earlier. The second type of fucosylated modification is characterized by the addition of a fucose to the core glycan structures, and is thus known as core fucosylation, which like sialylation terminates glycan elongation (*220*). Elevated expressions of fucose modifications have been detected in the serum of prostate cancer patients (*221*). The observed upregulation in fucosylated glycans has been suggested to be due to the overexpression of fucosyltransferases, such as FUT8, 3, 6 and 7 (*182*).

1.3 The chemistry and biology of mucin glycoproteins and their role in cancer

1.3.1 The two classes of mucins

Mucins are a family of heavily glycosylated proteins, and are characterized by what is referred to as the mucin domain known as variable number of tandem repeats (VNTR) (*222, 223*), which is the unique characteristic region that allows to differentiate mucins from other glycoprotein types (Figure 1.7). Human MUC family includes 21 different members named from MUC1 to MUC21 (*224*). Mucins encompass the largest portion of the protein component found in the mucus (*225*). Mucus is a hydrophilic mixture that creates the layer that is known to cover or protect the epithelial lining of the human body, and it acts as a barrier that helps in protecting the body from invading pathogens (*226*). Differential regulation of this mixture is involved in the development of various diseases such as cystic fibrosis (CF), asthma, and cancer (*227-229*). Mucins can be divided into two subfamilies, dependent on their location in relation to the cellular surface. These are: a) secreted mucins, which are the extracellular type mucins that lack a transmembrane

domain, and b) mucins that are found on the cell surface and located either on the transmembrane or tethered to the membrane (Figure 1.7), The later mucins are characterized by having a transmembrane domain (TM), a COOH terminal cytoplasmic tail (CT), and an NH_2 extracellular domain (Figure 1.7 and 1.8). Secreted type mucins include MUC2, MUC5AC, MUC5B, MUC6-8 and MUC19, and the transmembrane mucins include MUC1, MUC3, MUC4, MUC12-17 and MUC20 (Figure 1.8).

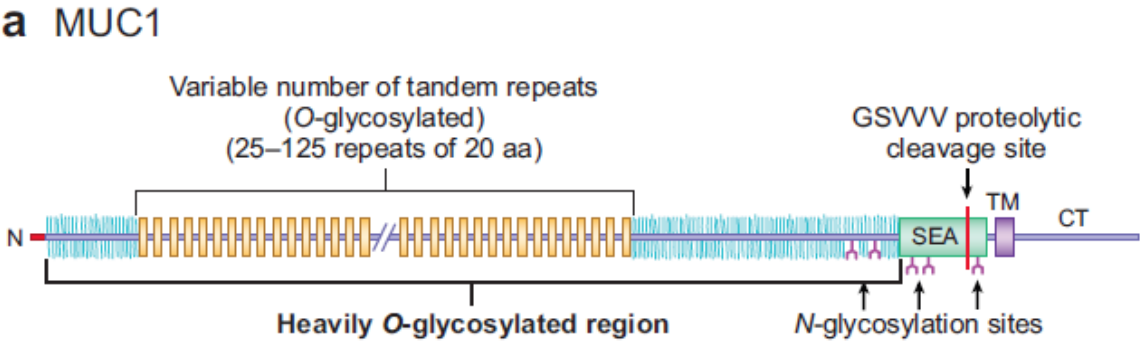


Figure 1.7. Structure of the major transmembrane protein Mucin1
Mucin 1 consists of key domains that include the N-terminal signal sequence shown in red; sperm protein, enterokinase, and agrin (SEA) sequence; transmembrane (TM) domains; cytoplasmic tail (CT). The figure also illustrates the *N*-glycosylation sites and
Variable tandem repeat region that is heavily glycosylated in addition to the GSVVV cleavage site. Adapted with permission from (*230*).

Studies of mucins structures were complicated due to the extensive type of PTMs found on their surface, but recent advances on proteomic analysis have allowed for their better structural identification (*230*). Mucins have a protein backbone known as ‘‘apomucin’’, and this backbone is most often covered with O-linked oligosaccharides and N-glycan chain structures (Figure 1.7 and 1.8) (*230*). Mucin glycoproteins undergo various forms of PTMs such as glycosylation, sialylation and sulfation, and this occurs in a cell-type specific manner (*230*). As seen from Figure 1.8, mucins VNTR, contain serine and threonine residues representing the known O-glycosylation sites (Figure 1.8). VNTR also contain proline residues which were shown to play a role in allowing for protein conformational changes, as well as for packing the carbohydrate structure found on the mucin backbone (*230*). Importantly, the sequence and number of these tandem repeats is highly variable

among the different mucin types, and these tandem repeats provide the presence of a wide array of oligosaccharide structures. This variability has allowed for the large degree of what is referred to as stoichiometry in mucin type glycoproteins (Figure 1.8) (*231*). In addition, the glycosylated region of mucins contains glycans which make up more than 80% of a mucin protein molecular mass, as these structures allow the binding of mucin proteins with their protein interacting partners (*230*). Moreover, the non-glycosylated part of the mucin backbone is important for mucins' specific pathophysiological roles both in normal and in diseased cells (*230*). For example, the epidermal growth factor (EGF) domain that is unique to membrane associated mucins (Figure 1.7 and 1.8) allows for the heterodimerization with ErbB receptors, while the CT domain also unique to membrane associated mucins (Figure 1.8), allows mucins to interact with kinases and other cytoskeleton linked proteins (*232*). These interactions allow for signalling between the inside and outside of the cells, and suggest the important role of these mucins in cell survival, proliferation, migration and cell differentiation (*229, 230, 233*). Although there are several structural similarities amongst the different mucins, their primary backbone structure carries unique domains that have been found to be specific to each individual mucin type (*230*). These structures allow for mucins to interact with unique and distinct protein partners found outside or inside the cell, increasing the specificity associated with these various cellular events (*230*).

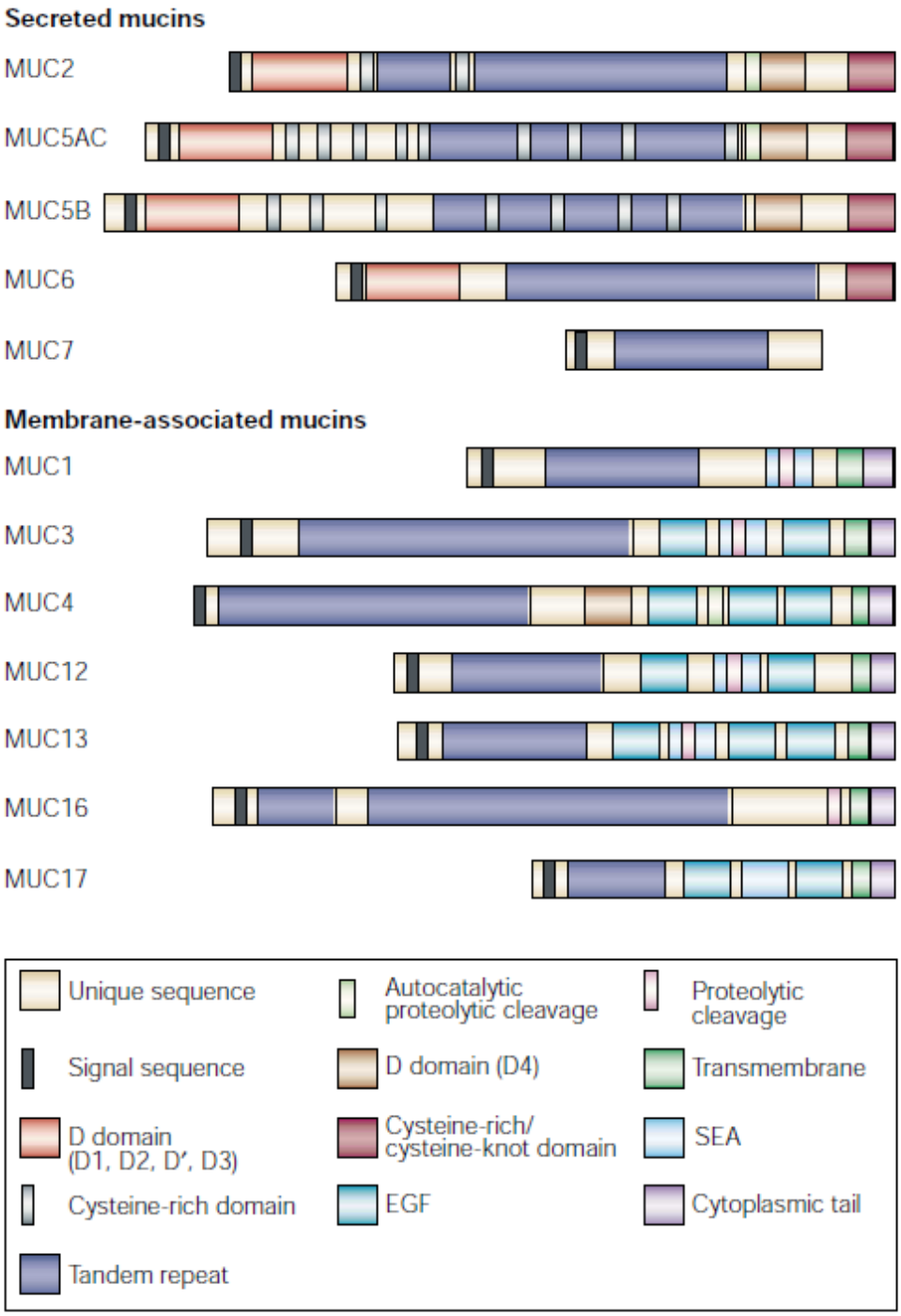


Figure 1.8. The different secreted and membrane bound mucin type proteins.
The list of all secreted mucins, and membrane associated mucins with a detailed depiction of their different domain sequences, Tandem repeat also referred to as variable number of tandem repeats (VNTR) (*229*).

1.3.2 Normal functions of mucins

The different and various structures of mucin proteins strongly suggest that these proteins play very complex roles in normal cells. In normal conditions, mucins production is usually maintained by regulatory mechanisms, and the level of maintenance and regulation of these proteins is rather complex. Regulatiory mechanisms involve signal transduction pathways, transcriptional mechanisms, in addition to epigenetic mechanisms such as (DNA methylation, histone modifications, and microRNA silencing), all suggested to play critical roles in the proper expression of mucins under normal conditions (*234*) (*235*) .

Mucins are expressed by epithelial cells, which are found in systems that are directly exposed to different environmental surfaces. Such examples include the surface of the respiratory system, the stomach acidic layer, the intestinal tracts, and also the epithelial surfaces of the salivary glands, liver, gall bladder, kidney, mammary glands and the female reproductive tract (*225*). Thus, as it appears mucins are found at the point of interaction between the epithelial surface and air, food, pH, bacteria, viruses; and these areas are where the gel-forming mucins generally accumulate (*230*). At the epithelial cell surface the mucin proteins are suggested to act as layers to form a physicochemical protection from the different conditions the organs might be subjected to, such as fluctuations in ionic concentration, pH, hydration, and the surface exposure to proteases, lipases, glycosidases and toxins (*230, 236*). Mucins work by essentially maintaining a proper level of homeostasis, which in turn can promote better cell survival and maintenance, when exposed to extreme conditions (*230, 236*). The role of mucins in forming the proper physicochemical barrier is also associated with the way they act as cellular sensors or receptors, allowing for proper signal transmission within the cell (*230, 236*). This form of communication allows for the proper coordination of functional characterises such as cell differentiation, signal for apoptosis and proliferation (*230, 236*). Additionally, mucin-forming gels have the capacity to capture biologically active molecules such as cytokines, growth factors, differentiation factors and mediators of inflammation (*230, 236*). Captured molecules can act as indicators of changes in the epithelial layer, and when released they act as indicators of cellular distress such as inflammation, and in turn help in promoting the initiation of cellular repair (*230, 236*).

1.3.3 The roles of mucins in cancer development

The aberrant expression of mucins has been linked with various malignancies, and these aberrant forms play essential roles in cancer pathogenesis and progression (*211, 229*). Cancer cells are shown to express aberrant forms of mucins causing deregulation of mucin core proteins and the enzymes that are known to modify them (*237*). For instance, malignant type epithelial cells gradually lose the correct polarity of membrane bound mucin proteins expressed on their surface (*237*). Similarly, secreted mucins in malignant cells have been shown to enter the extracellular space and circulate in body fluids, such as the blood (*237*). The aberrant glycosylation of mucins confers novel combinations and forms of attached oligosaccharide structures (*211*). These changes have been shown to contribute to the number of potential sites on ligands, and thus affecting the survival and invasive capacity of tumor cells (*211*). Several studies, focused on examining human tumor malignancies observed an overexpression of membrane bound or secreted proteins. One very well studied mucin is *MUC1*, found to be involved in breast cancer pathogenesis, and it is found overexpressed in more than 90% of breast carcinomas (*238-242*). MUC1 has been suggested to act as an oncoprotein through its interactions with the EGFR and ErbB2 receptors, which activates signalling pathways such as PI3K-AKT and MEK-ERK, leading to enhanced breast cancer metastasis (*243, 244*). MUC1 is also currently used as a serum marker of breast cancer (*243*), and has also been found to be highly overexpressed in ovarian, pancreatic and gastric cancer (*245-247*). Similarly, other mucins are useful serum markers of cancers such as MUC16 (CA-125), which is the largest transmembrane mucin examined today, and is now being extensively used as a clinical serum marker for proper monitoring of patient response to chemotherapy, detecting recurrence, and for differentiating malignant from benign masses in EOC (*248, 249*). MUC16 acts as a pro-tumorigenic protein by playing a protective role by evading cytotoxic responses by tumor cells from natural killer cells (*250, 251*). CA-125 is not only a serum marker for EOCs, but is also being used as a marker for breast cancer, pelvic masses, clear cell colorectal cancer, endometroid cancer, lung cancer, pancreatic cancer and breast cancer (*126, 252-258*). Another transmembrane mucin is MUC13; although this mucin is not as extensively examined as other members of the mucin family, it has nonetheless been observed in several epithelial cancers including ovarian, gastric, colorectal and pancreatic cancers (*259-*

261). MUC4 is another transmembrane protein known for its role in the pathobiology of cancers such as pancreatic, breast, lung, cervical, ovarian and prostate cancers (*262*). Similar to transmembrane mucins, secreted type mucins also have a role in malignancies because of their dysregulated expression and functions. The most highly examined secreted mucin is MUC5AC, and its overexpression has been linked to inflammation (*263, 264*) and different cancer types including colon, clear-cell renal cell carcinoma, lung, pancreatic, and colorectal cancers (*265-269*). Similarly, MUC5B is known its deregulation in various malignancies such as lung, pancreatic, breast, lung and colorectal cancers (*270-274*). MUC2 expression has been associated with positive outcomes in cancer patients as observed in pancreatic, colorectal and gastric cancers (*275-277*). Collectively, these studies highlight the important role mucins in the development and progression of various malignancies.

1.3.3.1 Role of mucin protein alterations in inducing cellular invasion, migration and metastasis in cancer

Tumor cell invasion and metastasis is a complex process, and is predicted to require loss of contact between the cell surface and the surrounding cell stroma. For metastasis to occur, tumor cells detach from their primary location causing anti-adhesion between the cells in the original site, and then cells invade by passing through normal stromal elements and then form new contact (re-establish adhesion) at the new site to establish a new colony of cells. Mucins are believed to mediate process of adhesion and anti-adhesion in the tumor environment. Several studies strongly suggest that glycan structures and certain of their alterations can have a direct effect on cellular invasion by promoting cells to disseminate and metastasise into distant sites/organs (*278*). The most extensively studied mucin playing a role in inducing tumor invasive capacity is MUC1 (*279, 280*). As previously mentioned, MUC1 is under-glycosylated, which allows for interactions between the MUC1 protein and transmembrane receptors and other components of the extracellular matrix (ECM) (*279*). Some of these interactions are suggested between MUC1 and ICAM-1, which is an adhesion receptor found on the surface of endothelial and peritumoral stromal cells (*281*). Another interaction of MUC1 is with E-selectin, which is another receptor found on the cell

surface of endothelial cells. E-selectin interacts with epithelial cells through its binding with the under-glycosylated form of MUC1. These types of interactions between MUC1/E-selection and ICAM-1 promote increased migratory capacity of tumor cells by allowing for easier interactions between epithelial and endothelial cells, thus enabling cancer cells circulating in the microenvironment to adhere to blood vessel lining and/or to escape from the blood vessels and to promote secondary site invasion and metastasis. Others also suggest for direct interactions between MUC1 and Src (*279, 282*), which is a nonreceptor tyrosine kinase playing a key role in mediating tumor metastatic progression by regulating integrin activation and cytoskeletal functions (*283*). Therefore, interactions between MUC1 and these membrane receptors were shown to promote the invasive capacity in tumor cells (*282*). Thus, these studies support the role MUC1 plays in driving intercellular interactions that promote metastatic spread, as well as intracellular interactions that promote migratory behaviour.

1.3.3.2 Role of mucin protein alterations in inflammation and immune suppression

Mucins play a role in regulating inflammatory and immune responses by mediating interactions with immune cells such as macrophages, natural killer cells and cytotoxic T cells, as these interactions allow for tumor cells to evade the immune system (*237*). Mucins in cancer cells display surface glycan epitopes characterized by glycan elongations beyond the Tn glycan antigen. Such glycan alterations were shown to alter tumor cells susceptibility toward natural killer (NK)- and cytotoxic T lymphocyte-mediated killing (*284*). This type of immune evasion has been shown in breast and pancreatic cancer cells (*284*). Knocking out the Cosmc protein inhibited the type of Tn glycan elongation in these tumor cells, and this suppression has led to increased sensitivity of the tumor cells to NK-mediated and T-lymphocytes cytotoxicity (*285*). Additionally, the susceptibility of pancreatic and breast cancer cells to NK cell mediated cytotoxicity was directly linked to the expression levels of MUC16 and MUC1, suggesting that the upregulation of these mucins is also involved in immune evasion (*285*). Leukocyte and macrophage motility and activation at tumor sites have also been affected by glycan structures and their elongation

type (*286-288*). Protein interactions such as selectins and siglecs have been shown to interact with certain mucins (*289-291*), and these interactions play a role in mediating leukocyte adherence and motility (*292*). This promotes tumor cells to evade the immune system by producing high numbers of soluble mucins, which prevents NK cells or effector cells from approaching the tumor microenvironment (*293*). Although the direct effects of tumor cells immune response and increasing levels of aberrant mucins have not been examined, the data gathered thus far clearly suggest that high concentrations of secreted mucins found in patients with advanced disease can predict their inflammatory and immune responses (*293*). These observations can indeed explain the observed immune suppression that is found frequently in malignancies displaying increased mucin expression.

1.3.3.3 Role of mucins in EOC development (focus on MUC1)

Several tumor-associated mucins play a very important role in EOC metastasis. An exhaustive review of all mucins associated with EOC tumorigenicity, including a discussion summarizing the new findings that suggest for the application of Mucins as targeted therapies for metastatic EOC is found in the Annex section. As previously mentioned, EOC is asymptomatic until advanced stages, thus EOC markers found in human biological fluids like MUC1, could be very useful for the early screening of the disease at the early or later stages. In EOC, the tumor associated antigen (TAA) MUC1 is overexpressed on the cell surface in more than 90% of EOCs (*294, 295*). Several studies have examined MUC1 overexpression in EOC, showing strong support for the use of MUC1 as a biomarker in EOC serous tumors (*296-302*). MUC1 in EOC is found to be aberrantly glycosylated and is characterized for having O-glycan chains that are shorter and containing structural elements not present on normal cells (*295*). These structural features allow MUC1 to have a higher number of revealed epitopes that are usually found hidden in normal cell and/or tissues (*295*). As mentioned, this property does not only induce the protein tumorigenic capacity, but can also be utilized as a feature for antibody or vaccine targeting antigens. An example of such epitopes are the MUC1-associated Tn / STn-epitopes, which are generated by early O-glycosylation termination (*295*). These early glycan terminations have been linked to two types of alterations: **1)** alterations in the ST6GalNAc-I transferase, playing a role in enhancing STn antigen formation (*303*) and **2)**

the possibility of an induced somatic mutation of the core 1 β3-Gal-T-specific molecular chaperone (Cosmc) gene, known to regulate the addition of (Gal) to the first GalNAC (*304*). Importantly, the MUC1-associated Tn / STn-epitopes have been suggested as major targets for immunotherapy and diagnostic imaging in EOC patients. Indeed, Elssen *et al*. (2010) have indicated that these MUC1 epitopes are highly expressed by all histotypes of primary ovarian adenocarcinomas, in addition to the metastatic and EOC precursor lesions, but were not detected by normal ovarian tissue (*295*). Recent work published from the lab has shown that MUC1 maybe overexpressed in EOC and this overexpression may play a role in the progression of the disease (*305*). Most interestingly, a new monoclonal antibody PankoMab-GEX, a glyco-optimised humanised IgG1 antibody, was found to bind with high affinity to a novel carbohydrate-induced conformational epitope (TA-MUC1) and was recently used in a phase I open-label dose-escalation study (*306*). The application of PankoMab-GEX showed excellent preclinical anti-tumor activity, as the clinical benefit of the use of this antibody was observed in a broad variety of primary tumors, but more frequently in EOC patients (*306*).

Overall, these compounding data on the role of mucins helps confirm that these proteins play an essential role in the pathobiology of cancers such as EOC.

1.3.4 The role of aberrant O-glycosylation in cancer development

Aberrant alterations in glycan structures and compositions on mucin type proteins have been found in multiple cancer types (*179*). Indeed, it has been recently suggested that alterations in glycosylation is associated with the acquisition of cancer hallmark capabilities and that altered glycans are playing an essential role in tumor development and progression (*307, 308*) (Figure 1.9). During metastasis, glycans either appear on new sites on glycoproteins such as mucins, or have irregular glycan forms (*307, 308*). These alterations appear to be due to a shift in the glycosylation patterns or incomplete glycosylation of proteins, which results in truncations of glycan structures that are now considered to be hallmarks of cancer metastasis (*307*). Interestingly, the aberrant glycosylation, also referred as ‘‘oncogenic glycosylation" (*308*), is not random but is regulated since distinct subset of glycans that become modified, enriched or decreased on the tumor cell surface, thus

playing a role in either promoting or inhibiting tumor progression (*202, 292*). As previously mentioned, truncated glycan structures such as Tn, sTn, T, ST, or aberrant extension of glycan chains (sLe^x, sLe^a) were shown to be highly expressed/detected in several cancer types such as breast, pancreatic, gastric and colon cancers (*212, 309-315*). The pathways that induce glycan changes and aberrant synthesis of mucin cell surface glycans have been associated with one or more of the following alterations:

1) The under- or overexpression of glycosyltransferases at the level of epigenetic changes of the transferases transcription or post-transcriptional modifications (*316*).
2) Alterations in the glycosidases activity, which are responsible for catalysing the hydrolysis of glycosidic linkages important in N-glycoprotein processing (*164*).
3) Alterations in the expression of glycoconjugate acceptors together with the availability and abundance of the sugar nucleotide donors. Thus for example, the blockade of donor biosynthesis or the observed functional loss of donor-specific transporters has appeared to abolish cellular glycans that contain fucose or sialic acid linkages (*317, 318*).
4) Alterations in the sugar nucleotide transporter activity (*319-321*).
5) Alterations in the function of the Golgi structure where glycosyltransferases are synthesized (*322-324*).

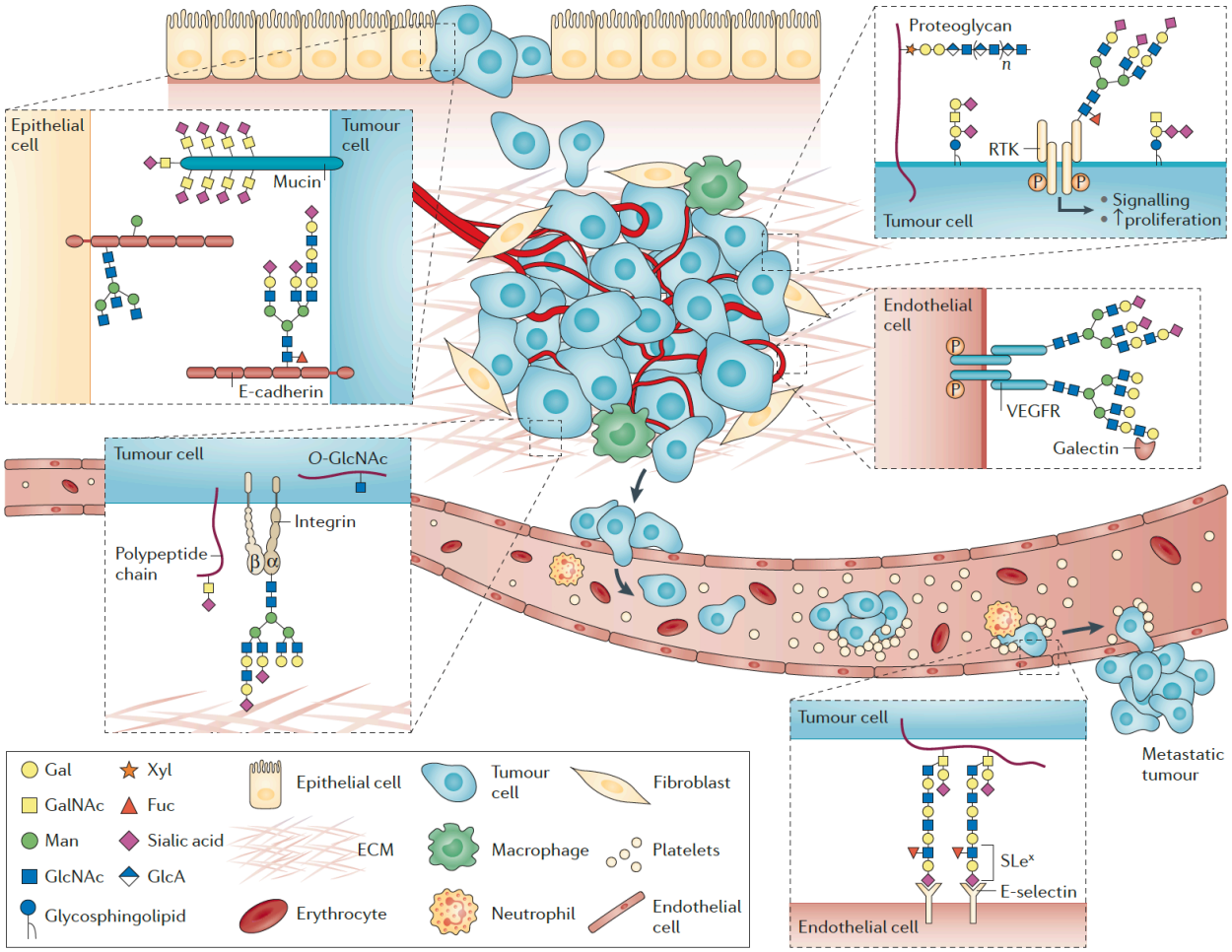


Figure 1.9. Role of glycans in the metastatic cascade

This figure is a representation of the multiple fundamental roles glycans are known to play in the pathology of tumor development and progression. One of the examined processes is the key role of tumor cell detachment and invasion. Additionally, glycans have also been shown to inhibit cell–cell adhesion in many tumor types. Aberrant O-glycosylation of these glycans, such as the expression of the sialyl Tn (STn) antigen has also been associated with tumor cell invasion. As the figure shows, the process of tumor growth and proliferation is characterized by alterations of the glycosylation pathway of key growth factors receptors, which modulate their activity and signalling. Another example is the alterations of O-GlcNAcylation, which is also associated with cancer progression. Moreover, in the process of tumor cell migration, integrins show altered glycosylation in both O-linked and N-linked glycans. Finally, terminal sialylation has also been suggested to interfere with cell–extracellular matrix (ECM) interactions, promoting an upregulation in the migration and invasive phenotype of cancer cells. (*278*).

1.4 GalNAc-Transferases (GalNAc-Ts)

1.4.1 Regulation of Mucin type O-glycosylation

As mentioned earlier, mucin O-glycosylation is characterized by the addition of a GalNAc sugar to a Ser or a Thr residue forming the Tn antigen, and this linkage is controlled and initiated by a family of 20 glycosyltransferases known as GalNAc-Ts (Figure 1.10) (*316*). The addition of the first GalNAc sugar from the nucleotide sugar UDP-GalNAc to the hydroxyl group of serine or threonine in protein substrates is what initiates the first step in the proper synthesis of membrane bound or secreted mucin glycoproteins (Figure 1.10) (*316*). O-glycosylation is unique compared to other types of glycosylation mainly due to the fact that it is controlled by a large subset of enzymes compared to a few number of enzymes required for N-glycosylation, and interestingly, it seems that not one but multiple of these GalNAc-Ts are required for the proper glycosylation of their substrate proteins (*175, 316*). There are 20 transferases encoded by the human genome, 17 of which have been extensively studied and characterized (*176*), and the remaining three were more recently characterized for their enzymatic activity (*325*). These comprehensive studies on this large subset of GalNAc-Ts, emphasizes their important role in controlling mucin O-glycosylation. The complexity of these GalNAcTs is represented by their protein substrate specificity, suggesting that these enzymes have unique and overlapping roles in fine tuning of the glycosylation state of mucins, and thus controlling their functions (*325*).

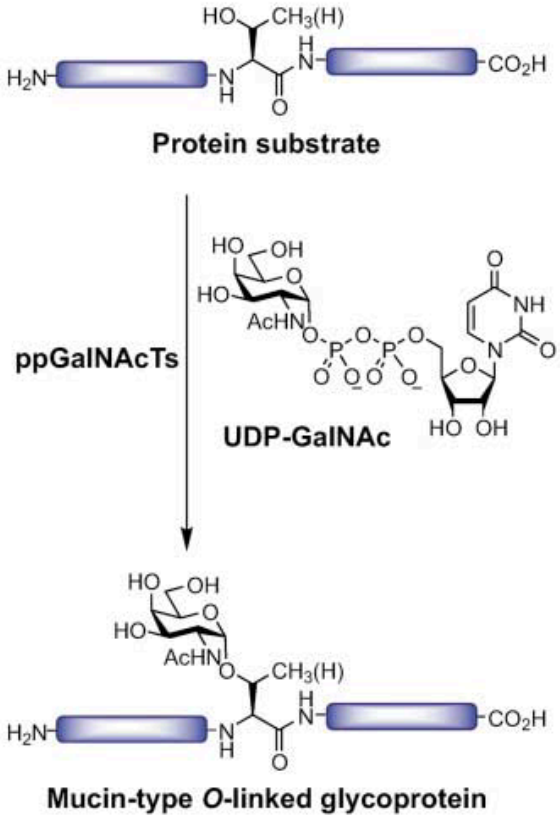


Figure 1.10. Mucin-type O-linked glycosylation and the initiation process that is controlled by a family of 20 enzymes known as GalNAc-Ts
Mucin-type O-linked glycosylation of proteins and a description of the glycosylation initiation pathway that is controlled by the family of enzymes (polypeptide *N*-α-acetylgalactosaminyltransferases (GalNAcTs)). The GalNAcTs make use of a UDP-GalNAc acting as their donor substrate, which allows them to modify the serine or threonine residues of their substrate proteins (*326*).

1.4.2 The human GalNAc-T gene family

Several studies have been focused on analyzing the functional characteristics of this enzymes family, specifically aimed at understanding their structural domains, distinct acceptor peptide substrate specificities, and their expression during development in cells and tissues.

The GalNAc-Ts are a highly conserved family of proteins, and their protein architecture reveals conservation in specific domains (*327*). The GalNAc-Ts are characterized by sharing the following structures:

1) A type II TM structure with an N-terminal cytosolic tail.

2) A stem region with variable lengths.
3) A catalytic luminal domain (GT1).
4) A Gal/GalNAc recognition domain that shows around 60% amino acid sequence similarity among the different GalNAc-Ts (*327*).
5) A C-terminal Ricin like (identified by homology to ricin) lectin domain that is suggested to be the region for acceptor substrate binding and was also found to show 40% amino acid similarity among the different GalNAc-Ts. (*176, 327, 328*) (Figure 1.11).

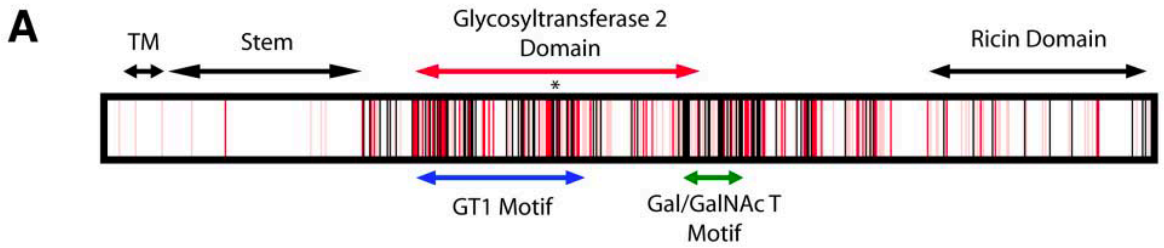


Figure 1.11. Detailed domain structure of the multiple GalNAc-Ts
Regions in white, pink, red, and black represent, the different sequence identity amongst the different GalNAc-Ts (*175*). The positions of the transmembrane (TM), stem, catalytic and ricin domains are indicated by the colored arrows above or under the boxed figure. Regions in white, pink, red, and black respectively indicate, 0±29%, 30±69%, 70±99%, and 100% sequence identity (*175*).

The 20 human GalNAc-T genes localize to different chromosomal loci, in addition to having large variability in genomic structures and organization (*316*) (Figure 1.12). Analysis of the GalNAc-Ts transcripts have suggested that a few of these transferases show a degree of ubiquitous expression in human tissues, while the majority show a specific and unique distribution in human and mouse tissues (*176*). Several analytical methods using protein expression analysis support data on the different and overlapping expression patterns of this family of enzymes, and transcriptional analysis of the expression of these transferases reveals the degree of differences in their spatial expression patterns during development in mouse embryos (*175, 329*). Monoclonal antibodies against these isoforms have been proven to be useful in better examining their expression patterns in a small subset of tissues, and interestingly changes in the normal expression patterns of these isoforms have been highly examined in diseases such as cancer (*175*).

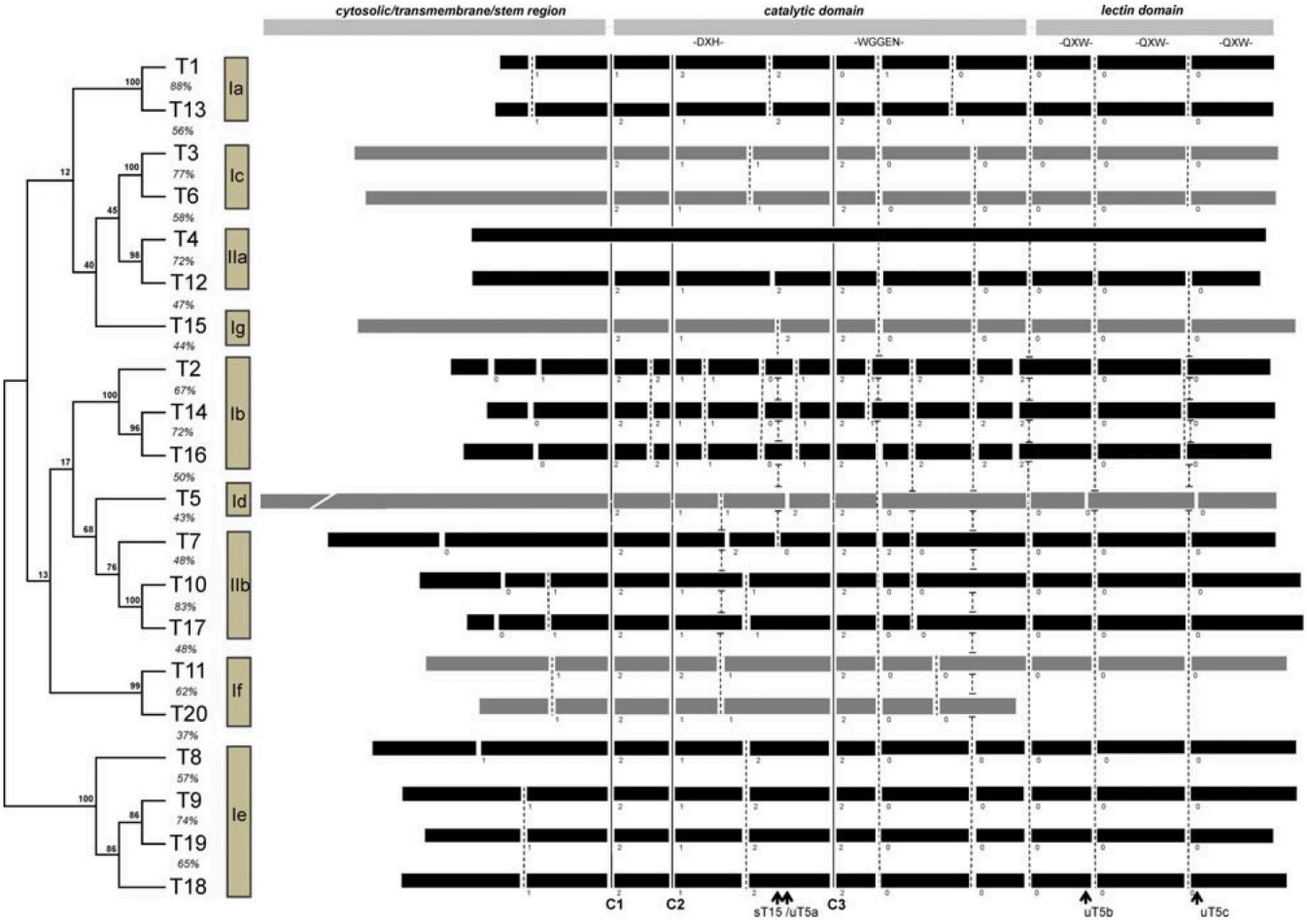


Figure 1.12. Phylogenetic and genomic tree of the human GalNAc-T gene family
The left panel shows the genomic tree of the 20 GalNAc-T members, which was derived from molecular phylogenetic anaylses by the maximum likelihood method of Gblock (*330*). The right panel depicts genomic organization of the ORF for all 20 human GalNAc-T genes. Exons are shown as boxes. Conserved intron/exon are indicated by a solid line and labeled. The GalNAc-T genes are based on the similarities of their genomic organizations. The top of the figure describes the regions encoding the different domains of GalNAc-Ts (cytosolic/transmembrane/stem regions, catalytic and lectin domains) (*316*).

1.4.3 Peptide and glycopeptide substrate specificities of the GalNAc-Ts family

As previously discussed, it has been determined that each GalNAc-T isoform has a unique acceptor peptide substrate specificities, with a certain level of overlap amongst the different isoforms (*175*). Thus, what determines the specific GalNAc-Ts activity has been extensively examined, and substrate preferences indicate that these enzymes may play unique roles in the addition of the first GalNAc sugar to a certain subset of sxyamino acid,

as well as having special features that dictate which serine and threonine will have an O-glycan linkage (*316*). *In vitro* studies have revealed that the different family members have unique roles in GalNAc additions that function by explicitly adding this sugar form either early, mid or late to their specified substrates (*316*). Unlike N-glycosylation that occurs at a defined consensus sequence (AsnXaaSer/Thr) (*331*), no specified consensus sequence has been defined for O-glycosylation of mucin glycoproteins. Several computational algorithms have been used to solely predict the possible O-linked site based on amino acid sequences alone (*332*), and similar algorithms have been used in our study, which is discussed in depth in Chapter 2. Interestingly, there seems to be multiple factors that play a role in defining the specified substrate glycosylated by these isoforms, which include secondary structure motifs of the peptide substrate, suggesting that it is not solely dependent on the amino acid sequence of the underlying peptide (*333, 334*). Other observations suggest that the density of O-linked glycans found on certain mucins could affect the level or the possibility of glycosylating neighboring amino acid residues and thus affecting the level of activity of the GalNAc-Ts (*335, 336*). Moreover, it also appears that not only peptides can be considered as potential substrates, since certain GalNAc-Ts use glycopeptides as their potential acceptor substrates: such examples were reported with GalNAc-T-17 and -10 (*337-339*). In addition, some GalNAc-Ts show substrate preference to what is referred to as "naked" peptides, and GalNAc-T16 has been characterized to glycosylate such substrates, which suggest that it may be an early acting glycosyl transferase (*340*). This kind of substrate preference gave rise to the categories these transferases belong to (Figure 1.12). Overall this demonstrates that transferases that show preference to non-glycosylated peptides may represent early-acting transferases, compared to those showing preference to glycopeptides, which may be associated with mid and late mucin glycosylation. Moreover, the localization of these enzymes in the Golgi apparatus could also play a role in the glycosylation pathways. Although all these transferases have been suggested to initiate the biosynthesis of O-linked structures in the Golgi, it seems that they have specific localizations within the Golgi compartments (191). Each compartment of the Golgi has a unique subset of GalNAc-Ts, and one factor that may control their localization is their structural organization that defines which compartment in the Golgi will retain specific GalNAc-Ts (*211, 341, 342*). Another factor includes GalNAc-Ts lectin domain

specificities, which determine the acceptor peptide or glycopeptide substrates (*343, 344*). These required features mainly suggest that these transferases work in a coordinated fashion driven by certain criteria not only linked to amino acid sequences of the peptide. This further demonstrates the higher level of complexity in the initiation of O-glycosylation. All information gathered on substrate preferences of these GalNAc-Ts is scarce, and data pertaining to the specific localization of each enzyme can further help determine the point of attachment of O-linked structure. Interestingly, aberrations in glycosylation in various diseases such as cancer have been largely attributed to the localization and activity of these GalNAc-Ts.

1.4.5 *In vivo* models of GalNAc transferases and the concept of GalNAc-T redundancy

The different members of the GalNAc-Ts family have also been shown to be differentially expressed in malignant tumors compared to normal tissue (*278, 345, 346*). The deregulation in the expression of the different GalNAc-Ts allows them to play diverse roles in carcinogenesis.

In vivo mouse models have been the main approach used for analyzing GalNAc-Ts function, and to better understand their role in cancer development. This approach is focused on knocking out or ablating the expression of one or multiple GalNAc-Ts in mouse and other living models to better depict the role each of these transferases play in the metastatic cascade, in addition to understanding the redundant roles that these transferases play in normal development and cancer progression. The first *in vivo* experiment targeted the ablation of *GALNT13*; animals following *GALNT13* suppression appear to develop normally, but presented a reduced expression in the Tn antigen in the cerebellum (*347*). Another study used an *in vivo* mouse model lacking the expression of the *GALNT1* gene, data showed that these mice exhibited a bleeding disorder linked with deficiency in B-cell maturation (*348*). Recent studies examined the use of animals that are genetically modified for multiple GalNAc-Ts (*GALNT4 and GALNT5*), and those animals presented with phenotypes similar to those observed when ablating for the *GALNT13* gene, showing normal phenotypes (*175, 349*). Mice deficient in the *GALNT8* gene also appeared to grow and survive normally (*350*). Similarly, mice deficient in *GALNT10* and *GALNT14* did not

present any obvious phenotypes (*316*). All these reports point to the important hypothesis that single gene knockouts do not seem to present with significantly altered or lethal phenotypes since there may be a compensatory regulation induced by the other GalNAc-T genes (*351, 352*). Moreover, multiple GalNAc-Ts may be co-expressed and thus act as back-up for a loss of one or multiple genes in these mouse models, especially since these enzymes display overlapping substrate specificities (*316*). Thus far, no *in vivo* study has been published to confirm the GalNac-Ts acceptor sequence, their protein substrates, or correlate any biological function linked to GalNac-Ts gene overexpression or loss. Evolutionary studies cannot predict why the GalNac-T family is so diverse and if this diversity is linked to some degree of redundancy between the genes. It is believed that their large number may reflect an essential functional requirement of the different GalNAc-Ts isoforms (*175*). A study of mannosyl- O-glycosylation on yeast cells has supported this hypothesis, showing that multiple and not one O-mannosyltransferases are needed for the survival of yeast cells (*353*).

1.5 Glycoproteomics approaches in studying alterations in glycoproteins expression and structure

1.5.1 Challenges in studying the process of protein glycosylation

As discussed throughout this section, it is clear that glycosylation is a PTM that has a plethora of effects on the structure and function of glycosylated proteins. Alterations and modifications in the glycosylation pattern of proteins play a very important role in different diseases including cancer. The exact understanding on how glycosylation affects mechanisms important for the proper function of the protein has not been very well studied, and this is mainly due to the fact that there are no proper methods developed to comprehend the process of how glycan structures get synthesized on proteins. Indeed, oligosaccharides are not easily monitored, since their monomers are usually branched and are synthesized with the help of a large list of enzymes (*354*). The monomers that make up those oligosaccharides have very similar chemical reactivity, and there is not enough information available to determine what encodes a particular oligosaccharide sequence (*354*). Additionally, these polymers are heterogeneous both in their sequence and spatial

arrangement thus creating large molecular diversity assembled from the nine common monosaccharides found in humans (*355*). Interestingly, there exists a large number of potential glycosylation combinations on a glycoprotein, where a single protein can undergo a number of N- and O-glycosylation on multiple or single sites (*278*). Additionally, the same glycosylation site can be occupied by different glycans in variable copies, a phenomenon called microheterogeneity (*356*). Therefore, creating libraries of saccharides for proper examination of their function and synthesis is not an easy task. Moreover, although oligosaccharide heterogeneity by itself presents a challenge for studying their branching properties, the identification of the enzymes responsible for their synthesis can be a bigger challenge. As previously mentioned, there are multiple enzymes implicated in the proper assembly of each class of glycoconjugate in addition to a family of enzymes responsible for each type of glycosidic linkage for these glycan structures (*175*). With all these factors, it becomes clear why research in glycobiology is slow, and that the process of developing new synthetic tools for understanding novel functions linked with carbohydrates is still not fully developed. In order to completely characterize the glycoproteome, it is important to identify and analyze the glycan structure found bound to the glycoprotein, the protein that expresses the glycan structure or structures, and the site on the protein where the glycan binds (*357, 358*). Therefore, several methods are often needed to provide proper detailed characterization of glycans and the proteins they are found on. In a typical glycoproteomic experiment there is usually the need for **1)** Glycoprotein detection and enrichment, **2)** Methods to determine the glycosylation type, **3)** Glycan release from protein for glycan structure analysis, **4)** Glycan structure analysis, **5)** Glycan site mapping, and finally **6)** Glycoprotein quantification (*359*) (Figure 1.13). The remainder of this section will focus on reviewing the major glycoproteomic approaches used for analysing protein glycosylation.

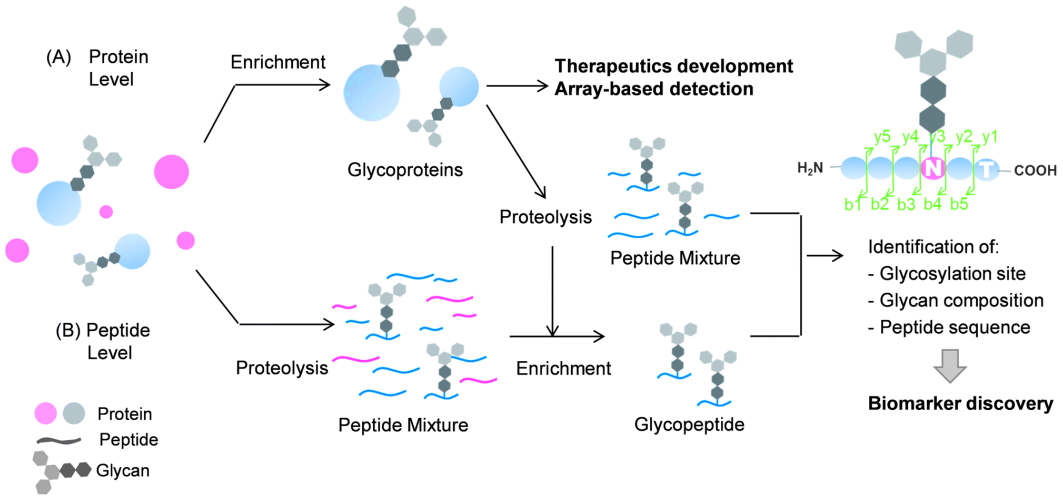


Figure 1.13. General strategy of glycoproteins detection and analysis
A. Protein level enrichment is required for array-based detection of glycoproteins that can be further used as biomarkers. B. Peptide level enrichment is utilized for identification of glycosylation site, glycan composition and peptide sequencing for biomarker discovery (*360*).

1.5.2 Glycoproteomic techniques for glycoprotein analysis

There are several high-throughput methods that have been developed to examine changes in glycoproteins in a disease setting such as cancer dissemination. These methods have been focused on providing "omics" research with sensitive technologies for better glycoproteomic analysis. Some of these high-throughput techniques include and are not limited to: liquid chromatography/mass spectrometry (LC/MS)-based glycoproteomics (*361*), stable isotope labeling and mass spectrometry (*362*), glycopeptide capture approaches coupled with mass spectrometry (*363*), and lectin-affinity purification of metabolicly labeled glycoproteins (*364*). Based on the targeted molecule subjected for analysis these approaches are categorized into two categories: **1)** glycoprotein based and **2)** glycopeptide based analyses (section A and B from Figure 1.13). Recently, a new technology has been developed – the Isotope-targeted glycoproteomics (IsoTaG) platform, allowing both these categories to be combined in a single experimental approach (*354, 365, 366*). The IsoTaG platform represents a mass-independent platform for intact N- and O-glycopeptide discovery and analysis, which currently is the only method for the analysis of fully elaborated glycopeptides in addition to revealing the inherent structural details of the glycoproteome (*354*). This section will discuss some of the glycoproteomic methods

currently in use, by discussing their advantages and disadvantages and the way they helped pave a new possibility for the field of glycobiology in providing more concrete information on cancer cell properties, including the identification of prognostic and therapeutic response biomarkers in cancer research.

Tandem MS is currently the most powerful proteomics tool, and is becoming more and more useful as a tool in the glycoproteomics field due to it high sensitivity and accuracy. The challenges that come with the application of MS in glycoproteomic studies are related to the level of abundance, since a lot of tumor-specific biomarkers are usually low-abundant proteins, and these proteins are normally masked by more abundant proteins found in the large dynamic range of protein concentrations in human biological samples. Thus, a lot of glycoproteins are sometimes difficult to detect by conventional MS methods (*367*). In glycoproteomics, a process of enriching for low-abundant glycoproteins is necessary prior to performing MS analyses, since such enrichment will increase the detection sensitivity of glycoproteins and decrease the level of complexity of protein identification for future analyses (*359, 368*). The most commonly used enrichment methods are lectin affinity capture, chemical capture or boronic acid beads, which are methods dependent on the use of affinity reagent, and are used to enrich glycoproteins and then digest them into peptides (*367*). There are other enrichment methods that first digest glycoproteins into peptides and then enrich for the digested peptides. These approaches are known as glycopeptide enrichment by chromatography, which includes size exclusion, ion exchange chromatography, and chemical immobilization techniques (*357, 363, 369-371*).

1.5.3 Other technologies for glycan profiling

Although the above glycoproteimics approaches are well-established methods used in the field, none of them can capture every single glycoprotein or glycopeptide and thus capture glycoproteins in an unbiased fashion. For these purposes, in the field of cancer glycoproteomics, research has become focused on the use of more specific methods targeting cancer glycan epitopes, such as sialic acids, or sialylated Lewis epitopes previously discussed in former sections. These methods include the application of antibodies such as the application of anti-glycan antibody for affinity capture (*372*), (*373*),(*372, 374, 375*). Other methods include the application of glycoengineered cell line

models that are focused on using site-directed mutagenesis (*376, 377*). Moreover, research has been also focused on the application of isogenic cell systems, which has permitted the precise determination of protein O-glycosylation in several cell types allowing for a revolution in the field of glycobiology (*178*). The genomic editing approaches created the so-called SimpleCell models, which are cellular models genetically engineered to produce a glycoproteome with homogenous and not heterogeneous glycan structures, hence the name "Simple" (*378*). Finally the more recent enrichment approach in glycopretomics is focused on using affinity purification of metabolic labeled glycoproteins, named the Staudinger capture reaction (*379*).

All the above mentioned revolutionary techniques applied in the field of glycobiology are dependent on the application of MS-based glycomic and glycoproteomic analyses; however, these types of analyses are expensive and require a certain level of expertise for proper glycoprotein examination. There is now a focus on applying MS-independent methods to analyze glycomic alterations (glycan structure alterations and/or glycoprotein alterations) in cancer cells or even in cancer tissue samples. Such applications include: antibody-lectin-sandwich, glycan/glycoprotein arrays, lectin arrays, proximity ligation assay (PLA) and imaging mass spectrometry (IMS) as reviewed in (*380*).

1.5.3.1 Affinity metabolic labeling-a new engineering method for the incorporation of unnatural substrates

In addition to the aforementioned techniques used in glycoproteomics studies, metabolic affinity tagging is now proven to be the best approach used to facilitate glycan labeling, separation and detection. Visualizing glycans is not easy, especially since glycans are not gene products that can be encoded with tags. Labeling of glycans has been recently achieved with the application of metabolic labeling, through a bioorthogonal reaction using chemical tags, an approach developed by the Bertozzi group (*381*). This approach utilizes bioorthogonal chemical reporters that can be incorporated into glycans using the cells' own metabolic machinery, as these chemical reporters are designed to be non-perturbing and thus to not interfere with the normal elongation process of the glycan structure (*381*). One major advantage for the application of bioorthogonal reactions is that these reactions can

normally follow the biological system process without interaction or interfering with the inside biomolecules (*382*). The labeling process is divided into two steps. The 1st step is dependent on the addition of a reactive handle with a bioorthogonal functional group to the glycan which is accomplished by feeding growing cells (*in vitro* process) (*382*) or living animals (*in vivo* process) (*383*) with an unnatural monosaccharide, and this monosaccharide gets incorporated by following the process of the glycan biosynthetic pathway (Figure 1.14). The 2nd step is based on the conjugation process using a specifically designed chemically engineered affinity tag or probe that can label or tag the glycans by applying the chemoenzymatic approach known as the bioorthogonal reaction (*384*). For proper labeling and incorporation, there are several criteria that need to be fulfilled, in order to assure proper covalent reaction between the probe and the reporter: **a)** the reporter and the probe have to be reactive in a specified physiological environment (37°C, and pH 6–8); **b)** the reporter and the probe have to be inactive and not react or affect the surrounding biological environment of a cell or organ, and **c)** the two partners (probe and reporter group) must be stable and should be bioavailable for proper use with different cell types and/or organisms (*381*). When such criteria are met, the metabolic labeling approach will properly enable the visualization of the tagged glycoconjugates through imaging (*385*) and/or will allow tracking the glycan structure and identifying the peptide site of attachment, especially when applying the above mentioned IsoTag approach that uses isotopic signature for glycan structure tracking (*354*).

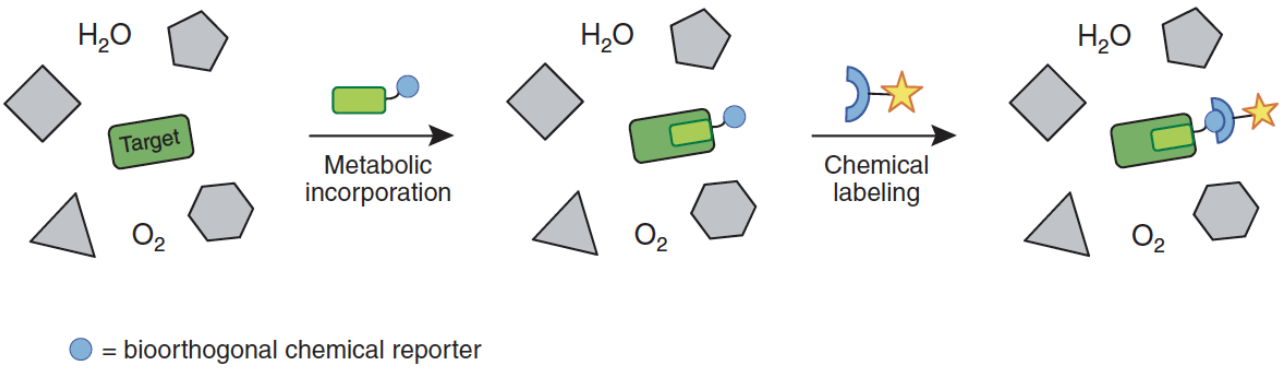


Figure 1.14. The bioorthogonal chemical reporter strategy
The bioorhogonal reporter system is based on a chemical reporter (blue circle) that is directly linked to a substrate (light green box), which gets introduced into a specified target

biomolecule through normal cellular metabolism. In the second step, the reporter gets covalently tagged with an exogenously delivered probe (blue arc). Both the chemical reporter and exogenous probe are designed to avoid any reactions with nontarget biomolecules (gray shapes). Adapted with permission from (*381*).

There are a few bioorthogonal chemical reporters designed to handle and fulfill the above-mentioned criteria. Some of them that are listed in Table 1.6 consist of activated alkynes and terminal alkynes that can be ligated with azides for protein, glycan and peptide labeling (*381, 382*). Azides have only been recently used as chemical reporters, and their advantage is that they can label any biological molecule as shown in Table 1.6. One example of biorthogonal labeling is presented in Figure 1.15, where the azido GalNAc analog (GalNAz) is used for metabolic labeling of mucin-type O-linked glycoproteins. An interesting characteristic of azides is their absence in natural conditions, and thus they do not react with abundant molecules found in living systems. Azides are now being exploited more for the rigorous development of bioorthogonal reactions (Table 1.6) through the application of chemoselective ligation processes such as the Staudinger ligation of azides (*379*), and click chemistry (*384*).

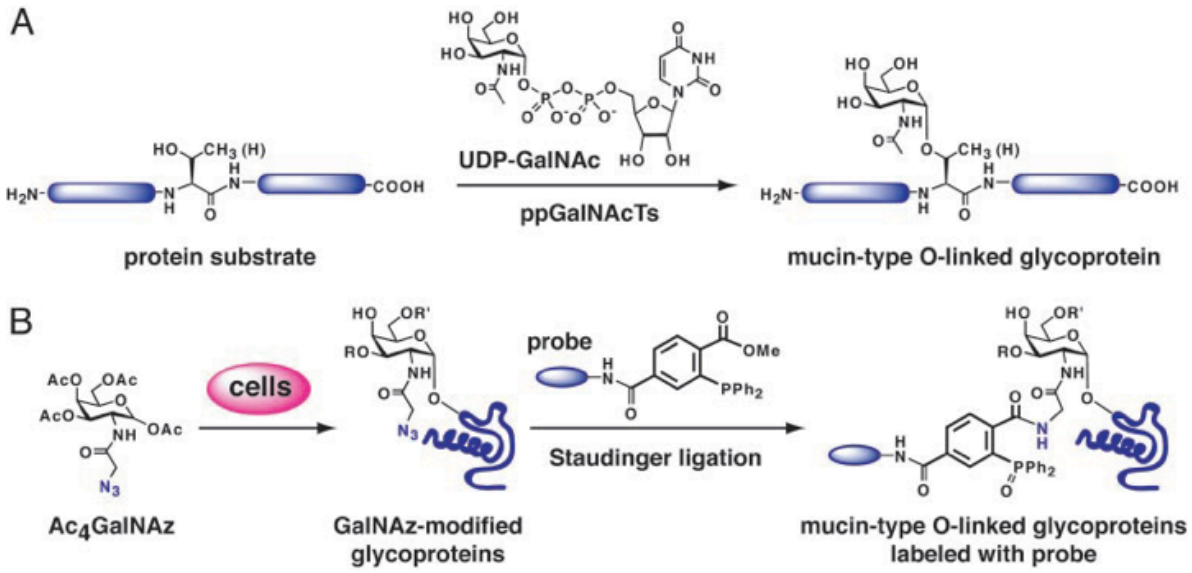


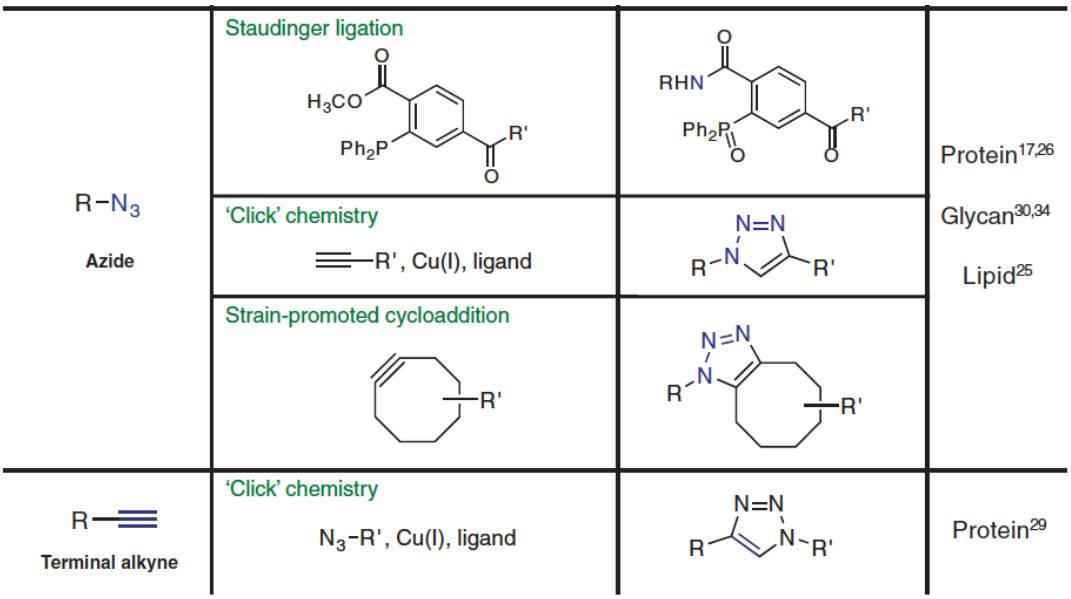
Figure 1.15. Strategy for metabolic labeling of mucin-type O-linked glycoproteins with an azido GalNAc analog (GalNAz) for proteomic analyses
A. Initiation of mucin-type O-linked glycan biosynthesis by the ppGalNAcTs. UDP–GalNAc is the nucleotide sugar donor substrate for the ppGalNAcT family. B. Optimized strategy for metabolic labeling of mucin-type O-linked glycoproteins with an azido GalNAc analog (GalNAz) for proteomic analysis with phosphine probes. This strategy is

dependent on the chemoselective Staudinger ligation reaction between azides and phosphines. The strategy provides a means to distinguish proteins bearing O-glycans from the remainder of the proteome (*386*).

1.5.3.2 Chemoselective ligation to metabolically installed unnatural cell surface oligosaccharides

The Staudinger ligation has been be used to covalently attach probes to azide-bearing molecules such as metabolically labeled glycans, by forming a link between azides and phosphines, and more specifically, by creating a link between a nitrogen atom of the azide group and the triarylphosphine (*381*). The more advantageous type of ligation with azides is the application of click-chemistry known as Copper-catalyzed [3+2] azide-alkyne cycloaddition (CuAAC) (*384*). Click chemistry is a cycloaddition between azides and terminal alkynes with the presence of catalytic amounts of copper (Cu) (*387*). The copper catalyzed click-reaction has been used to tag azides, and in many cases the ligation process has been used with alkyne groups as the chemical reporter for the click reaction (Table 1.6). This reaction also has some advantages since it allows for labeling without introducing any background (*381*). Click chemistry has been used for glycoproteomics labeling since it allows the detection of very small amounts of azide-labeled molecules such as glycans, and it was suggested to be a much faster ligation process compared to Staudinger ligation (*388*). The major disadvantage that comes with the use of the CuAAC is the cellular toxicity of copper (*381*), but there are more recent applications of click chemistry known as copper-free click chemistry that are now being used in *in vivo* studies (*389*).

Table 1.6. Chemical reporters and bioorthogonal reactions used in living systems.



Few of the different chemical reports used for the bioorthogonal reactions such as: Staudinger ligation, Click chemistry, and strain-promoted cycloaddition. Adapted with permission from (*381*).

1.5.4 The new glycoproteomic approach – the isotopic targeted glycoproteomics (IsoTaG) platform

As mentioned, enzymatic digestion has been achieved for N-glycans (*390*), however for O-glycans, gene editing through genetic engineering for glycan simplification was achieved only when applying the Simplecell strategy (*391*). Indeed, truncation of the glycan structure represents a major disadvantage, since information on glycosylation sites may be lost, which in turn will not allow for proper correlation of the glycan structure to the correct protein site (*176*). Chemically labeling glycans using chemical reporters have shown to be very advantageous in the field of glycobiology, since they can be used to monitor glycans in their natural environment. This approach allowed the proper examination of glycan changes caused by disease induced changes and/or mutations at the level of truncation and aberrant glycosylation. As previously mentioned, glycans determine the biological functions of all glycoproteins. Thus, it seems that acquiring the ability to examine glycan structures in conjugation with their attached protein and peptide site can most likely

provide great insight into their biological and fundamental role in cellular and molecular biology.

The IsoTaG approach developed by the Bertozzi group is the first proteome-scale platform capable of identifying peptides and glycan structures of both N- and O-glycopeptides (*354*). This approach comprises: **1)** metabolic labeling of the glycoproteome, **2)** chemical enrichment and isotopic recoding of glycopeptides, and finally **3)** directed mass spectrometry (*354*). The first part of the platform protocol involves the application of a method used to metabolically label glycoproteins in cell culture. This platform makes use of bioorthogonal labeling, accomplished by studying glycoconjugates containing the monosaccharides: N-acetylgalactosamine (GalNAc), N-acetylglucosamine (GlcNAc), and sialic acid (Sia) (*354*), in addition to other monosaccharides types that are now in the process of being developed for monitoring changes in glycan modifications such as fucosylation and others. Studying these glycoconjugates was made possible due to the permisssiveness of the approach to follow the GalNAc and GlcNAc salvage pathways, as well as the sialic acid biosynthetic pathway, as these pathways can tolerate the introduction of unnatural sugars that bear bioorthogonal functionality. Some of the analogs examined to follow such routes are N-azidoacetylgalactosamine (GalNAz), and N-azidoacetylglucosamine (GlcNAz) (*392*). GalNAz is an unnatural GalNAc derivative that is used to incorporate GalNAz to tag mucin type O-linked glycans (*392*). The GalNAz derivative is used as part of a method to specifically label proteins with O-linked glycans (*392*). It has been demonstrated that GalNAz is transformed to UDP-GalNAz *in vitro* and *in vivo* by exploiting the endogenous GalNAc salvage pathway, enabling the labeling for proteins with O-linked glycans (Figure 1.16) (*392*). More importantly, Ac4GalNAz can be metabolized in the cell to both GalNAz and GlcNAz by the mammalian enzyme UDP-galactose 4′-epimerase (GALE) (*393*), and is used as a reporter for intracellular O-GlcNAc and extracellular N-GlcNAc and O-GalNAc (*392*). In contrast, GlcNAz works by binding to cytosolic and nuclear proteins at sites occupied by O-GlcNAc (*394*).

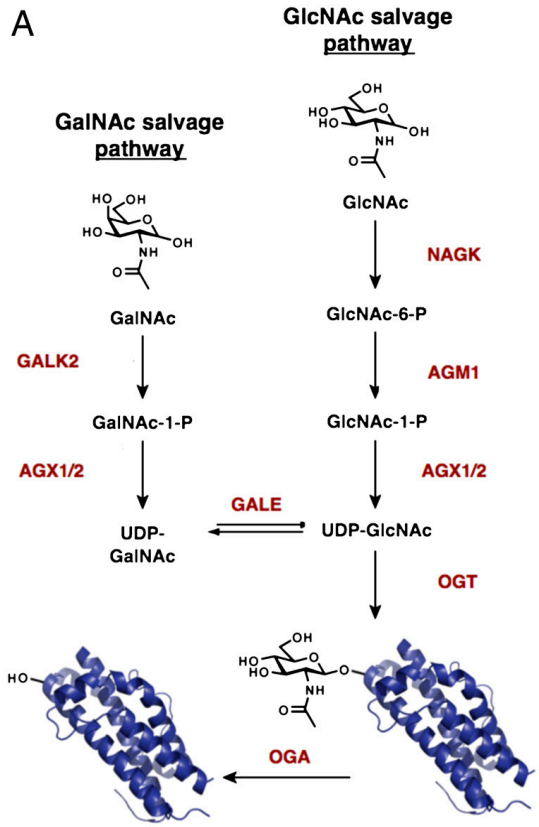


Figure 1.16. The GlcNAc and GalNAc salvage and O-GlcNAc signaling pathway
A detailed description of both the human GalNAc salvage and GlcNAc salvage pathways, including the different salvage pathway enzymes (*392*).

These azide-labeled glycoproteins can then be tagged by applying a chemical glycoproteomics enrichment procedure, which includes tagging glycoproteins with Click chemistry (*395*). The Click chemistry reaction applied in this method is briefly described in Figure 1.17. The reaction is a two-step process that involves the application of two functional groups. The first part of this chemical process involves the chemical reporter, which is here referred to as the bioorthogonal functional group that is incorporated into a substrate. The approach is followed by a covalent linkage to a probe through a click reaction, which allows the detection and isolation of a gene or protein target. The most used bioorthogonal click reagent is azide, which allows for modification of proteins, as well as various classes of biomolecules, such as nucleotides, lipids, and glycans, as well as other metabolites. Once installed into the target substrate, azides can covalently combine naturally to specified complementary partners by the azide-alkyne cycloaddition (*354*).

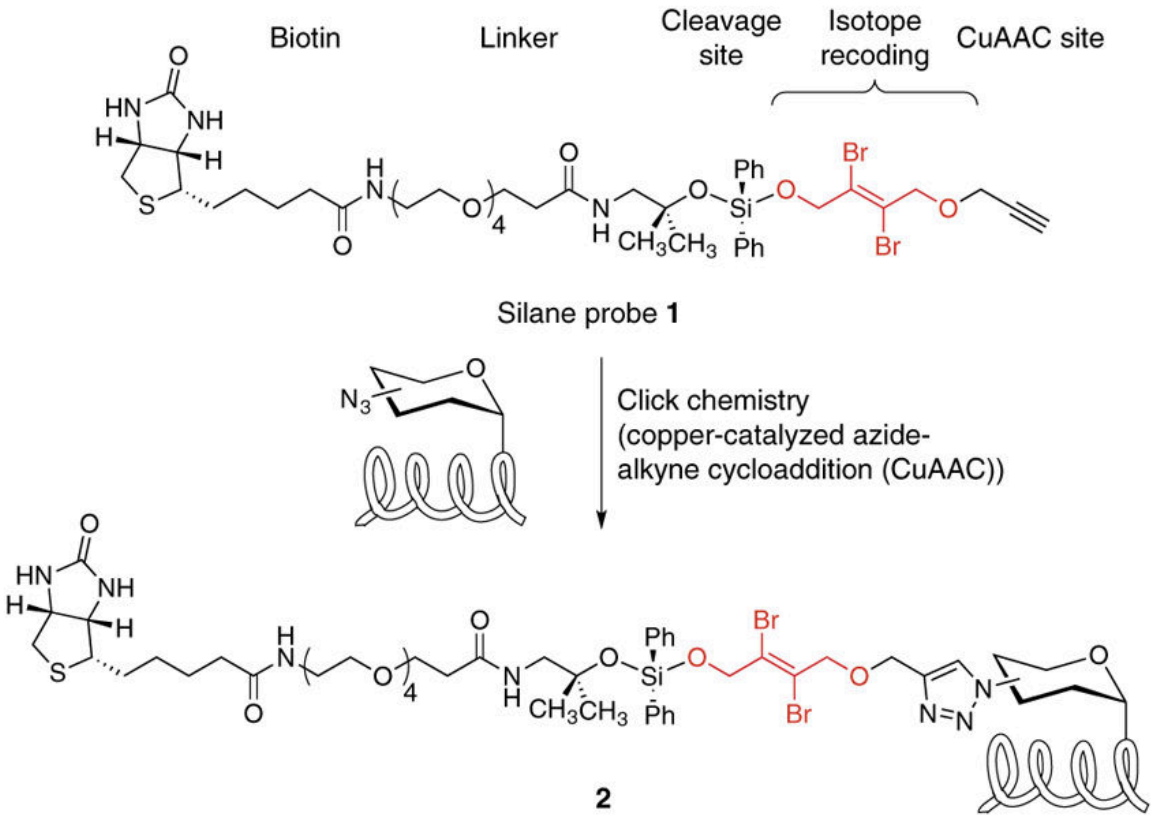


Figure 1.17. Structure of cleavable probe for mass-independent glycoproteomics
The silane probe (1) tags glycoproteins via copper-catalyzed azide-alkyne [3+2] cycloaddition Click chemistry. The tagged glycoproteins (2) are then enriched by streptavidin–agarose affinity chromatography. Enriched glycoproteins get digested on beads with trypsin to release all non-conjugated peptides for glycoprotein identification. Treatment with 2% formic acid is then used to recover the glycopeptides from the beads. The isotopically recoded glycopeptides are then recovered by mild acid hydrolysis. The isotope pattern is then retained on the glycopeptide and used for targeted glycoproteomics (*354*).

After tagging, glycoproteins are getting enriched by streptavidin–agarose affinity chromatography. Enriched glycoproteins are then denatured, which is achieved by subjecting them to reduction and alkylation to break the disulfide bonds that hold the glycoprotein structure together (*396*). Consecutively, glycoproteins get digested on beads with trypsin to release all non-conjugated peptides for glycoprotein identification (*397*). This approach first allows for analysis of glycoproteins, followed by steps of cleavage and recovery of glycopeptides in treatment with 2% formic acid to recover the glycopeptides from the beads. Released glycopeptides are then analyzed by reversed-phase nanoflow liquid chromatography coupled to MS, which allows the proper identification of

glycoproteins. The principle behind this mass-independent assignment of intact glycopeptides from the cleaved fraction is that the IsoTaG strategy makes use of an encoded chemical tag on a glycopeptides isotopic envelope to create a unique pattern that is visually and computationally recognizable (Figure 1.18) (*398*). The pattern is then used to direct tandem MS to tagged glycopeptides, irrespective of abundance, followed by targeted computational analysis for high confidence identification of the glycopeptide, including structural assignment of the attached glycan. The application of this novel IsoTaG technique has shown improved knowledge acquisition on aberrant glycosylation and its implication in disease progression (*366, 399*). This technology allows attainment of essential information relating to the structures and the sites of attachment of glycans to their protein targets. Thus, a more comprehensive understanding of the effects of altered glycan expression in cancer can be easily addressed. Indeed, the knowledge of the glycan alterations and specificities could be important for drug discovery.

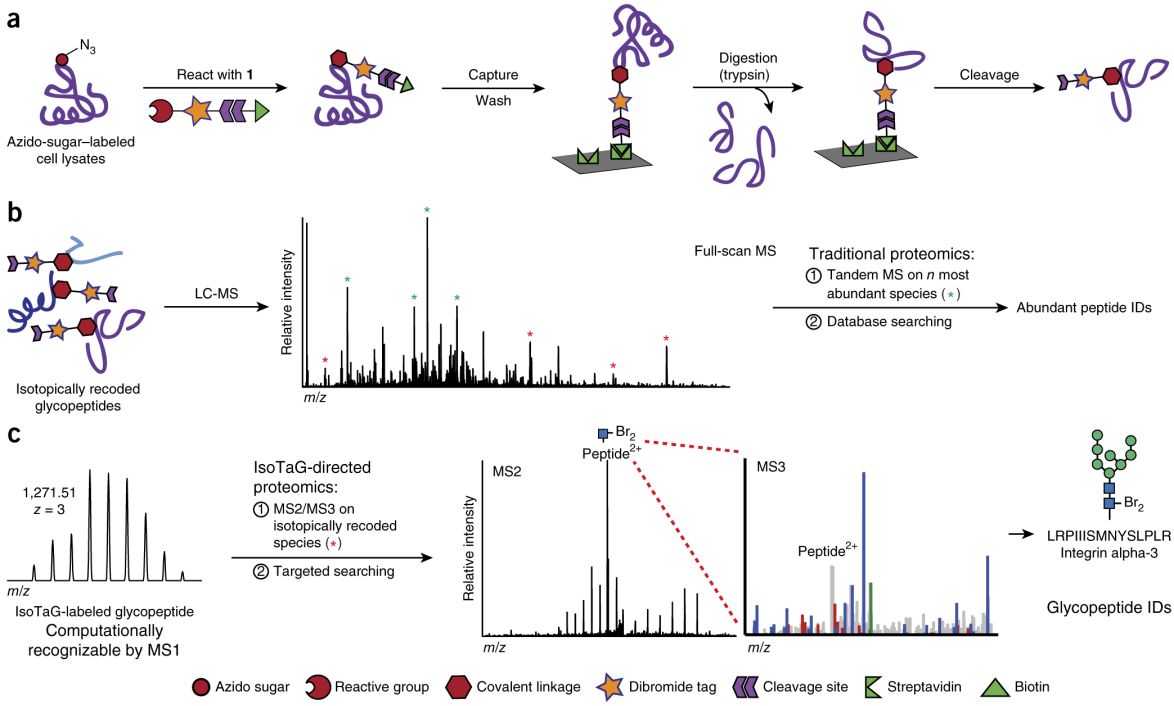


Figure 1.18. The IsoTag approach
A. The first part of the approach consists of an enrichment strategy using a dibrominated silane probe 1. The metabolically labeled glycoproteins are then enriched using a streptavidin-agarose affinity chromatography technique. On-bead digestion is then used to

remove all nonglycosylated peptides. The isotopically recoded glycopeptides are then cleaved and eluted from the resin. B. The collected isotopically recoded glycopeptides are then analyzed by the application of reversed-phase liquid chromatography coupled to mass spectrometry. C. The technique IsoTaG uses a pattern-searching algorithm to allow for a direct tandem MS (such as the recognition of MS2 and MS3) analysis of the isotopically recoded species. The IsoTag approach later focuses on performing targeted searching of the selected species, this allows for high-confidence glycopeptide identification. The tagged glycan is denoted with Br_2, and this structure can be designed with a different isotope depending on the targeted glycan structure and probe (*354*).

1.6 Problematic, hypothesis and objectives

Problematic

EOC is a disease that is responsible for more cancer deaths among women in the Western world than all other gynaecologic malignancies. EOC lethality primarily stems from the inability to detect the disease at an early, organ-confined stage, and the lack of effective therapies for the advanced-stages of the disease. EOC treatment includes the combination of surgery and platinum based chemotherapy, these approaches in addition to hormonal therapy show up to 80% responsiveness in patients, but unfortunately more than 50% of EOC patients eventually relapse. It is well established that cancer metastasis still represents as the major cause of the failure of cancer treatment. Thus, management of the metastatic disease becomes a crucial problem for the treatment of EOC. One possible way to resolve this problem is to target metastasis-specific pathways with novel therapies. Hence, focused identification of novel pro-metastatic mechanisms, target pathways and molecules could help identify new effective therapies.

Using an epigenomics approach, our lab has previously shown that DNA hypermethylation occurs in all (including less invasive/early) stages of EOC tumorigenesis, while advanced disease was exclusively associated with DNA hypomethylation of a number of oncogenes, implicated in EOC progression, including invasion/metastasis. These experiments identified the *GALNT3* gene as hypomethylated and overexpressed in HG EOC tumors, when compared to LMP EOC tumors and normal ovarian tissue. Additionally, the *GALNT3* expression significantly correlated with shorter PFS intervals in women with advanced EOC. Consecutive functional analyses, performed following shRNA-mediated GALNT3 KD in the EOC cell line A2780s, were strongly indicative of *GALNT3* implication in EOC cell proliferation, migration, invasion and cell cycle control. Gene expression profiling and

successive network and pathway analyses confirmed these findings, as numerous genes and pathways known previously to be implicated in EOC tumor invasion and metastasis were found to be downregulated upon *GALNT3* suppression, while some tumor suppressor genes were induced. Moreover, *GALNT3* KD induced the reduction of MUC1 protein in EOC cells, although the transcriptional level of MUC1 was unchanged. A significantly diminished protein expression of other glycosylated proteins in EOC cells upon *GALNT3* suppression was also observed, suggesting that *GALNT3* may influence the PTM and stabilization of MUC1 and other O-glycosylated mucin-like targets in EOC cells. Additionally, the *GALNT3* KD-mediated MUC1 suppression promoted the expression of the cell adhesion molecules E-cadherin and β-catenin, suggesting that the involvement of the *GALNT3*-MUC1 pathway in EOC invasion could include the destabilization of the E-cadherin/ β-catenin complex formation. The data collected specify mechanisms of abnormal O-glycosylation in EOC carcinogenesis and the identification of the *GALNT3* enzyme as an EOC oncogene and a novel EOC biomarker/possible molecular target for EOC therapy. The ability to distinguish differences in the glycosylation of proteins and identifying modification changes in the glycan profiles between A2780s *GALNT3* KD and control clones can identify potential glycoproteins and glycan structures as targets to improve on existing EOC biomarkers.

Moreover, not much is known about the specific GalNAc-T involved in MUC1 glycosylation. Understanding the specific role of the *GALNT3* enzyme in regulating differential glycosylation of MUC1 in EOC cancer may be important for developing therapeutic strategies for modulating MUC1 function in EOC. We believed that complementing our *in vitro* experiments with *in vivo* experimental models could be a critical component in elucidating the functional implications of *GALNT3* and possibly of other members of the GalNAc-T gene family in EOC tumorigenesis.

Some of the mammalian GalNAc-Ts isoforms have been grouped into subfamilies based on their high homology. One such example is the human subfamily comprised of *GAlNT3* and -*T6*, and these two transferases have around 65% homology in their coding sequence. *In vitro* functional studies have examined the roles and enzymatic functions of the different GalNAc-Ts, unfortunately a very limited number of *in vivo* studies were able to confirm or support data observed *in vitro*. So far, no *in vivo* study has been published to confirm the

GalNac-Ts acceptor sequences, their protein substrates, or report any biological function linked to GalNac-Ts gene overexpression or loss.

Hypothesis

We hypothesized that *GALNT3* overexpression contributes to EOC dissemination through aberrant mucin O-glycosylation. Thus, we believed that specifying some of the putative mechanisms of *GALNT3*-mediated abnormal glycosylation implicated in EOC carcinogenesis is an essential first step to be investigated. We also wanted to examine the concept of redundancy observed between the members of the GalNac-T family, as we hypothesized that there might be a cooperative catalytic action of *GALNT3* with other GalNac-Ts, and especially with its closest homolog *GALNT6*, in regulating mechanisms of EOC dissemination.

Objectives

1. Apply a targeted isotopic glycoproteomics approach (IsoTaG) to identify altered expression of O-glycoproteins and/or alterations in their glycan structures following *GALNT3* gene KD in EOC cells. This objective aimed at performing comparative analysis of the altered expression of metabolically labeled glycoproteins between *GALNT3* KD A2780s cells and control A2780s cells, after tryptic digestion and mass spectrometry analysis. The following aim focused on performing structural assignments of intact glycopeptides using a mass-independent chemical glycoproteomics approach to identify altered glycan structures between the *GALNT3* KD and A2780s control cells.

2. Analyze the implication of other members of the GalNAc-Ts gene family in aberrant O-glycosylation associated with EOC dissemination. We examined this objective, since the human GalNAc-Ts form a large evolutionary conserved family of enzymes that are responsible for the transfer of GalNAc to proteins, but they seem to have unique but yet overlapping substrate specificities. We thus hypothesize that it is probable that *GALNT3* is not the only GalNAc-T implicated in aberrant O-glycosylation in EOC. Therefore, we aimed at examining by immunohistochemistry (IHC) protein expression levels of different members of the GalNAc-T family in HG EOC tumor samples in order to examine their correlation with patient clinical data. We believe such gathered data could possibly aid in

identifying aberrantly expressed GalNAc-Ts and their implication in molecular mechanisms of abnormal O-glycosylation associated with EOC dissemination.

3. Defining the role of *GALNT3* in EOC tumor expansion and metastasis *in vivo* in an intraperitonial mouse model, and analysis of the possible redundant implication of other members of the GalNac-T gene family in aberrant O-glycosylation associated with EOC dissemination.

This objective aimed at characterizing if *GALNT3* gene KD is sufficient to induce effects on protein glycosylation, and in turn cause significant phenotypic changes in our animal models. This objective was also aimed at investigating if other GalNac-Ts members could be possibly playing compensatory roles in EOC cancer cells upon *GALNT3* gene KD.

Chapter 2:
Chapter 2.A: A metabolic labeling approach for glycoproteomic analysis reveals altered glycoprotein expression upon *GALNT3* knockdown in ovarian cancer cells

Razan Sheta[1,2], Christina M. Woo[3], Florence Roux-Dalvai[4], Frédéric Fournier[4], Sylvie Bourassa[4], Arnaud Droit[1,4], Carolyn R. Bertozzi[3,5] and Dimcho Bachvarov[1,2]

[1]Department of Molecular Medicine, Laval University, Québec PQ, Canada
[2]Centre de recherche du CHU de Québec, L'Hôtel-Dieu de Québec, Québec PQ, Canada
[3]Department of Chemistry, Stanford University, Stanford, CA, USA
[4]Centre de recherche du CHU de Québec, CHUL, Québec PQ, Canada
[5]Howard Hughes Medical Institute, Stanford University, Stanford, CA USA

(Originally published in Journal of Proteomics, August 2016, PMID: 27095597)

Keywords

Epithelial ovarian cancer; *GALNT3*; glycosylation; metabolic labeling; glycoproteomics; pathway and network analysis

Chapter 2.B: Proteomic dataset for altered glycoprotein expression upon *GALNT3* knockdown in ovarian cancer cells

Razan Sheta[1,2§], Florence Roux-Dalvai[3§], Christina M. Woo[4], Frédéric Fournier[3], Sylvie Bourassa[3], Carolyn R. Bertozzi[4,5] Arnaud Droit[1,3], and Dimcho Bachvarov[1,2]

[1]Department of Molecular Medicine, Laval University, Québec PQ, Canada

[2]Centre de recherche du CHU de Québec, L'Hôtel-Dieu de Québec, Québec PQ, Canada

[3]Centre de recherche du CHU de Québec, CHUL, Québec PQ, Canada

[4]Department of Chemistry, Stanford University, Stanford, CA, USA

[5]Howard Hughes Medical Institute, Stanford University, Stanford, CA USA

[§]These authors have equally contributed to the generation of the presented data.

Contact email: dimtcho.batchvarov@crhdq.ulaval.ca

(Originally published in Data Brief, May 2016, PMID: 27331112)

2.1 Preface

The following chapter presents my scientific findings as a first author, which were published in the Journal of Proteomics on August 2016, PMID: 27095597, and the second part of the chapter presents my scientific data, which was published in the Data in Brief on May 2016, PMID: 27331112. I executed almost all experiments presented in this manuscript and written the manuscript. The application of IsoTag was made possible through the time, help and teaching of Dr. Christina Woo who provided us with the probe, and taught me all the steps required for the application of this approach. All the Mass-Spec sample injections were made by Ms. Florence Roux-Dalvai, who also provided me with the required knowledge needed for the collection and analysis of the acquired data. Mr. Frédéric Fournier helped with some of the data analysis. Ms. Sylvie Bourassa and Dr. Arnaud Droit provided us with guidance on how to manipulate the data. Dr. Carolyn Bertozzi has introduced us to the application of glycoproteomics and provided us with help in correcting the manuscript. Dr. Dimcho Bachvarov is the principal investigator of this study. He was involved in experimental design, data analysis and article writing.

.

2.2 Résumé en français

Même si le cancer épithélial de l'ovaire est la forme la plus létale des cancers chez la femme, les mécanismes cellulaires et moléculaires de ce cancer restent méconnus. De plus, la compréhension des changements moléculaires associés aux métastases du cancer ovarien peut mener à l'identification de potentielles cibles pour de nouvelles interventions thérapeutiques. Dans notre laboratoire, nous avons décrit le rôle potentiel du gène polypeptidique N-acétylgalactosaminyltransférase 3 (*GALNT3*), membre de la famille des GalNAc-Transférases (GalNAc-Ts) dans la tumorigenèse de CEO. Nos données ont démontré que la surexpression de *GALNT3* contribue à la dissémination du CEO, via une O-glycosylation de type mucine aberrante. Dans ce chapitre, nous décrivons l'utilisation d'une approche en protéomique visant à mieux caractériser le rôle de *GALNT3* dans la O-glycosylation aberrante de type mucine dans les cellules du CEO. Nous avons utilisé la plateforme "mass-independent glycoproteomics", IsoTaG, afin d'identifier les glycopeptides intacts à partir de l'ensemble du glycoprotéome de la cellule. Les protéines candidates ont par la suite été identifiées en utilisant la banque de protéome humain Swiss-Prot, et les données ont montré l'implication de quelques centaines de gènes reliés au métabolisme cellulaire.

Les résultats ont été obtenus après application de "bioorthogonal chemical reporter strategy" pour l'analyse de l'expression des gycoprotéines dans un modèle cellulaire du CEO. Les cellules A2780s étaient déplétées en *GALNT3* (*GALNT3* KD A2780s). L'analyse par spectrométrie de masse LC-MS/MS a pu identifier plusieurs protéines différemment exprimées entre les cellules contrôles et les cellules *GALNT3* KD A2780s. Cette étude a permis de découvrir de nouvelles glycoprotéines représentant de nouvelles cibles pharmacologiques pour le traitement du CEO.

2.3 Abstract

Epithelial ovarian cancer (EOC) is a disease responsible for more deaths among women in the Western world than all other gynecologic malignancies. There is urgent need for new therapeutic targets and a better understanding of EOC initiation and progression. We have previously identified the polypeptide N-acetylgalactosaminyltransferase 3 (*GALNT3*) gene, a member of the GalNAc-transferases (GalNAc-Ts) gene family, as hypomethylated and overexpressed in high-grade serous EOC tumors, compared to low malignant potential EOC tumors and normal ovarian tissues. This data also suggested for a role of *GALNT3* in aberrant EOC glycosylation, possibly implicated in disease progression. To evaluate differential glycosylation in EOC caused by modulations in *GALNT3* expression, we used a metabolic labeling strategy for enrichment and mass spectrometry-based characterization of glycoproteins following *GALNT3* gene knockdown (KD) in A2780s EOC cells. A total of 589 differentially expressed glycoproteins were identified upon *GALNT3* KD. Most identified proteins were involved in mechanisms of cellular metabolic functions, post-translational modifications, and some are implicated in EOC etiology.

The *GALNT3*-dependent glycoproteins identified by this metabolic labeling approach support the oncogenic role of *GALNT3* in EOC dissemination and may be pursued as novel EOC biomarkers and/or therapeutic targets.

2.4 Introduction

Epithelial ovarian cancer (EOC) is a disease that is responsible for more cancer deaths among women in the Western world than all other gynecologic malignancies [1]. EOC lethality primarily stems from the inability to detect the disease at an early, organ-confined stage, and the lack of effective therapies for advanced-stage (metastatic) disease. Indeed, despite advances in cytotoxic therapies [2, 3], only 30% of patients with advanced-stage EOC survive 5 years after initial diagnosis [1]. Thus, management of metastatic disease is a crucial problem for EOC treatment. One way to resolve this problem is to target metastasis-specific pathways with novel therapies. Hence, focused identification of novel pro-metastatic EOC pathways and molecules could improve the chances of discovering new and more effective EOC therapies.

Protein glycosylation is a highly heterogeneous and ubiquitous post-translational modification (PTM) representing glycan structures attached at asparagine (N-linked) or serine or threonine (O-linked) side chains [4]. Analysis of the SWISS-PROT database predicted that more than 50% of all proteins are glycosylated [5]. Glycoproteins are involved in many biological functions, such as intercellular communication [6], intracellular signaling [7] and protein stability [8]. Glycosylation changes characterized in cancer cells follow a variety of forms, as glycan alterations can be associated with loss or gain of expression, depending on the cell type and the specific glycan's structure [9]. Aberrant glycosylation in cancer cells could affect certain ligand-receptor interactions and more importantly, could favor cancer cell proliferation, migration and invasion/metastasis [4, 10]. Some of the most extensively examined mechanisms of aberrant glycosylation in cancer represent the alterations observed in the expression of glycotransferases, which are the enzymes responsible for the biosynthesis of glycoprotein and glycolipid sugar chains [11]. These enzymes are located in the endoplasmic reticulum and the Golgi apparatus [5]. Multiple studies have investigated the expression profiles of glycotransferases in different cancer types [11-14], and showed that oncogenic transformations of these enzymes are essentially regulated at their transcriptional level, thus emphasizing the importance of analyzing the glycotransferases expression patterns as inducers/markers for tumor progression and disease prognosis [12].

We have previously identified the polypeptide N-acetylgalactosaminyltransferase 3 (*GALNT3*) gene, a member of the GalNAc-transferases (GalNAc-Ts) gene family, as hypomethylated and overexpressed in high-grade serous EOC tumors, compared to low-malignant potential EOC tumors and normal ovarian tissues, as *GALNT3* expression was functionally related to EOC progression [14]. Moreover, a short hairpin RNA- (shRNA)-mediated *GALNT3* knockdown (KD) in the EOC cell line A2780s was associated with reduced mucin-1 (MUC1) protein expression, probably related to destabilization of the MUC1 protein due to lack of *GALNT3* glycosylation activity [14]. Taken together, our data are indicative of a strong oncogenic potential of the *GALNT3* gene in advanced EOC, and suggest that *GALNT3* overexpression may contribute to EOC dissemination through aberrant mucin O-glycosylation. Studying the role of individual GalNAc-T isoforms has been difficult in the past due to the small number of identified glycoproteins. O-Glycosylation directed by distinct GalNAc-Ts has been demonstrated to emerge as an important regulator of protein function [15], and thus a better understanding of the O-glycoproteome is vital for advancing our knowledge on the large GalNAc-T gene family and their role during cell differentiation and malignant progression.

In the present study, we performed a glycoproteomic analysis aimed to characterize *GALNT3*-dependent glycoprotein expression in EOC and its role in disease progression. Our methodological approach included the application of a bioorthogonal chemical reporter strategy for probing glycans [16, 17]. The method is implemented by metabolical labeling of glycoproteins with a monosaccharide precursor attached to a functional azido group [18]. This approach allows the labeling of glycans bearing azido sugars followed by selective capture from cell lysates [18]. Following metabolic labeling of *GALNT3*-KD and control A2780s EOC cells with a peracetylated azido GalNAc analog (peraceetylated N-azidoacetylgalactosamine, Ac4GalNAz), glycoproteomics was performed by mass spectrometry (MS) analysis of enriched glycoproteins, to better understand the molecular mechanisms of GALNT3 in EOC dissemination.

We show herein that *GALNT3* KD in A2780s cells was associated with the down-regulation of glycoproteins having significant functions in cancer progression including PTMs, cell adhesion and proliferation, regulation of protein transport, cholesterol biosynthesis, cellular metabolism, as well as with alterations of some cancer-related signaling pathways,

including the TGF-β, MAPK and NF-kappa-B pathways. Our glycoproteomic analysis led to the identification of novel glycoproteins potentially relevant to EOC etiology and confirmed our previous data of the oncogenic role of *GALNT3* in EOC dissemination.

2.5 Materials and methods

2.5.1 Cell culture

The *GALNT3* KD and control A2780s clones used in this study were previously generated, as described [14]. Briefly, a shRNA targeting the *GALNT3* sequence 5'-TACTGCTGAAGGAAATCAT-3' was designed using the siRNA Ambion Target Finder software (http://www.ambion.com/techlib/misc/siRNA_finder.html), and subcloned into the pSilencer 4.1 puro vector (Ambion). A2780s cells were stably transfected with this construct and clones were isolated as previously shown [14]. In addition, A2780s cells were mock-transfected with the pSilencer 4.1 puro vector, and the stably-transfected clones were isolated and used as controls.

2.5.2 Metabolite labeling and protein enrichment

Four cell culture dishes (150 mm) were seeded with 100 uM Ac4GalNAz metabolite. Another four cell culture dishes (150 mm) were seeded with 100 uM Ac4GalNAc as a control. Cell suspensions of control and *GALNT3* KD cells (2 × 105 cells/ml) were added to each tissue culture plate in Media 1 (RPMI 1640 media containing 10% FBS, 1% penicillin-streptomycin and 1 μg/ml puromycin). The cell culture dishes were incubated for 48 h at 37°C in a humidified 5% CO2 incubator. Following incubation, the cell culture plates were washed with PBS; then Media 2 (RPMI 1640 media containing 1% penicillin-streptomycin and 1 μg/ml puromycin) was added to the aspirated plates. The cell culture dishes were incubated for an additional 48 h at 37°C in a humidified 5% CO2 incubator. Media and cell lysates were harvested and concentrated as described previously [19, 20]. Briefly, to obtain the "conditioned media fraction", media from all conditions were harvested, cleared by centrifugation and spin concentrated (using Amicon, 15-ml 10-kDa spin filters) to a final volume of 1 ml. The concentrated residue was washed with PBS and transferred to an Eppendorf microcentrifuge tube. To obtain the "soluble" and "insoluble" fractions, adherent cells were washed with PBS and trypsinized for 5 min at 37 °C. Cells were harvested, centrifuged (300g, 3 min), and washed with PBS. Cell pellets were resuspended in lysis buffer (2 ml: 10 mM HEPES, pH 7.9, 1.5 mM MgCl2, 10 mM KCl,

0.5% Triton X-100, 1× protease inhibitors, 1 μM thiamet G), swelled for 5 min on ice, and disrupted by Dounce homogenization using a tight glass hand pestle (20-30 strokes) as described previously [19, 20]. Insoluble material was pelleted by centrifugation (3,700g, 10 min, at 4°C). The supernatant was collected as the "soluble fraction" while the pellet was kept as the "insoluble fraction." The insoluble fraction was resuspended in 1% RapiGest/PBS and briefly sonicated. The conditioned medium and soluble fractions were adjusted to 1% RapiGest/PBS as described before [19, 20]. Protein concentration from the three fractions was measured by bicinchonic acid assay (Pierce). Western blots were performed as described in [13]. Affinity purification via copper-catalyzed azide-alkyne cycloaddition (CuAAC) with a biotin–alkyne probe was next performed as previously shown ([19, 20]; see also [13]). Briefly, following CuAAC, protein aliquots were precipitated and solubilized. Streptavidin-agarose resin was added, and the resulting mixture was incubated for 12 h at 24°C with rotation. Beads were pelleted by centrifugation, and the supernatant containing uncaptured proteins was separated as described in [13]. The beads were then washed, reduced and alkylated, followed by on-bead trypsin digestion as shown in [13].

2.5.3 Mass spectrometry

Peptides resulting from trypsin digestion were resuspended in 15 μl of acetonitrile 2%, trifluoroacetic acid 0.05%, and 1 μl of each sample was analyzed by nanoLC/MS/MS. Peptides were injected and separated by online reversed-phase (RP) nanoscale capillary liquid chromatography (nanoLC) and analyzed by electrospray mass spectrometry (ESI MS/MS). The experiments were performed with a Dionex UltiMate 3000 nanoRSLC chromatography system (Thermo Fisher Scientific/Dionex Softron GmbH, Germering, Germany) connected to an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) equipped with a nanoelectrospray ion source. Peptides were trapped at 20 μl/min in loading solvent (2% acetonitrile, 0.05% trifluoroacetic acid) on a 5 mm x 300 um C18 pepmap cartridge pre-column (Thermo Fisher Scientific / Dionex Softron GmbH, Germering, Germany) during a 5 min period. Then, the pre-column was switched online with a self-made 50 cm x 75 μm internal diameter separation column packed with ReproSil-Pur C18-AQ 3-μm resin (Dr. Maisch HPLC GmbH, Ammerbuch-Entringen,

Germany) and the peptides were eluted with a linear gradient from 5-40% solvent B (A: 0.1% formic acid, B: 80% acetonitrile, 0.1% formic acid) for 90 mins, at 300 nl/min. Mass spectra were acquired using a data dependent acquisition mode using Thermo XCalibur software version 3.0.63. Full scan mass spectra (350 to 1800 m/z) were acquired in the Orbitrap using an Automatic Gain Control (AGC) target of 4e5, a maximum injection time of 50 ms and a resolution of 120 000. Internal calibration using lock mass on the m/z 445.12003 siloxane ion was used. Each MS scan was followed by acquisition of fragmentation MS spectra of the most intense ions for a total cycle time of 3 sec (top speed mode). The selected ions were isolated using the quadrupole analyzer in a window of 1.6 m/z and fragmented by Higher-energy Collision-induced Dissociation (HCD) with 35% of collision energy. The resulting fragments were detected by the linear ion trap in rapid scan rate with an AGC target of 1e4 and a maximum injection time of 50 ms. Dynamic exclusion of previously fragmented peptides was set for a period of 20 sec and a tolerance of 10 ppm.

2.5.4 Database searching and Label Free Quantification

Spectra were searched against a human proteins database (Uniprot Complete Proteome – taxonomy Homo sapiens – 69165 sequences) using the Andromeda module of MaxQuant software v. 1.5.2.8 [11, 21], as shown in [13]. For protein validation, a maximum false discovery rate of 1% at peptide and protein level was used based on a target/decoy search. MaxQuant was also applied for Label Free Quantification (LFQ), as described in [13].

2.5.5 Bioinformatic annotation & analysis

O- and N-glycoprotein prediction analysis using the NetOGlyc 4.0 and the NetNGlyc 1.0 servers was performed, as described in [13]. Gene ontology (GO) enrichment analysis was performed using AmiGO (http://amigo.geneontology.org), as shown in [13]. The GO term enrichment tool was used to determine the observed level of annotations for the set of proteins from our study and determine the significance in the context of all proteins annotated in the human proteome.

Functional, canonical pathway and network analyses were generated using the Ingenuity pathways analysis (IPA) software v7.1 (Ingenuity Systems, www.ingenuity.com). Proteins that met the expression ratio with a cut-off ≥ 2.0, and a *p*-value cut-off ≤ 0.05 for differential expression were considered for the analyses. Swiss-Prot accession numbers were inserted into the software along with corresponding comparison fold change ratios between our control and *GALNT3* KD protein data set. The corresponding Gene ID of each protein accession ID was then generated using the IPA software.

2.6 Results

2.6.1 Purification and LC/MS analysis of glycosylated proteins from control and *GALNT3* KD A2780s cells

In order to analyze the differential expression of mucin-type glycoproteins upon *GALNT3* KD in EOC cells, control and *GALNT3* KD A2780s cells were separately labeled with Ac4GalNAz or tetraacetylated N-acetylgalactosamine (Ac4GalNAc, negative control) (See Figure 2.7). Ac4GalNAz is intracellularly deacetylated and used to incorporate N-azidoacetylgalactosamine (GalNAz) to tag mucin type O-linked glycans [5]. It has been demonstrated that GalNAz is transformed to UDP-GalNAz *in vitro* and *in vivo*, enabling the labeling for proteins with O-linked glycans [5]. The labeled control and *GALNT3* KD A2780s cells were then subjected to subcellular fractionation to obtain the three fractions (conditioned media fraction, soluble fraction and insoluble fraction) followed by glycoprotein enrichment (see Figure 2.7). Incorporation of azido-sugar into glycoproteins from all three fractions was examined by Western blot analysis of the collected lysates. Anti-biotin signal was checked before affinity-capture (Load) and after affinity-capture on the fraction not bound to the beads (Supernatant), as well as on the fractions that included the bead after washing (Capture) from all there fractions collected from both the control and *GALNT3* KD A2780s EOC cells (see Figure 2.8). Glycoproteins from the conditioned media fraction incubated with Ac4GalNAz showed a strong robust labeling, whereas labeling in the soluble and insoluble fractions were not as efficient, as evidenced in lane C in both control and *GALNT3* KD blots (see Figure 2.8). Trypsin digestion was then performed and the released peptides were analyzed by LC-MS/MS mass spectrometry. The analysis of the three subcellular fractions from the control and *GALNT3* KD A2780s cells cultured with Ac4GalNAz lead to the identification of a total of 5111 proteins by searching the data against the UniProt Complete Proteome Human database. Subtraction of proteins found in the Ac4GalNAc fraction, identified a total of 2120 proteins in the Ac4GalNAz treated samples. Prediction analysis of the 2120 proteins dataset using the NetOGlyc 4.0 server indicated that the majority of these proteins (1870 proteins, representing ~88%) could be O-glycosylated (Figure 2.1A). Moreover, we examined if our list of 2120 proteins

also contain N-glycosylation prediction sites; for that we used the NetNGlyc 1.0 server, and interestingly 1576 (~74%) of the annotated proteins were predicted to be N-glycosylated (Figure 2.1A). Comparison of our list of annotated glycoproteins to previously published assigned glycopeptides from key references [19, 22-31], focused on isolating and/or sequencing a profound number of glycoproteins, indicated that 638 (~30%) of the 2120 proteins identified in our study have been formerly characterized as glycoproteins (Figure 2.1B), while the remaining proteins have not been previously known to be glycosylated, suggesting that many more glycoproteins, and particularly O-glycoproteins, remain to be discovered.

2.6.2 Comparative proteomic analysis of differentially regulated glycoproteins identified between control and *GALNT3* KD A2780s EOC cells.

We further examined the differential regulation of glycoproteins upon *GALNT3* KD in EOC cells, based on our identified set of 2120 proteins. Metabolic labeling and high mass accuracy MS were applied to retrieve a maximum level of proteomic data. We investigated the three subcellular fractions by quantitative proteomics using MaxQuant software and Andromeda search engine (included in MaxQuant) [11, 32]. Proteins were considered as differentially expressed between control and *GALNT3* KD A2780s cells if they met the following criteria: Welch test p-value ≤ 0.05 and fold change in relative expression of ≥ 2. Based on these stringent selection criteria, we found a total of 275 glycoproteins to be up-regulated and 314 glycoproteins to be down-regulated following *GALNT3* KD in A2780s cells. These criteria were similar to those described in the literature on the use of label-free quantification methods to define biological regulation evaluated at the proteome level [33-36]. Alternative statistical analysis applying z-score statistics (as described in [34]) gave almost identical results with 95% overlap of the data obtained with the previous analysis. Thus, data obtained with the initial selection criteria (Welch test p-value ≤ 0.05 and fold change ≥ 2), were used for all consecutive analyses (see section 2.4 for a more detailed description of the statistical tests performed on the datasets). Following these analyses, a quantifiable difference between control A2780s and *GALNT3* KD cells was found for 327,

118, and 144 proteins in the conditioned media, the soluble and the insoluble fractions, respectively. In the conditioned media fraction, 104 of the identified proteins were up-regulated and 223 were down-regulated. In the soluble fraction comparison, 63 of the identified proteins were up-regulated and 55 were down-regulated. In the insoluble fraction comparison, 108 of the identified proteins were up-regulated and 36 were down-regulated. Volcano plots -log10 (*p*-value) vs. log2 (fold change of *GALNT3* KD/Control) were also constructed to graphically display the quantitative data from each of the three fractions (Figure 2.8A-C).

2.6.3 Cellular classification of differentially regulated proteins identified between control and *GALNT3* KD A2780s EOC cells

Cellular component GO analysis of the differentially regulated glycoproteins identified between the control and *GALNT3* KD A2780s cells was performed on each of the identified fractions (conditioned media fraction, cytosolic fraction and nuclear fraction), and data were compared to the entire human proteome using the GO Consortium for enrichment analysis [12]. Analysis from the conditioned media fraction showed a clear overrepresentation in membrane type and extracellular, as well as organelle compartment proteins (Figure 2.3; also see Figure 3.3). Interestingly, cytoplasmic and ER proteins were also clearly represented in this fraction (Figure 2.3; also see Figure 3.3). O-glycosylation of ER type proteins has not been extensively examined, as our metabolic labeling approach has revealed a strong overrepresentation of glycoproteins in the ER organelle. In contrast, in addition to membrane-bound organelle proteins, proteins localized to the intracellular, organelle and cytoplasm compartments were significantly enriched in the soluble fraction (Figure 2.3; also see Figure 3.3). Finally, protein localization analysis from the insoluble fraction indicated enrichment in intracellular and organelle compartments, as expected (Figure 2.3; also see Figure 3.3). These observations indicate that *GALNT3* KD was predominantly and significantly associated with glycoproteins alterations in the membrane, cytoplasmic and secreted/extracellular space compartments (Figure 2.3; also see Figure 3.3). Western blot assays confirmed the enrichment of the glycoproteins in the conditioned media fraction (see Figure 2.8). Overall our data indicate that most of the labeled

glycoproteins arising from the Ac4GalNAz treatment maybe of the O-GalNAz-type, as judged by the relative enrichment of membrane glycoproteins in the conditioned media fraction (Figure 2.3; also see Figure 3.3).

2.6.4 Pathways and network analyses following *GALNT3* gene KD in EOC cells

The functional and canonical pathway and network analyses of differentially expressed glycoproteins identified upon *GALNT3* KD in A2780s cells were explored using the IPA software. Here, the 589 identified up- and down-regulated glycoproteins were uploaded in IPA and their matched gene symbols were used for consecutive analyses.

We illustrated the functional changes induced upon KD of the *GALNT3* gene by separately comparing the functional categories that were up- and down-regulated in the three studied fractions (conditioned media, soluble and insoluble fractions). The top 15 most significantly perturbed up- and down-regulated molecular and cellular functions are presented in Figure 2.4. As seen from Figure 2.4A, the major up-regulated functional categories were related to essential cellular mechanisms including molecular transport, small molecule biochemistry, cell-to-cell signaling and interaction, cellular assembly and organization, cellular function and maintenance, cellular development, cellular movement, cell morphology, cellular compromise, protein trafficking, while notably, PTM-related pathways were found to be both induced and suppressed (Figure 2.4A and 2.4B). Down-regulated functional pathways were mostly implicated in energy production and metabolism (including nucleic acid, lipid, amino acid and carbohydrate metabolism), and protein synthesis (Figure 2.4B).

IPA canonical pathway analysis supported these findings, as the top 18 up-regulated canonical pathways were mostly related to sphingosine 1-phosphate signaling, epithelial adherens junction signaling, *ILK* signaling, integrin signaling, ovarian cancer signaling, *p38 MAPK* signaling and *AMPK* signaling (Figure 2.5A), while the top 18 down-regulated canonical pathways were associated with mitochondrial dysfunction, oxidative phosphorylation, cholesterol biosynthesis, leucine degradation, isoleucine degradation,

valine degradation, autophagy, and UDP-N-acetyl-D-galactosamine biosynthesis I (Figure 2.5B).

Network analyses confirmed the major functionally related gene groups that were found to be differentially expressed in the *GALNT3* KD clone compared to the control A2780s cells from the three fractions examined. We identified 25 highly significant networks with score ≥ 13. The five top-scoring networks were associated with functions linked to PTM, cardiovascular disease, inflammatory disease, lipid metabolism, embryonic development, cell morphology, and cellular assembly and organization (Table 2.1). A common network obtained upon merging the five top-scoring networks (Figure 2.6) recognized several important nodes linked with numerous interaction partners, including transforming growth factor beta (TGF-β1), mitogen-activated protein kinases (*MAPK*), phosphoinositide 3-kinase complex (*PI3K*), sphingomyelin phosphodiesterase 1 (*SMPD1*), tubulin beta 3 (*TUBB3*), magnesium superoxide dismutase (*SOD2*), mitochondrial folate-coupled dehydrogenase (*MTHFD2*), autophagy protein 5 (*ATG5*), microtubule-associated protein 1B (*MAP1B*), dolichyl-diphosphooligosaccharide-protein glycosyltransferase (*DDOST*), dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3B (*STT3B*), in addition to several prominent cancer-related genes stemming from the nuclear factor kappa-light-chain-enhancer of activated B complex (*NF-Kappa B*) such as metadherin (*MTDH*), annexin A4 (*ANXA4*), and ribonucleotide reductase M2 (*RRMP2*).

2.7 Discussion

Aberrant glycosylation has been previously described and investigated in EOC (reviewed in [37]). Thus, a number of studies suggest the involvement of N-glycosylation in EOC spreading and metastasis. By using a quantitative glycoproteomic analysis, Tian *et al*. (2011) have demonstrated that different N-linked glycoproteins are associated with/could represent biomarkers for the specific EOC subtypes, including high grade and low grade serous, mucinous, high grade and low grade endometrioid, clear-cell, transitional-cell, squamous-cell, mixed and undifferentiated subtypes, as some N-glycoproteins are exclusively expressed in normal ovaries [38]. It was also shown that the N-glycan-dependent mesothelin-MUC16 binding facilitates EOC peritoneal dissemination [39]. Two recent studies have described metastasis-related N-glycan alterations in EOC secretomes *in vitro* [40, 41], as definite N-glycan substructures and their complexes were found to be associated with specific epigenetic programming of synthetic enzymes in EOC [41]. Moreover, specific sialylated N-glycoproteins were exclusively identified in biological fluids from EOC patients [42]. Aberrant N-glycosylation has been also involved in EOC chemoresistance, as N-glycosylation inhibitors were shown to markedly reduce multidrug resistance [43].

Aberrant O-linked (mucin-type) glycosylation has been less extensively studied and is poorly defined in EOC. Several studies have demonstrated that MUC1 is overexpressed and aberrantly O-glycosylated in most ovarian adenocarcinomas [44], including primary and metastatic EOC tumors [45]. The major structural difference observed in tumor-associated MUC1, when compared to its normal counterpart, is that tumor-associated MUC1 contains shorter and less dense O-glycan chains, thus exposing more regions of the MUC1 protein core [46]. This feature of the tumor-associated MUC1 makes it a more accessible EOC therapeutic target, and thus delineate novel approaches for more effective EOC therapy. To date, 20 GalNAc-Ts have been identified that are responsible for the initiation of mucin type O-glycosylation. The differential glycosylation of mucins in cancer is caused by the dysregulation of the glycosylation machinery that include expression alterations of specific glycotransferases, including GalNAc-Ts, which consecutively leads to the formation of aberrant glycan structures implicated in cancer growth and metastasis.

Alterations of the expression of the large family of 20 GalNAc-Ts have been investigated using RNA interference approaches including small interfering RNA (siRNA), and shRNA in a number of cancers (reviewed in [37, 47]), which reinforces the conviction that GalNAc-Ts may play essential roles in the mechanisms leading to tumor dissemination and disease progression. Indeed, a direct correlation has been revealed between elevated expression of *GALNT6* in breast cancer cells and increased glycosylation and surface expression of MUC1 [48]. It has been also shown that *GALNT14* contributes to the aberrant glycosylation and overexpression of MUC13 in EOC cells, suggesting a role of *GALNT14* in EOC carcinogenesis [49]. We have previously demonstrated that *GALNT3* KD is associated with increased adhesion and decreased MUC1 stability in EOC cells, suggesting for *GALNT3* involvement in EOC progression via aberrant MUC1 glycosylation [14]. Overall, these data provide evidence that altered expression of different members of the GalNAc-Ts gene family can contribute to aberrant O-glycosylation associated with tumor progression and dissemination, also emphasizing the potential use of these enzymes as cancer biomarkers and/or therapeutic targets.

Global analysis of proteins and their modifications has been majorly dependent on mass spectrometry tools for sequencing N- and O-glycans from cells and tissues. Multiple factors have contributed to difficulties in specifically performing O-glycoprotein analysis and have thus complicated the advancement in determining and understanding the functions of O-glycosylation in normal or disease conditions. More recently, advances in mass-spectrometry have overcome some of the challenges accompanied with the complexity of glycosylation, including enrichment protocols aimed at labeling O-linked glycosylation. Thus, extensive research has introduced some enrichment methods including lectin affinity chromatography, which has been used for specifically analyzing O-GlcNAc and O-GalNAc type glycoproteins [33, 50], while other approaches have been developed towards a more targeted O-glycan selection including Vicia villosa agglutinin (VVA) enrichment for O-glycans combined with serial lectin chromatography methods [34].

In this study, we have characterized the role of the *GALNT3* enzyme in aberrant O-glycosylation of mucin-like targets in EOC cells using a bioorthogonal chemical reporter strategy that involves metabolic labeling of glycans with a monosaccharide precursor (Ac4GalNAz) carrying a bioorthogonal azido group, which allows a linker attachment to

glycoproteins (including mucin-type O-linked glycoproteins) for consecutive labeling and isolation/purification ([36]; also see Figure 2.7). To the best of our knowledge, this is the first study that has used metabolic labeling to identify the impact of a GalNAc-T inhibition on altering the glycoproteins pattern in an EOC cell line. Moreover, the approach used allowed us to identify numerous novel glycoproteins differentially regulated in EOC cells.

To date, only limited numbers of glycoproteins have been characterized, which reflects the low percentage of glycoprotein entries annotated in SWISS-PROT [5]. This fact is mainly due to difficulties in performing experimental determination of glycosylation sites [5]. Here, we used the NetOGlyc 4.0 server, which is a developed glycosylation site prediction method for O-glycosylation [51]. Our prediction analysis estimated that more than 88% of the identified proteins may be O-glycosylated (Figure 2.1A). In addition to performing NetOGlyc type analysis, we further compared our list of annotated proteins to previously published studies [19, 22-31] (Figure 2.1B), and results indicated that more than 30% of the identified proteins are O-glycosylated (Figure 2.1B). Together, our comparative analyses suggest that the O-glycoproteome is much larger than previously predicted, and a more extensive examination may reveal more interesting numbers of glycosylated proteins, than those currently found in SWISS-PROT.

Additionally, using cellular fractionation, glycoprotein enrichment and high-resolution LC/MS-based quantitative proteomics, we quantified over 2000 proteins. Among these proteins, we found several hundreds to be differentially expressed (≥ 2 fold, p-value ≤ 0.05) between the control and *GALNT3* KD A2780s EOC cells (Figure 2.2). GO cellular component classifications revealed that these differentially regulated proteins were predominantly mapped to membrane, extracellular, and cytoplasmic and intracellular organelle compartments (Figure 2.3; also see Figure 3.3). More interestingly, we observed a relative enrichment in proteins identified in the ER compartment, and recent evidence suggests that the GalNAc-Ts can relocate to the ER, indicating a possible role of the GalNAc-Ts in protein O-glycosylation of ER proteins [52].

It has been previously demonstrated that Ac4GalNAz treatment of human cells results in the metabolic labeling of mucin-type glycoproteins, in addition to the labeling of extracellular N-glycan and intracellular O-GlcNAcylated proteins [19, 20, 50, 53]. The biosynthesis of both UDP- GalNAz and GlcNAz from Ac4GalNAz is conducted by the

mammalian enzyme GALE, which interconverts UDP-N-acetylglucosamine (UDP-GlcNAc) and its C4 epimer, UDP-N-acetylgalactosamine (UDP-GalNAc) [18]. UDP-GalNAc is the common substrate of the GalNAc-Ts, which is derived primarily from UDP-GlcNAc via GALE and alternatively, UDP-GalNAc can be generated from GalNAc by the action of GalNAc 1-kinase and UDP-GalNAc pyrophosphorylase, representing enzymes of the salvage pathway [47]. Thus, metabolic labeling with Ac4GalNAz results in heterogeneous labeling of nuclear and cytoplasmic proteins, in addition to membrane-bound mucin-type glycoproteins; such a robust labeling by Ac4GalNAz can allow for a better comprehensive measurement of the GalNAc-Ts activities on diverse aspects of cellular functions, and especially in human diseases such as cancer. GO and Western blot analysis were indicative that in both control and *GALNT3* KD cells, the majority of the Ac4GalNAz labeling occurred in the conditioned media fraction, as compared to lower labeling in the soluble and insoluble fractions (Figure 2.3; also see Figure 2.8, and Figure 2.9). This data indicates that EOC cells may have an inactive or defective *GALE* activity, as *GALE* deficiency has been reported to decrease the production of UDP-GalNAc from UDP-GlcNAc, which in turn affects glycolipid and glycoprotein biosynthesis [54]. Moreover, IPA pathway analyses illustrated that the *GALE* enzyme is significantly down-regulated in the *GALNT3* KD cells, suggesting that *GALE* activity has been additionally perturbed by *GALNT3* KD in A2780s EOC cells. Collectively, our data are indicative for *GALNT3*-mediated modulation of *GALE* expression in EOC, possibly associated with alterations in the GALE glycosylation pattern.

The IPA functional pathway analyses of our proteomics data have revealed that the *GALNT3* gene suppression resulted in reduced expression of essential pathways related to lipid, carbohydrate and amino acid metabolism, also demonstrating that changes in glycogen metabolism can provoke crucial alterations in various EOC metabolic pathways (Figure 2.4B). This was further confirmed by IPA canonical pathway analyses (Figure 2.5), as the top 5 canonical pathways (PTM, cardiovascular and inflammatory disease, lipid metabolism, embryonic development and cell morphology, and cellular assembly) were significantly perturbed upon *GALNT3* KD (Table 2.1). Our data also support others' findings for the implications of members of the GalNAc-Ts gene family in controlling cellular metabolism. Indeed, a role of *GALNT2* in controlling lipid metabolism was

postulated, as *GALNT2* was functionally associated with the glycosylation of high-density lipoprotein cholesterol and/or triglycerides [55]. Interestingly, the canonical pathway associated with degradation of the branched-chain amino acids (BCAAs) leucine, isoleucine and valine was found to be among the significantly suppressed canonical pathways following *GALNT3* KD (Figure 2.5B). BCAAs degradation is regulated by the *BCAT1* and *BCAT2* enzymes [34]. We have recently shown that the suppression of the *BCAT1* gene in EOC cells resulted in the down-regulation of numerous genes implicated in lipid production and protein synthesis, suggesting its important role in controlling EOC metabolism [56]. Our data suggest a similar effect imposed by the suppression of the *GALNT3* gene.

Additionally, *GALNT3* KD was associated with the down-regulation of the metabolism related gene nodes *SOD2* and *MTHFD2* (Figure 2.6), previously shown to be overexpressed in different cancer types, including EOC [33, 57]. *SOD2* encodes the enzyme manganese superoxide dismutase (*MnSOD*), which is a mitochondrial enzyme responsible for the reduction of free radicals and thus acts by protecting cells from oxidative stress, and *SOD2* over-expression was demonstrated to be involved in EOC development [33]. *MTHFD2* represents an enzyme of the mitochondrial metabolic pathway implicated in cancer cell proliferation, as *MTHFD2* has displayed strong overexpression in different tumors including EOC tumors [57]. Our findings suggest that metabolic pathways regulated via genes of the oxidative stress and/or mitochondrial enzymes might be controlled by PTMs such as glycosylation. *GALNT3* KD also resulted in the upregulation of the *SMPD1* gene node. *SMPD1* represents an acid sphingomyelinase, responsible for the hydrolysis of sphingomyelin to phosphorylchole ceramide, and the increase in ceramide production leads to cell death showing a link between ceramide expression and EOC [58]. *SMPD1* expression correlates with enhanced cell death in *GALNT3* KD cells compared to their corresponding controls [14].

Moreover, two oligosaccharyltransferases (OST) were found to be down-regulated upon *GALNT3* KD. The main function of OSTs is the transfer of lipid-linked oligosaccharides to selected asparagine residues within the consensus sequence Asn-Xaa-Ser/Thr on nascent polypeptides [59]. OST mediated glycosylation is shared between N- and O-glycosylation pathways [60]. The OST subunits examined from our network analysis include DDOST

and STT3B (Figure 2.6), which were significantly down-regulated upon *GALNT3* KD in EOC cells.

Network analysis also demonstrated that *GALNT3* KD directs the down-regulation of several major gene nodes and corresponding pathways previously shown to be associated with EOC etiology, including *TGF-β, MAPK*, and *NF-kappa-B* (Figure 2.6). Accordingly, some members of the *NF-kappa-B* pathway (*MTDH, ANXA4* and *RRM2*; see Figure 2.6) were found to be significantly suppressed in *GALNT3* KD EOC cells. *MTDH* (also known as *AEG1*) was repeatedly shown to be implicated in EOC dissemination and chemoresistance [61-63] and has been recognized as an independent prognostic biomarker for metastatic EOC tumors [64]. *ANXA4* is up-regulated in EOC, and is being used as molecular marker to distinguish the different EOC subtypes [65]. Similarly, *RRM2* was shown to be over-expressed in EOC patients with advanced disease [66], and was recently identified as a potential target for pro-senescence EOC therapy [67]. Moreover, several gene nodes, implicated in EOC chemoresistance (including *TUBB3, MAP1B* and *ATG5*; see Figure 2.6), displayed a significant decrease upon *GALNT3* KD. Indeed, augmented expression of *TUBB3* and *MAP1B* has been associated with worse prognosis and paclitaxel resistance in EOC patients [68, 69]. Also, knockdown of autophagy-related *ATG5* gene was shown to increase cisplatin-induced growth inhibition by enhancing cisplatin-induced apoptosis in EOC [70].

Finally, the lipid kinases *PI3K* complex, known to be highly activated in human cancers [71], was up-regulated upon *GALNT3* KD. The PI3K complex controls multiple cellular signaling pathways that play key regulatory roles in cell survival, differentiation, and proliferation [71, 72]. Interestingly, recent reports have shown that alterations in protein glycosylation can be implicated in the activation of multiple oncogenic signaling pathways including the *PI3K* pathway in different cancer types [71]. These glycosylation alterations are usually caused by a premature stop in proteins O-glycosylation due to sialylation, mediated by the cancer-associated sialyl-Tn antigen. Notably, it has been shown that in cancer, the enhanced expression of sialyltransferases is often associated with shorter O-GalNAc glycan structures [5]. Another gene node *B4GALT1*, known as an oncogene, was also found to be up-regulated following *GALNT3* KD. *B4GALT1* synthesizes Galβ1-4GlcNAc (N-acetyllactosamine) by transferring galactose from UDP-Gal to an acceptor

sugar molecule, and thus represents a key enzyme in glycobiology [73]. Several studies have reported *B4GALT1* to play an essential role in the EOC dissemination [74, 75]. The observed up-regulations of the *PI3K* complex and the *B4GALT1* gene may be due to the frequently observed redundancy associated with overlapping functions of other GalNAc-Ts enzymes [76], partially compensating for the KD of *GALNT3* in EOC cells.

Conclusions

In this study, we performed a quantitative proteomic analysis of the GalNAz-metabolically labeled proteins in both control and *GALNT3* KD A2780s cells. The use of the azido analog GalNAz for metabolic labeling and enrichment of glycoproteins in control and *GALNT3* KD A2780s cells, coupled with MS analysis, has allowed us to more quantitatively characterize the role of the *GALNT3* enzyme in aberrant glycosylation of mucin-like targets in EOC cells and better understand the role of this transferase in EOC dissemination. This study has led to the identification of novel glycoproteins expressed in EOC and more importantly, sets of glycoproteins whose expression is altered by *GALNT3* KD, indicative for a potential role of *GALNT3* in modulating PTMs and metabolism pathways in EOC cells, which could significantly impact disease development. However, we do not exclude that certain compensatory changes may occur upon *GALNT3* KD due to some level of redundancy amongst the family of 20 GalNAc-Ts. Further studies focused on these genes may provide valuable insights into the pathophysiological roles of the GalNAc-T enzymes in EOC. Consecutive analyses should also include validation of these data in EOC tumor sets displaying high or low *GALNT3* expression.

Furthermore, changes in glycosylation patterns in cancer are associated not only with differential expression of glycoproteins [77], but also with alterations in their glycan structures [78], as a common feature of tumors is the overexpression of truncated O-glycans. We anticipate further expansion of our analyses by the application of targeted O-glycoproteomic strategies that do not reduce O-glycan structures, so that information about the site and glycan structures can be studied simultaneously [19, 20]. The application of such an innovative approach could allow for a better comprehension of the role of the *GALNT3* enzyme and other members of the GalNAc-T isoforms in aberrant EOC glycosylation, associated with disease dissemination.

2.8 Acknowledgements

This study was sustained by grant from the Cancer Research Society of Canada to D.B., as well as by grants from Jane Coffin Childs Fund to C.M.W., Burroughs Wellcome Fund CASI to C.M.W., National Institutes of Health (CA200423) to C.R.B. and Howard Hughes Medical Institute to C.R.B.

2.9 References

[1] Siegel R, Ward E, Brawley O, Jemal A. Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. CA: a cancer journal for clinicians. 2011;61:212-36.

[2] Fruscio R, Corso S, Ceppi L, Garavaglia D, Garbi A, Floriani I, et al. Conservative management of early-stage epithelial ovarian cancer: results of a large retrospective series. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO. 2013;24:138-44.

[3] Alouini S. Management of ovarian cancer has changed. Gynecologic oncology. 2012;126:313; author reply 4.

[4] Yarema KJ, Bertozzi CR. Characterizing glycosylation pathways. Genome biology. 2001;2:REVIEWS0004.

[5] Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochimica et biophysica acta. 1999;1473:4-8.

[6] Paszek MJ, DuFort CC, Rossier O, Bainer R, Mouw JK, Godula K, et al. The cancer glycocalyx mechanically primes integrin-mediated growth and survival. Nature. 2014;511:319-25.

[7] Zachara NE, Hart GW. Cell signaling, the essential role of O-GlcNAc! Biochimica et biophysica acta. 2006;1761:599-617.

[8] Parodi AJ. Protein glucosylation and its role in protein folding. Annual review of biochemistry. 2000;69:69-93.

[9] Fuster MM, Esko JD. The sweet and sour of cancer: glycans as novel therapeutic targets. Nature reviews Cancer. 2005;5:526-42.

[10] Pinho SS, Reis CA. Glycosylation in cancer: mechanisms and clinical implications. Nature reviews Cancer. 2015;15:540-55.

[11] Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: a peptide search engine integrated into the MaxQuant environment. Journal of proteome research. 2011;10:1794-805.

[12] Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, Lewis S. AmiGO: online access to ontology and annotation data. Bioinformatics. 2009;25:288-9.

[13] Sheta R, Roux-Dalvai F, Woo CM, Fournier F, Bourassa S, Bertozzi CR, et al. Proteomic dataset for altered glycoprotein expression upon GALNT3 knockdown in ovarian cancer cells. Data in Brief. 2016;In press.

[14] Wang ZQ, Bachvarova M, Morin C, Plante M, Gregoire J, Renaud MC, et al. Role of the polypeptide N-acetylgalactosaminyltransferase 3 in ovarian cancer progression: possible implications in abnormal mucin O-glycosylation. Oncotarget. 2014;5:544-60.

[15] Bennett EP, Mandel U, Clausen H, Gerken TA, Fritz TA, Tabak LA. Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. Glycobiology. 2012;22:736-56.

[16] Sletten EM, Bertozzi CR. Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. Angew Chem Int Ed Engl. 2009;48:6974-98.

[17] Laughlin ST, Bertozzi CR. Metabolic labeling of glycans with azido sugars and subsequent glycan-profiling and visualization via Staudinger ligation. Nature protocols. 2007;2:2930-44.

[18] Hang HC, Yu C, Kato DL, Bertozzi CR. A metabolic labeling approach toward proteomic analysis of mucin-type O-linked glycosylation. Proceedings of the National Academy of Sciences of the United States of America. 2003;100:14846-51.

[19] Woo CM, Iavarone AT, Spiciarich DR, Palaniappan KK, Bertozzi CR. Isotope-targeted glycoproteomics (IsoTaG): a mass-independent platform for intact N- and O-glycopeptide discovery and analysis. Nature methods. 2015;12:561-7.

[20] Woo CM, Bertozzi CR. Isotope Targeted Glycoproteomics (IsoTaG) to Characterize Intact, Metabolically Labeled Glycopeptides from Complex Proteomes. Current protocols in chemical biology. 2016;8:59-82.

[21] Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature biotechnology. 2008;26:1367-72.

[22] Trinidad JC, Schoepfer R, Burlingame AL, Medzihradszky KF. N- and O-glycosylation in the murine synaptosome. Molecular & cellular proteomics : MCP. 2013;12:3474-88.

[23] Hornbeck PV, Kornhauser JM, Tkachev S, Zhang B, Skrzypek E, Murray B, et al. PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. Nucleic acids research. 2012;40:D261-70.

[24] Halim A, Ruetschi U, Larson G, Nilsson J. LC-MS/MS characterization of O-glycosylation sites and glycan structures of human cerebrospinal fluid glycoproteins. Journal of proteome research. 2013;12:573-84.

[25] Trinidad JC, Barkan DT, Gulledge BF, Thalhammer A, Sali A, Schoepfer R, et al. Global identification and characterization of both O-GlcNAcylation and phosphorylation at the murine synapse. Molecular & cellular proteomics : MCP. 2012;11:215-29.

[26] Hahne H, Sobotzki N, Nyberg T, Helm D, Borodkin VS, van Aalten DM, et al. Proteome wide purification and identification of O-GlcNAc-modified proteins using click chemistry and mass spectrometry. Journal of proteome research. 2013;12:927-36.

[27] Zaro BW, Yang YY, Hang HC, Pratt MR. Chemical reporters for fluorescent detection and identification of O-GlcNAc-modified proteins reveal glycosylation of the ubiquitin ligase NEDD4-1. Proceedings of the National Academy of Sciences of the United States of America. 2011;108:8146-51.

[28] Chuh KN, Zaro BW, Piller F, Piller V, Pratt MR. Changes in metabolic chemical reporter structure yield a selective probe of O-GlcNAc modification. Journal of the American Chemical Society. 2014;136:12283-95.

[29] Alfaro JF, Gong CX, Monroe ME, Aldrich JT, Clauss TR, Purvine SO, et al. Tandem mass spectrometry identifies many mouse brain O-GlcNAcylated proteins including EGF domain-specific O-GlcNAc transferase targets. Proceedings of the National Academy of Sciences of the United States of America. 2012;109:7280-5.

[30] Hubbard SC, Boyce M, McVaugh CT, Peehl DM, Bertozzi CR. Cell surface glycoproteomic analysis of prostate cancer-derived PC-3 cells. Bioorganic & medicinal chemistry letters. 2011;21:4945-50.

[31] Khidekel N, Ficarro SB, Peters EC, Hsieh-Wilson LC. Exploring the O-GlcNAc proteome: direct identification of O-GlcNAc-modified proteins from the brain. Proceedings of the National Academy of Sciences of the United States of America. 2004;101:13132-7.

[32] Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. The EMBO journal. 2013;32:1478-88.

[33] Sun N, Pan C, Nickell S, Mann M, Baumeister W, Nagy I. Quantitative proteome and transcriptome analysis of the archaeon Thermoplasma acidophilum cultured under aerobic and anaerobic conditions. Journal of proteome research. 2010;9:4839-50.

[34] Ramus C, Hovasse A, Marcellin M, Hesse AM, Mouton-Barbosa E, Bouyssie D, et al. Spiked proteomic standard dataset for testing label-free quantitative software and statistical methods. Data Brief. 2016;6:286-94.

[35] Zhang P, Guo Z, Zhang Y, Gao Z, Ji N, Wang D, et al. A preliminary quantitative proteomic analysis of glioblastoma pseudoprogression. Proteome science. 2015;13:12.

[36] Gautier V, Mouton-Barbosa E, Bouyssie D, Delcourt N, Beau M, Girard JP, et al. Label-free quantification and shotgun analysis of complex proteomes by one-dimensional SDS-PAGE/NanoLC-MS: evaluation for the large scale analysis of inflammatory human endothelial cells. Molecular & cellular proteomics : MCP. 2012;11:527-39.

[37] Abbott KL. Glycomic analysis of ovarian cancer: past, present, and future. Cancer biomarkers : section A of Disease markers. 2010;8:273-80.

[38] Tian Y, Yao Z, Roden RB, Zhang H. Identification of glycoproteins associated with different histological subtypes of ovarian tumors using quantitative glycoproteomics. Proteomics. 2011;11:4677-87.

[39] Gubbels JA, Belisle J, Onda M, Rancourt C, Migneault M, Ho M, et al. Mesothelin-MUC16 binding is a high affinity, N-glycan dependent interaction that facilitates peritoneal metastasis of ovarian tumors. Molecular cancer. 2006;5:50.

[40] Zhang X, Wang Y, Qian Y, Wu X, Zhang Z, Liu X, et al. Discovery of specific metastasis-related N-glycan alterations in epithelial ovarian cancer based on quantitative glycomics. PloS one. 2014;9:e87978.

[41] Sheta R, Woo CM, Roux-Dalvai F, Fournier F, Bourassa S, Droit A, et al. A metabolic labeling approach for glycoproteomic analysis reveals altered glycoprotein expression upon GALNT3 knockdown in ovarian cancer cells. J Proteomics. 2016;In press.

[42] Kuzmanov U, Musrap N, Kosanam H, Smith CR, Batruch I, Dimitromanolakis A, et al. Glycoproteomic identification of potential glycoprotein biomarkers in ovarian cancer proximal fluids. Clinical chemistry and laboratory medicine. 2013;51:1467-76.

[43] Wojtowicz K, Szaflarski W, Januchowski R, Zawierucha P, Nowicki M, Zabel M. Inhibitors of N-glycosylation as a potential tool for analysis of the mechanism of action and cellular localisation of glycoprotein P. Acta biochimica Polonica. 2012;59:445-50.

[44] Taylor-Papadimitriou J, Burchell J, Miles DW, Dalziel M. MUC1 and cancer. Biochimica et biophysica acta. 1999;1455:301-13.

[45] Wang L, Ma J, Liu F, Yu Q, Chu G, Perkins AC, et al. Expression of MUC1 in primary and metastatic human epithelial ovarian cancer and its therapeutic significance. Gynecologic oncology. 2007;105:695-702.

[46] Price MR, Hudecz F, O'Sullivan C, Baldwin RW, Edwards PM, Tendler SJ. Immunological and structural features of the protein core of human polymorphic epithelial mucin. Molecular immunology. 1990;27:795-802.

[47] Sheta R, Bachvarov D. Role of aberrant glycosylation in ovarian cancer dissemination. Biochemical Revies. 2014;25:83-92.

[48] Park JH, Nishidate T, Kijima K, Ohashi T, Takegawa K, Fujikane T, et al. Critical roles of mucin 1 glycosylation by transactivated polypeptide N-acetylgalactosaminyltransferase 6 in mammary carcinogenesis. Cancer research. 2010;70:2759-69.

[49] Wang R, Yu C, Zhao D, Wu M, Yang Z. The mucin-type glycosylating enzyme polypeptide N-acetylgalactosaminyltransferase 14 promotes the migration of ovarian cancer by modifying mucin 13. Oncology reports. 2013;30:667-76.

[50] Boyce M, Carrico IS, Ganguli AS, Yu SH, Hangauer MJ, Hubbard SC, et al. Metabolic cross-talk allows labeling of O-linked beta-N-acetylglucosamine-modified proteins via the N-acetylgalactosamine salvage pathway. Proceedings of the National Academy of Sciences of the United States of America. 2011;108:3141-6.

[51] Julenius K, Molgaard A, Gupta R, Brunak S. Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. Glycobiology. 2005;15:153-64.

[52] Gill DJ, Clausen H, Bard F. Location, location, location: new insights into O-GalNAc protein glycosylation. Trends in cell biology. 2011;21:149-58.

[53] Vocadlo DJ, Hang HC, Kim EJ, Hanover JA, Bertozzi CR. A chemical approach for identifying O-GlcNAc-modified proteins in cells. Proceedings of the National Academy of Sciences of the United States of America. 2003;100:9116-21.

[54] Lai K, Elsas LJ, Wierenga KJ. Galactose toxicity in animals. IUBMB life. 2009;61:1063-74.

[55] Schjoldager KT, Vakhrushev SY, Kong Y, Steentoft C, Nudelman AS, Pedersen NB, et al. Probing isoform-specific functions of polypeptide GalNAc-transferases using zinc finger nuclease glycoengineered SimpleCells. Proceedings of the National Academy of Sciences of the United States of America. 2012;109:9893-8.

[56] Sheta R, Woo CM, Roux-Dalvai F, Fournier F, Bourassa S, Droit A, et al. A metabolic labeling approach for glycoproteomic analysis reveals altered glycoprotein expression upon GALNT3 knockdown in ovarian cancer cells. Journal of proteomics. 2016;Submitted.

[57] Nilsson R, Jain M, Madhusudhan N, Sheppard NG, Strittmatter L, Kampf C, et al. Metabolic enzyme expression highlights a key role for MTHFD2 and the mitochondrial folate pathway in cancer. Nature communications. 2014;5:3128.

[58] Smith EL, Schuchman EH. Acid sphingomyelinase overexpression enhances the antineoplastic effects of irradiation in vitro and in vivo. Molecular therapy : the journal of the American Society of Gene Therapy. 2008;16:1565-71.

[59] Pan S, Chen R, Tamura Y, Crispin DA, Lai LA, May DH, et al. Quantitative glycoproteomics analysis reveals changes in N-glycosylation level associated with pancreatic ductal adenocarcinoma. Journal of proteome research. 2014;13:1293-306.

[60] Dell A, Galadari A, Sastre F, Hitchen P. Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes. International journal of microbiology. 2010;2010:148178.

[61] Zhou B, Yang J, Shu B, Liu K, Xue L, Su N, et al. Overexpression of astrocyte-elevated gene-1 is associated with ovarian cancer development and progression. Molecular medicine reports. 2015;11:2981-90.

[62] Haug S, Schnerch D, Halbach S, Mastroianni J, Dumit VI, Follo M, et al. Metadherin exon 11 skipping variant enhances metastatic spread of ovarian cancer. International journal of cancer Journal international du cancer. 2015;136:2328-40.

[63] Li C, Liu J, Lu R, Yu G, Wang X, Zhao Y, et al. AEG -1 overexpression: a novel indicator for peritoneal dissemination and lymph node metastasis in epithelial ovarian cancers. International journal of gynecological cancer : official journal of the International Gynecological Cancer Society. 2011;21:602-8.

[64] Li C, Chen K, Cai J, Shi QT, Li Y, Li L, et al. Astrocyte elevated gene-1: a novel independent prognostic biomarker for metastatic ovarian tumors. Tumor biology : the journal of the International Society for Oncodevelopmental Biology and Medicine. 2014;35:3079-85.

[65] Kim A, Enomoto T, Serada S, Ueda Y, Takahashi T, Ripley B, et al. Enhanced expression of Annexin A4 in clear cell carcinoma of the ovary and its association with chemoresistance to carboplatin. International journal of cancer Journal international du cancer. 2009;125:2316-22.

[66] Wang LM, Lu FF, Zhang SY, Yao RY, Xing XM, Wei ZM. Overexpression of catalytic subunit M2 in patients with ovarian cancer. Chinese medical journal. 2012;125:2151-6.

[67] Aird KM, Li H, Xin F, Konstantinopoulos PA, Zhang R. Identification of ribonucleotide reductase M2 as a potential target for pro-senescence therapy in epithelial ovarian cancer. Cell Cycle. 2014;13:199-207.

[68] Ferrandina G, Zannoni GF, Martinelli E, Paglia A, Gallotta V, Mozzetti S, et al. Class III beta-tubulin overexpression is a marker of poor clinical outcome in advanced ovarian cancer patients. Clinical cancer research : an official journal of the American Association for Cancer Research. 2006;12:2774-9.

[69] Yu Y, Gaillard S, Phillip JM, Huang TC, Pinto SM, Tessarollo NG, et al. Inhibition of Spleen Tyrosine Kinase Potentiates Paclitaxel-Induced Cytotoxicity in Ovarian Cancer Cells by Stabilizing Microtubules. Cancer cell. 2015;28:82-96.

[70] Wang J, Wu GS. Role of autophagy in cisplatin resistance in ovarian cancer cells. The Journal of biological chemistry. 2014;289:17163-73.

[71] Liu P, Cheng H, Roberts TM, Zhao JJ. Targeting the phosphoinositide 3-kinase pathway in cancer. Nature reviews Drug discovery. 2009;8:627-44.

[72] Engelman JA, Luo J, Cantley LC. The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. Nature reviews Genetics. 2006;7:606-19.

[73] Pennell DJ, Underwood SR, Ell PJ. Symptomatic bradycardia complicating the use of intravenous dipyridamole for thallium-201 myocardial perfusion imaging. International journal of cardiology. 1990;27:272-4.

[74] Yamashita H, Kubushiro K, Ma J, Fujii T, Tsukazaki K, Iwamori M, et al. Alteration in the metastatic potential of ovarian cancer cells by transfection of the antisense gene of beta-1,4-galactosyltransferase. Oncology reports. 2003;10:1857-62.

[75] Ma J, Kubushiro K, Tashima Y, Tsukazaki K, Udagawa Y, Nozawa S, et al. Expression of human beta 1,4-galactosyltransferase in gynecological cancer cell lines. International journal of oncology. 1997;11:117-22.

[76] Plantner JJ, Lentrichia BB, Kean EL. Biogenesis and content of rhodopsin in the retina of the chick during development. Current eye research. 1988;7:503-10.

[77] Tousi F, Hancock WS, Hincapie M. Technologies and strategies for glycoproteomics and glycomics and their application to clinical biomarker research. Analytical Methods. 2011;3:20-32.

[78] Kolarich D, Jensen PH, Altmann F, Packer NH. Determination of site-specific glycan heterogeneity on glycoproteins. Nature protocols. 2012;7:1285-98.

2.10 Figure and table legends

Figure 2.1. Glycoprotein predictions. A. Comparative analysis of O-glycoprotein and N-glycoprotein predictions using the NetOGlyc 4.0 and NetNGlyc 1.0 servers. B. Comparative analysis of the proteins identified in our study to glycoproteins previously identified in the literature.

Figure 2.2. Volcano plots illustrate differentially regulated proteins. The quantified proteins in A. Conditioned media fraction, B. Soluble fraction and C. Insoluble fraction were plotted on Volcano graphs – The x-axis represents the $\log_2$ (ratio) where 'ratio' is the ratio of the averages of LFQ Intensities of *GALNT3* KD over control. The y-axis represents the $-\log_{10}$ (*p*-value) where '*p*-value' is the *p*-value associated to the statistical Welch test applied on the three technical replicates of each *GALNT3* and control samples. The non-axial vertical dotted lines denote ±2-fold change while the non-axial horizontal line denotes $p \leq 0.05$. The colored dots showed proteins considered as up- (red) or down- (blue) regulated, while the non-regulated proteins are display as gray dots.

Figure 2.3. Cellular component Gene Ontology analysis. Bar graphs showing the cellular component GO terms that are significantly over represented from the differentially regulated proteins in our study, compared to the entire human proteome. Data were submitted to the Gene Ontology Consortium for enrichment analysis. The analysis was performed on the differentially regulated proteins identified from each of the three fractions: Conditioned media fraction (red bars), Soluble fraction (green bars) and Insoluble fraction (purple bars) as compared to the entire human proteome (blue bars). All identified proteins annotated with GO cellular component terms are compared against the annotated human proteome. Categories were scored by a combination of percentage of enrichment and *p*-value ($p \leq 0.05$). The enrichment *p*-values are corrected by Benjamini's methods.

Figure 2.4. Functional pathway analysis. Functional pathway analysis for a dataset of differentially expressed glycoproteins (≥ 2-fold; as based on their matched gene symbols) following *GALNT3* suppression in A2780s cells, as observed in the three fractions (conditioned media, soluble and insoluble fractions). A. Functional pathway analysis of up-

regulated genes; B. Functional analysis of down-regulated genes. Dark blue indicates the conditioned media fraction functional network gene expression values, light blue indicates the soluble fraction functional network gene expression values; cyan blue indicates the insoluble fraction functional network gene expression values. The functional pathways included in this analysis are shown along the x-axis of the bar chart. The y-axis indicates the statistical significance on the left. Calculated using the right-tailed Fisher exact test, the *p*-value indicates which biologic annotations are significantly associated with the input molecules relative to all functionally characterized mammalian molecules. The yellow threshold line represents the default significance cutoff at $p = 0.05$.

Figure 2.5. Canonical pathway analysis. Canonical pathway analysis for a dataset of differentially expressed genes (≥ 2-fold; as based on their matched gene symbols) following *GALNT3* suppression in A2780s cells, as observed in the three fractions (conditioned media, soluble and insoluble fractions). A. Canonical pathway analysis of up-regulated genes; B. Canonical pathway analysis of down-regulated genes. Dark blue indicates the conditioned media fraction canonical network gene expression values, light blue indicates the soluble fraction canonical network gene expression values; cyan blue indicates the insoluble fraction canonical network gene expression values. The canonical pathways included in this analysis are shown along the x-axis of the bar chart. The y-axis indicates the statistical significance on the left. Calculated using the right-tailed Fisher exact test, the *p*-value indicates which biologic annotations are significantly associated with the input molecules relative to all functionally characterized mammalian molecules. The yellow threshold line represents the default significance cutoff at $p = 0.05$.

Figure 2.6. Network analysis of dynamic gene expression in A2780s cells based on the 2-fold glycoprotein expression list obtained following *GALNT3* KD. The five top-scoring networks of up- and down-regulated genes were merged and are displayed graphically as nodes (genes/gene products) and edges (the biological relationships between the nodes). Intensity of the node color indicates the degree of up- (red) or down-regulation (green). Nodes are displayed using various shapes that represent the functional class of the gene product (square, cytokine, vertical oval, transmembrane receptor, rectangle, nuclear

receptor, diamond, enzyme, rhomboid, transporter, hexagon, translation factor, horizontal oval, transcription factor, circle, other). Edges are displayed with various labels that describe the nature of relationship between the nodes: __ binding only, → acts on. The length of an edge reflects the evidence supporting that node-to-node relationship, in that edges supported by article from literature are shorter. Dotted edges represent indirect interaction.

Table 2.1. IPA networks identified via analysis of focus genes. Analysis of differentially expressed genes using IPA revealed a set of gene networks based on known interactions. Up-regulated genes are shown in (red) and down-regulated genes are shown in (green). Genes not altered in our signature are indicated in plain text. The table represents networks 1 to 5 as illustrated in Figure 2.7. The *p*-values are calculated using Fisher's exact test.

2.11 Tables

Table 2.1. The five top-scoring genetic networks affected following glycoproteomics analysis of *GALNT3* KD in A2780s EOC cells

Molecules in network	Focus molecules	Top diseases and functions	p–Value
AAGAB,ALDH2,Ap2alpha,BIN1,CHGA,CHMP4B,CLPTM1L,CLTA,CNBP,DLGAP4,DUB,FDXR,FGFR1,GALC,Importin beta,MLLT1,MTCH2,MTHFD2,NUP107,NUP153,PI3K (complex),PSAP,RAB4A,RABEP2,RING1,SEC61A1,SEC61B,SNX4,STAMBP,SV2A,USP11,USP25,USP28,USP30,USP32	31	Post–translational modification	(1.49E–09 – 4.73E–03)
		Nervous system development and function	(1.16E–05 – 4.90E–02)
		Neurological disease	(1.16E–05 – 4.70E–02)
AASS,ACAA2,ALDH3A2,Alphatubulin,BCAT2,CPQ,FSTL3,GNS,GPAA1,H1FX,HADH,KDM5B,LPCAT3,MAN2B2,MAP1B,MARCH5,MBD1,MBD3,MFN1,NLGN1,PCDH10,PLCB1,RAB18,RAD23B,Rnr,SLC4A2,Smad2/3,SNTB2,STOML2,SUSD5,TDG,TGFB1,TUBB3,TUBB2A,tubulin (family)	31	Cardiac Inflammation	(8.27E–06 – 5.24E–01)
		Cardiovascular disease	(8.27E–06 – 5.32E–01)
		Inflammatory disease	(8.27E–06 – 2.23E–01)
ANXA4,B4GALT1,C1R,CC2D1A,CCDC22,COMMD8,COMMD9,Cpla2,CSTF2,CTSC,DDOST,DDRGK1,DOHH,ERC1,FUCA1,Igh (family),Ikb,IL–1R,KBTBD7,MLEC,MTDH,NFkB (complex),PIP5K1A,PLA2G3,PLOD2,POM121/POM121C,RALGAPA1,SCLY,SMPD1,SMPD2,sphingomyelinase,STT3B,Tnf receptor,TOM1,TRIP6	28	Lipid metabolism	(4.27E–07 – 4.62E–02)
		Molecular transport	(4.27E–07 – 4.62E–02)
		Small molecule biochemistry	(4.60E–06 – 4.73E–02)
ANXA4,B4GALT1,C1R,CC2D1A,CCDC22,COMMD8,COMMD9,Cpla2,CSTF2,CTSC,DDOST,DDRGK1,DOHH,ERC1,FUCA1,Igh (family),Ikb,IL–1R,KBTBD7,MLEC,MTDH,NFkB (complex),PIP5K1A,PLA2G3,PLOD2,POM121/POM121C,RALGAPA1,SCLY,SMPD1,SMPD2,sphingomyelinase,STT3B,Tnf receptor,TOM1,TRIP6	27	Embryonic development	(4.00E–05 – 1.95E–01)
		Nervous system development and Function	(4.00E–05 – 2.62E–02)
		Organ development	(4.00E–05 – 4.01E–01)
ACAD9,ADAM15,Akt,ATG5,ATP synthase,ATP5F1,ATP5H,ATPAF2,CLPB,Collagen type VI,ECHS1,EPHA7,FOLH1,IRS,LAMTOR1,LAMTOR2,MT–ATP6,MTORC2,NCBP2,NFkB (family),Pak,PDGF (family),Pdgf Ab,PDGFC,PIK3R2,PMPCB,RPIA,RRM2,SARNP,SLC25A4,TAGLN2,TIMM13,TIMM17B,TOMM22,WTAP	26	Cell morphology	(4.16E–04 – 4.26E–02)
		Cellular assembly and organization,	(4.16E–04 – 4.72E–02)
		Hereditary disorder	(5.61E–04 – 1.05E–02)

.

2.12 Figures

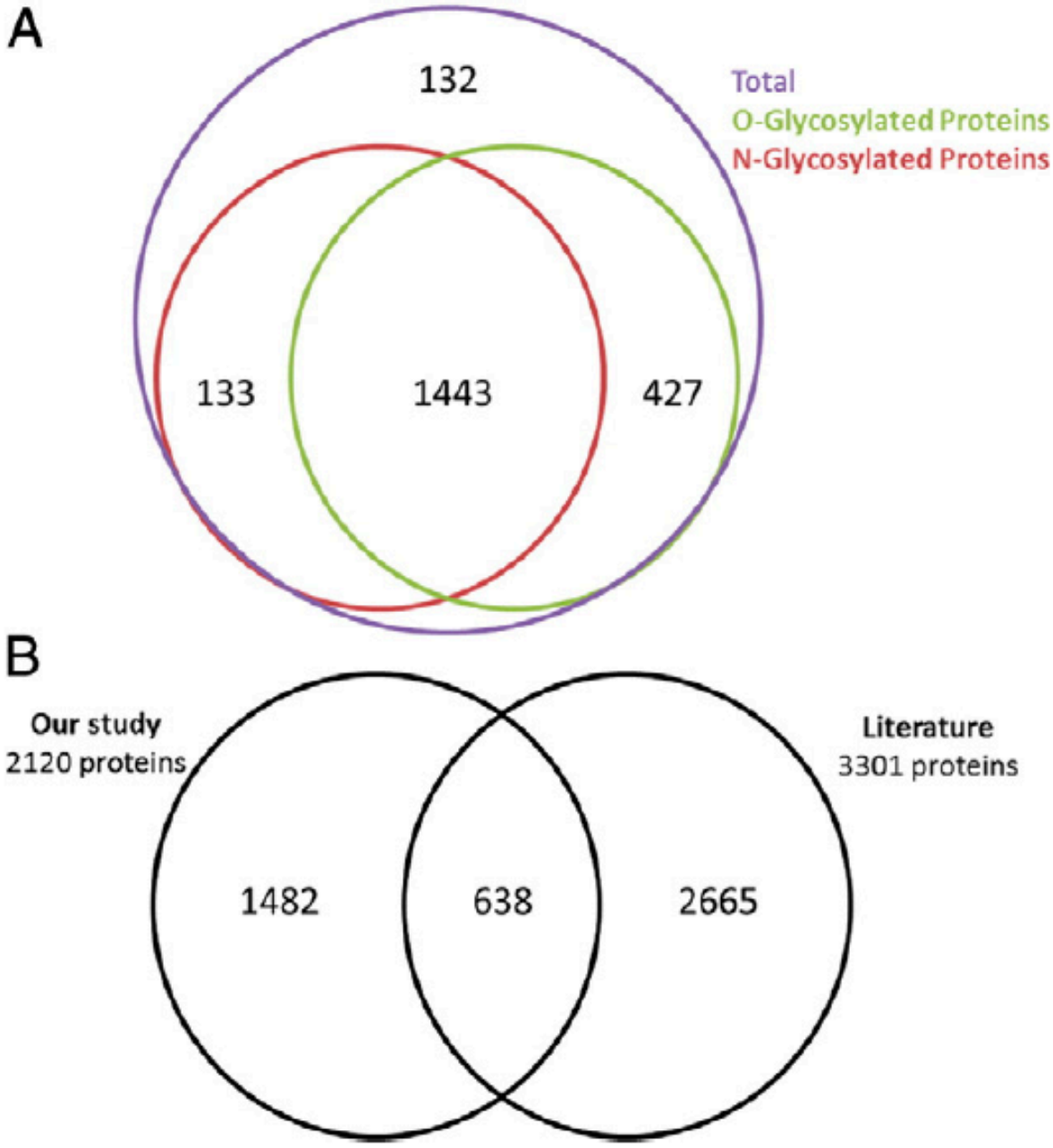


Figure 2.1. Glycoproteinwss predictions

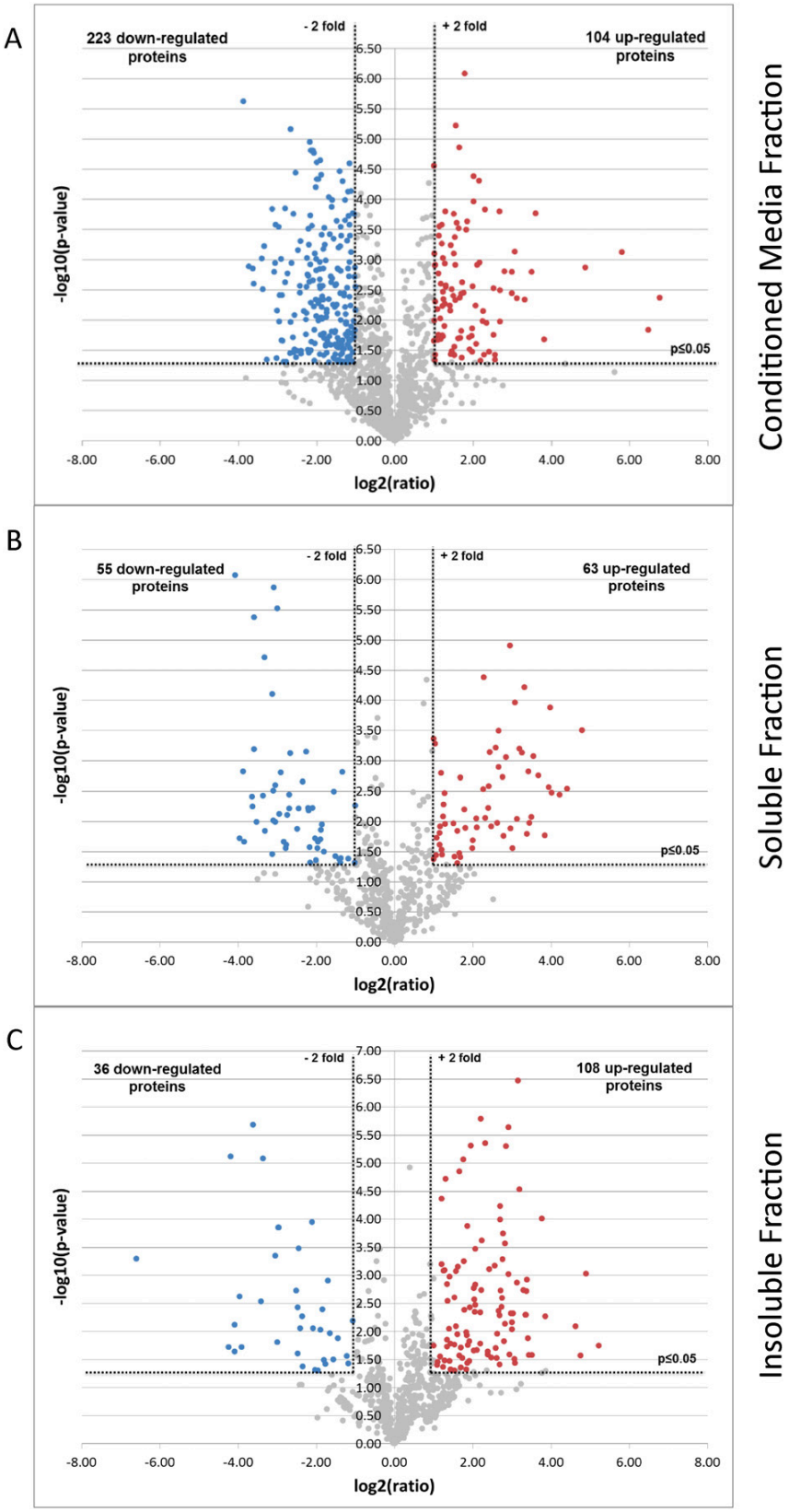


Figure 2.2. Volcano plots illustrate differentially regulated proteins

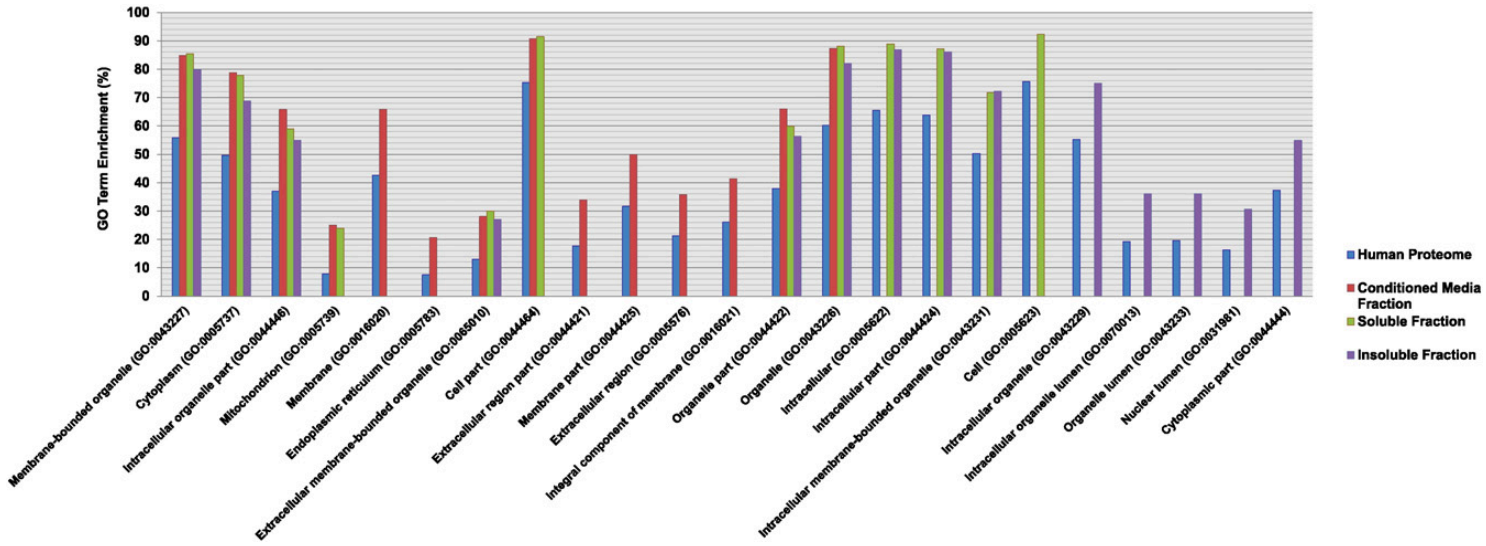
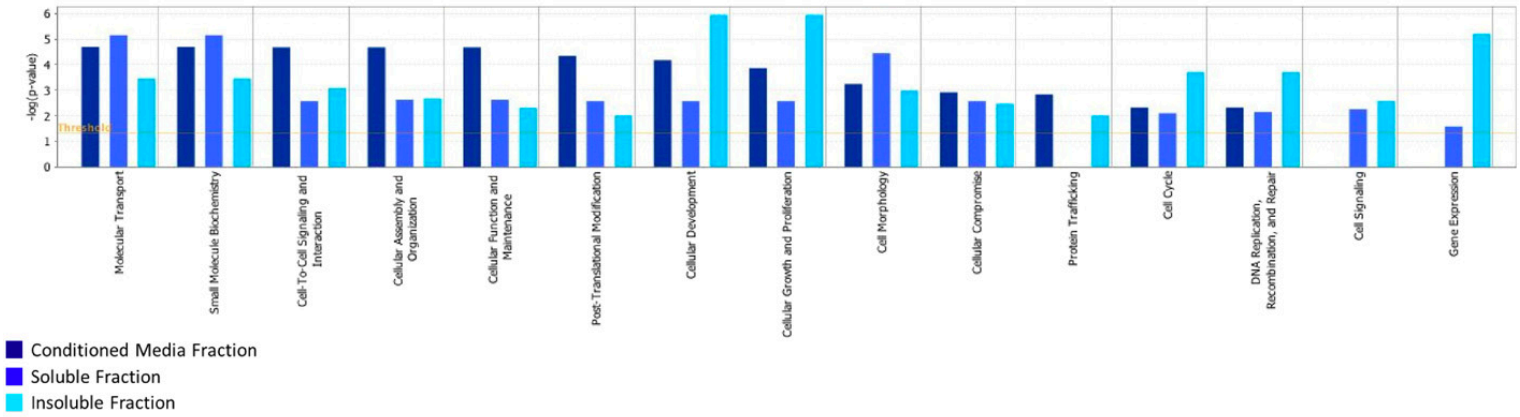


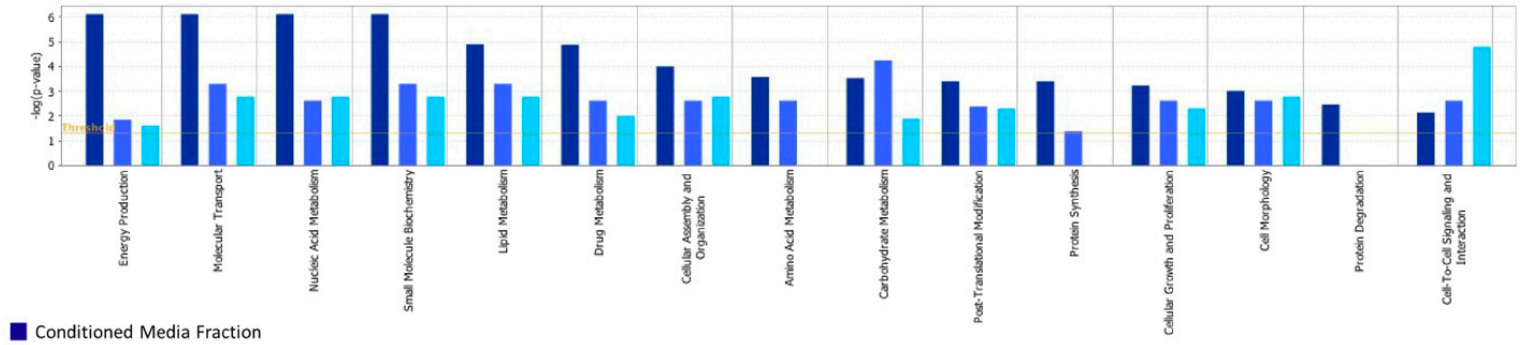
Figure 2.3. Cellular component Gene Ontology analys

A. Functional pathways, up-regulated upon *GALNT3* knockdown in A2780s cells.



Conditioned Media Fraction
Soluble Fraction
Insoluble Fraction

B. Functional pathways, down-regulated upon *GALNT3* knockdown in A2780s cells.



Conditioned Media Fraction
Soluble Fraction
Insoluble Fraction

Figure 2.4. Functional pathway analysis

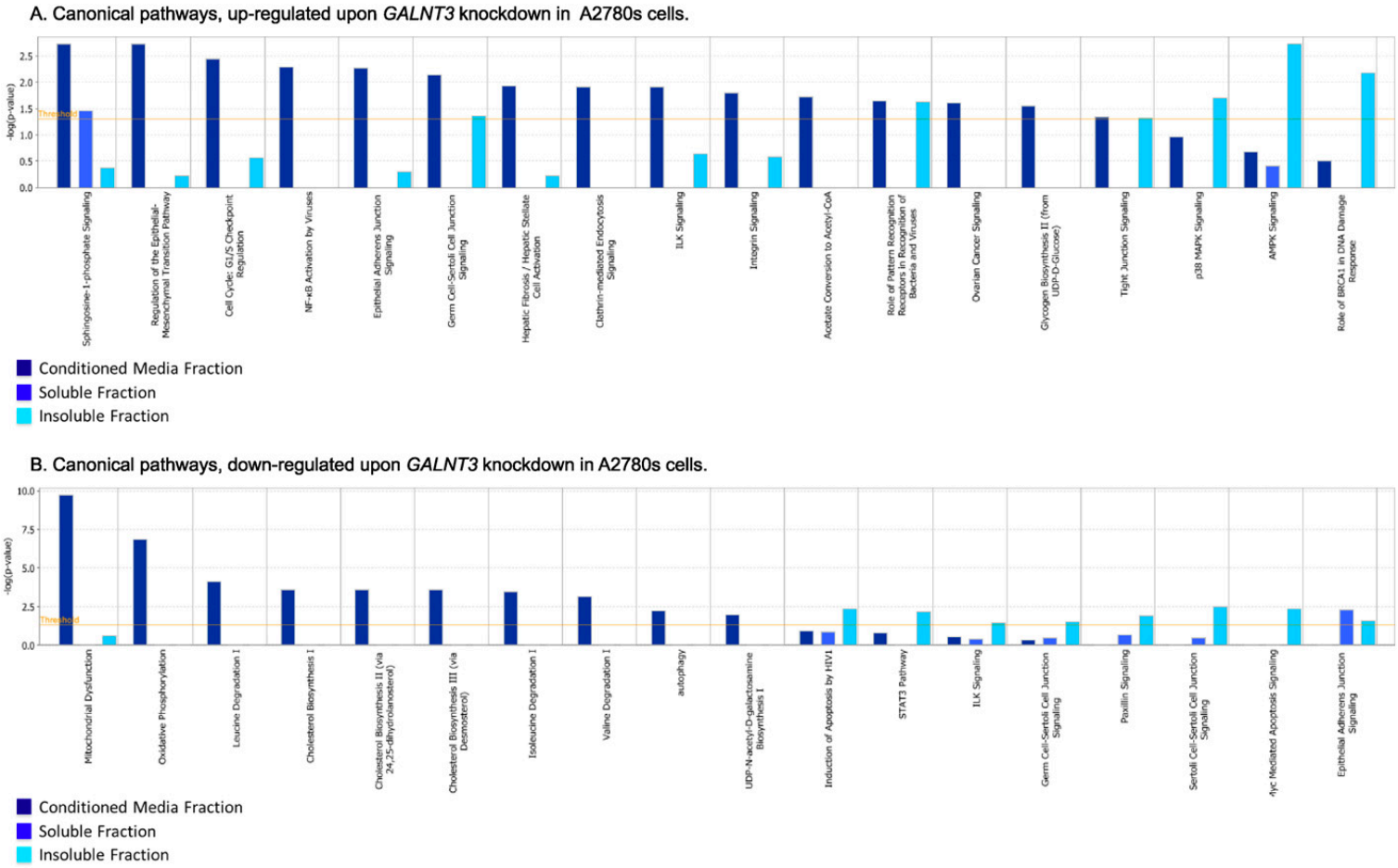


Figure 2.5. Canonical pathway analysis

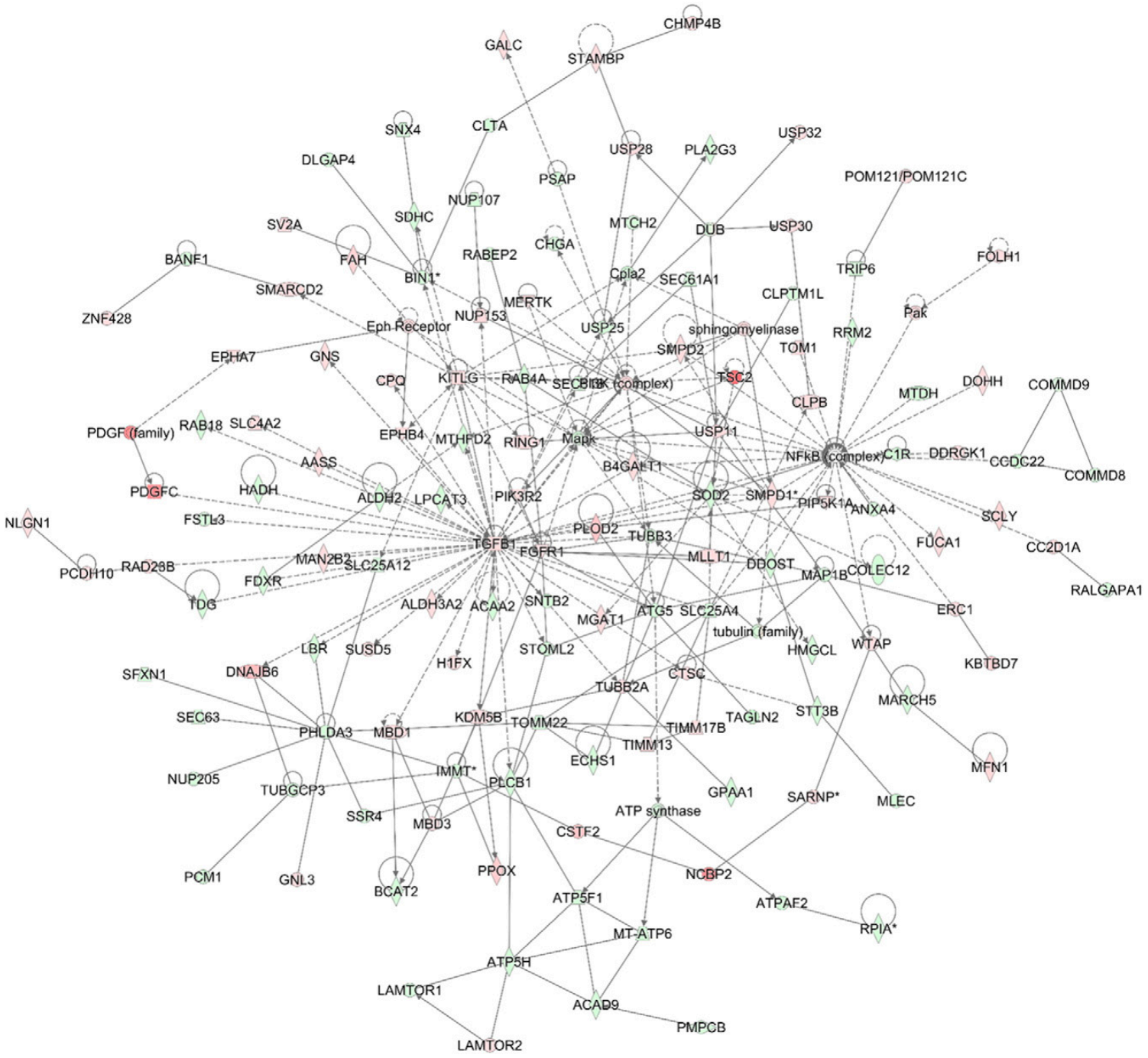


Figure 2.6. Network analysis of dynamic gene expression in A2780s cells based on the 2-fold glycoprotein expression list obtained following *GALNT3* KD

Chapter 2.B: Proteomic dataset for altered glycoprotein expression upon *GALNT3* knockdown in ovarian cancer cells

(Originally published in Data Brief, May 2016, PMID: 27331112)

Specifications Table

Subject area	Biology
More specific subject area	Oncology, Proteomics
Type of data	Table, figure
How data was acquired	Mass spectrometry, Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA)
Data format	Raw, analyzed
Experimental factors	Metabolite labeling of glycoproteins from extracellular/membrane bound, cytoplasmic and nuclear fractions of ovarian cancer cells, followed by trypsin digestion and glycoproteomics enrichment using Click Chemistry, and subjected to nanoLC and analyzed by ESI MS/MS.
Experimental features	Subcellular fractionation, ESI MS/MS peptide identification, data analysis in MaxQuant followed by data predictions using the NetOGlyc 4.0 and NetNGlyc 1.0 servers, and enrichment analysis using GO Consortium for enrichment analysis.
Data source location	Quebec City, Canada
Data accessibility	Data is with this article

Value of the data

- The presented list of differentially regulated glycoproteins identified upon *GALNT3* KD in EOC cells could represent novel putative biomarkers/molecular targets involved in EOC metastasis and thus the data presented here can be a useful resource to examine some of these biomarkers.

- The metabolic labeling approach applied in this study followed by the MS analysis could be a useful tool/guide for the quantification and the identification of glycoproteins from different cell lines.

The data presented herein provide a comprehensive list of newly identified glycoproteins, which strongly suggests that the metabolic labelling approach applied can essentially

increase the magnitude of recognized glycoproteins by comparing to organism-specific database for a more complete level of identification.

2.13 Data

The datasets provided in this article represent the entire list of identified glycoproteins after GalNAz metabolic labeling in both control and *GALNT3* KD A2780s cells, in addition to the processed data identifying the quantitatively significant list of differentially regulated proteins between the control and *GALNT3* KD cells. Data represented here also include comparative analysis of identified glycoproteins with previously published glycoproteins' data. This was also supported by data presenting predictive analysis performed by investigating for possible glycosylation sites from the list of the identified proteins for further confirmation. Finally, data of protein enrichment analysis performed were included as a representation of the cellular localization of the assigned glycoproteins from our list of differentially regulated proteins.

2.14 Experimental design, Materials and Methods

We applied a bioorthogonal chemical reporter strategy [2, 3] for analyzing differential glycoprotein expression following *GALNT3* KD in the EOC cell line A2780s. Figure 2.7 represents a schematic overview of the glycoproteomic workflow used. The method is explicitly used for metabolically labeling glycans with a monosaccharide precursor attached to a functional azido group [4]. Control and *GALNT3* KD A2780s cells were separately labeled with tetraacetylated N-azidoacetylgalactosamine (Ac4GalNAz) or tetraacetylated N-acetylgalactosamine (Ac4GalNAc, negative control). The labeled control and *GALNT3* KD A2780s cells were then subjected to subcellular fractionation (conditioned media fraction, soluble fraction and insoluble fraction) followed by glycoprotein enrichment (Figure 2.7). A Western blot analysis was performed to examine the enrichment efficiency (Figure 2.8). Trypsin digestion was then performed and the released peptides were analyzed by LC-MS/MS mass spectrometry. For each sample labeled with Ac4GalNaz, three technical replicates were performed in order to get statistical values on intensity measurements, while single injections of Ac4GalNAc samples were done for evaluation of non-specific binding on streptavidin-agarose resin.

The data displays the total number of proteins identified in the three subcellular fractions of the control and *GALNT3* KD A2780s cells cultured with Ac4GalNAz, as well as the subtracted proteins, exclusively found in the Ac4GalNAc (negative control) fraction. Analyses of these data using the NetOGlyc 4.0 and the NetNGlyc 1.0 servers generated a list of proteins with predicted O- and N-glycosylation sites. Additionally, The data contains a list of proteins identified in our study that have been previously characterized as glycoproteins in the literature.

The MaxQuant software and Andromeda search engine (included in MaxQuant) [5, 6] were consecutively used to generate a list of differentially regulated proteins identified in the three subcellular fractions of A2780s cells upon *GALNT3* KD, as based on the following criteria: Welch test p-value ≤ 0.05 and fold change in relative expression of ≥ 2 similar to that applied in [7] and [8]. Cellular component Gene Ontology (GO) analysis of the differentially regulated glycoproteins identified between the control and *GALNT3* KD A2780s cells was additionally performed on each of the identified fractions (conditioned

media fraction, cytosolic fraction and nuclear fraction), and data were compared to the entire human proteome using the GO Consortium for enrichment analysis (Figure 2.9).

2.14.1 Chemical Glycoproteomics Enrichment using Click Chemistry

The first part of the platform protocol applied in this study represents a method used to metabolically label glycoproteins in cell culture, as described in [1]. Briefly, cells were separately labeled with Ac4GalNAz and Ac4GalNAc (Figure 2.7). We started by isolating and enriching glycoproteins from different biological fractions, including proteins secreted into the media (conditioned media fraction), as well as proteins enriched in the cytosolic and nuclear and fractions (soluble and insoluble fractions respectively), as shown in [1].

The next step of the protocol was the chemical glycoproteomics enrichment procedure, which included tagging glycoproteins with Click Chemistry (Figure 2.7). The Copper-Catalyzed Azide-Alkyne Cycloaddition (CuAAC) enrichment was performed, as previously described [9, 10]. GalNAz and GalNAc labeled cell fractions were divided into aliquots, each containing 3 mg of protein. Click-chemistry reagents (200 μM alkynyl biotin probe, 300 μM copper sulfate, 600 μM BTTP, and 2.5 mM sodium ascorbate) were pre-mixed and added, and the reaction was incubated for 3.5 h at 24°C. Proteins were precipitated, resuspended and then solubilized as previously described [9, 10]. Briefly, protein pellets were resuspended in 400 μl 1% RapiGest/PBS and solubilized by probe sonication. Streptavidin-agarose resin was first washed with PBS and then added to the samples, and the resulting mixture was incubated for 12 h at 24°C with rotation. The beads were then pelleted by centrifugation, and the supernatant containing uncaptured proteins was removed as a separate fraction (Figure 2.7). The beads were then washed with 1% Rapigest, 6 M urea and PBS; beads were then pelleted by centrifugation and resuspended in PBS. Samples were subjected to reduction and alkylation, as previously described [9, 10]. Briefly, proteins were reduced by the addition of 5 mM DTT followed by alkylation completed by the addition of 10 mM iodoacetamide. Trypsin was then added to the slurry of beads and the resulting mixture was incubated for 12 h at 37ºC. The beads were pelleted and the supernatant digest was collected (Figure 2.7). The trypsin fraction was concentrated to dryness using a Speedvac set to 40ºC. Samples were desalted by ZipTip P10 for subsequent MS analysis (Figure 2.7).

2.14.2 Western Blot Analysis

Western blot analyses were performed on protein lysates collected from both the control and *GALNT3* KD A2780s EOC cells. Whole-cell lysates labeled with 100 μM Ac4GalNAz were incubated with a biotinylated bioorthogonal probe. Biotinylated glycoproteins were enriched from the supernatant by affinity-capture with streptavidin–agarose beads. To each aliquot collected during the enrichment procedure, 3 μl of 4X SDS buffer was added and the aliquots were loaded to 5% polyacrylamide gels. Proteins were then transferred to nitrocellulose membranes, which were consecutively incubated with Ponceau stain (Figure 2.8). The membranes were then blocked with 2% bovine serum albumin in Tris-buffered saline with 0.1% Tween-20 for 1 h at 24ºC with gentle shaking and washed 3x with PBS-Tween. The blots were stained with streptavidin–HRP (1:1000) (Pierce, Streptavidin Poly-HRP) overnight at 4ºC with gentle shaking. Upon washing with PBS-Tween, the membranes were developed using the ECL Chemiluminescent Substrate (OriGene). Figure 2.8 shows a Western blot demonstrating the incorporation of GalNAz into glycoproteins from protein lysates collected from the three fractions (conditioned media fraction, soluble and insoluble fractions). Anti-biotin signal was checked before affinity-capture (Load) and after affinity-capture on the fraction not bound to the beads (Supernatant) and on the fractions that included the bead after washing (Capture), as performed in [9] (Figure 2.8).

2.14.3 Database searching and Label Free Quantification

The released glycopeptides were consecutively analyzed by reversed-phase nanoflow liquid chromatography coupled to a Thermo Linear Trap Quadrupole (LTQ)–Orbitrap fusion mass spectrometer, as described in [1] (also see Figure 2.7). Spectra were searched against a human proteins database (Uniprot Complete Proteome – taxonomy Homo sapiens – 69165 sequences) using the Andromeda module of MaxQuant software v. 1.5.2.8 [6, 11]. Trypsin/P enzyme parameter was selected with two possible missed cleavages. Carbamidomethylation of cysteins was set as fixed modification, and methionine oxidation and acetylation of protein N-terminus were set as variable modifications, similar to that applied in [12]. Search mass tolerances were defined at 5 ppm and 0.6 Da for MS and MS/MS respectively. For protein validation, a maximum false discovery rate of 1% at

peptide and protein level was used based on a target/decoy search. MaxQuant was also applied for Label Free Quantification (LFQ), as shown in [13]. The 'match between runs' option was used with a 20 min value as the alignment time window and 3 min as match time window. Only unique and razor peptides were used for quantification. The LFQ intensity values (normalized values) extracted by MaxQuant for each protein in each sample replicate were used to calculate a ratio between two samples to compare as well as a *p*-value based on a Welch's test similar to that applied in [14]. When LFQ intensity values were missing, they were replaced either by the average of the values of the two other replicates, or, if less than two replicate values were present, by a noise value corresponding to the first percentile of LFQ values of all proteins of the sample replicate, as described in [14]. A protein was considered as quantifiable only if at least two of the replicate values in one of the two samples to compare were present before performing the missing values replacement.

Differentially regulated proteins between *GALNT3* KD and control A2780s cells were defined based on the following selection criteria: 2-fold change in expression level and t-test *p*-value cutoff of ≤ 0.05, as described [15-18]. A *z*-score was also calculated for each protein based on the statistical approach described in [16], where z-score = {(Welch t-test difference) - Median (Welch t-test difference) for all quantified proteins}/Standard deviation (Welch t-test difference) for all quantified proteins as described in [16].

To classify proteins as variant, different combinations of stringent filtering criteria were tested:

1. Filtering 1 (Welch *p*-value ≤ 0.05 and fold change of ≥ 2)
2. Filtering 2 (Welch *p*-value ≤ 0.05 and z-score > 1)

2.14.4 Bioinformatic annotation & analysis

Glycoprotein prediction analysis: The NetOGlyc 4.0 server (http://www.cbs.dtu.dk/services/NetOGlyc-4.0/) was used to identify the O-glycosylated proteins identified from our control and *GALNT3* KD A2780s EOC cells, (using G-score > 0.5), as described in [5].

The NetNGlyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc/) was used to find the N-glycosylated proteins identified from our control and *GALNT3* KD A2780s EOC cells. Sequences having N-glycosylation potential > 0.5 were considered as cut-off value [19].

An additional prediction approach used in our study was essentially focused on reviewing recent literature for previously identified glycoproteins.

Protein enrichment analysis: GO enrichment analysis of the cellular localization of the identified differentially regulated proteins was performed using information from AmiGO (http://amigo.geneontology.org). The GO term enrichment tool was used to determine the observed level of annotations for the set of proteins from our study and determine the significance in the context of all proteins annotated in the human proteome [20]. Data was presented as percent of enrichment. The GO terms found to be over/under represented by a two-tailed Fisher Exact test with a p-value ≤ 0.05 were presented, p-values were corrected using Bonferroni statistics correction (See Figure 2.3). P-values were additionally transformed to scores (–log10 (p-value)), to determine whether the fold enrichment is significant based on the relative abundance of each GO term in our data sets ($p \leq 0.05$ is considered significant) (see Figure 2.9). The GO terms based on the gene list of our study were compared to the background distribution of annotation based on the genes in the whole genome that are annotated to the GO Term similar to that applied in [20] (see Figure 2.9).

2.15 Acknowledgements

This work has been supported by a grant to D.B. from the Cancer Research Society of Canada, as well as by grants from Jane Coffin Childs Fund to C.M.W., Burroughs Wellcome Fund CASI to C.M.W., National Institutes of Health (CA200423) to C.R.B. and Howard Hughes Medical Institute to C.R.B.

2.16 References

[1] Sheta R, Woo CM, Roux-Dalvai F, Fournier F, Bourassa S, Droit A, et al. A metabolic labeling approach for glycoproteomic analysis reveals altered glycoprotein expression upon GALNT3 knockdown in ovarian cancer cells. J Proteomics. 2016 Apr 17. doi: 10.1016/j.jprot.2016.04.009. [Epub ahead of print].

[2] Sletten EM, Bertozzi CR. Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. Angew Chem Int Ed Engl. 2009;48:6974-98.

[3] Laughlin ST, Bertozzi CR. Metabolic labeling of glycans with azido sugars and subsequent glycan-profiling and visualization via Staudinger ligation. Nature protocols. 2007;2:2930-44.

[4] Hang HC, Yu C, Kato DL, Bertozzi CR. A metabolic labeling approach toward proteomic analysis of mucin-type O-linked glycosylation. Proceedings of the National Academy of Sciences of the United States of America. 2003;100:14846-51.

[5] Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. The EMBO journal. 2013;32:1478-88.

[6] Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: a peptide search engine integrated into the MaxQuant environment. Journal of proteome research. 2011;10:1794-805.

[7] Bergemann TL, Wilson J. Proportion statistics to detect differentially expressed genes: a comparison with log-ratio statistics. BMC bioinformatics. 2011;12:228.

[8] Cervello M, Bachvarov D, Lampiasi N, Cusimano A, Azzolina A, McCubrey JA, et al. Novel combination of sorafenib and celecoxib provides synergistic anti-proliferative and pro-apoptotic effects in human liver cancer cells. PloS one. 2013;8:e65569.

[9] Woo CM, Iavarone AT, Spiciarich DR, Palaniappan KK, Bertozzi CR. Isotope-targeted glycoproteomics (IsoTaG): a mass-independent platform for intact N- and O-glycopeptide discovery and analysis. Nature methods. 2015;12:561-7.

[10] Woo CM, Bertozzi CR. Isotope Targeted Glycoproteomics (IsoTaG) to Characterize Intact, Metabolically Labeled Glycopeptides from Complex Proteomes. Current protocols in chemical biology. 2016;8:59-82.

[11] Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature biotechnology. 2008;26:1367-72.

[12] Genoux A, Pons V, Radojkovic C, Roux-Dalvai F, Combes G, Rolland C, et al. Mitochondrial inhibitory factor 1 (IF1) is present in human serum and is positively correlated with HDL-cholesterol. PloS one. 2011;6:e23949.
[13] Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Molecular & cellular proteomics : MCP. 2014;13:2513-26.

[14] Vargas A, Roux-Dalvai F, Droit A, Lavoie JP. Neutrophil-derived Exosomes: A New Mechanism Contributing to Airway Smooth Muscle Remodeling. American journal of respiratory cell and molecular biology. 2016.

[15] Sun N, Pan C, Nickell S, Mann M, Baumeister W, Nagy I. Quantitative proteome and transcriptome analysis of the archaeon Thermoplasma acidophilum cultured under aerobic and anaerobic conditions. Journal of proteome research. 2010;9:4839-50.

[16] Ramus C, Hovasse A, Marcellin M, Hesse AM, Mouton-Barbosa E, Bouyssie D, et al. Spiked proteomic standard dataset for testing label-free quantitative software and statistical methods. Data Brief. 2016;6:286-94.

[17] Zhang P, Guo Z, Zhang Y, Gao Z, Ji N, Wang D, et al. A preliminary quantitative proteomic analysis of glioblastoma pseudoprogression. Proteome science. 2015;13:12.

[18] Gautier V, Mouton-Barbosa E, Bouyssie D, Delcourt N, Beau M, Girard JP, et al. Label-free quantification and shotgun analysis of complex proteomes by one-dimensional SDS-PAGE/NanoLC-MS: evaluation for the large scale analysis of inflammatory human endothelial cells. Molecular & cellular proteomics : MCP. 2012;11:527-39.

[19] Gupta R, Brunak S. Prediction of glycosylation across the human proteome and the correlation to protein function. Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing. 2002:310-22.

[20] Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, Lewis S. AmiGO: online access to ontology and annotation data. Bioinformatics. 2009;25:288-9.

2.17 Figure legends

Figure 2.7 Schematic overview of the glycoproteomic workflow used.

Figure 2.8 Western blot analysis of glycoproteins enrichment in control and *GALNT3* KD A2780s cells. Whole-cell lysates labeled with 100 μM Ac4GalNAz were incubated with biotinylated bioorthogonal probe. Anti-biotin signal was checked before affinity-capture (Load) and after affinity-capture on the fraction not bound to the beads (Supernatant) and on the fractions that included the bead after washing (Capture).

Figure 2.9 GO cellular component analysis of significantly enriched proteins found upon *GALNT3* KD. Bar graphs showing the cellular component GO terms that are significantly enriched from the differentially regulated proteins in our study, compared to the entire human proteome. Data were submitted to the GO Consortium for enrichment analysis [20]. The analysis was performed on the differentially regulated proteins identified from each of the three fractions: Conditioned media fraction (blue bars), Soluble fraction (red bars) and Insoluble fraction (green bars). All identified proteins annotated with GO cellular component terms were compared against the annotated human proteome. The enrichment *p*-value (≤ 0.05) of each term was transformed to a –log10 (*p*-value).

2.18 Figures

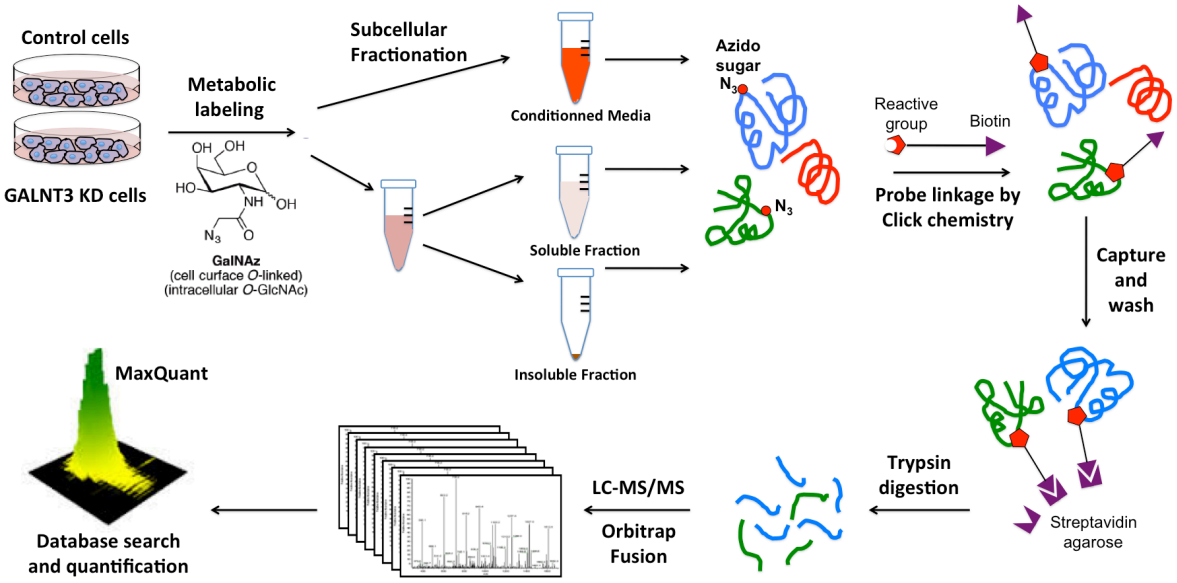


Figure 2.7. Schematic overview of the glycoproteomic workflow used

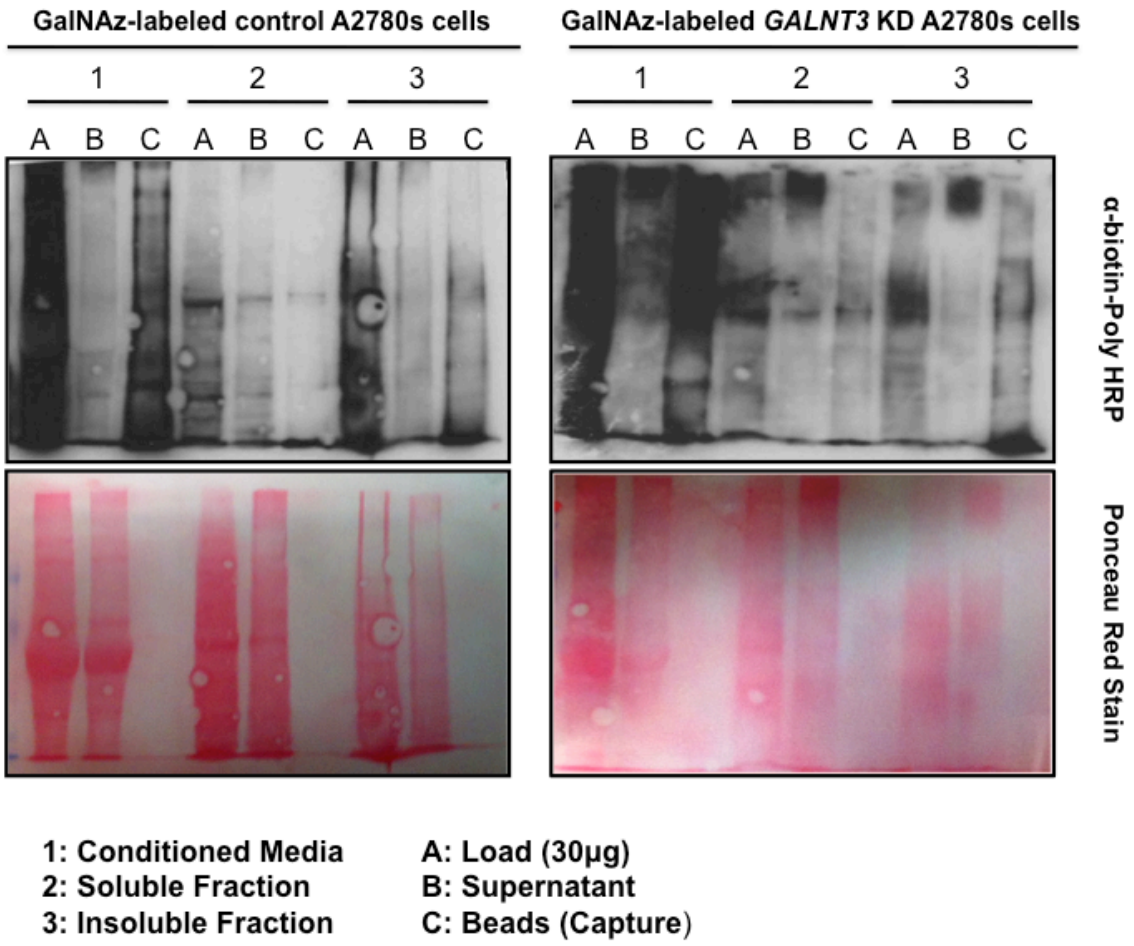


Figure 2.8. Western blot analysis of glycoproteins enrichment in control and *GALNT3* KD A2780s cells

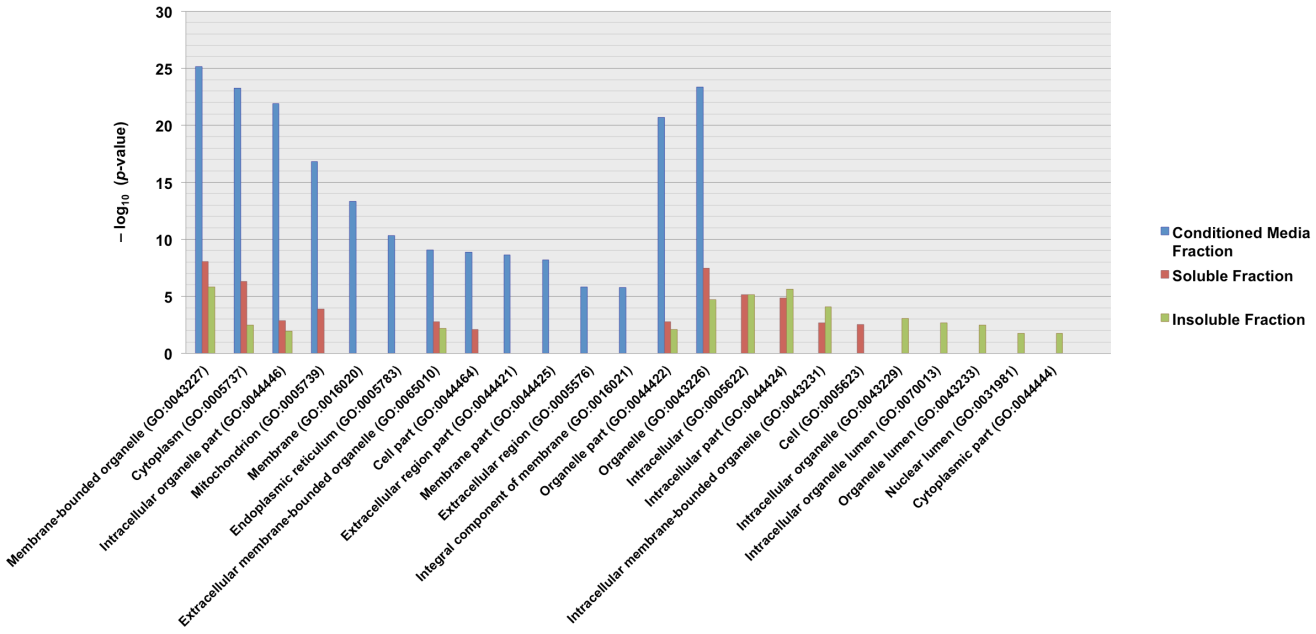


Figure 2.9. GO cellular component analysis of significantly enriched proteins found upon *GALNT3* KD

Chapter 3: Altered expression of different GalNAc-Transferases (GalNAc-Ts) is associated with disease progression and poor prognosis in women with high-grade serous ovarian cancer

Razan Sheta[1,2], Magdalena Bachvarova[2], Marie Plante[2,3], Jean Gregoire[2,3], Marie-Claude Renaud[2,3], Alexandra Sebastianelli[2,3], Ion Popa[4,5] and Dimcho Bachvarov[1,2]*

[1]Department of Molecular Medicine, Laval University, Quebec (Quebec), Canada

[2]CHU de Québec Research Center, Hotel-Dieu de Québec, Quebec (Quebec), Canada

[3]Department of Obstetrics and Gynecology, Laval University, Quebec (Quebec), Canada

[4]Department of Molecular Biology, Medical Biochemistry, and Pathology, Laval University, Quebec (Quebec), Canada

[5]CHU de Québec Research Center, Hospital Saint-Sacrement, Quebec (Quebec), Canada

* Corresponding author:

CHU de Québec Research Center, Hotel-Dieu de Québec

9 rue McMahon, Quebec (Quebec) Canada G1R 3S3

dimtcho.batchvarov@crhdq.ulaval.ca

Tel.: (418) 525-4444, ext. 15590

Fax: (418) 691-5439

List of abbreviations

EOC: Epithelial ovarian cancer; GalNAc-Ts: GalNAc transferases; HGSC: High-grade serous carcinoma; HOSE: Human ovarian surface epithelial; IHC: immunohistochemistry; KD: knockdown; LMP: low-malignant potential; MD: moderately differentiated; MCA: Multiple Correspondence Analyses; PD: Poorly differentiated; PTM: Post-translational modifications; PFS: Progression-free survival; SOC: Serous ovarian carcinomas; SNO: Superficial scrapings from normal ovaries; TMAs: Tissue microarrays.

Keywords: Epithelial ovarian cancer, Mucin O-glycosylation, Polypeptide GalNAc transferases, Immunohistochemistry, Tissue microarrays, Kaplan–Meier survival curve analysis.

(Originally published in International Journal of Oncology, Ocotber 2017, PMID: 29039611)

3.1 Preface

The following chapter presents my scientific findings as a first author, which was published in the International Journal of Oncology on December 2017, PMID: 29039611.

I was involved in all the experiments performed in this study, including the preparation of tissue microarrays (TMA), performing IHC and data analysis. I also assisted in the scoring of protein staining in the TMA used in this study. Magdalena Bachvarova, research assistant in our laboratory, was involved in the collection, storage and archiving of ovarian tissues for the tumor bank. Drs. Marie Plante, Jean Gregoire, Marie Claude Renaud and Alexandra Sebastenilli were involved in the recruitment of different patient cohorts and providing extensive clinical data. Dr. Ion Popa was involved in the protein stain scoring for all TMAs used in this study. Dr. Dimcho Bachvarov, the principal investigator of this study was involved in experimental design, data analysis and article writing.

3.2 Résumé en français

Le cancer épithélial de l'ovaire (CEO) est une pathologie hautement métastatique et la principale cause de décès parmi toutes les tumeurs malignes gynécologiques. Les taux élevés de létalité du CEO découlent principalement de notre incapacité à détecter la maladie à un stade précoce ou à un stage où l'atteinte est confinée à l'organe primaire et aussi au manque de traitements efficaces pour la maladie à des stades avancés. Ainsi, la gestion des stades avancés représente un problème crucial pour un traitement efficace du CEO et l'identification de nouvelles voies et molécules pro-métastatiques pourrait améliorer le diagnostic précoce de cette pathologie.

La glycosylation des protéines est une modification post-traductionnelle (PTM) hautement hétérogène et omniprésente, attachée aux chaînes latérales de l'asparagine (N-linked), de la serine ou de la thréonine (O-linked). La glycosylation aberrante dans le cancer est fréquemment attribuée à l'expression modifiée des GalNAc-Transférases polypeptidiques - les enzymes initiant l'O-glycosylation de type mucine. Jusqu'à présent, vingt membres de la famille humaine de gènes GalNAc-Ts ont été identifiés. Ce chapitre traite de l'examen de plusieurs membres de cette famille de gènes et de leur rôle dans la progression du CEO. Ce chapitre présente principalement des données cliniques du profil d'expression de cinq membres de la famille de gènes GalNAc-Ts dans les différents types de cancer de l'ovaire. Les données de ce chapitre fournissent plus de preuves sur l'implication de l'O-glycosylation aberrante dans les mécanismes de progression du CEO, ce qui pourrait identifier de nouveaux biomarqueurs glycosylés du CEO et / ou des cibles thérapeutiques pour un traitement plus efficace du CEO.

3.3 Abstract

Protein glycosylation perturbations are implicated in a variety of diseases, including cancer. Aberrant glycosylation in cancer is frequently attributed to altered expression of polypeptide GalNAc transferases (GalNAc-Ts) - enzymes initiating mucin-type O-glycosylation. Previous work from our group demonstrated that one member of this family (*GALNT3*) is overexpressed in epithelial ovarian cancer (EOC), and *GALNT3* expression correlated with shorter progression-free survival (PFS) in EOC patients with advanced disease. As considerable degree of redundancy between members of the GalNAc-Ts gene family has been frequently observed, we decided to investigate whether other members of this family are essential in EOC progression.

In silico analysis based on publically available data was indicative for altered expression of five GalNAc-Ts (*GALNT2, T4, T6, T9* and *T14*) in ovarian high-grade serous carcinoma (HGSC) samples compared to non-tumoral ovarian tissue. We analyzed protein expression of these GalNAc-Ts in EOC cells and tumors by Western blots, followed by immunohistochemical (IHC) evaluation of their expression in EOC tumor and control samples using tissue microarrays (TMAs). Western blot analyses were indicative for low expression of GALNT2 and strong expression of GALNT6, T9 and T14 in both EOC cells and tumors. These observations were confirmed by IHC. GALNT2 displayed significantly lower expression, while GALNT6, GALNT9 and GALNT14 showed significantly higher expression in HGSC tumors compared to control tissue. Importantly, GALNT6 and GALNT14 expression correlated with poor prognosis of serous EOC patients. Moreover, our results suggest for overlapping functions of some GalNAc-Ts, more specifically *GALNT3* and *GALNT6*, in directing EOC progression.

Our results are indicative for a possible implication of different members of the GalNAc-T gene family in modulating EOC progression, and the potential use of *GALNT6* and *GALNT14* as novel prognostic EOC biomarkers. These data warrant future studies about the role of members of the GalNAc-Ts gene family in ovarian tumorigenesis.

3.4 Introduction

Epithelial ovarian cancer (EOC) accounts for 4% of all cancers in women and is the leading cause of death from gynecologic malignancies (1). Despite medical and surgical improvements, long-term survival rates for EOC patients with advanced disease remain disappointing, primarily due to its asymptomatic nature and the lack of reliable methods for early diagnosis (2, 3). Indeed, most women are diagnosed with EOC when micro- and macro-metastases are already present and currently, the 5-year EOC survival rate is less than 40% (1). Thus, the identification of novel pro-metastatic EOC molecules and associated pathways could provide additional therapeutic targets for improved management of this deadly disease.

There are many documented studies investigating different post-translational modifications (PTMs) and their association with cancer; among these, aberrant glycosylation has important roles in cancer progression (4, 5), including EOC dissemination (6, 7). Mucin-type O-glycosylation of proteins represents the most diverse PTM form, and is considered to be a conserved type of glycosylation found in many species and organ types (8). This type of PTM is complex and involves the transfer of different monosaccharides to each of the six O-linked glycans (9). O-glycosylation is initiated by 20 GalNAc-Transferases, a family of enzymes known as the UDP N-acetylgalactosamine: polypeptide N-acetyl galactosaminyl transferases (GalNAc-Ts) (9), which are responsible for the transfer of the monosaccharide GalNAc from UDP-GalNAc to the hydroxyl group of the serine, threonine or tyrosine residues found in the target protein substrate (10). Interestingly, the genes comprising this family of enzymes show tissue-specific expression (11, 12), but also have overlapping substrate specificities, suggestive for redundancy in their function (11). Several *in vitro* studies have shown that GalNAc-Ts are differentially expressed in cells and tissues during development (9). Aberrant GalNAc-Ts expression patterns have been frequently observed in cancer (13), which warrants further studies on the possible implications of these enzymes in tumor progression and their potential use as therapeutic targets and/or prognostic markers (14).

We have previously identified the GalNAc-T3 (*GALNT3*) gene as a potential EOC oncogene, highly expressed in advanced disease, as *GALNT3* expression was significantly associated with poor outcome (15). We also demonstrated that in EOC cells, *GALNT3*

might direct altered biosynthesis and/or aberrant glycosylation of different O-glycoproteins that could play an essential role in EOC dissemination (15-17). However, to date, no other studies have comprehensively assessed the implications of other members of the GalNAc-Ts gene family in EOC dissemination. We thus decided to examine the role of some of the GalNAc-Ts known to play essential roles in the etiology of different cancer types. Five members of this gene family (*GALNT2, T4, T6, T9* and *T14)* display significant alterations in their expression in different cancer types, as reviewed in (7). *In silico* analysis of the GalNAc-Ts expression profiles from publicly available data were also indicative for overexpression of some of these GalNAc-Ts (including *GALNT4*, *T6*, and *T14*), as well as *GALNT3*, in ovarian HGSC samples.

The above data prompted us to investigate the expression levels of these five GalNAc-Ts (*GALNT2, T4, T6, T9* and *T14*) in EOC cells and EOC tumor tissues, and to correlate their expression with the corresponding clinicopathological characteristics. To our knowledge this is the first report that comprehensively examines the expression profiles of several GalNAc-Ts in EOC with the perspective of identifying novel prognostic factors for this deadly disease. We show that most of the GalNAc-Ts analyzed (including *GALNT2, T6, T9* and *T14*) displayed altered expression in EOC cells and tumors and more importantly, our data are indicative for a significant association of two GalNAc-Ts (*GALNT6* and *T14*) with progression-free survival (PFS) values of EOC patients, suggestive for oncogenic functions of these two enzymes in EOC, and their potential use as novel EOC prognostic biomarkers.

3.5 Material and Methods

3.5.1 Patient cohort

Patients included in this study were operated between January 1998 to December 2015 for advanced ovarian cancer at the CHUQ Hôtel-Dieu Hospital in Quebec City, Canada. Inclusion criteria were: serous papillary carcinoma histology, FIGO Stages II, III or IV and chemotherapy received after the surgery (see Table 3.1 for detailed clinicopathological characteristics). All tumors were histologically classified according to the criteria defined by the World Health Organization (18). Disease progression was evaluated following the guidelines of the Gynecology Cancer Intergroup (18). PFS was defined as the time from surgery to the first observation of disease progression, recurrence or death. The follow-up was available until death or to the date the study was closed (30 December 2016). Nineteen non-tumoral (control) ovarian samples were derived from women undergoing hysterectomy with oophorectomy due to non-ovarian pathologies.

3.5.2 Cell culture

The EOC cell lines SKOV3 and CaOV3 were purchased from the American Tissue Type Collection; OV-90, OV2008, TOV-112 and TOV-21 cell lines were a kind gift from Dr. Anne-Marie Mes-Masson (Montreal University), A2780s and A2780cp cell lines were a kind gift from Dr. Benjamin Tsang (Ottawa University), the OVCAR4 cell line was a kind gift from Dr. Stephan Gobeil (Laval University), and the two human ovarian surface epithelial (HOSE) cell lines; HOSE 6.3 and 17.1 were a kind gift from Dr. Francis Jacob (University Hospital Basel). The cell lines were passaged in different culture media supplemented with 10% fetal bovine serum, and 1% Penicillin Streptomycin or 50µg/ml of gentamicin as described previously (15).

3.5.3 Western blotting

Western blot analysis was performed as previously described (15). Briefly, protein lysates of the EOC and HOSE cell lines, were prepared by suspending cell pellets in Laemmli

sample buffer containing 5% β-mercaptoethanol. While ovarian tumor tissue and non-tumoral tissues were homogenized and sonicated in RIPA buffer (50 mM Tris (pH 7.4), 150 mM NaCl, 0.5% Sodium deoxycholate, 0.1% SDS, 1% Triton X-100) containing protease and phosphatase inhibitors, samples were then incubated on ice for 15 min. Protein samples from cells and tissues were measured using a BCA protein assay kit (Thermo Scientific Pierce, Rockford, IL, USA). BSA standard or samples (20 µl) were transferred to a 96 well plate to which 200 µl working reagent was added (working reagent 50:1 ratio of assay reagents A and B). The plate was incubated for 30 min at 37°C, the plates were then analyzed with a spectrophotometer at 560 nm (iMARK microplate reader (Bio-Rad) Hercules, CA, USA). After centrifugation at 13,300 rpm for 15 min at 4°C, the supernatant was taken and 20-30 µg of the protein were used for sample preparation. Protein lysates were separated by 6–12% Tris-glycine gel electrophoresis and transferred onto a polyvinylidene difluoride membrane using a semidry apparatus (Bio-Rad Laboratories, Mississauga, ON). The membrane was blocked with 4% nonfat dry milk in TBST (20 mmol/l TRIS-HCl, 0.5 M NaCl and 0.1% Tween 20), and the membranes were incubated overnight with the primary antibody at 4°C or 1 h at RT (For a list of all the antibodies used in this study refer to Table 3.3). After 3× 15 min washes with TBST (20 mmol/l TRIS-HCl, 0.5 M NaCl and 0.1% Tween 20) at room temperature, the membranes were incubated with horseradish peroxidase-conjugated secondary antibody and detected with ECL solution (Thermo Fisher Scientific, Waltham, MA).

3.5.4 Tissue Micro arrays (TMAs) and immunohistochemistry (IHC)

TMAs were constructed as previously described (15). Briefly, one representative block of each ovarian tumor and control ovarian tissue was selected for the preparation of the tissue arrays. Three 0.6mm cores of tumor were taken from each tumor block and placed, 0.4mm apart, on a recipient paraffin block using a commercial tissue arrayer (MTA-II arrayer) (Beecher Instruments, Silver Spring, MD, USA). The cores were randomly placed on one of three recipient blocks to avoid evaluation biases.

IHC analyses were performed, as previously described (15). Briefly, 4 μm tissue sections were deparaffinised and rehydrated in graded alcohols, then incubated with blocking serum for 20 min. Following treatment with 3% H_2O_2 for 10 min to quench the endogenous peroxidise activity, sections were incubated with the primary antibody overnight at 4°C. (For a list of all the antibodies used in this study refer to Table 3.3). Incubation and detection with SignalStain 3,3′-diaminobenzidine (DAB) Substrate Kit (IDetect Universal Mouse Kit HRP – DAB, IDLABS, Buffalo NY) were done according to the manufacturer's manual. Sections were then counterstained with Hematoxylin. Images were acquired using a Leica Confocal Scope (TCS SP5 X (Leica Microsystems, Exton, PA) and analyzed via the Leica Application Suite Software (Leica Microsystems, Exton, PA).

3.5.5 Scoring and Statistical analysis

Protein expression was scored according to intensity (value of 0 for absence, 1 for low, 2 for moderate, and 3 for high) of staining based on manual visualization. A composite score was defined as the product of staining intensity (nuclear, cytoplasmic, or membranous depending on the expected staining). All slides were independently scored in a blind manner by 2 observers, and the integration was greater than 85%. In case of differences between the 2 scorings, the core was re-evaluated to reach a consensus. The relationship between the protein expression of the listed genes in HGSC and LMP tumors, and control ovarian tissues was evaluated by the Wilcoxon two-sample test. A significant association was considered when *p*-values were less than 0.05.

Multiple Correspondence Analyses (MCA) was used to produce 2-dimensional displays of similarities of the relative expression patterns amongst the five GalNAc-Ts staining intensity in the different patient samples. MCA statistical analyses were carried out using (SPSS software, version 13.0). Survival analysis results were visualized using Kaplan-Meier survival curve analysis (SPSS software, version 13.0). A Kaplan Meier curve and the log-rank test were performed based on PFS values to test the effect of the intensity of the gene (3, 2 vs. 0, 1) on disease progression. The relationship between the GalNAc-Ts staining intensity and PFS was determined by the non-parametric Mantel-Cox log rank test to compare survival distributions (SPSS software, version 13.0). Bivariate and Multivariate analyses, taking into account standard or strongly associated prognostic variables, were

performed to identify independent prognostic factors and a statistic test *p*-value less than 0.05 was considered significant.

3.6 Results

3.6.1 Analyses of the expression profiles of different members of the GalNAc-Ts family in both ovarian HGSC tumors and EOC cell lines

Initially, we investigated GalNAc-Ts expression profiles from publicly available data generated from samples found in the independently derived Affymetrix GeneChip Human Genome U133 Plus 2.0 [HG-U133_Plus_2] (19). The in silico data examined were based on global gene expression analysis of differentially expressed candidate mRNAs in 21 moderately differentiated (MD) and poorly differentiated (PD) serous ovarian carcinomas (SOC) (MD/PD SC), 13 serous ovarian borderline (SOBT), and seven superficial scrapings from normal ovaries (SNO) samples (19). These analyses showed significant correlation and high fold change in the overexpression of several GalNAc-Ts in the SOC samples compared to both the SBOT and SNO samples including: *GALNT3, GALNT4, GALNT6*, and G*ALNT14*. Based on the above in silico analyses, as well as on literature data (see Introduction), we decided to further proceed with investigating the expression levels of five GalNAc-Ts (GALNT2, T4, T6, T9 and T14) in EOC specimens by Western blot. The protein expression levels of these GalNAc-Ts were initially analyzed in nine EOC cell lines and in two HOSE cell lines (see Figure 3.1A). As seen from Figure 3.1A, GALNT2 displayed lack of expression in seven EOC cell lines studied and weak expression in SKOV3 and TOV112 EOC cells, compared to strong expression in the two control HOSE cell lines, indicative for a suppression of this enzyme in EOC cells. GALNT4 showed very weak, or lack of expression in all (both EOC and HOSE) cell lines analyzed. Strong GALNT6 expression was observed in three EOC cell lines (TOV21, CAOV3 and OVCAR4) while no, or weak expression was observed in the two HOSE cell lines. Finally, both GALNT9 and GALNT14 displayed very high expression in all EOC cell lines and almost undetectable expression in the HOSE cell lines (Figure 3.1A).

Further analyses of the protein expression levels of the selected GalNAc-Ts in four ovarian HGSC tumor samples and four control ovarian tissue samples were confirmed (see Figure 3.1B). Indeed, GALNT2 showed relatively high expression in all control samples, while no expression was observed in the HGSCs, and GALNT4 showed an analogous pattern of weak or no expression between control tissues and HGSC tumors. Correspondingly, GALNT6 showed relatively high expression in the HGSCs tumor samples compared to control tissues, and GALNT9 and GALNT14 displayed the strongest overexpression levels in HGSCs, relative to the controls (Figure 3.1B).

3.6.2 IHC analysis of the expression patterns of five GalNAc-Ts in numerous EOC tumors using TMAs

We next proceeded with evaluating the GalNAc-Ts expression levels by IHC in TMAs comprising triplicate cores of 131 HGSCs and 13 LMP tumors, and including 18 control ovarian tissue specimens Table 3.1 shows the major clinical characteristics of these patients for whom extensive follow-up clinical data were available. The age ranged from 38 to 91 years (median: 60 years). HGSC tumors were grade 3 (126 patients (96%) and grade 2 (5 patients (4%)). Grade 3 tumors included stage III (70%) and stage IV (26%) tumors, while grade 2 tumors included stage II (4%). Most patients (92%) received standard treatment of cytoreductive surgery followed by platinum/taxane chemotherapy regimens. The median baseline CA-125 was approximately 800. Forty three percent of the patients had a progression or a recurrence within the first 6 months of follow-up; for 35% of the patients the PFS interval was in the range of 7 to 24 months, and 22% of the patients displayed PFS values higher than 25 months. The mean score for expression was defined by the extent and intensity of immunohistochemistry staining for GALNT2, T4, T6, T9 and T14 (Figure 3.2). As expected, GALNT2 displayed highly significant expression (strong cytoplasmic granular staining) in control tissues compared to negative staining in both LMP ($p < 0.0001$) and HGSC tumors ($p < 0.0001$) (Figure 3.2A). As shown previously, GALNT4 was not expressed in all cases (tumor samples and control tissues), despite repeated attempts with various retrieval systems, antibody concentration, incubating time, or signal enhancements systems (Figure 3.2B). GALNT6 showed a strong diffuse cytoplasmic

staining which was exclusively present in HGSCs compared to control tissues ($p = 0.0001$) and LMP tumors ($p < 0.0001$) (Figure 3.2C). A similar expression pattern was also observed for GALNT9 and GALNT14, as both these GalNAc-Ts were significantly overexpressed in HGSC tumors, when compared to control tissues ($p < 0.0001$), and ($p < 0.0001$) respectively (see Figure 3.2D and 2E); however, GALNT9 also displayed significant overexpression in LMP tumors compared to control samples ($p = 0.0146$) (Figure 3.2D), while no significant difference in GALNT14 protein expression was observed between LMP tumors and control tissues ($p = 0.0859$) (Figure 3.2E).

The above results, as well as data from our previous study (15) are indicative for the simultaneous overexpression of several GalNAc-Ts in EOC cell lines and tumor tissues (including GALNT3, T6, T9 and T14). Since a considerable degree of redundancy between the different members of the GalNAc-Ts gene family has been frequently observed (20), we decided to apply the MCA approach in order to more deeply investigate the extent of overlapping expression of these GalNAc-Ts among the EOC tumor samples included in this study. Two quantitative variables are included in the analysis based on staining intensity; (1) for positive staining in patient samples, and (2) for negative staining in patient samples. The two dimensions of the MCA explained 42.27% (Dimension 1) and 19.32% (Dimension 2) of the total data variability, respectively (Figure 3.3A). These analyses were indicative for a strong and a highly overlapping relationship between GALNT3 and GALNT6, as additionally, some overlapping relationship between GALNT6 and GALNT9 cannot be excluded (Figure 3.3A). GALNT14 showed the highest diversity between staining 1 and 2, with no observed overlap with the other 4 genes (Figure 3.3A). GALNT2 confirmed our previous data showing a complete inverse relationship to the other 4 genes, as GALNT2 staining expression 1 and 2 showed a high inverse correlation to expression pattern observed for GALNT14, and overall differential expression to those of GALNT3, T6 and T9 (Figure 3.3A). The MCA-suggested GALNT3/T6 redundancy was further confirmed by examining protein expression profiles of GALNT6, T9 and T14 in the GALNT3 knockdown (KD) clone, previously generated in the EOC cell line A2780s (15). As shown in Figure 3.3B, we observed a clear upregulation in GALNT6 protein expression in the GALNT3 KD clone, while no expression alterations were observed for both GALNT6 and T14.

3.6.3 Association of GalNAc-Ts expression with clinicopathological data

We further proceeded to investigate the prognostic values of the four differentially expressed GalNAc-Ts (GALNT2, T6, T9 and T14) in EOC patients, as we evaluated the relationship between the TMA immunostaining of the 4 GalNAc-Ts and patients' PFS data using Kaplan–Meier survival curve analyses. PFS follow up data were available for 124 patients (see Table 3.1). The analyses were based on the staining intensities: low (includes staining levels of < 2) and high (includes staining levels of > 2) (Figure 3.4). No significant differences were observed between GALNT2 and T9 expression levels and patients' PFS values (log rank = 2.715, p = 0.099) and (log rank =0.995, p = 0.319) respectively (Figure 3.4A and 4C), which suggests that staining intensities for both GALNT2 and T9 in pre-treatment surgical EOC specimens are not predictive of PFS. However there was a significant association between PFS and the expression of the two GalNAc-Ts T6 and T14 (log rank = 5.119, p = 0.024) and (log rank = 5.770, p = 0.016) respectively (Figure 3.4B and 4D).

Bivariate and multivariate analyses to predict PFS were also performed on all the 4 genes. Multivariate analyses taking into account standard or strongly associated prognostic variables (age, grade, stage, and initial CA-125) were performed to identify independent prognostic factors. Multivariate analyses showed a significant association for both GALNT6 and GALNT14 and the higher risk of progression (HR: 1.672; CI: 1.043–2.682; p = 0.033) and (HR: 2.163; CI: 1.032–4.534; p = 0.041) respectively (Table 3.2), but not for GALNT2 and GALNT9 (HR: 0.339; CI: 0.103–1.119; p = 0.076) and (HR: 1.470; CI: 0.939–2.300; p = 0.092) respectively (Table 3.2). Taken together, our findings indicate that GALNT6 and T14 may represent useful markers to predict PFS of women with high grade serous EOC.

3.7 Discussion

Accumulating data are indicative for the important role of glycosylation perturbations in cancerogenesis, including also the abnormal expression of glycans, exclusively involved in embryonic development under normal conditions (21). Glycan alterations can be associated with cancer cell signaling and communication, tumor cell dissociation, migration and invasion, cell–matrix interactions, tumor angiogenesis, immune modulation and metastasis formation (22, 23), and can serve as important biomarkers and/or therapeutic targets (24).

Altered expression of glycans in cancer is frequently attributed to aberrant expression of different members of the GalNAc-Ts family in malignant tumors compared to non-tumoral tissue (7, 14). The deregulation in the expression of the different GalNAc-Ts allows them to play diverse roles in cancerogenesis (14). We have previously reported that one member of these GalNAc-Ts (*GALNT3*) represents a potential EOC oncogene, as its expression significantly correlated with shorter PFS intervals in EOC patients with advanced disease (15); however, data concerning the implication of other members of the GalNAc-Ts gene family in EOC dissemination were scarce. Indeed, only one study was suggestive for a possible role of *GALNT14* in mediating the malignant behavior of EOC cells (25). Otherwise, some of the most extensively studied GalNAc-Ts in cancer include *GALNT2*, *GALNT4, GALNT6, GALNT9* and *GALNT14* (7), which represent the transferases included in the present study. Thus, concerning *GALNT2*, only one study is suggestive for a possible oncogenic role of this transferase in oral cancer (26), while other studies in different cancer types, including neuroblastoma, liver and gastric cancer, are strongly supportive for a role of *GALNT2* in suppressing tumorigenesis (27-29). *GALNT4* has also displayed discrepant roles in different cancers, as an implication of *GALNT4* in breast cancerogenesis has been repeatedly demonstrated (30-32); however low *GALNT4* expression was associated with poor PFS and overall survival of clear cell renal cell carcinoma patients (33). *GALNT6* has been extensively studied for its implication in the malignant transformation and metastasis of epithelial cancers (34-37), and especially in breast cancer (38-40), where *GALNT6* has been suggested as a novel marker for breast cancer detection and potential therapeutic target (41-44). Similarly to *GALNT2* and *GALNT4*, literature data for the role of *GALNT9* in cancerogenesis is contradictory, since a protective role of *GALNT9* was suggested in neuroblastoma and breast cancer dissemination (45, 46), while a recent publication was

indicative for an oncogenic function of this enzyme in colorectal cancer (47). Interestingly, *GALNT9* was among the genes identified as potentially hypomethylated and overexpressed in advanced EOC (48). Finally, *GALNT14* has been characterized as an oncogene in breast and lung cancer (49-51); moreover, recent data are strongly indicative for a role of this transferase in mediating chemotherapy resistance in different cancer types (52-58).

These reports are indicative for tumorigenic roles of the different GalNAc-Ts in cancer, suggesting a highly complex and specific O-glycosylation pattern of glycoproteins in the different cancer types. In this study we initially analyzed the expression levels of the selected five members of the polypeptide GalNAc-Ts family in EOC cell lines and EOC tumor samples (HGSCs) by Western blot. Similar, if not identical patterns of expression were observed for these GalNAc-Ts in both EOC cell lines and EOC tumors, as GALNT2 displayed very low or lack of expression in all EOC specimens studied, compared to control samples (including HOSE cells and non-tumoral ovarian tissues), while the expression levels of GALNT6, GALNT9 and GALNT14 were significantly higher in EOC cells and EOC tumors, as compared to very low/lack of expression in the corresponding controls (1). Only GALNT4 showed very subtle, or absence of any expression in all EOC and control specimens analyzed, indicative for no implications of this enzyme in EOC progression (Figure 3.1). These observations were further confirmed by performing IHC examination of the expression levels in numerous EOC tumors and control tissues, using TMAs. Indeed, GALNT6, GALNT9 and GALNT14 exhibited very strong staining in HGSC tissues and very weak/no staining in the LMP tumors and control tissues, while GALNT2 displayed an inverse staining pattern indicative for a subtle or no expression in both HGSC and LMP tumors, compared to relatively strong expression in control tissues (Figure 3.2). Again, lack of any staining was observed for GALNT4 in all (EOC tumor and control) tissue samples analyzed (Figure 3.2). With the exception of GALNT4, our data are in accordance with previous studies for the roles of these GalNAc-Ts in different carcinoma types as summarized above, and suggest for a possible correlation of *GALNT6, GALNT9* and *GALNT14* expression with EOC progression, while a putative protective role of *GALNT2* in EOC dissemination cannot be excluded.

Moreover, Kaplan–Meier survival curves and consecutive cox-regression analyses revealed that the expression levels of two of the GalNAc-Ts analyzed (GALNT6 and T14) were

significantly related with poor PFS in the studied cohort of HG serous EOC patients (see Figure 3.4B and 3.4D and Table 3.2), which suggest the putative use of these two transferases as novel prognostic serous EOC biomarkers. These results, as well as data from our previous study (15), are indicative for the significant inverse association of three members of the GalNAc-Ts gene family (GALNT3, T6 and T14) with disease progression in serous ovarian adenocarcinoma patients.

Furthermore, the simultaneous overexpression of four GalNAc-Ts (GALNT3, T6, T9 and T14) in advanced EOC raises the question about the specific and/or redundant functions of the members of this gene family in cancers, including EOC. Notably, it is known that the different GalNAc-Ts do have partially overlapping but distinct substrate specificities, which may result in these GalNAc-Ts having partial functional redundancy (59). So far, twenty members of the human GalNAc-Ts gene family have been identified (20), as such an abundance of GalNAc-Ts could provide substantial biosynthetic back-up. Although different GalNAc-Ts are differentially expressed within tissues, between cells within a single tissue, and in different patterns at different stages in the development and differentiation, it is now becoming clear that a subset of the GalNAc-Ts display both distinct and overlapping substrate specificities (11, 13). Some of the mammalian GalNAc-transferase isoforms have been grouped into subfamilies based on their high homology (11, 20). One such example is the human subfamily comprised of *GALNT3* and *GALNT6*, displaying 65% homology in their coding sequence, although this homology does not provide complete functional redundancy (59). Interestingly, *GALNT6* displayed quite similar oncogenic functions in breast cancer (modulating aberrant O-glycosylation and MUC1 stabilization) (39), as those found by us for *GALNT3* in EOC (15). Increased *GALNT3* and *GALNT6* co-expression has been detected in pancreatic (60) and renal (61) carcinomas, suggestive for *GALNT3/T6* complementary correlations. As shown (60, 61), the upregulation of *GALNT3* within the malignant transformation and progression of these cancers could partly depend on that of *GALNT6* in both synergistic and compensatory ways. Moreover, not a separate *GALNT3* or *GALNT6* KD, but only double *GALNT3/T6* KD was shown to inhibit TGF-β-induced epithelial-to-mesenchymal transition (EMT) in prostate cancer cells (35). Similarly, our Western blot and consecutive MCA analyses described herein (see Figure 3.3) are strongly suggestive for a possible functional overlap

and redundancy of GALNT3 and GALNT6 in EOC. Based on the above considerations we can suggest that different GalNAc-Ts be performing redundant and/or overlapping functions in disease progression of women with HGSC. We believe that these relationships need to be examined further, both *in vitro* and *in vivo* to better understand how do these transferases function together in initiating the biosynthesis of specific target glycoproteins, and if they can demonstrate compensatory methods that may enhance their implication in disease progression.

In conclusion, we have shown that four members of the GalNAc-Ts gene family are differentially expressed in EOC specimens, as GALNT6, GALNT9 and GALNT14 were significantly overexpressed in both EOC cell lines and HGSCs, while GALNT2 displayed an inversed expression pattern indicative for very weak or no expression in both EOC cells and EOC tumors, compared to relatively strong expression in HOSE cells and control tissues. Importantly, GALNT6 and GALNT14 expression significantly correlated with poor prognosis of EOC patients with advanced disease. These data and our previously published data are indicative for a possible implication of different members of the GalNAc-Ts gene family (including *GALNT2, T3, T6, T9* and *T14*) in modulating EOC progression, as GALNT6 and GALNT14, together with the previously characterized *GALNT3*, could represent novel prognostic EOC biomarkers. Moreover, our results are suggestive for overlapping functions of some GalNAc-Ts, and especially *GALNT3* and *GALNT6*, in EOC, in conformity with *GALNT3/T6* functional redundancy described in other cancer types. Further functional studies are warranted to more completely elucidate *in vitro* and *in vivo* the individual and/or synergistic implications of the members of the GalNAc-Ts gene family in ovarian tumorigenesis.

3.8 Acknowledgements

Clinical specimens were provided by the Banque de tissus et de données of the Réseau de recherche sur le cancer of the Fonds de recherche du Québec - Santé (FRQ-S), associated with the Canadian Tumor Repository Network (CTRNet). We would also like to thank Ms. Anne-Sophie Julien from the Plateforme de Recherche Clinique du CHU de Québec-Université Laval for helping us with the statistical analyses and interpretation of the data generated. In addition, we would like to thank Mr. Khaly Mbodji from the department of radi-oncologie and the Centre de recherché du Chu de Québec for his help with the statistical analyses.

Declarations

Ethics approval and consent to participate

The study was approved by the Clinical Research Ethics Committee of the Hotel-Dieu de Québec Hospital, and all patients signed an informed consent for voluntary participation and agreed to report individual patient data.

Competing interests

The authors declare no conflict of interests.

Funding

This study was sustained by a grant to D.B. from the Cancer Research Society of Canada.

3.9 References

1. Siegel R, Ward E, Brawley O and Jemal A: Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. CA: a cancer journal for clinicians 61: 212-236, 2011.

2. Fruscio R, Corso S, Ceppi L, et al.: Conservative management of early-stage epithelial ovarian cancer: results of a large retrospective series. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO 24: 138-144, 2013.

3. Alouini S: Management of ovarian cancer has changed. Gynecologic oncology 126: 313; author reply 314, 2012.

4. Pinho SS and Reis CA: Glycosylation in cancer: mechanisms and clinical implications. Nature reviews. Cancer 15: 540-555, 2015.

5. Stowell SR, Ju T and Cummings RD: Protein glycosylation in cancer. Annual review of pathology 10: 473-510, 2015.

6. Abbott KL: Glycomic analysis of ovarian cancer: past, present, and future. Cancer biomarkers : section A of Disease markers 8: 273-280, 2010.

7. Sheta R and Bachvarov D: Role of aberrant glycosylation in ovarian cancer dissemination. Biomdecial Reviews 25: 83-92, 2014.

8. Tabak LA: The role of mucin-type O-glycans in eukaryotic development. Seminars in cell & developmental biology 21: 616-621, 2010.

9. Ten Hagen KG, Fritz TA and Tabak LA: All in the family: the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases. Glycobiology 13: 1R-16R, 2003.

10. Hang HC and Bertozzi CR: The chemistry and biology of mucin-type O-linked glycosylation. Bioorganic & medicinal chemistry 13: 5021-5034, 2005.

11. Bennett EP, Mandel U, Clausen H, Gerken TA, Fritz TA and Tabak LA: Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. Glycobiology 22: 736-756, 2012.

12. Bennett EP, Hassan H, Mandel U, et al.: Cloning of a human UDP-N-acetyl-alpha-D-Galactosamine:polypeptide N-acetylgalactosaminyltransferase that complements other GalNAc-transferases in complete O-glycosylation of the MUC1 tandem repeat. The Journal of biological chemistry 273: 30472-30481, 1998.

13. Beaman EM and Brooks SA: The extended ppGalNAc-T family and their functional involvement in the metastatic cascade. Histology and histopathology 29: 293-304, 2014.

14. Brooks SA, Carter TM, Bennett EP, Clausen H and Mandel U: Immunolocalisation of members of the polypeptide N-acetylgalactosaminyl transferase (ppGalNAc-T) family is consistent with biologically relevant altered cell surface glycosylation in breast cancer. Acta histochemica 109: 273-284, 2007.

15. Wang ZQ, Bachvarova M, Morin C, et al.: Role of the polypeptide N-acetylgalactosaminyltransferase 3 in ovarian cancer progression: possible implications in abnormal mucin O-glycosylation. Oncotarget 5: 544-560, 2014.

16. Sheta R, Woo CM, Roux-Dalvai F, et al.: A metabolic labeling approach for glycoproteomic analysis reveals altered glycoprotein expression upon GALNT3 knockdown in ovarian cancer cells. Journal of proteomics 145: 91-102, 2016.

17. Sheta R, Roux-Dalvai F, Woo CM, et al.: Proteomic dataset for altered glycoprotein expression upon GALNT3 knockdown in ovarian cancer cells. Data in brief 8: 342-349, 2016.

18. Taylor PT and Haverstick D: Re: New guidelines to evaluate the response to treatment in solid tumors (ovarian cancer). Journal of the National Cancer Institute 97: 151; author reply 152, 2005.

19. Elgaaen BV, Olstad OK, Sandvik L, et al.: ZNF385B and VEGFA are strongly differentially expressed in serous ovarian carcinomas and correlate with survival. PloS one 7: e46317, 2012.

20. Marth JD: Complexity in O-linked oligosaccharide biosynthesis engendered by multiple polypeptide N-acetylgalactosaminyltransferases. Glycobiology 6: 701-705, 1996.

21. Dube DH and Bertozzi CR: Glycans in cancer and inflammation--potential for therapeutics and diagnostics. Nature reviews. Drug discovery 4: 477-488, 2005.

22. Hakomori S: Glycosylation defining cancer malignancy: new wine in an old bottle. Proceedings of the National Academy of Sciences of the United States of America 99: 10231-10233, 2002.

23. Christiansen MN, Chik J, Lee L, Anugraham M, Abrahams JL and Packer NH: Cell surface protein glycosylation in cancer. Proteomics 14: 525-546, 2014.

24. Fuster MM and Esko JD: The sweet and sour of cancer: glycans as novel therapeutic targets. Nature reviews. Cancer 5: 526-542, 2005.

25. Wang R, Yu C, Zhao D, Wu M and Yang Z: The mucin-type glycosylating enzyme polypeptide N-acetylgalactosaminyltransferase 14 promotes the migration of ovarian cancer by modifying mucin 13. Oncology reports 30: 667-676, 2013.

26. Lin MC, Huang MJ, Liu CH, Yang TL and Huang MC: GALNT2 enhances migration and invasion of oral squamous cell carcinoma by regulating EGFR glycosylation and activity. Oral oncology 50: 478-484, 2014.

27. Ho WL, Chou CH, Jeng YM, et al.: GALNT2 suppresses malignant phenotypes through IGF-1 receptor and predicts favorable prognosis in neuroblastoma. Oncotarget 5: 12247-12259, 2014.

28. Wu YM, Liu CH, Hu RH, et al.: Mucin glycosylating enzyme GALNT2 regulates the malignant character of hepatocellular carcinoma by modifying the EGF receptor. Cancer research 71: 7270-7279, 2011.

29. Liu SY, Shun CT, Hung KY, et al.: Mucin glycosylating enzyme GALNT2 suppresses malignancy in gastric adenocarcinoma by reducing MET phosphorylation. Oncotarget 7: 11251-11262, 2016.

30. Wright PK, May FE, Darby S, Saif R, Lennard TW and Westley BR: Estrogen regulates vesicle trafficking gene expression in EFF-3, EFM-19 and MCF-7 breast cancer cells. International journal of clinical and experimental pathology 2: 463-475, 2009.

31. Niang B, Jin L, Chen X, et al.: GalNAc-T4 putatively modulates the estrogen regulatory network through FOXA1 glycosylation in human breast cancer cells. Molecular and cellular biochemistry 411: 393-402, 2016.

32. Zhang J, Zhang Z, Wang Q, Xing XJ and Zhao Y: Overexpression of microRNA-365 inhibits breast cancer cell growth and chemo-resistance through GALNT4. European review for medical and pharmacological sciences 20: 4710-4718, 2016.

33. Liu Y, Liu W, Xu L, et al.: GALNT4 predicts clinical outcome in patients with clear cell renal cell carcinoma. The Journal of urology 192: 1534-1541, 2014.

34. Li Z, Yamada S, Wu Y, et al.: Polypeptide N-acetylgalactosaminyltransferase-6 expression independently predicts poor overall survival in patients with lung adenocarcinoma after curative resection. Oncotarget 7: 54463-54473, 2016.

35. Freire-de-Lima L, Gelfenbeyn K, Ding Y, et al.: Involvement of O-glycosylation defining oncofetal fibronectin in epithelial-mesenchymal transition process. Proceedings of the National Academy of Sciences of the United States of America 108: 17690-17695, 2011.

36. Tarhan YE, Kato T, Jang M, et al.: Morphological Changes, Cadherin Switching, and Growth Suppression in Pancreatic Cancer by GALNT6 Knockdown. Neoplasia 18: 265-272, 2016.

37. Gomes J, Marcos NT, Berois N, et al.: Expression of UDP-N-acetyl-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase-6 in gastric mucosa,

intestinal metaplasia, and gastric carcinoma. The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society 57: 79-86, 2009.

38. Liesche F, Kolbl AC, Ilmer M, Hutter S, Jeschke U and Andergassen U: Role of N-acetylgalactosaminyltransferase 6 in early tumorigenesis and formation of metastasis. Molecular medicine reports 13: 4309-4314, 2016.

39. Park JH, Nishidate T, Kijima K, et al.: Critical roles of mucin 1 glycosylation by transactivated polypeptide N-acetylgalactosaminyltransferase 6 in mammary carcinogenesis. Cancer research 70: 2759-2769, 2010.

40. Park JH, Katagiri T, Chung S, Kijima K and Nakamura Y: Polypeptide N-acetylgalactosaminyltransferase 6 disrupts mammary acinar morphogenesis through O-glycosylation of fibronectin. Neoplasia 13: 320-326, 2011.

41. Berois N, Mazal D, Ubillos L, et al.: UDP-N-acetyl-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase-6 as a new immunohistochemical breast cancer marker. The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society 54: 317-328, 2006.

42. Freire T, Berois N, Sonora C, Varangot M, Barrios E and Osinaga E: UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 6 (ppGalNAc-T6) mRNA as a potential new marker for detection of bone marrow-disseminated breast cancer cells. International journal of cancer 119: 1383-1388, 2006.

43. Patani N, Jiang W and Mokbel K: Prognostic utility of glycosyltransferase expression in breast cancer. Cancer genomics & proteomics 5: 333-340, 2008.

44. Kolbl AC, Hiller RA, Ilmer M, et al.: Glycosyltransferases as marker genes for the quantitative polymerase chain reaction-based detection of circulating tumor cells from blood samples of patients with breast cancer undergoing adjuvant therapy. Molecular medicine reports 12: 2933-2938, 2015.

45. Pangeni RP, Channathodiyil P, Huen DS, et al.: The GALNT9, BNC1 and CCDC8 genes are frequently epigenetically dysregulated in breast tumors that metastasise to the brain. Clinical epigenetics 7: 57, 2015.

46. Berois N, Gattolliat CH, Barrios E, et al.: GALNT9 gene expression is a prognostic marker in neuroblastoma patients. Clinical chemistry 59: 225-233, 2013.

47. Tuupanen S, Hanninen UA, Kondelin J, et al.: Identification of 33 candidate oncogenes by screening for base-specific mutations. British journal of cancer 111: 1657-1662, 2014.

48. Keita M, Wang ZQ, Pelletier JF, et al.: Global methylation profiling in serous ovarian cancer is indicative for distinct aberrant DNA methylation signatures associated

with tumor aggressiveness and disease progression. Gynecologic oncology 128: 356-363, 2013.

49. Huanna T, Tao Z, Xiangfei W, et al.: GALNT14 mediates tumor invasion and migration in breast cancer cell MCF-7. Molecular carcinogenesis 54: 1159-1171, 2015.

50. Wu C, Guo X, Wang W, et al.: N-Acetylgalactosaminyltransferase-14 as a potential biomarker for breast cancer by immunohistochemistry. BMC cancer 10: 123, 2010.

51. Kwon OS, Oh E, Park JR, et al.: GalNAc-T14 promotes metastasis through Wnt dependent HOXB9 expression in lung adenocarcinoma. Oncotarget 6: 41916-41928, 2015.

52. Wagner KW, Punnoose EA, Januario T, et al.: Death-receptor O-glycosylation controls tumor-cell sensitivity to the proapoptotic ligand Apo2L/TRAIL. Nature medicine 13: 1070-1077, 2007.

53. Wu C, Shan Y, Liu X, et al.: GalNAc-T14 may be involved in regulating the apoptotic action of IGFBP-3. Journal of biosciences 34: 389-395, 2009.

54. Liang KH, Lin CL, Chen SF, et al.: GALNT14 genotype effectively predicts the therapeutic response in unresectable hepatocellular carcinoma treated with transcatheter arterial chemoembolization. Pharmacogenomics 17: 353-366, 2016.

55. Lin WR, Chiang JM, Liang KH, et al.: GALNT14 Genotype Predicts Postoperative Outcome of Stage III Colorectal Cancer With Oxaliplatin as Adjuvant Chemotherapy. Medicine 95: e3487, 2016.

56. Liang KH, Lin CC and Yeh CT: GALNT14 SNP as a potential predictor of response to combination chemotherapy using 5-FU, mitoxantrone and cisplatin in advanced HCC. Pharmacogenomics 12: 1061-1073, 2011.

57. Soria JC, Mark Z, Zatloukal P, et al.: Randomized phase II study of dulanermin in combination with paclitaxel, carboplatin, and bevacizumab in advanced non-small-cell lung cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 29: 4442-4451, 2011.

58. Thorburn A, Behbakht K and Ford H: TRAIL receptor-targeted therapeutics: resistance mechanisms and strategies to avoid them. Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy 11: 17-24, 2008.

59. Bennett EP, Hassan H, Mandel U, et al.: Cloning and characterization of a close homologue of human UDP-N-acetyl-alpha-D-galactosamine:Polypeptide N-acetylgalactosaminyltransferase-T3, designated GalNAc-T6. Evidence for genetic but not functional redundancy. The Journal of biological chemistry 274: 25362-25370, 1999.

60. Li Z, Yamada S, Inenaga S, et al.: Polypeptide N-acetylgalactosaminyltransferase 6 expression in pancreatic cancer is an independent prognostic factor indicating better overall survival. British journal of cancer 104: 1882-1889, 2011.

61. Kitada S, Yamada S, Kuma A, et al.: Polypeptide N-acetylgalactosaminyl transferase 3 independently predicts high-grade tumors and poor prognosis in patients with renal cell carcinomas. British journal of cancer 109: 472-481, 2013.

3.10 Figure and table legends

Figure 3.1. GalNAc-Ts expression in EOC cells and EOC tumors. (A) Western blot analysis of endogenous GALNT2, T4, T6, T9 and T14 protein expression in different EOC cell lines and the HOSE cell lines. (B) Western blot analysis of endogenous GALNT2, T4, T6, T9 and T14 protein expression in 4 HG ovarian samples (OG1-OG4) and 4 non-tumoral ovarian tissue (Control tissue) samples (NOV1-NOV4). β-actin was used as a loading control. EOC: Epithelial ovarian cancer, HOSE: Human ovarian surface epithelial, NOV: normal ovarian, OV: ovarian tumor.

Figure 3.2. GalNAc-Ts protein expression in HG samples compared to LMP and non-tumoral ovarian samples. (A) a, GALNT2 staining patterns in representative cores in control ovarian tissues, LMP tumors and HG tumors. b, Box-plot presentation of GALNT2 protein expression levels in control ovarian tissues, LMP tumors and HG tumors. (B) a, GALNT4 staining patterns in representative cores in control ovarian tissues, LMP tumors and HG tumors. b, Box-plot presentation of GALNT4 protein expression levels in control ovarian tissues, LMP tumors and HG tumors. (C) a, GALNT6 staining patterns in representative cores in control ovarian tissues, LMP tumors and HG tumors. b, Box-plot presentation of GALNT6 protein expression levels in control ovarian tissues, LMP tumors and HG tumors. (D) a, GALNT9 staining patterns in representative cores in control ovarian tissues, LMP tumors and HG tumors. b, Box-plot presentation of GALNT9 protein expression levels in control ovarian tissues, LMP tumors and HG tumors. (E) a, GALNT14 staining patterns in representative cores in control ovarian tissues, LMP tumors and HG tumors. b, Box-plot presentation of GALNT14 protein expression levels in control ovarian tissues, LMP tumors and HG tumors. All *p*-values were derived from log-rank tests. Error bars denote standard deviation of each mean calculation. HG, high-grade; LMP, low-malignant potential.

Figure 3.3. (A) Multiple Correspondence Analyses (MCA) of the relationship of the GalNAc-Ts expression in the HG ovarian tissue samples. Two major groups, separated by the first and second components, are observed. The first group includes the genes GALNT6 and T9 showing a somewhat overlapping expression pattern in the cohort of

patients examined in the study. The second group indicates the inverse correlation observed between GALNT2 and GALNT14. For each analysis, the percentage of variance explained by the first two dimensions is indicated in parenthesis. (B) Western blot analysis of GALNT3, T6, T9 and T14 expression in control (Ctrl) and GALNT3 knockdown (KD) A2780s clones. β-actin was used as a loading control. HG: High-grade.

Figure 3.4. Kaplan-Meier progression-free survival curves showing the association between GalNAc-Ts expression patterns and prognosis in HG ovarian cancer. Kaplan–Meier survival curve analysis of HG cases for progression free survival (PFS), presented as survival probability, of patients whose tumors show a defined staining pattern.
For evaluation of GalNAc-Ts staining, HG ovarian cancer subjects were divided into two groups, those with positive staining (3 or 2 gene staining intensity) defined here as (1,00) on the Kaplan-Meier curves and those with negative staining (1 or 0 gene staining intensity) defined here as (,00) on the Kaplan-Meier curves. The staining pattern for (A) GALNT2 protein: negative (n = 101), and positive (n = 15). (B) GALNT6 protein: negative (n = 60), and positive (n = 56). (C) GALNT9 protein: negative (n = 52), and positive (n = 64). (D) GALNT14 protein: negative (n = 19), and positive (n = 97). All *p*-values were derived from log-rank tests. HG: High-grade.

3.11 Tables

Table 3.1. Detailed patients' clinicopathological characteristics

Variable	Range	N/Total	%
Age (years)	≥65	54/162	33
	<65	108/162	67
Median age	60		
Tissue/tumor type	Normal	18/162	11
	LMP	13/162	8
	High-grade	131/162	81
Grade	3	126/131	96
	2	5/131	4
Stage	II	5/131	4
	III	92/131	70
	IV	34/131	26
Chemotherapy[a]	Platinum+Taxol	120/131	92
	Other	11/131	8
CA125[b] (U/ml)	≥800	59/118	50
	<800	59/118	50
PFS[c] (months)	0-6	53/124	43
	7-24	44/124	35
	>25	27/124	22

[1]All patients were subjected to adjuvant therapy.

[2] Extended follow-up, including CA-125 values, were available for 118 patients. CA-125 values correspond to the baseline serum concentrations of EOC patients.

[3] Extended follow-up, including progression-free survival (PFS) values, were available for 124 patients. PFS is defined as the time from surgery to the first observation of disease progression, recurrence or death. The follow-up was available until death or to the date the study was closed. LMP: Low-malignant potential.

Table 3.2. Cox regression analysis to predict progression-free survival (PFS)

Marker	Value	Frequency N (%)	Event n (%)	Crude			Adjusted		
				HR	95% CI	P-value	HR[a]	95% CI[b]	P-value
GALNT2	Negative	101 (87.1%)	78 (77.2%)	1.0			1.0		
	Positive	15 (12.9%)	11 (73.3%)	0.599	0.317-1.130	0.113	0.339	0.103-1.119	0.076
GALNT6	Negative	60 (51.7%)	40 (66.7%)	1.0			1.0		
	Positive	56 (48.3%)	49 (87.5%)	1.596	1.047-2.434	**0.030**	1.672	1.043-2.682	**0.033**
GALNT9	Negative	52 (44.8%)	37 (71.2%)	1.0			1.0		
	Positive	64 (55.2%)	52 (81.3%)	1.232	0.807-1.880	0.334	1.470	0.939-2.300	0.092
GALNT14	Negative	20 (17.2%)	11 (55.0%)	1.0			1.0		
	Positive	96 (82.8%)	78 (81.3%)	2.185	1.117-4.275	**0.022**	2.163	1.032-4.534	**0.041**

[a]Note: adjusted for age, stage, grade, and baseline CA-125. The hazards models were applied to adjust for the baseline patients' characteristics.

[b]Note: The confidence interval was based on the arithmetic mean of the gene expression values of the selected cases from this study.

HR: Hazard Ratio; CI: confidence interval.

Table 3.3. Dilution and technique used for each antibody in IHC and Western blot analyses.

Antibody	Species	Dilution TMA	Company	Retrieval	Incubation TMA	Dilution WB	Incubation WB
Anti-GALNT2	Rabbit	1:50	Abcam	Microwave	4˚C overnight	1:1,000	4˚C overnight
Anti-GALNT4	Rabbit	1:100, 1:50, 1:25	Proteintech	Microwave/ pronase	4˚C overnight	1:500	4˚C overnight
Anti-GALNT6	Mouse	1:75	Abcam	Microwave	4˚C overnight	1:1,000	4˚C overnight
Anti-GALNT9	Rabbit	1:100	LifeSpan BioScience	Microwave	4˚C overnight	1:1,000	4˚C overnight
Anti-GALNT14	Rabbit	1:100	Abcam	Microwave	4˚C overnight	1:1,000	4˚C overnight
Anti-β-actin	Mouse	N/A	Santa Cruz	N/A	N/A	1:2,000	1 h/room temperature

IHC: Immunohistochemistry; TMA: Tissue microarray; WB: Western blot.

3.12 Figures

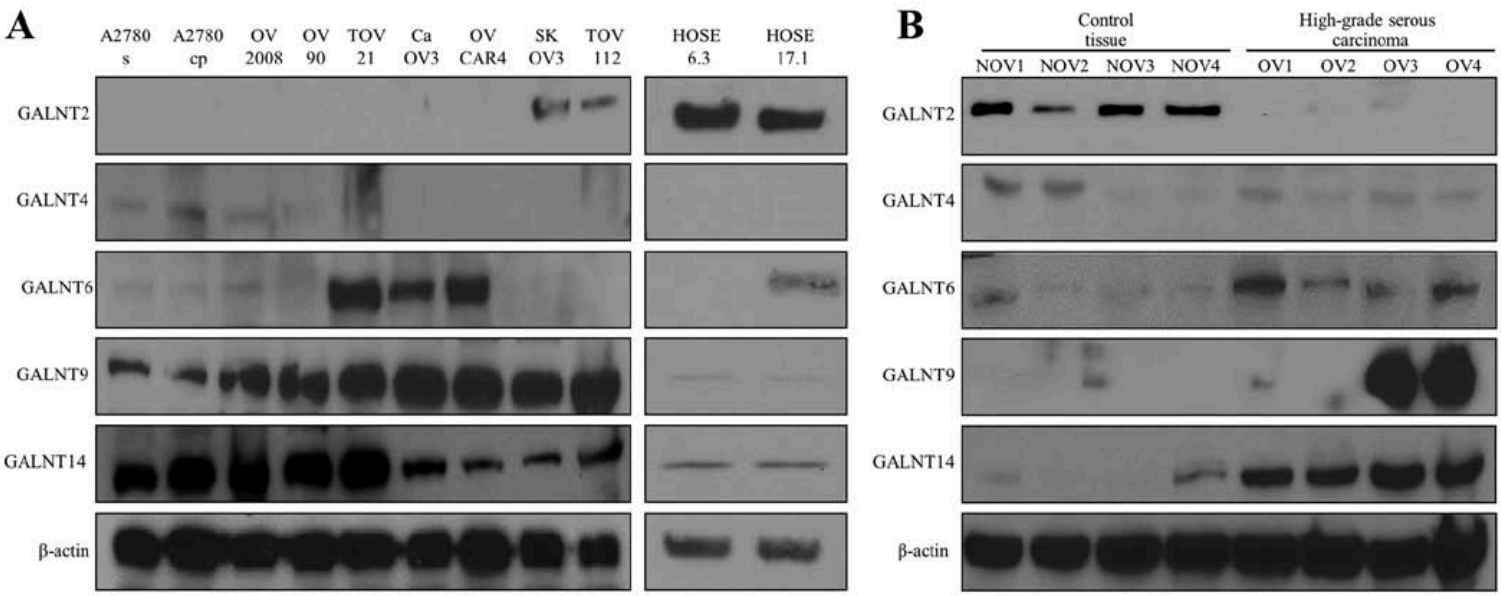


Figure 3.1 GalNAc-Ts expression in EOC cells and EOC tumors

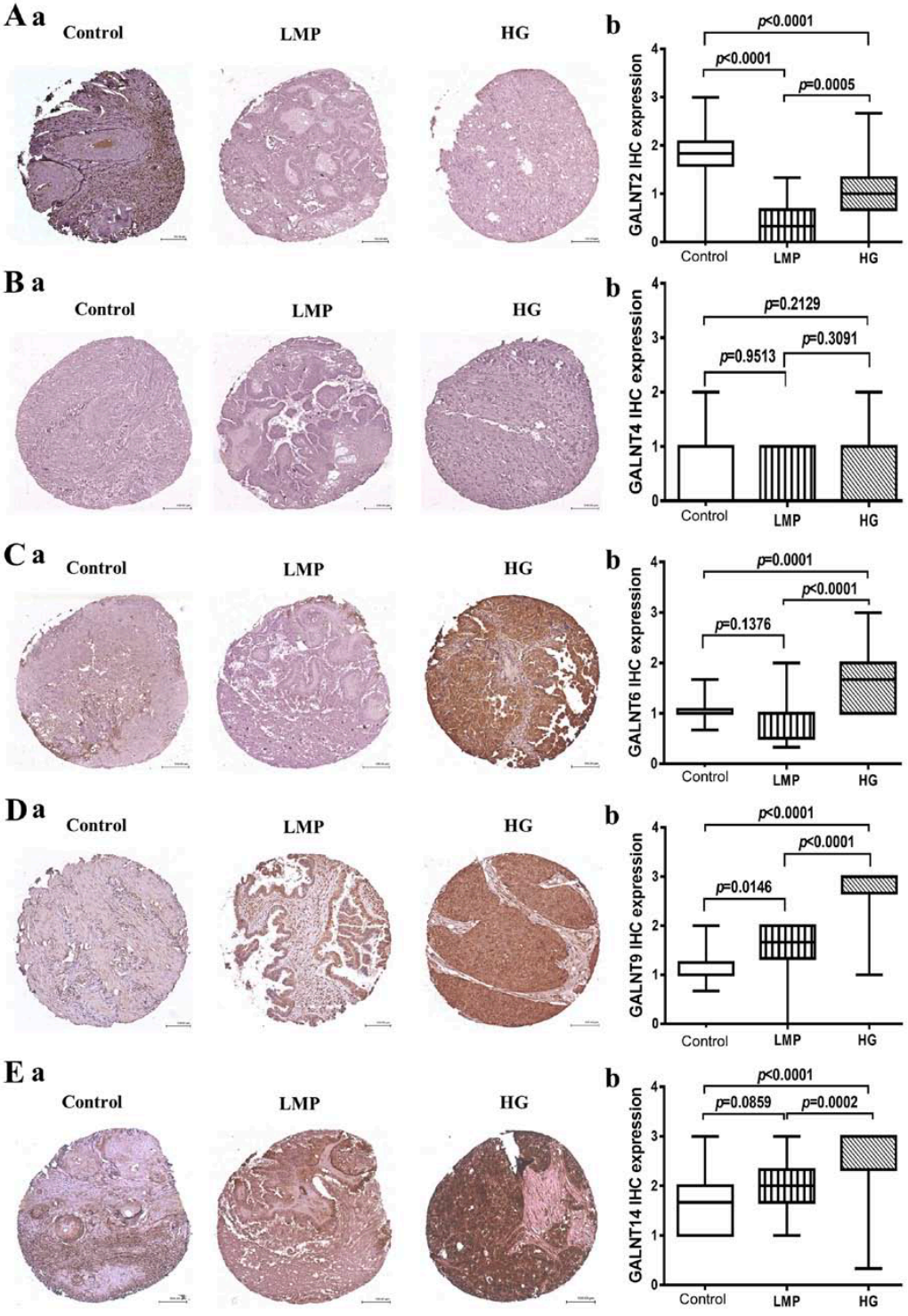


Figure 3.2 GalNAc-Ts protein expression in HG samples compared to LMP and non-tumoral ovarian samples

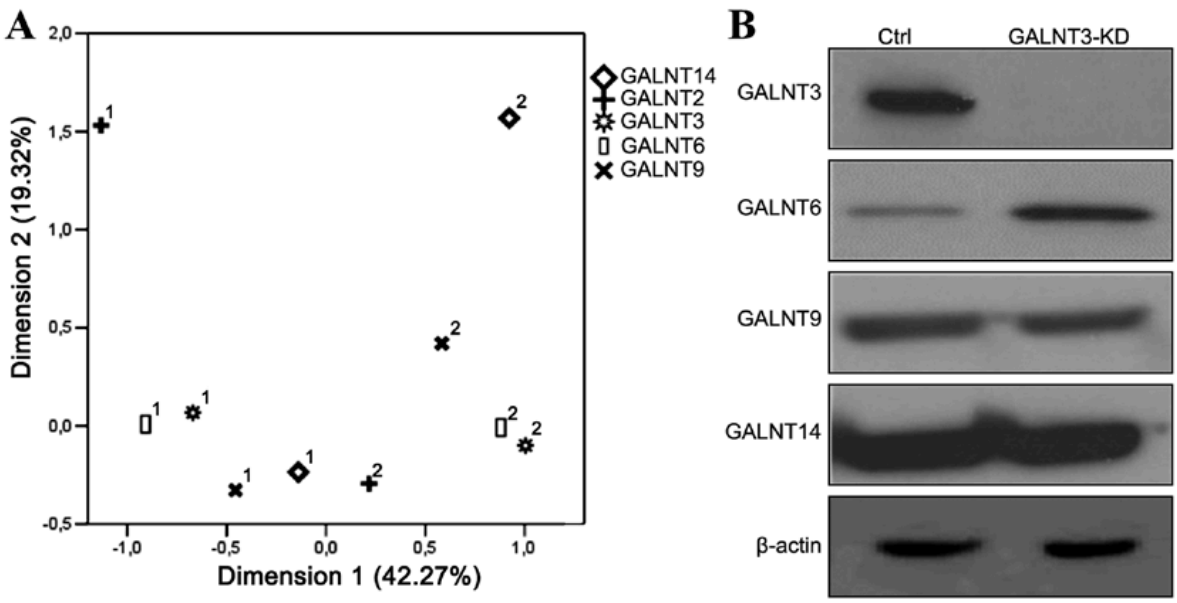


Figure 3.3. Multiple Correspondence Analyses (MCA) of the relationship of the GalNAc-Ts expression in the HG ovarian tissue samples, and Western blot analysis of GALNT3, T6, T9 and T14 expression in control (Ctrl) and *GALNT3* knockdown (KD) A2780s clones

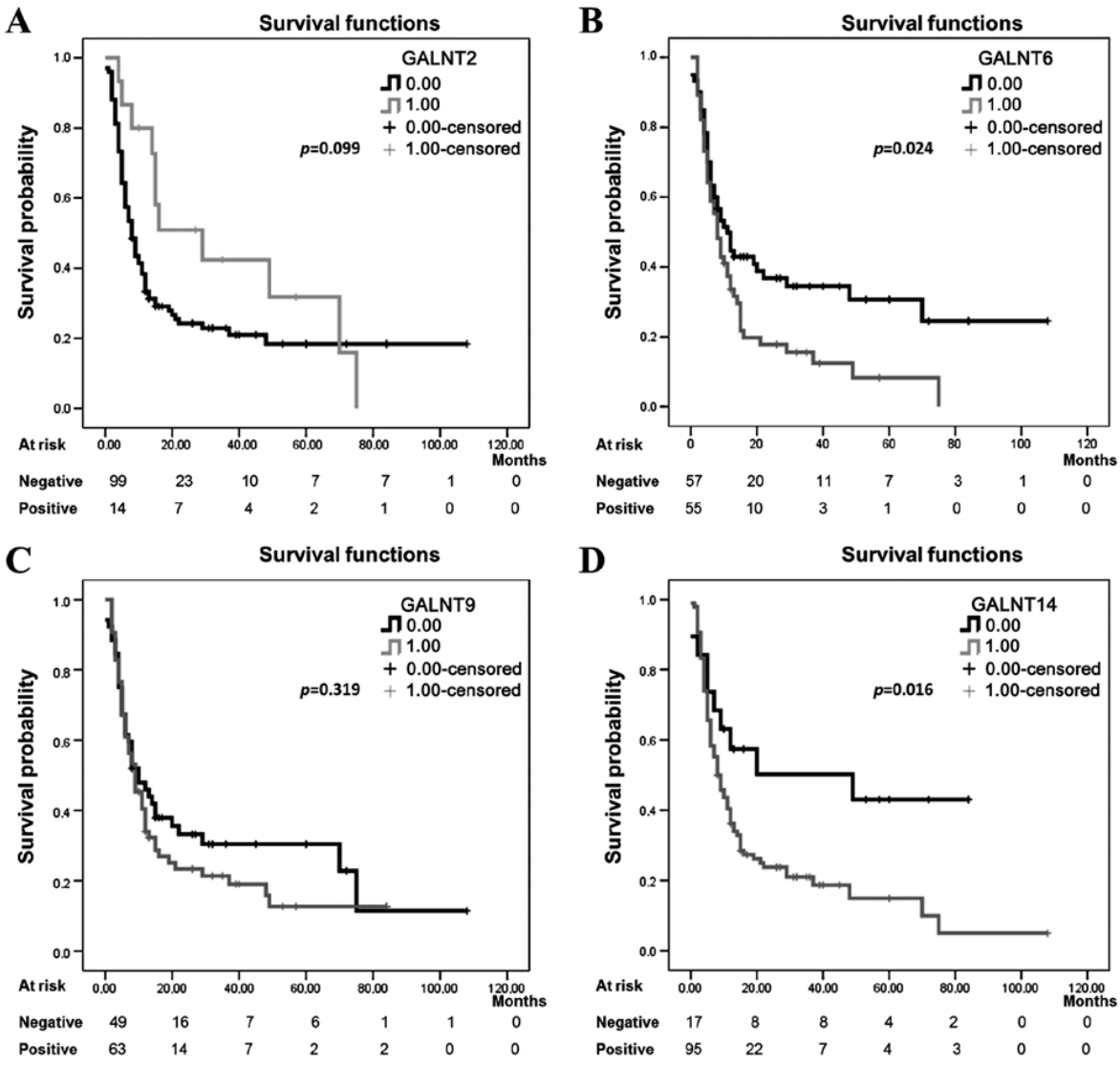


Figure 3.4 Kaplan-Meier progression-free survival curves showing the association between GalNAc-Ts expression patterns and prognosis in HG ovarian cancer

Chapter 4: Elucidating the role of the polypeptide N-acetylgalactosaminyltransferase 3 (*GALNT3*) and its closest homolog-*GALNT6* in mediating aberrant O-glycosylation associated with ovarian cancer dissemination

Razan Sheta[1,2], Magdalena Bachvarova[2], Elizabeth Macdonald[3], Stephane Gobeil[1,4], Barbara Vanderhyden[3] , Dimcho Bachvarov[1,2]

[1] Department of Molecular Medicine, Université Laval, Québec (Québec), Canada,

[2] Centre de recherche du CHU de Québec, L'Hôtel-Dieu de Québec, Québec (Québec),

[3] Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, ON, Canada,

[4] Centre de recherche du CHU de Québec, CHUL, Québec (Québec), Canada

(In preparation)

4.1 Preface

This paper is currently in preparation for submission on September 2017. I am listed as the first author on the paper. I was involved in performing and analyzing almost all experiments presented in this chapter. Magdalena Bachvarova, research assistant in our laboratory, was specifically involved in performing DNA microarrays gene expression analyses. Ms. Elizabeth Macdonald was involved in performing all *in vivo* experiments and collecting data for the study. Dr. Stephane Gobeil kindly provided us with shRNA constructs for the *GALNT6* gene. All *in vivo* studies were performed in the laboratory of Dr. Barbara Vanderhyden. Dr. Dimcho Bachvarov is the principal investigator of this study, he was involved in experimental design, data analysis and article writing.

4.2 Résumé en français

Jusqu'à présent, vingt membres de la famille des gènes GalNAc-Ts ont été identifiés et ce grand nombre de GalNAc-Ts pourrait fournir un soutien substantiel à la biosynthèse des protéines. Bien que différentes GalNAc-Ts soient différentiellement exprimées au sein des tissus, entre les cellules d'un même tissu et dans différents modèles à des stades variés de développement, il est aujourd'hui admis que quelques GalNAc-Ts ont des spécificités de substrats particulières mais également redondantes. Certaines des isoformes de GalNAc-Ts chez les mammifères, ont été regroupées en sous-familles en fonction de leur grande homologie. Par exemple, la sous-famille composée de *GALNT3* et *GALNT6* présente une homologie de 65% dans leur séquence codante. Fait intéressant, *GALNT6* a montré des fonctions oncogènes assez similaires dans le cancer du sein (modulant l'O-glycosylation aberrante et la stabilisation de MUC1), comme celles que nous avons trouvées pour *GALNT3* dans le CEO. L'augmentation de la co-expression de *GALNT3* et -*T6* a été détectée dans de multiples carcinomes, suggérant des effets complémentaires des *GALNT3* et -*T6.* La régulation positive de *GALNT3* dans la transformation maligne et la progression de ces cancers pourrait dépendre en partie de celle de *GALNT6* à la fois de manière synergique et compensatoire. Cela nous a incité à vérifier si la déplétion de *GALNT3*, dans les cellules CEO, induit une régulation compensatoire des autres gènes GalNAc-Ts, y compris *GALNT6*. Nous avons constaté que, en effet, la perte de *GALNT3* dans les cellules CEO A2780s était associée à une forte induction de l'expression de *GALNT6* à la fois *in vitro* et *in vivo*. Dans ce chapitre, nous avons voulu voir si la déplétion simultanée en *GALNT3* et en *GALNT6* a un effet profond *in vivo* correspondant à une suppression de l'expansion tumorale et des métastases du CEO et à une amélioration du taux de survie dans le modèle de souris xénogreffé. Les données ont montré des résultats positifs *in vitro* et *in vivo* en soutenant le concept de redondance fonctionnelle entre les deux gènes dans le CEO.

4.3 Abstract

Epithelial ovarian cancer (EOC) is the most lethal gynecologic malignancy, thus understanding molecular changes associated with EOC metastasis could lead to the identification of essential therapeutic targets. Glycosylation is a post-translational modification (PTM) that participates in major pathophysiology events during tumor progression. Aberrant glycosylation in cancer is frequently attributed to altered expression of different polypeptide N-acetylgalactosaminyltransferases (GalNAc-Ts) - the enzymes initiating mucin-type O-glycosylation. Our lab has identified the N-acetylgalactosaminyltransferase 3 (*GALNT3*) gene as a potential EOC oncogene, highly expressed in advanced disease. We demonstrated that *GALNT3* contributes to EOC dissemination through aberrant glycosylation of O-glycoproteins. Interestingly, the GalNAc-Ts family of isoenzymes reveal genetic and functional redundancy. The GalNAc-Ts exhibit overall sequence similarity of around 45%, while *GALNT3* and *GALNT6* display 65% homology. Redundancy has been suggested as a mechanism of cooperation amongst these genes. We hypothesized that the *GALNT6* gene may be providing biosynthetic backup in cells ablated for the *GALNT3* gene. Our results show that when knocking out the *GALNT3* gene in A2780s cells, an upregulation of the *GALNT6* expression occurs. We consecutively demonstrated that GALNT6 provides functional redundancy when knocking out the *GALNT3* gene. Cell migration, invasion, proliferation and cell cycle analysis were significantly reduced in double *GALNT3/T6* gene KO cells compared to single *GALNT3* gene KO. These observations were supported when examining the expression pattern of fibronectin (FN1), a suggested glycoprotein substrate of *GALNT6*, which displays a significant increase in *GALNT3* KO clones as a result of the observed induction of GALNT6 expression. Furthermore, MUC1 expression was downregulated in both *GALNT3* KO and *GALNT3/T6* KO, with a more prominent decrease observed in the double KO clone. *In vivo* studies supported the observed functional redundancy imposed by *GALNT6*, as animals injected with double *GALNT3/T6* KO EOC cell clones displayed a significant increase in survival rates compared to those injected with the control and the *GALNT3* KO clones.

Our study is the first to report a significant positive effect on mice survival upon knocking out two GalNAc-Ts genes, showing great promise in using both of these genes as EOC markers and/or therapeutic targets. Collectively our data provide strong evidence for the possible genetic redundancy of the two polypeptide GalNAc-Ts (*GALNT3* and *T6*) in EOC.

4.4 Introduction

Epithelial ovarian cancer (EOC) is a highly aggressive disease, and is responsible for more cancer deaths among women in the Western world than all other gynecologic malignancies (1, 2). EOC lethality primarily stems from the difficulty in detecting the disease at an early, organ-confined stage, and the lack of effective therapies for the advanced-stages of the disease, which represent the challenges associated with abrogating the spread of EOC (1). Thus, there is an urgent need for new therapeutic targets and a better understanding of the mechanisms involved in the spread of ovarian carcinoma. EOC is histologically classified into different carcinoma subtypes, including low-grade serous, high-grade serous, endometrioid, mucinous and clear cell carcinoma (3). EOC treatment includes the combination of surgery and platinum based chemotherapy (4); these approaches in addition to hormonal therapy have been associated with up to 80% responsiveness in patients, but unfortunately more than 50% of EOC patients eventually relapse. In order to improve the survival rates of EOC patients, research on the mechanism of EOC pathogenesis is essential.

It is well established that cancer invasion and metastasis represent the two major causes of the failure of cancer treatment (1). Thus, management of the metastatic disease is a major problem for the treatment of EOC. One possible way to resolve this problem is to target metastasis-specific pathways with novel therapies. Hence, focused identification of novel pro-metastatic mechanisms, target pathways and molecules could help identify new effective therapies.

Protein O-glycosylation is a highly heterogeneous enzymatic process that allows for the proper assembly of glycan sugars to proteins and lipids. Glycoproteins are involved in many biological functions, such as intercellular communication (5), intracellular signaling (6), and protein stability (7). Glycosylation perturbations such as aberrant glycosylation of the glycan sugars affect cellular processes such as cell proliferation, differentiation, migration, invasion, cell cycle and drug sensitivity (8, 9). The most commonly examined type of O-glycosylation is the mucin type, and this type of glycosylation is mediated by the family of 20 N-acetylgalactosaminyltransferases (GalNAc-Ts) (10). Altered expression of glycans in cancer is frequently attributed to under- or overexpression of

glycosyltransferases (11-14), and these studies denote the important role these GalNAc-Ts play in the regulation of cancer-associated cellular characteristics. Interestingly, very few studies have examined the role of GalNAc-Ts in EOC. We have previously identified the GalNAc-T3 (*GALNT3*) gene as a potential EOC oncogene, highly expressed in advanced disease, as *GALNT3* expression was significantly associated with poor outcome (15). We also demonstrated that *GALNT3* might contribute to EOC dissemination through aberrant glycosylation of different O-glycoproteins, including the *MUC1* oncogene (15). Subsequently, we analyzed the altered expression of glycoproteins following *GALNT3* gene knockdown (KD) in EOC cells by using a metabolic labeling strategy for enrichment and glycoprotein analysis, coupled with MS characterization of enriched glycoproteins (16, 17). This glycoproteomics approach led to the identification of numerous O-glycoproteins differentially expressed upon *GALNT3* KD, including O-glycoproteins implicated in mechanisms of cellular metabolism, PTMs and EOC dissemination and which could represent novel EOC biomarkers/therapeutic targets (16, 17). Furthermore, another study has suggested that *GALNT14* plays a role in modulating MUC13 glycosyaltion by promoting ovarian cancer cell migration. Moreover, one recent study has examined the role of *GALNT6* expression levels and its association with the negative prognosis in EOC patients, which has been suggested to be linked with the role *GALNT6* plays in modifying EGFR O-glycosylation contributing to the observed malignant behavior of EOC cells (18).

Although different GalNAc-Ts are differentially expressed within tissues, between cells, within a single tissue, and in different patterns at different stages in the development and differentiation, it is now becoming clear that a subset of the GalNAc-Ts display both distinct and overlapping substrate specificities (11, 19). For instance, the *GALNT3* and *GALNT6* gene members display 65% homology in their coding sequence (20). Interestingly, *GALNT6* exhibited very similar oncogenic functions in breast cancer (modulating aberrant O-glycosylation and MUC1 stabilization) (21), as those found from our previous work examining the role of *GALNT3* in EOC dissemination (15).

Understanding the role that the *GALNT3* gene plays in the aberrant glycosylation of MUC1 was briefly elaborated in our previous study (15). Some of the mammalian GalNAc-T isoforms have been grouped into subfamilies based on their high homology (69).

Evolutionary studies cannot predict why the GalNac-T family is so diverse and if this diversity can be linked to some degree of redundancy between the genes. Data from our previous study examining the expression of GalNac-Ts in high-grade serous carcinomas (HGSC) showed that there is simultaneous overexpression of four GalNAc-Ts (*GALNT3, T6, T9* and *T14*) in advanced EOC (Chapter 3). Our study also strongly suggested for a possible functional overlap and redundancy of *GALNT3* and *GALNT6* in EOC (Chapter 3). When examining protein-protein interactions of GALNT3 and GALNT6 using the STRING database (http://string-db.org) (Figure 4.1), it is clear that GALNT3 and GALNT6 have a direct interaction in their role in glycosylating MUC1.

Additionally, multiple *in vitro* studies of the GalNAc-T isoenzymes suggest that this enzyme family has a wide variety of roles in the process of proteins' glycosylation, but their regulation and specified role in normal development and cancer biology have not been examined. Moreover, *in vivo* examination of deficiencies in the GalNAc-Ts genes appears to cause moderate or subtle phenotypic alterations in animal models when examining cancer development or metastasis (19). This is related to the seemingly broad and partly overlapping roles of the different GalNAc-Ts in the protein glycosylation, since many GalNAc-Ts can serve to fine-tune protein functions as observed in development and different pathologies, including cancer. A report by Bennett *et al*. (1999), demonstrated the genetic redundancy of the GalNAc-transferases (*GALNT3* and *GALNT6*), but suggested that this genetic overlap does not provide full functional redundancy (20). In this study, we showed that *GALNT6* provides functional redundancy when knocking out (KO) the *GALNT3* gene. This was observed when performing functional analyses comparing the effect of a single gene *GALNT3* KO versus a double gene KO for both *GALNT3* and *GALNT6*. Cell migration, invasion, proliferation and cell cycle analysis were significantly reduced in the double *GALNT3/T6* gene KO cells compared to the single *GALNT3* gene KO, which suggested that *GALNT6* expression compensates for the lack of *GALNT3* in the single *GALNT3* gene KO clones. This genetic activity was supported when examining the expression pattern of fibronectin (FN1). Our protein expression analysis showed a significant increase in FN1 expression in the *GALNT3* KO clones, supporting the stabilizing effect of *GALNT6* expression on FN1. Furthermore, MUC1 protein expression was more strongly downregulated in the *GALNT3/T6* KO cells compared to *GALNT3* KO

clones and controls. Finally, *in vivo* studies very well supported the observed functional redundancy of *GALNT6*, as animals injected with the double *GALNT3/T6* gene KO clones showed significant increase in survival rates compared to those injected with the single *GALNT3* gene KO or the control clone, respectively. Our study is the first to report a positive effect on mice survival upon knocking out two GalNAc-Ts genes, showing great promise in using both these genes as EOC markers/therapeutic targets.

4.5 Results

4.5.1 Overexpression of *GALNT6* in *GALNT3* KO EOC cells

Our previously published data (Chapter 3), in addition to available information mined from the literature are strongly indicative for a possible functional overlap of *GALNT3* and *GALNT6* genes in cancer. We thus wanted to examine the role these two genes play in EOC dissemination. *GALNT3* KO clones were generated in A2780s cells using the CRISPR/Cas9 system (Figure 4.2A). Interestingly, around 80% of the *GALNT3* KO clones showed an increase in GALNT6 protein expression as observed using Western blot (Figure 4.2A). These protein expression data confirm the contribution of the *GALNT6* gene in compensating for *GALNT3* absence in this EOC cell line. Western blots analysis was performed to check if other GalNAc-T members could also compensate for the loss of the *GALNT3* gene in the A2780s cells. We have previously presented data examining protein levels of *GALNT9* and *GALNT14* in several EOC cell lines (Chapter 3), and both of these GalNAc-T members show relatively high expression in the A2780s cell line (Chapter 3), while *GALNT2* does not show any expression in the A2780s cell line (Chapter 3). However, protein expression analyses did not show any significant overexpression in the levels of GALNT2, GALNT9 or GALNT14 in the *GALNT3* KO clones (Supplementary Figure 4.1A). Subsequently, double gene *GALNT3/T6* KOs were generated in the A2780s cell line, as short hairpin (shRNA) approach was used to target *GALNT6* in the *GALNT3* KO clones. CaOV3 is one of the cell lines that show high expression of *GALNT6* (Chapter 3), and we thus sought to test the efficiency of shRNA targeting *GALNT6* in this EOC cell line. Protein expression analysis showed very high downregulation in the expression of GALNT6 in the CaOV3 cell line following shRNA KD (Supplementary Figure 4.1B).

Similarly qPCR data confirmed the absence of expression of *GALNT6* after shRNA KD (Supplementary Figure 4.1C). However, no differences in *GALNT3* expression were observed both at the gene or protein expression levels upon *GALNT6* KD (Supplementary Figure 4.1B and 4.1C). The GALNT3 protein levels show relatively strong expression in the CaOV3 cell line (Supplementary Figure 4.1D), and we hypothesize that the genetic redundancy of the *GALNT6* gene was a result of the loss of function of the *GALNT3* gene. These results may indicate that this type of compensation is not required by the *GALNT3* gene in the CaOV3 cell line.

In support of this data, we looked to confirm if *GALNT6* is not only overexpressed but is also activated in our studied *GALNT3* KO clones. We examined the expression pattern of fibronectin (FN1). Earlier studies showed FN1 to be glycosylated by *GALNT3* (77), but more recent studies disproved the specificity of *GALNT3* to the glycosylation site of FN1 and showed FN1 as a candidate substrate of *GALNT6* (70, 78). A significant increase in FN1 expression in the *GALNT3* KO clones was observed (Figure 4.2B), supporting the stabilizing effect of *GALNT6* expression on FN1. Collectively our data provide strong evidence for the possible genetic redundancy of the two polypeptide GalNAc-Ts (*GALNT3* and *T6*) in EOC.

We next created double *GALNT3/T6* KO clones in the A2780s cell line by using the shRNA targeting *GALNT6* to abolish its gene expression in *GALNT3* KO A2780s clones. We were able to establish such clones in the A2780s cell line and our protein expression data show a very strong KD for the *GALNT6* gene in the A2780s *GALNT3* KO clones referred to hereinafter as *GALNT3/T6* double KO clones (Figure 4.2B).

4.5.2 *GALNT3* and *T6* KO reduced cell proliferation rate, migration, invasion and cell cycle in EOC cells

To verify the observed compensatory role of *GALNT6* in the absence of *GALNT3*, multiple functional assays were employed in order to examine what kind of cancer related phenotypic changes maybe acquired upon knocking out the two members of the GalNAc-T family in EOC. It is well established that O-glycosylation plays a very important role in

regulating cell migration and the level of cellular invasiveness. These functional characteristics are dependent on the expression of members of these GalNAc-Ts, where inhibition of one or multiple GalNAc-Ts have been shown to either induce or reduce the level of invasion and metastasis formation in different cancer cells (22-25). One *GALNT3* KO A2780s clone (GAL3 KO) and one *GALNT3/T6* double KO (GAL3/6 KO) A2780s clone were selected for further functional analyses (Figure 4.3), as the impact of the *GALNT3* KO and the double *GALNT3/T6* depletion was investigated on A2780s migration, invasion, cell proliferation, and cell cycle control. Knocking out the *GALNT3* gene inhibited both migration and invasion of A2780s cells (Figure 4.3A and 5.3B), but interestingly, as shown in Figure 4.3A (migration) and Figure 4.3B (invasion), the numbers of A2780s cells that passed through the filter in the double *GALNT3/T6* KO A2780s clones were remarkably smaller when compared to both the single *GALNT3* KO clone and the Ctrl clone (Figure 4.3). Moreover, the *GALNT3* gene ablation led to a sharp decrease in the number of viable adherent A2780s cells (represented by cell index), compared to Ctrl cells (Figure 4.3C), and interestingly a stronger decrease was observed between the Ctrl and the double *GALNT3/T6* KO clone (Figure 4.3C); these observations were further supported by the reduced number of colony formation upon double *GALNT3/T6* suppression compared to *GALNT3* KO and controls (Figures 4.3D). Finally, the *GALNT3* depletion also induced G1 cell cycle arrest (Figure 4.4), and this observation was similarly more prominent when examining cell cycle analysis in the double *GALNT3/T6* KO clones. The observed reduction in the G1 phase was evident starting from time 0h, where the percentage of cells accumulating at time 0h for the A2780s Ctrl cells were at 41.5%, compared to 34.8% and 32.8% for the *GALNT3* KO and *GALNT3/T6* KO clones (Figure 4.4). The most evident differences were observed at 6h, the percentage of cells accumulating at time 0h for the A2780s Ctrl cells were at 32.3.5%, compared to 25.4% and 18.4% for the *GALNT3* KO and *GALNT3/T6* KO clones (Figure 4.4).

4.5.3 Evidence of full functional redundancy in protein glycosylation by *GALNT6* in *GALNT3* KO EOC cells

To further understand the biologic function of *GALNT3* and *GALNT6* and the effect of their ablation in the A2780s cells, we focused on studying MUC1 and FN1 as substrates of these two enzymes, since as previously discussed, MUC1 has been reported to be a candidate substrate of *GALNT6* in breast cancer, and we have also previously supported that *GALNT3* has a similar capacity to glycosylate MUC1 in EOC. In order to better examine glycosylation of MUC1 and FN1 by both *GALNT6* and *GALNT3*, protein expression of both MUC1 and FN1 were evaluated by Western blot. Data presented a large but not complete reduction in MUC1 protein expression in the *GALNT3* single KO clone, while MUC1 protein expression was interestingly completely reduced in the *GALNT3/T6* double KO clones (Figure 4.5A), and as previously demonstrated FN1 was highly upregulated in the *GALNT3* KO clone while absent in both the control and *GALNT3/T6* KO clone (Figure 4.2B and 4.5A). Moreover, we complemented this data by performing VVA lectin pull down assays, and interestingly we detected a glycosylated band of MUC1 protein by VVA-lectin Western blot in the Ctrl clone, which was slightly reduced in the *GALNT3* KO clone, but completely absent in the double *GALNT3/T6* clone (Figure 4.5B). MUC1 has been suggested as a direct substrate of *GALNT6*, and it appears that the compensation level of *GALNT6* in the A2780s EOC cell line was relatively sufficient to contribute in the O-glycosylation and stabilization of MUC1 (Figure 4.5B). Moreover, the VVA lectin blots also showed a slight increase in the FN1 glycosylated band in the *GALNT3* KO clone, which was not observed in either the Ctrl or the *GALNT3/T6* double KO clone (Figure 4.5B). Collectively this data confirms the activity and compensatory role that *GALNT6* plays in the *GALNT3* KO clone, proving the functional redundancy imposed by *GALNT6* in the A2780s EOC cell line. We also examined the level of *MUC1* and *FN1* at the gene transcript level in both KO clones, and no differences were observed in the mRNA levels between Ctrl, *GALNT3* KO and *GALNT3/T6* KO clones (Figure 4.5C), which confirms that alterations of MUC1 and FN1 protein expression are due to glycosylation modifications.

4.5.4 Molecular mechanisms of *GALNT3* and *GALNT6* action in EOC cells

To better understand the molecular mechanisms of *GALNT3* and *GALNT6* action in EOC cells, we employed the Agilent Whole Human Genome 4x44K microarrays (containing 44, 000 genes) to identify global gene expression changes upon *GALNT3* KO and *GALNT3/T6* KO in A2780s cells. The gene expression patterns of the selected clones were compared against the corresponding Ctrl clone. All microarray experiments were performed in duplicates, as two hybridizations were carried out for each of *GALNT3* and *GALNT3/T6* KO clones against the corresponding Ctrl, using a fluorescent dye reversal (dye-swap) technique. For all comparisons, a subset of common differentially expressed genes was selected by initial filtering on confidence at p-value ≤ 0.05, followed by filtering on expression level ($\geq$ 1.5 fold). Using these selection criteria, we found 417 genes to be upregulated and 868 genes to be downregulated in A2780s cells following *GALNT3* KO, while a much larger number of differentially regulated proteins were detected the *GALNT3/T6* double KO clone, where 2503 genes were found to be upregulated and 4553 genes were found to be downregulated.

Canonical, functional and network analyses generated through the use of the Ingenuity Pathway Analysis (IPA) software reflected the cancer-related phenotypic changes observed in A2780s and when comparing the *GALNT3* KO clone to the *GALNT3/T6* double KO clone. The top downregulated canonical pathways observed when comparing the double gene KO to the single gene KO were related to *BRCA1* role in DNA damage response, molecular mechanisms of cancer, TGF-β signaling, p-53 signaling, serine biosynthesis, cell cycle regulation, in addition to major cancer related signaling pathways such as (AMPK signaling, Protein Kinase A signaling, EIF2 signaling, and the mTOR signaling pathway) (Figure 4.6A), while upregulated canonical pathways were predominantly associated with regulation of epithelial-mesenchymal transition (EMT) pathway, interferon signaling, protein regulation pathways such as unfolded protein response, and protein ubiquitination pathway, in addition to an upregulation in the antigen presentation pathway (Figure 4.6B). Collectively this data confirms that the suppression of two members of the GalNAc-T family has a stronger and a more profound effect on many of the functional characteristics of EOC cells. The above data support our *in vitro* observations about the role these two genes play in regulating numerous pathways such as cell cycle regulation and other major pathways that have shown strong implication in cancer cell migration and invasion.

Furthermore, functional pathway analysis was also confirmative of the observations we have made *in vitro* in addition to some of the major canonical pathways listed above (Figure 4.6 and 4.7). The top downregulated functional pathways observed when comparing the double *GALNT3/T6* gene KO to the single *GALNT3* gene KO were related to cell cycle, DNA replication, recombination and repair, gene expression, post-translational modification and protein synthesis (Figure 4.7A), while interestingly when comparing the *GALNT3* gene KO to the double *GALNT3/T6* gene KO it is clear that there is a predominant donwregulation in metabolic pathways related to carbohydrate metabolism, lipid metabolism and amino acid metabolism (Figure 4.7A), all confirmative of the role the *GALNT3* gene has shown to play in regulating metabolic pathways in EOC which was pointed in our previous studies (17, 26). Moreover, when examining upregulated functional pathways upon comparing the double *GALNT3/T6* gene KO to the single *GALNT3* gene KO, our data show that the major upregulated pathways are related to cell death and survival, cell morphology, cell signaling, cellular compromise, and post-translational modification (Figure 4.7B). Collectively this data emphasizes and complements initial canonical pathways observations in these KO clones in the A2780s EOC cell line.

Common IPA networks, obtained upon merging the five top-scoring networks following both *GALNT3* suppression and GALNT3/T6 suppression were indicative for important gene nodes linked to the suppression of these GalNAc-Ts enzymes in EOC cells, as most of these substantiate and/or complement the functional data obtained. As shown in Figure 4.8, GALNT3 KO in A2780s cells led to a strong downregulation in several gene nodes known to be implicated in EOC tumorigenesis including *VEGF, PI3K* complex, members of the *PRC2* complex including *EZH2*, and *EED*. Moreover, a noted upregulation in some homeobox gene members, such as *HOXD10*, was observed (Figure 4.8) as these genes play role in reducing cancer cell invasiveness including EOC spreading (27). Likewise, the top 5 common IPA networks following *GALNT3/T6* suppression were indicative for the downregulation of major gene nodes known to have major implications in EOC, such as members of the *PRMT* family, *BMI-1* and its multiple documented functions in enhancing the progression of EOC, in addition to other gene nodes showing great promise as EOC therapeutic targets, such as *CUL1, EEF2, PTN, UCHL1, CBX5* and *USP14* (Figure 4.9).

To validate microarray results, we arbitrarily selected several differentially expressed genes from each of the two experimental conditions (*GALNT3* KO and *GALNT3/T6* KO in A2780s cells) and quantified their expression by qPCR upon comparison to the corresponding controls. Supplemental Figure 4.2 summarizes the gene expression measurements of all validated genes. We found that both methods (microarray analysis and qPCR) detected similar patterns for the up- and down-regulated genes selected for validation.

4.5.5 Double *GALNT3 GALNT6* gene KO reduces EOC metastasis *in vivo*

Our *in vitro* studies provided evidence for a strong oncogenic capacity of *GALNT3* and *GALNT6* in EOC, including their potential role in EOC cell proliferation, cell cycle control and cell migration/ invasion. Several studies have focused on performing genetic approaches to understand the roles of GalNac-Ts in mammalian O-glycosylation, and many found that the loss or KD of one GALNT gene may not be enough to produce detectable phenotypic changes or affect the survival of experimental animals. Synthetic deficiencies of the GalNac-Ts have been produced for animal model studies, and only a few have showed signs of effects on animal survival or development. These studies confirm that redundancy may exist amongst the big GalNac-T family. By performing *in vivo* studies we aimed to investigate if the demonstrated redundant oncogenic function of the *GALNT3* and *GALNT6* genes can be confirmed *in vivo*, and thus help emphasize their strong implication in EOC tumorigenesis. EOC spreads by intraperitonial (IP) sloughing, lymphatic invasion, and hematogenous dissemination (28). IP dissemination takes place by malignant cells evading from the ovarian capsule (29). Several reports have demonstrated that inoculation of tumor cells through IP injection can best mimic EOC metastasis (30). Hence, IP injection of cancer cells in animal models can accurately model advanced disease, as EOC metastases frequently appear disseminated throughout the peritoneum (31). We used a similar *in vivo* approach; thus, Ctrl, *GALNT3* KO and *GALNT3/T6* KO A2780s clones were IP injected in SCID mice (n=8 Ctrl group, n=7 *GALNT3* KO group, and n=7 *GALNT3/T6* KO group; two animals were consecutively removed from the experiment since the cause of death was not

related to the experimental treatment). Mice injected with Ctrl cells displayed a significantly shorter survival (p = 0.0193) than those injected with *GALNT3* KO cells, reaching endpoint on average 31 (+/- 4.80 SEM) and 38 (+/- 4.35 SEM) days post injection respectively (Figure 4.10A). However, the double *GALNT3/T6* KO cell line resulted in a more significant difference in survival when compared to mice injected with the Ctrl cells, as well as those injected with the *GALNT3* KO cells (p = 0.0005) and (p = 0.0010) respectively, reaching endpoint on average 57 (+/- 8.16) days post injection (Figure 4.10A). There were no significant differences in the tumor burden between mice injected with *GALNT3* KO cells and those injected with the Ctrl cells (p = 0.2343) reaching an average tumor mass of (7.9g +/- 2.53), compared to (6.18g +/- 1.95) respectively (Supplementary Figure 4.3A). In comparison, mice injected with the *GALNT3/T6* KO cells displayed significantly smaller tumor masses when compared to both the Ctrl and *GALNT3* KO groups (p = 0.0012) and (p = 0.0411) respectively, having tumor mass of (3.66g +/-0.73) at endpoint (Supplemental Figure 4.3A).

Consistent with our previous *in vitro* data, tumor specimens derived from *GALNT3* KO and *GALNT3/T6* KO cells showed very low IHC staining intensity for the GALNT3 protein (Figure 4.10B). Moreover, tumors injected with the *GALNT3* KO cells showed strong staining for GALNT6 and FN1, and low staining for MUC1 when compared with Ctrl A2780s derived tumors (Figure 4.10B). Whereas *GALNT3/T6* KO clones showed no, or reduced staining intensity for all of GALNT6, FN1 and MUC1 when compared to the Ctrl and *GALNT3* KO cells (Figure 4.10B). These data sustain our *in vitro* findings for the role of *GALNT3* and *GALNT6* in glycosylating and stabilizing protein substrates in EOC cells. The IHC results were further supported by Western blot data (see Supplemental Figure 4.3B).

4.6 Discussion

So far, twenty members of the human GalNAc-Ts gene family have been identified (32), as this abundant number of GalNAc-Ts could provide substantial biosynthetic back-up. Although different GalNAc-Ts are differentially expressed it is now becoming clear that a subset of the GalNAc-Ts display both distinct and overlapping substrate specificities (11, 19). Some of the mammalian GalNAc-T isoforms have been grouped into subfamilies based on their high homology (19, 32). One such example is the human subfamily comprised of *GALNT3* and *GALNT6* (20). *GALNT3* has been often suggested as a possible therapeutic target in multiple cancer types, and in recent years more studies have examined its role in cancer such as in pancreatic (33, 34), gastric (35, 36), colorectal cancers (37). Similarly, *GALNT6* has been also extensively studied for its role in the development of different cancers such as breast (21, 38, 39), pancreatic (40, 41) and gastric cancer (42). Moreover, some recent studies have examined the expression of *GALNT6* in ovarian cancer, where one study suggested that *GALNT6* overexpression in EOC is associated with poor prognosis and outcome of ovarian cancer, and this is accomplished through the regulation of EGFR glycosylation (18). Interestingly, the expression of *GALNT6* has also been suggested to be dependent on the type and stage of ovarian cancer, since one study has correlated improved long-term survival with *GALNT6* expression in low-grade serous carcinoma (43). These studies suggest that these GalNAc-Ts do indeed show differential expression between cancer types and even between different cancer stages.

Importantly, increased *GALNT3* and *GALNTT6* co-expression has been detected in pancreatic (41) renal (44), prostate carcinomas (45). Furthermore, showed a strong correlation between the expression of both *GALNT3* and *GALNT6* in different stages of EMS (46). Overall, these data suggest that the upregulation of *GALNT3* within the malignant transformation and progression of cancer could partly depend on that of *GALNT6* in both synergistic and compensatory ways. Such hypotheses are supported by studies documenting that *GALNT6* displays quite similar oncogenic functions in breast cancer (modulating aberrant O-glycosylation and MUC1 stabilization) (21), as those found by us for *GALNT3* in EOC (15).

Multiple *in vitro* studies on different GalNAc-Ts suggest that these enzymes display a wide variety of roles in the process of proteins' glycosylation, but their regulation and specified role in normal development and cancer biology and have not been suggested. In this study we closely examined the possible compensatory role of *GALNT6* upon knocking out the *GALNT3* gene in the A2780s EOC cell line. Data gathered here indicate the presence of an overexpression in *GALNT6* when ablating *GALNT3* in A2780s cells. This was further supported when performing functional analyses comparing the effect of a single gene *GALNT3* KO versus a double gene KO for both *GALNT3* and *GALNT6*. Cell migration, invasion, proliferation and cell cycle analysis were significantly reduced in the double *GALNT3/T6* gene KO cells compared to the single *GALNT3* gene KO, which was suggestive that *GALNT6* expression provides a strong backup genetic activity in the single *GALNT3* gene KO clones. This genetic functionality was supported when examining the expression pattern of FN1, a glycoprotein that is synthesized within the cell matrix surrounding tumors and fetal tissue (47). A significant increase in FN1 expression was observed in the *GALNT3* KO clones, which was evidently associated with the compensatory *GALNT6* expression in these clones. This was further proven by VVA lectin pull down assays, confirming the activity of *GALNT6* in the *GALNT3* KO cells. Furthermore, MUC1 expression was also shown to be relatively downregulated in the *GALNT3* KO, but displayed much strongert downregulation in the *GALNT3/T6* KO cells, when examined using both protein expression analysis and VVA lectin pull down assays.

The above experimental data were further confirmed upon analyzing global gene expression variations observed upon knocking out *GALNT3*, or both *GALNT3* and *GALNT6* in the A2780s EOC cell line. Thus, canonical and functional pathways analysesd were strongly confirmative of the major effects the *GALNT3* and *GALNT6* double KO had on cancer related pathways. Moreover, network analyses further supported the implication these GalNAc-Ts had in modulating EOC tumorigenesis. Major gene nodes were shown to be affected when examining *GALNT3* KO A2780s cells, The *VEGF* gene node, which is a gene known for its role in the dissemination of ovarian cancer was shown to be downregulated (48-51). The *PI3K* complex was also shown to be affected by the *GALNT3* gene KO, and in ovarian cancer *PI3K* is a relevant pathway to target (52-54). *EZH2* was also downregulated, and is another gene known to promote ovarian cancer dissemination

(55-58). Likewise *IQGAP1* and *HOXD10* were two other gene nodes differentially regulated upon *GALNT3* KO, and studies have also documented their implication in ovarian cancer progression (59-61). Similarly, multiple gene nodes were also affected by the *GALNT3/T6* double KO in A2780s cells, some include members of the *PRTM* family, previously shown to be implicated in cancer development (62-64). Also a major gene node affected was the *BMI-1* gene, as this gene has been extensively examined for promoting aggressive EOC phenotypes (65-70). Other gene nodes also included perturbations in *CUL1* (71), *EEF2* (72), *PTN* (73), *UCHL1* (74), *CBX5* (75) and finally *USP14* (76), all involved in EOC progression.

In vivo examinations of deficiencies in the *GALNT* genes cause moderate or subtle phenotypic alterations in animal models when examining cancer development or metastasis (19). This has been possibly related to the seemingly broad and partly overlapping roles of the different GalNAc-Ts in the protein glycosylation. The concept of genetic and functional redundancy has been explored in many biological systems, and this phenomenon has been examined in genetic variants ranging from isoenzymes to transcription factors (77). The term genetic redundancy can also be referred to as functional redundancy and is defined as partial or complete overlap in function that occurs upon deleting or knocking out a gene from a genome. In this case, this ablation shows minimal impact on phenotype due to functional compensation that has been conferred by one or more genes with often-similar function (78, 79). Genetic studies of gene redundancy in S. cerevisiae (80-82) have shown that about 25% of redundant gene duplicates are metabolic enzymes (77). Interestingly, the GalNAc-Ts family of isoenzymes has been reported for their genetic and functional redundancy, as examined through their high sequence similarity and their evolutionary identity (19). Nonetheless, explanations behind the preservation of functional redundancy may stem from the fact that system biology favours synthetic backup by these genes in case of mutations that may have occurred in other members of the GalNAc-Ts throughout evolution (19). The GalNAc-Ts exhibit overall sequence similarity of around 45%, while the *GALNT3* and *GALNT6* display around 65% homology in their coding sequence (19, 20). Genomic sequencing studies have shown that although the genomic structures of *GALNT3* and *GALNT6* are highly homologous, these genes are localized at different locus (2q31 and 12q13, respectively (19, 20)). Additionally, studies have shown that they have

both similar and different substrate acceptor specificities (20). Nevertheless, redundancy has also been suggested as a mechanism of cooperation amongst these two genes, since glycosylation of specific acceptor sites in the MUC1 tandem repeat by one of them is required before other sites can be glycosylated by the other enzyme (20).

Several studies have focused on performing genetic approaches to understand the roles of GalNac-Ts in mammalian O-glycosylation. Importantly, a number of synthetic deficiencies of the GalNac-Ts have been produced in animal model studies, and only few have shown signs of effects on animal survival or development. The lack of presentable phenotypic differences is most likely not a result of very subtle changes in the genome, but is the outcome of effective mechanisms that have allowed decreasing any major perturbations that may have been acquired from a single gene knockout or deletion. We thus hypothesized that the *GALNT6* gene may be providing biosynthetic backup in those cells ablated for the *GALNT3* gene, and we proposed that it might occur through a feedback mechanism. We suggest that when one isoenyzme (*GALNT3*) is knocked out, the functional pathway of *GALNT3* gets reduced, which in turn triggers an increase in theexpression of its closest homolog *GALNT6* allowing for the reprograming of *GALNT3* pathway in these EOC cells. Our results show that when knocking out the *GALNT3* gene in A2780s cells, an upregulation of the *GALNT6* expression occurs. We consecutively demonstrated that *GALNT6* provides functional redundancy when knocking out the *GALNT3* gene. Cell migration, invasion, proliferation and cell cycle analysis were significantly reduced in double *GALNT3/T6* gene KO cells compared to single *GALNT3* gene KO. Moreover, our *in vivo* studies very well supported the observed functional redundancy imposed by *GALNT6* in EOC cells, as animals injected with the double *GALNT3/T6* gene KO clones showed significant increase in survival rates compared to those injected with the single *GALNT3* gene KO or the control clone, respectively. Our study is the first to report a positive effect on mice survival upon knocking out two GalNAc-Ts genes, showing great promise in using both these genes as EOC markers/therapeutic targets. Collectively our data provide strong evidence of the possible genetic redundancy of the two polypeptide GalNAc-Ts (*GALNT3* and *GALNT6*) in EOC and highlight the importance of integrating and examining multiple members of the GalNAc-T family into providing some prognostic information for EOC

patients. Moreover, screening for *GALNT3* and *GALNT6* inhibitors could be valuable for the development of novel therapeutic modalities against EOC.

4.7 Materials and methods

4.7.1 Cell culture

The EOC cell lines OVCAR4, CaOV3 and SKOV3 were purchased from American Tissue Type Collection (Manassas, VA); OV90, OV2008, TOV-112 and TOV-21 cell lines were a kind gift from Dr. Anne-Marie Mes-Masson (Montreal University), while A2780s and A2780cp cell lines were a kind gift from Dr. Benjamin Tsang (Ottawa University). The cell lines were passed in different culture media supplemented with 10% fetal bovine serum, as described previously (83, 84).

4.7.2 CRISPR/Cas9 mediated *GALNT3* KO in EOC cells

CRISPR/Cas9 mediated *GALNT3* KO in A2780s cells was done according to the manufacturer's guidelines (Santa Cruz Biotechnology, Inc). Briefly, 1.5 × 105 cells were plated onto 6 × 30-mm well plates and allowed to grow to 70% confluence. Five microliters of UltraCruz Transfection Reagent (sc-395739) were added to 2 μg of *GALNT3* CRISPR/Cas9 KO Plasmid (sc-407682) and 2 μg *GALNT3* HDR Plasmid (sc-407682-HDR). The complexes were incubated at room temperature for 20 min and then overlaid onto the cells. The plates were then incubated at 37°C, 5% CO2 for 48 h. Stably transfected clones were selected by adding puromycin (1 μg/ml). Selected clones were transfected with Cre vector (sc-418923) for the removal of genetic material flanked by LoxP sites. The control CRISPR/Cas9 Plasmid (sc-418922) contained a non-targeting 20 nt scramble guide RNA (gRNA) designed as a negative control. Thus, the Cas9/gRNA complex does not recognize any DNA sequence and will not bind or cleave genomic DNA.

4.7.3 Short hairpin RNA (shRNA) - mediated *GALNT6* knockdown in EOC cells

The shRNA-mediated *GALNT6* KD in EOC cells was done, as previously described (84). Briefly, two *GALNT6* shRNAs cloned into the pLKO.1-puro vector targeting the *GALNT6* mRNA sequences were retrieved from the Sigma Mission TRC human 1.5 shRNA library (clone numbers TRCN035450 and TRCN035489). Viral supernatants were generated by transfecting 293T cells with the shRNA constructs and the packaging vectors psPAX2 and pMD2.G (Addgene, Cambridge, MA). The high-titer lentiviral supernatants in the presence of 8 mg/ml polybrene were used to infect A2780s, and CaOV3 cells. Two days later, infected cells were treated with puromycin (1 μg/ml) for the selection of stably-transduced clones. The pLKO.1-puro vector encoding a scramble sequence not matching any mammalian sequence was used for the generation of mock-transduced (control) clones. Stable clones with inhibited *GALNT6* expression were evaluated and validated by quantitative RT-PCR, semi-quantitative PCR and Western blot.

4.7.4 Western blotting

Western blot analyses were performed as previously described (83, 84). Briefly, protein lysates were prepared by resuspending cell pellets in Laemmli sample buffer containing 5% β-mercaptoethanol. Protein lysates were separated by 6 to 12% Tris-glycine gel electrophoresis and transferred onto a polyvinylidene difluoride membrane. The membranes were blocked with 4% non-fat dry milk in TBST (20 mmol/l Tris-HCl, 0.5 M NaCl, and 0.1% Tween 20), incubated with the appropriate primary antibody at 4oC overnight: anti-GALNT3 (abgent), anti-MUC1, anti-FN1, anti-β-actin (Santa Cruz Biotechnology), anti-GALNT6 (abcam). After 3 x 15 min washes with TBST at room temperature, the membranes with the corresponding secondary antibody in TBST containing 4% non-fat dry milk for 1-2 h at room temperature. Upon washing, the signal was visualized using ECL solution (Thermo Fisher Scientific, Waltham, MA) and detected on blue sensitive autoradiography film (Marsh Bio Products, Rochester, NY).

4.7.5 Functional assays

Cell proliferation (cell index) was checked by the xCELLigence Real-Time Cell Analyzer (RTCA) instrument, as previously described (84). Colony formation assay was performed

as previously described (15). Cell migration and invasion assays were performed as previously described (15, 83, 84). Cell cycle flow cytometry analysis was performed as previously described (83), while the cell cycle phase distribution was calculated from the resultant DNA using the FlowJo software (v10). All statistical data were determined by a Student's t-test, where $p < 0.05$ was considered significant.

4.7.6 Semi-quantitative RT-PCR (sqRT-PCR)

Analysis of gene expression in stably transfected *GALNT3* and *GALNT3/T6* KO clones and the corresponding mock-transfected clone (Ctrl) was performed by sqRT-PCR as previously described (15). The *GUSB* gene was used as an internal standard. Primers were designed for these loci with the sequences freely available from the Entrez Nucleotide database and the Primer3 algorithm for primer design (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi).

4.7.7 Quantitative PCR (qPCR)

Quantitative PCR was performed as previously described (15). Briefly, total RNA was extracted by RNeasy Plus Mini Kit (QIAGEN) and cDNA was obtained by qScript™ cDNA SuperMix (Quanta BioSciences, Inc.). Primers were designed for these loci with the sequences freely available from the Entrez Nucleotide database and the Primer3 algorithm for primer design (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi). PerfeCTa® SYBR® Green FastMix® (Quanta BioSciences, Inc.) was used according to manufacturer's instructions. PCR reactions were performed on Rotor-Gene RG-3000 Real Time PCR System (Qiagen), with 18S ribosomal RNA used as endogenous control. PCR volume was 20 µl, and conditions were as follow: initial cycle 50°C, 2 min, 95°C, 15 min; 45 cycles at 95°C, 20 s, 60°C, 20 s and 72°C, 20 s; final cycle 72°C, 30 s. Data were analyzed by the Rotor-Gene software using the comparative ΔΔCt method. The relative copy number was calculated based on the target gene/18S RNA ratio.

4.7.8 VVA lectin pull-down assay for O-glycosylated (GalNAc-conjugated) proteins

VVA lectin pull-down assay was performed as briefly described (15). Briefly: 600 µg of cell lysate protein was incubated for 3 h at 4oC with 4 µg of biotinylated lectin VVA (EY Laboratories). Twenty µl of streptavidin-agarose (Sigma) was then added, and samples were incubated for an additional 2 h at 4°C with rotation. Lectin/glycoprotein complexes were collected by brief centrifugation (1400 rpm, 5 min), and washed three times with lysis buffer, followed by one wash with phosphate-buffered saline (PBS). Glycoproteins were released from the complexes by boiling in 30-50 µl SDS–PAGE sample buffers (5 min). The glycoproteins were resolved by SDS–PAGE, then immunoblotted to detect MUC1 and FN1 or GalNAc-conjugated proteins.

4.7.9 Gene expression profiling and data analysis

Gene expression analysis was carried out as previously described (84). Briefly, total RNA was extracted from the *GALNT3* KO clone, *GALNT3/T6* KO clone and their corresponding controls (Ctrl). The quality of the RNA samples was examined by capillary electrophoresis using the Agilent 2100 Bioanalyzer (Agilent). Fluorescently labeled cRNA targets were generated from 0.5 µg of total RNA from each corresponding cell line clone, using the Fluorescent Linear Amplification Kit (Agilent) and 10 mM Cyanine 3- or 5-labeled CTP (PerkinElmer), following the user's manual. Cyanine labeled cRNA from the clone suppressing *GALNT3* and *GALNT3/T6* in the A2780s cells were mixed with the same amount of reverse-color cyanine-labeled cRNA from their corresponding control (Ctrl) clone and hybridized on the Agilent Whole Human Genome microarrays, containing 44,000 genes. Array hybridization, washing, scanning, data extraction and analyses were performed as previously described (84). Network analysis of the microarray data was completed using the Ingenuity Pathway Analysis (IPA) software (see http://www.Ingenuity.com).

4.7.10 Peritoneal tumor formation in mice

Control A2780s cells, as well as cells from the *GALNT3* KO and *GALNT3/T6* KO clones (1 × 107 cells in 500 µl of PBS), were IP injected into 8 × 8 week old CB17 SCID female mice (CB17/Icr-Prkdcscid/IcrIcoCrl strain code 236, Charles River) using a 25G5/8 needle,

as previously describe (84). Mice were monitored daily by staff blinded to the cell type injected and euthanized when they reached a loss of wellness endpoint that was most often respiratory distress associated with ascites accumulation. The animals had free access to food and water and experiments were done in accordance with the Canadian Council on Animal Care's Guidelines for the Care and Use of Animals. Protocols were approved by the University of Ottawa Animal Care Committee.

4.8 Acknowledgements

This study was sustained by grants to D.B. and B.V. from the Cancer Research Society of Canada. Microarray experiments were performed in the Core Genomic Facility at the Cancer Research Centre at Laval University (https://www.crc.ulaval.ca/en/research/core-facilities/), supported by FRQ-S - Réseau de Recherche en Cancer.

4.9 References

1. R. Siegel, E. Ward, O. Brawley, A. Jemal, Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. CA: a cancer journal for clinicians 61, 212 (Jul-Aug, 2011).

2. B. M. Reid, J. B. Permuth, T. A. Sellers, Epidemiology of ovarian cancer: a review. Cancer biology & medicine 14, 9 (Feb, 2017).

3. K. R. Cho, M. Shih Ie, Ovarian cancer. Annual review of pathology 4, 287 (2009).

4. S. Bhoola, W. J. Hoskins, Diagnosis and management of epithelial ovarian cancer. Obstetrics and gynecology 107, 1399 (Jun, 2006).

5. M. J. Paszek et al., The cancer glycocalyx mechanically primes integrin-mediated growth and survival. Nature 511, 319 (Jul 17, 2014).

6. N. E. Zachara, G. W. Hart, Cell signaling, the essential role of O-GlcNAc! Biochimica et biophysica acta 1761, 599 (May-Jun, 2006).

7. A. J. Parodi, Protein glucosylation and its role in protein folding. Annual review of biochemistry 69, 69 (2000).

8. S. Hakomori, Glycosylation defining cancer malignancy: new wine in an old bottle. Proceedings of the National Academy of Sciences of the United States of America 99, 10231 (Aug 06, 2002).

9. S. R. Stowell, T. Ju, R. D. Cummings, Protein glycosylation in cancer. Annual review of pathology 10, 473 (2015).

10. K. G. Ten Hagen, T. A. Fritz, L. A. Tabak, All in the family: the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases. Glycobiology 13, 1R (Jan, 2003).

11. E. M. Beaman, S. A. Brooks, The extended ppGalNAc-T family and their functional involvement in the metastatic cascade. Histology and histopathology 29, 293 (Mar, 2014).

12. H. E. Miwa, Y. Song, R. Alvarez, R. D. Cummings, P. Stanley, The bisecting GlcNAc in cell growth control and tumor progression. Glycoconjugate journal 29, 609 (Dec, 2012).

13. D. L. Meany, D. W. Chan, Aberrant glycosylation associated with enzymes as cancer biomarkers. Clinical proteomics 8, 7 (Jun 03, 2011).

14. J. Ashkani, K. J. Naidoo, Glycosyltransferase Gene Expression Profiles Classify Cancer Types and Propose Prognostic Subtypes. Scientific reports 6, 26451 (May 20, 2016).

15. Z. Q. Wang et al., Role of the polypeptide N-acetylgalactosaminyltransferase 3 in ovarian cancer progression: possible implications in abnormal mucin O-glycosylation. Oncotarget 5, 544 (Jan 30, 2014).

16. R. Sheta et al., A metabolic labeling approach for glycoproteomic analysis reveals altered glycoprotein expression upon GALNT3 knockdown in ovarian cancer cells. J Proteomics, (Apr 17, 2016).

17. R. Sheta et al., Proteomic dataset for altered glycoprotein expression upon GALNT3 knockdown in ovarian cancer cells. Data Brief 8, 342 (Sep, 2016).

18. T. C. Lin et al., GALNT6 expression enhances aggressive phenotypes of ovarian cancer cells by regulating EGFR activity. Oncotarget 8, 42588 (Jun 27, 2017).

19. E. P. Bennett et al., Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. Glycobiology 22, 736 (Jun, 2012).

20. E. P. Bennett et al., Cloning and characterization of a close homologue of human UDP-N-acetyl-alpha-D-galactosamine:Polypeptide N-acetylgalactosaminyltransferase-T3, designated GalNAc-T6. Evidence for genetic but not functional redundancy. The Journal of biological chemistry 274, 25362 (Sep 03, 1999).

21. J. H. Park et al., Critical roles of mucin 1 glycosylation by transactivated polypeptide N-acetylgalactosaminyltransferase 6 in mammary carcinogenesis. Cancer research 70, 2759 (Apr 01, 2010).

22. D. J. Gill et al., 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 (Aug 20, 2013).

23. B. T. Hofmann et al., COSMC knockdown mediated aberrant O-glycosylation promotes oncogenic properties in pancreatic cancer. Molecular cancer 14, 109 (May 29, 2015).

24. R. Wang, C. Yu, D. Zhao, M. Wu, Z. Yang, The mucin-type glycosylating enzyme polypeptide N-acetylgalactosaminyltransferase 14 promotes the migration of ovarian cancer by modifying mucin 13. Oncology reports 30, 667 (Aug, 2013).

25. T. T. Nguyen et al., GalNAc-T6 in the relationship with invasion ability of endometrial carcinomas and prognostic significance. American journal of cancer research 7, 1188 (2017).

26. R. Sheta et al., A metabolic labeling approach for glycoproteomic analysis reveals altered glycoprotein expression upon GALNT3 knockdown in ovarian cancer cells. Journal of proteomics 145, 91 (Aug 11, 2016).

27. I. Nakayama et al., Loss of HOXD10 expression induced by upregulation of miR-10b accelerates the migration and invasion activities of ovarian cancer cells. International journal of oncology 43, 63 (Jul, 2013).

28. D. Amadori, E. Sansoni, A. Amadori, Ovarian cancer: natural history and metastatic pattern. Frontiers in bioscience : a journal and virtual library 2, g8 (Jan 01, 1997).

29. A. Feki et al., Dissemination of intraperitoneal ovarian cancer: Discussion of mechanisms and demonstration of lymphatic spreading in ovarian cancer model. Critical reviews in oncology/hematology 72, 1 (Oct, 2009).

30. T. J. Shaw, M. K. Senterman, K. Dawson, C. A. Crane, B. C. Vanderhyden, Characterization of intraperitoneal, orthotopic, and metastatic xenograft models of human ovarian cancer. Molecular therapy : the journal of the American Society of Gene Therapy 10, 1032 (Dec, 2004).

31. K. Garson, T. J. Shaw, K. V. Clark, D. S. Yao, B. C. Vanderhyden, Models of ovarian cancer--are we there yet? Molecular and cellular endocrinology 239, 15 (Jul 15, 2005).

32. J. D. Marth, Complexity in O-linked oligosaccharide biosynthesis engendered by multiple polypeptide N-acetylgalactosaminyltransferases. Glycobiology 6, 701 (Oct, 1996).

33. K. Taniuchi et al., Overexpression of GalNAc-transferase GalNAc-T3 promotes pancreatic cancer cell growth. Oncogene 30, 4843 (Dec 08, 2011).

34. S. Chugh, J. Meza, Y. M. Sheinin, M. P. Ponnusamy, S. K. Batra, Loss of N-acetylgalactosaminyltransferase 3 in poorly differentiated pancreatic cancer: augmented aggressiveness and aberrant ErbB family glycosylation. British journal of cancer 114, 1376 (Jun 14, 2016).

35. K. Onitsuka et al., Prognostic significance of UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase-3 (GalNAc-T3) expression in patients with gastric carcinoma. Cancer science 94, 32 (Jan, 2003).

36. M. Ishikawa, J. Kitayama, H. Nariko, K. Kohno, H. Nagawa, The expression pattern of UDP-N-acetyl-alpha-d-galactosamine: polypeptide N-acetylgalactosaminyl transferase-3 in early gastric carcinoma. Journal of surgical oncology 86, 28 (Apr 01, 2004).

37. K. Shibao et al., Expression of UDP-N-acetyl-alpha-D-galactosamine-polypeptide galNAc N-acetylgalactosaminyl transferase-3 in relation to differentiation and prognosis in patients with colorectal carcinoma. Cancer 94, 1939 (Apr 01, 2002).

38. J. Lin et al., GALNT6 Stabilizes GRP78 Protein by O-glycosylation and Enhances its Activity to Suppress Apoptosis Under Stress Condition. Neoplasia 19, 43 (Jan, 2017).

39. J. H. Park, T. Katagiri, S. Chung, K. Kijima, Y. Nakamura, Polypeptide N-acetylgalactosaminyltransferase 6 disrupts mammary acinar morphogenesis through O-glycosylation of fibronectin. Neoplasia 13, 320 (Apr, 2011).

40. Y. E. Tarhan et al., Morphological Changes, Cadherin Switching, and Growth Suppression in Pancreatic Cancer by GALNT6 Knockdown. Neoplasia 18, 265 (May, 2016).

41. Z. Li et al., Polypeptide N-acetylgalactosaminyltransferase 6 expression in pancreatic cancer is an independent prognostic factor indicating better overall survival. British journal of cancer 104, 1882 (Jun 07, 2011).

42. Y. Guo et al., Polypeptide N-acetylgalactosaminyltransferase-6 expression in gastric cancer. OncoTargets and therapy 10, 3337 (2017).

43. M. Murakami et al., Expression of Polypeptide N-Acetylgalactosaminyltransferase-6 in Epithelial Ovarian Carcinoma. Anticancer research 37, 3911 (Jul, 2017).

44. S. Kitada et al., Polypeptide N-acetylgalactosaminyl transferase 3 independently predicts high-grade tumors and poor prognosis in patients with renal cell carcinomas. British journal of cancer 109, 472 (Jul 23, 2013).

45. L. Freire-de-Lima et al., Involvement of O-glycosylation defining oncofetal fibronectin in epithelial-mesenchymal transition process. Proceedings of the National Academy of Sciences of the United States of America 108, 17690 (Oct 25, 2011).

46. X. Xu et al., Correlation of polypeptide N-acetylgalactosamine transferases-3 and -6 to different stages of endometriosis. Archives of gynecology and obstetrics 295, 1413 (Jun, 2017).

47. H. Matsuura et al., The oncofetal structure of human fibronectin defined by monoclonal antibody FDC-6. Unique structural requirement for the antigenic specificity provided by a glycosylhexapeptide. The Journal of biological chemistry 263, 3314 (Mar 05, 1988).

48. S. Masoumi Moghaddam, A. Amini, D. L. Morris, M. H. Pourgholami, Significance of vascular endothelial growth factor in growth and peritoneal dissemination of ovarian cancer. Cancer metastasis reviews 31, 143 (Jun, 2012).

49. T. J. Duncan et al., Vascular endothelial growth factor expression in ovarian cancer: a model for targeted use of novel therapies? Clinical cancer research : an official journal of the American Association for Cancer Research 14, 3030 (May 15, 2008).

50. A. D. Santin et al., Secretion of vascular endothelial growth factor in ovarian cancer. European journal of gynaecological oncology 20, 177 (1999).

51. X. Yin et al., A VEGF-dependent gene signature enriched in mesenchymal ovarian cancer predicts patient prognosis. Scientific reports 6, 31079 (Aug 08, 2016).

52. B. Cheaib, A. Auguste, A. Leary, The PI3K/Akt/mTOR pathway in ovarian cancer: therapeutic opportunities and challenges. Chinese journal of cancer 34, 4 (Jan, 2015).

53. Z. C. Dobbin, C. N. Landen, The importance of the PI3K/AKT/MTOR pathway in the progression of ovarian cancer. International journal of molecular sciences 14, 8213 (Apr 15, 2013).

54. S. Mabuchi, H. Kuroda, R. Takahashi, T. Sasano, The PI3K/AKT/mTOR pathway as a therapeutic target in ovarian cancer. Gynecologic oncology 137, 173 (Apr, 2015).

55. H. Li, R. Zhang, Role of EZH2 in Epithelial Ovarian Cancer: From Biological Insights to Therapeutic Target. Frontiers in oncology 3, 47 (2013).

56. X. Yi et al., EZH2-mediated epigenetic silencing of TIMP2 promotes ovarian cancer migration and invasion. Scientific reports 7, 3568 (Jun 15, 2017).

57. Z. Y. Rao et al., EZH2 supports ovarian carcinoma cell invasion and/or metastasis via regulation of TGF-beta1 and is a predictor of outcome in ovarian carcinoma patients. Carcinogenesis 31, 1576 (Sep, 2010).

58. B. G. Bitler et al., Synthetic lethality by targeting EZH2 methyltransferase activity in ARID1A-mutated cancers. Nature medicine 21, 231 (Mar, 2015).

59. L. Huang et al., IQGAP1 Is Involved in Enhanced Aggressive Behavior of Epithelial Ovarian Cancer Stem Cell-Like Cells During Differentiation. International journal of gynecological cancer : official journal of the International Gynecological Cancer Society 25, 559 (May, 2015).

60. P. X. Dong et al., Silencing of IQGAP1 by shRNA inhibits the invasion of ovarian carcinoma HO-8910PM cells in vitro. Journal of experimental & clinical cancer research : CR 27, 77 (Nov 27, 2008).

61. L. Y. Bourguignon, E. Gilad, K. Rothman, K. Peyrollier, Hyaluronan-CD44 interaction with IQGAP1 promotes Cdc42 and ERK signaling, leading to actin binding, Elk-1/estrogen receptor transcriptional activation, and ovarian cancer progression. The Journal of biological chemistry 280, 11961 (Mar 25, 2005).

62. N. Stopa, J. E. Krebs, D. Shechter, The PRMT5 arginine methyltransferase: many roles in development, cancer and beyond. Cellular and molecular life sciences : CMLS 72, 2041 (Jun, 2015).

63. H. Chen, B. Lorton, V. Gupta, D. Shechter, A TGFbeta-PRMT5-MEP50 axis regulates cancer cell invasion through histone H3 and H4 arginine methylation coupled transcriptional activation and repression. Oncogene 36, 373 (Jan 19, 2017).

64. M. Yoshimatsu et al., Dysregulation of PRMT1 and PRMT6, Type I arginine methyltransferases, is involved in various types of human cancers. International journal of cancer 128, 562 (Feb 01, 2011).

65. G. F. Yang et al., Intensive expression of Bmi-1 is a new independent predictor of poor outcome in patients with ovarian carcinoma. BMC cancer 10, 133 (Apr 08, 2010).

66. M. M. Gavrilescu, A. M. Todosi, M. G. Anitei, B. Filip, V. Scripcariu, Expression of bmi-1 protein in cervical, breast and ovarian cancer. Revista medico-chirurgicala a Societatii de Medici si Naturalisti din Iasi 116, 1112 (Oct-Dec, 2012).

67. A. Honig et al., Overexpression of polycomb protein BMI-1 in human specimens of breast, ovarian, endometrial and cervical cancer. Anticancer research 30, 1559 (May, 2010).

68. R. Bhattacharya et al., MiR-15a and MiR-16 control Bmi-1 expression in ovarian cancer. Cancer research 69, 9090 (Dec 01, 2009).

69. E. Wang et al., Enhancing chemotherapy response with Bmi-1 silencing in ovarian cancer. PloS one 6, e17918 (Mar 21, 2011).

70. F. Zhang, L. Sui, T. Xin, Correlations of BMI-1 expression and telomerase activity in ovarian cancer tissues. Experimental oncology 30, 70 (Mar, 2008).

71. A. Carlucci, V. D'Angiolella, It is not all about BRCA: Cullin-Ring ubiquitin Ligases in ovarian cancer. British journal of cancer 112, 9 (Jan 06, 2015).

72. Y. Oji et al., The translation elongation factor eEF2 is a novel tumorassociated antigen overexpressed in various types of cancers. International journal of oncology 44, 1461 (May, 2014).

73. G. Sethi et al., PTN signaling: Components and mechanistic insights in human ovarian cancer. Molecular carcinogenesis 54, 1772 (Dec, 2015).

74. C. Jin et al., UCHL1 Is a Putative Tumor Suppressor in Ovarian Cancer Cells and Contributes to Cisplatin Resistance. Journal of Cancer 4, 662 (2013).

75. G. Pongas et al., BRD4 facilitates DNA damage response and represses CBX5/Heterochromatin protein 1 (HP1). Oncotarget 8, 51402 (Aug 01, 2017).

76. Y. Wang et al., Ubiquitin-specific protease 14 (USP14) regulates cellular proliferation and apoptosis in epithelial ovarian cancer. Medical oncology 32, 379 (Jan, 2015).

77. R. Kafri, M. Springer, Y. Pilpel, Genetic redundancy: new tricks for old genes. Cell 136, 389 (Feb 06, 2009).

78. F. B. Pickett, D. R. Meeks-Wagner, Seeing double: appreciating genetic redundancy. The Plant cell 7, 1347 (Sep, 1995).

79. J. Zhang, Genetic redundancies and their evolutionary maintenance. Advances in experimental medicine and biology 751, 279 (2012).

80. E. J. Dean, J. C. Davis, R. W. Davis, D. A. Petrov, Pervasive and persistent redundancy among duplicated genes in yeast. PLoS genetics 4, e1000113 (Jul 04, 2008).

81. A. DeLuna et al., Exposing the fitness contribution of duplicated genes. Nature genetics 40, 676 (May, 2008).

82. R. Kafri, O. Dahan, J. Levy, Y. Pilpel, Preferential protection of protein interaction network hubs in yeast: evolved functionality of genetic redundancy. Proceedings of the National Academy of Sciences of the United States of America 105, 1243 (Jan 29, 2008).

83. A. Faddaoui et al., Suppression of the grainyhead transcription factor 2 gene (GRHL2) inhibits the proliferation, migration, invasion and mediates cell cycle arrest of ovarian cancer cells. Cell Cycle, 1 (Feb 22, 2017).

84. A. Faddaoui et al., The mannose receptor LY75 (DEC205/CD205) modulates cellular phenotype and metastatic potential of ovarian cancer cells. Oncotarget 7, 14125 (Mar 22, 2016).

4.10 Figure legends

Figure 4.1. Protein-protein interaction network of GALNT3 and GALNT6. The figure displays protein-protein interactions of the human GALNT3 and GALNT6 enzymes as illustrated by the confidence view of the STRING 8.3 server. The nodes of the graph represent human proteins and the connections illustrate their known or predicted, direct and indirect interactions. The connection between these two protein-nodes is based on the available information mined from relevant databases and literature. The network includes protein interactions of GALNT3 and GALNT6 to MUC1. These interactions have .0.9 estimated probability.

Figure 4.2. Western blot protein expression analysis of *GALNT3* KO and *GALNT3/T6* KO clones. A. The CRISPR/Cas9 system was used to generate *GALNT3* KO (GAL3 KO) clones in the A2780s cell line. For confirmation of protein KOs, we used both positive and negative controls, those that highly express GALNT3 and those cells that do not express GALNT3 (+ve control (A2780s cells), -ve control (OV2008 cells)). Seven clones show complete protein ablation upon *GALNT3* KO. Similarly Western blots confirm the compensation by GALNT6 in the *GALNT3* KO clones, likewise we used both positive and negative controls, those that highly express GALNT6 and those cells that do not express GALNT3 (+ve control (CaOV3 cells), -ve control (A2780s cells)). B. Western blot analysis of the double KO clones (GAL3/6 KO) generated using the CRISPR/Cas9 system followed by the shRNA system, A2780s cell were used as the control clone (A2780s Ctrl). FN1 was also evaluated in these clones, confirming the activity of *GALNT6* in the single *GALNT3* KO clone. β-Actin was used as the loading control.

Figure 4.3. Effect of *GALNT3* and *GALNT6* double KO on A2780s cell proliferation migration, invasion and cell cycle. A. Cell migration of GAL3 KO clone GAL3 KO and GAL3/6 KO clones was compared to the control clone (Ctrl). Migration was assessed using Boyden-chamber assay. C. Cell invasion of GAL3 KO clone GAL3 KO and GAL3/6 KO clones was compared to the control clone (Ctrl). The bar graphs represent quantitative determinations of migration and invasion data obtained by selecting 10 random fields per

filter under phase contrast microscopy and results are expressed as number of cell change (migration and invasion) of the GAL 3 KO and GAL3/6 KO clone compared to the Ctrl clone. C. Cell proliferation of GAL3 KO and GAL3/6 KO clone was compared to the control clone (Ctrl). Differences between the GAL 3 KO and GAL3/6 KO clones and vehicle- transfected A2780s cells were determined by a Student's t-test; error bars denote mean ± SEM; ($p < 0.05$). D. Representative images of colony forming assays following GAL3 KO in the A2780s cell line. Representative quantitative determinations of colony formation assay obtained by determining the colony area which focuses on adjusting for the percentage of the area covered by crystal violet stained cell colonies, and the intensity of the staining of the colonies per plate. Results are expressed as number of colony differences between the GAL3 KO and GAL3/6 KO clones compared to the Ctrl clone. Differences in colony numbers were determined by a Student's t-test; error bars denote mean ± SEM; *indicates statistical significance ($p < 0.05$).

Figure 4.4. Cell-cycle profiling of *GALNT3* and *GALNT3T/6* double KO clones. Cell cycle was examined by flow cytometry and percentages of cells in G0/G1, S, and G2/M phase in the GAL3 KO, GAL3/6 KO clones were compared to the mock-transfected control (Ctrl) clone. Propidium iodide staining shows a decreased fraction of cells in the G1-phase and a at 0 hr, and especially at 6hr and 12hr post hydroxyurea removal in the GAL3 KO clone, and more strongly in the GAL3/6 KO clone when compared with the control clone (Ctrl).

Figure 4.5. Western blot analysis of FN1 and MUC1 in the *GALNT3* KO and *GALNT3/T6* KO clones. A. Western blot protein expression data against FN1 and MUC1 in both GAL3 KO and GAL3/6 KO clones compared to Ctrl. A band of FN1 was observed in the GAL3 KO clone which was absent in the Ctrl and the GAL3/6 KO clones, additionally a MUC1 band was observed in the Ctrl clone which was diminished in the GAL3 KO clone and was completely abolished in the GAL3/6 KO clones. B. VVA-lectin-mediated immunoblot analysis of GalNAc-conjugated proteins in protein lysates of the Ctrl, the GAL3 KO and the GAL3/6 KO A2780s clones following VVA lectin pull-down assay (pull-down). The arrow indicates bands corresponding to possible GalNAc-conjugated FN1 and MUC1 peptides. β-actin was used as a loading control. C. Semi-quantitative RT-PCR analysis of

MUC1 and *FN1* mRNA levels in the Ctrl clone, GAL3 KO and GAL3/6 KO clones; the *GUSB* gene was used as an internal control.

Figure 4.6. Canonical pathway analysis for a dataset of differentially expressed genes (≥ 1.5-fold) following *GALNT3* KO and *GALNT3/T6* KO in the A2780s cells. A. Canonical analysis of downregulated genes; B. Canonical analysis of upregulated genes. Top functions that meet a Bonferroni-Holm multiple testing correction *p*-value of 0.05 are displayed.

Figure 4.7. Functional pathway analysis for a dataset of differentially expressed genes (≥ 1.5-fold) following *GALNT3* KO and *GALNT3/T6* KO in the A2780s cells. A. Functional analysis of downregulated genes; B. Functional analysis of upregulated genes. Top functions that meet a Bonferroni-Holm multiple testing correction *p*-value of 0.05 are displayed.

Figure 4.8. Network analysis of dynamic gene expression in A2780s cells based on the 1.5-fold gene expression list obtained following *GALNT3* KO. The five top-scoring networks of up- and down-regulated genes were merged and are displayed graphically as nodes (genes/gene products) and edges (the biological relationships between the nodes).

Figure 4.9. Network analysis of dynamic gene expression in A2780s cells based on the 1.5-fold gene expression list obtained following *GALNT3/T6* double KO. The five top-scoring networks of up- and down-regulated genes were merged and are displayed graphically as nodes (genes/gene products) and edges (the biological relationships between the nodes).

Figure 4.10. *In vivo* examination of the effect of *GALNT3* and *GALNT3/T6* KO in tumor formation and survival in immunodeficient (SCID) mice. A. Survival curves for mice injected with Ctrl, GAL3 KO and GAL3/6 KO cells. The median survival of mice injected with the Ctrl cells (31 days, n = 8). Survival of mice injected with the *GALNT3* KO cells was significantly shorter than the vector control (38 days; $p = 0.0193$, Log rank test), while mice injected with the GALNT3/T6 KO cells was significantly shorter than the vector

control (57 days; p = 0.0005, Log rank test). B. Representative IHC images of GALNT3, GALNT6, FN1, and MUC1 expression in tumor tissues extracted from mice injected with the Ctrl, GAL3 KO and GAL3/6 KO cells.

Supplemental figure legends

Supplemental Figure 4.1. Protein analysis of GALNT3 and GALNT6 expression in CaOV3 cell line. A. Western blot examination of GALNT protein expression (GALNT2, GALNT9 and GALNT14) in GAL3 KO and GAL3/6 KO clones. B. Western blot confirmation of *GALNT6* KD in CaOV3 cell line using shRNA, in addition to GALNT3 protein expression analysis in *GALNT 6* CaOV3 KO clones. C. Gene expression analysis of *GALNT6* and *GALNT3* mRNA expression in *GALNT6* CaOV3 KD clones. D. Expression analysis of GALNT3 expression in nine EOC cell lines. β-Actin was used as the loading control.

Supplemental Figure 4.2 Gene expression validation analysis of microarray data. A. The figure shows bar graphs presentation of the differential expression of the selected genes in A2780s cells following *GALNT3* KO compared to control A2780s cells. B. The figure shows bar graphs presentation of the differential expression of the selected genes in A2780s cells following *GALNT3/T6* KO compared to control A2780s cells. The relative copy number was calculated based on the target gene/18S ribosomal RNA ratio. Values more than or equal to 1 represent gene upregulation and less than 1 display gene downregulation.

Supplemental Figure 4.3. *In vivo* analysis of *GALNT3* and *GALNT3/T6* double KO in SCID mice. A. Tumor weight in SCID mice were measured. Data shown are the tumor weight averages with SD from 8 mice in Control and 7 mice in *GALNT3* KO and *GALNT3/T6* KO injected mice. B. Western blot analysis of protein expression of (GALNT3, GALNT6, FN1, and MUC1) from tumor tissue of the SCID mice used in the study. β-Actin was used as the loading control.

4.11 Figures

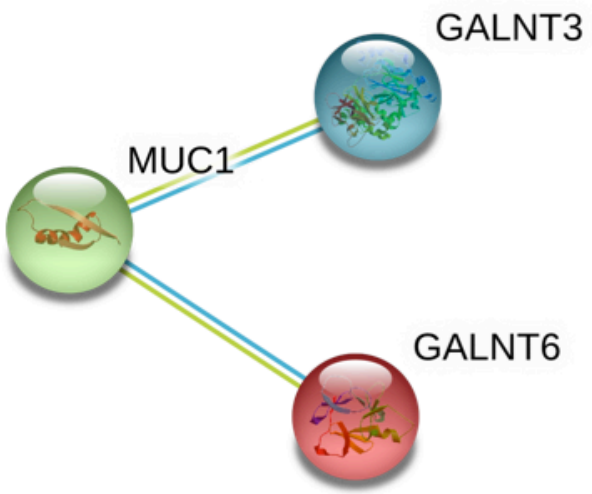


Figure 4.1. Protein-protein interaction network of GALNT3 and GALNT6

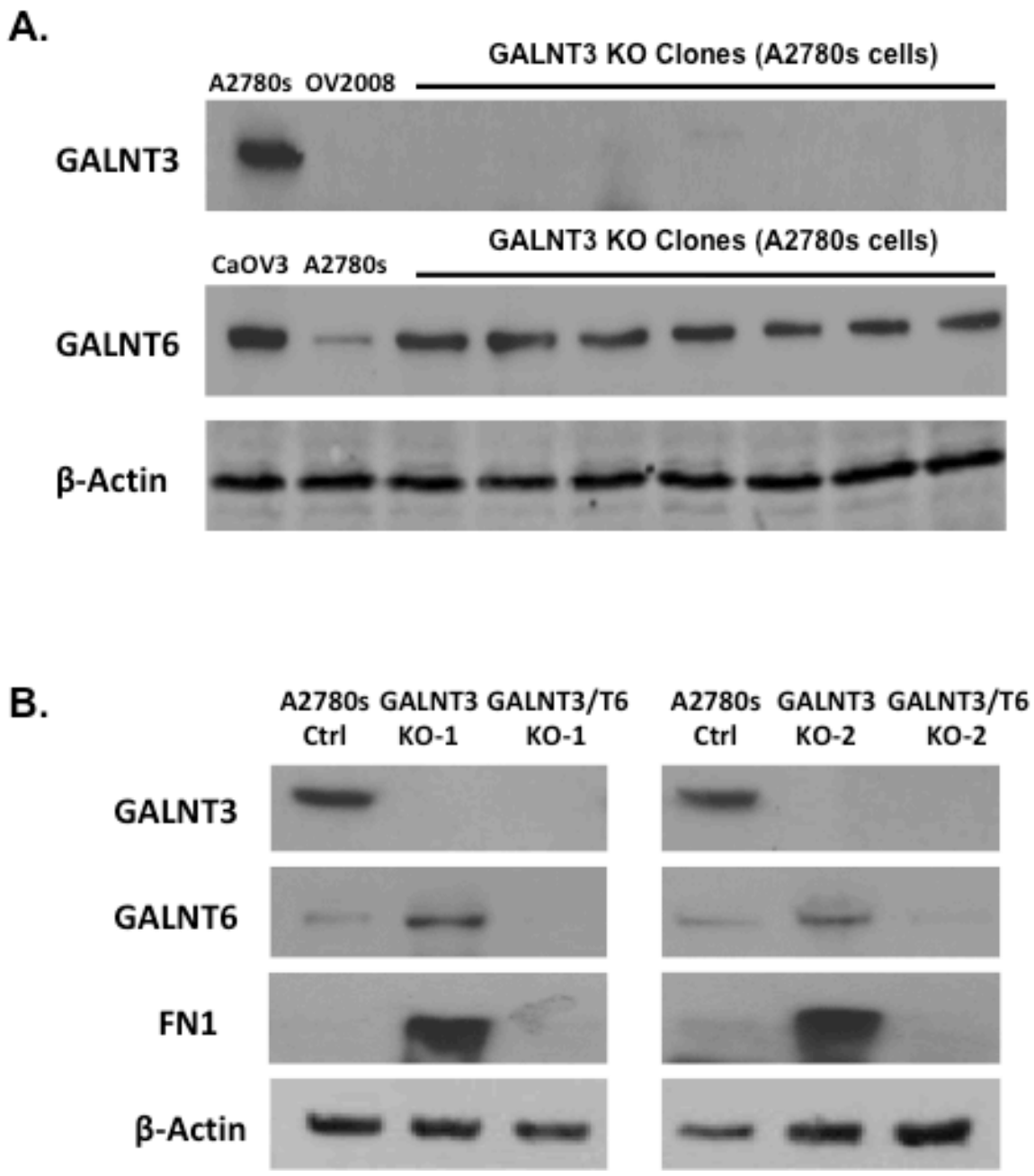


Figure 4.2. Western blot protein expression analysis of *GALNT3* KO and *GALNT3/T6* KO clones

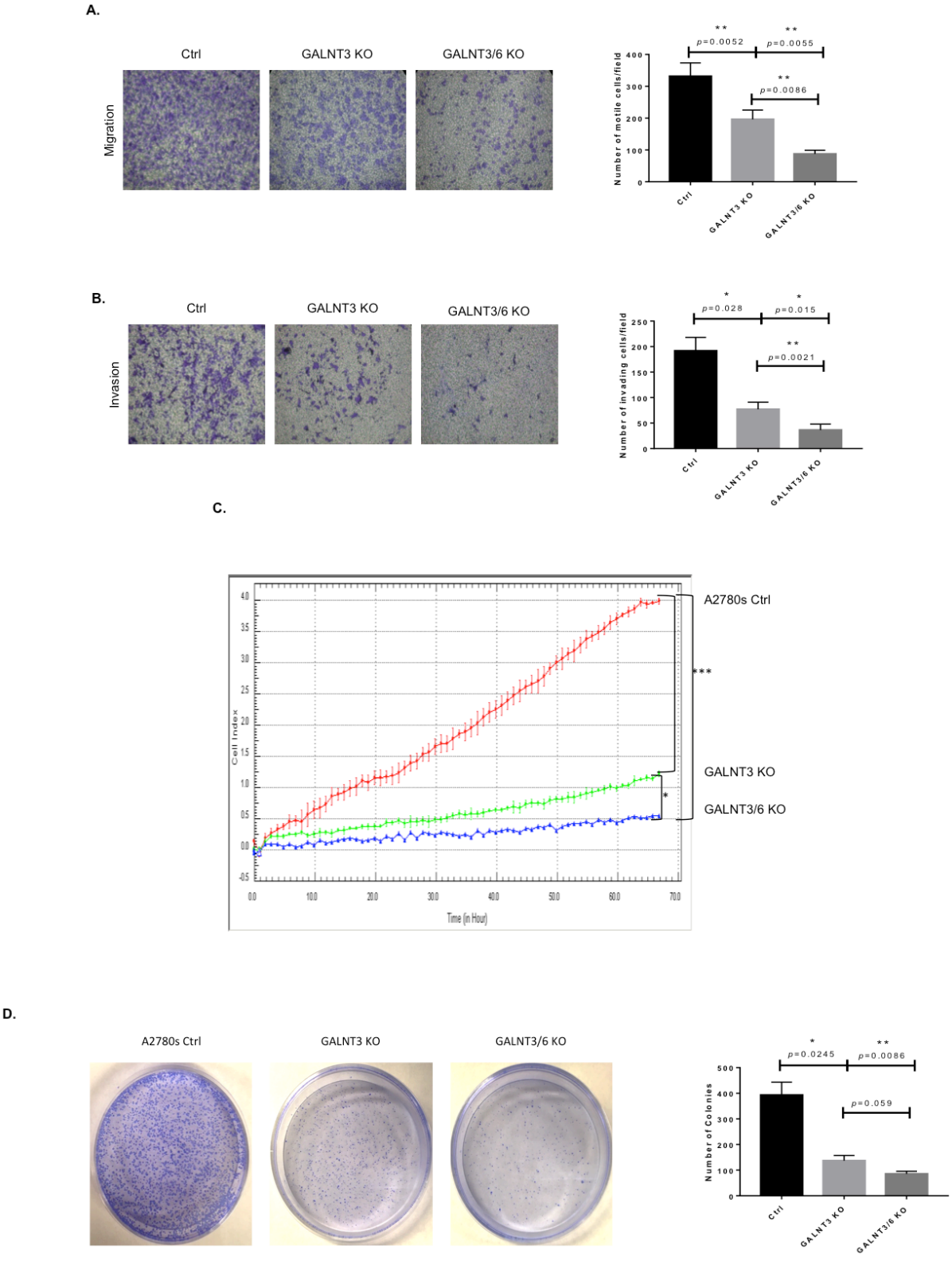


Figure 4.3. Effect of *GALNT3* and *GALNT6* double KO on A2780s cell proliferation migration, invasion and cell cycle

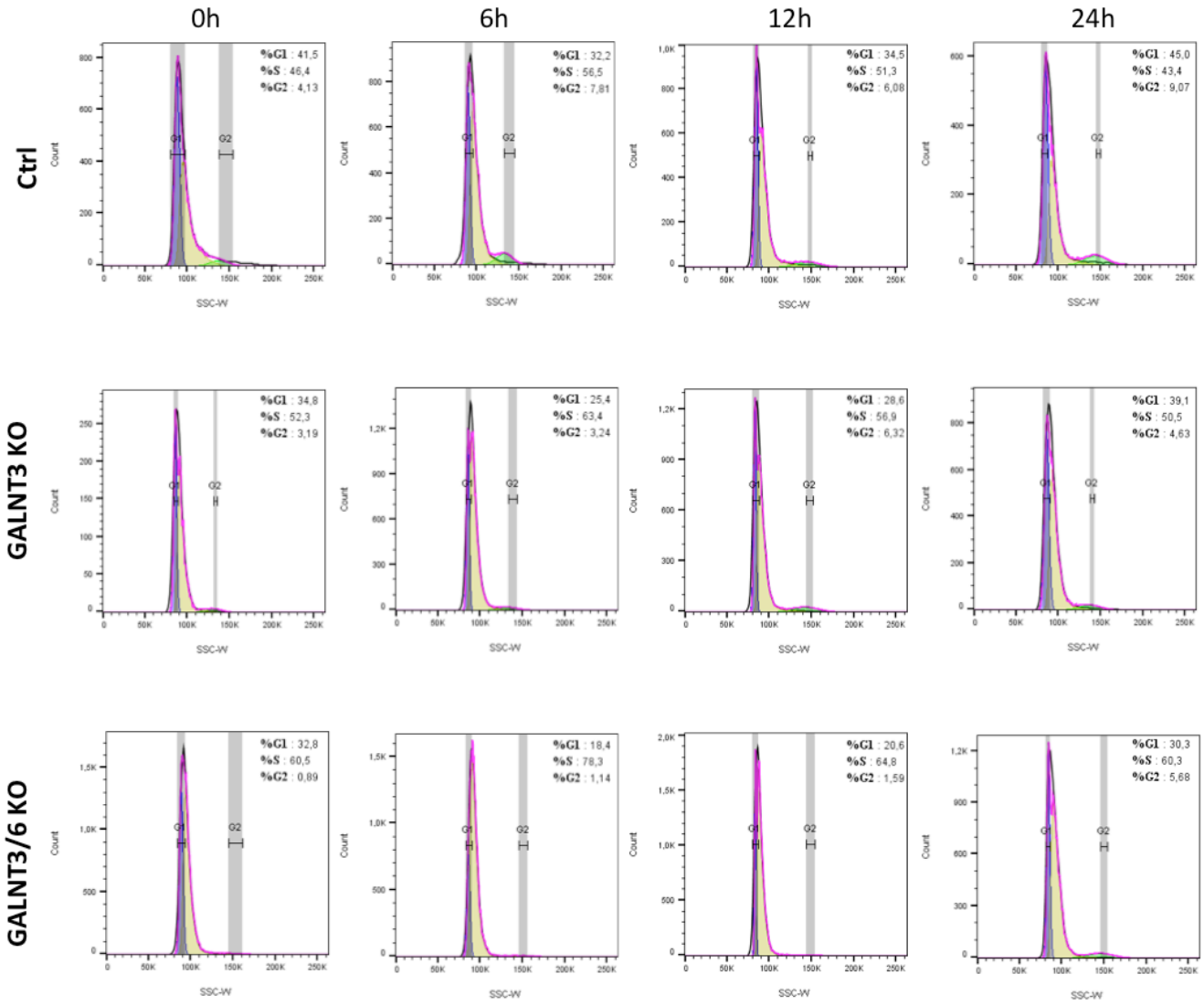


Figure 4.4. Cell-cycle profiling of *GALNT3* and *GALNT3/T6* double KO clones

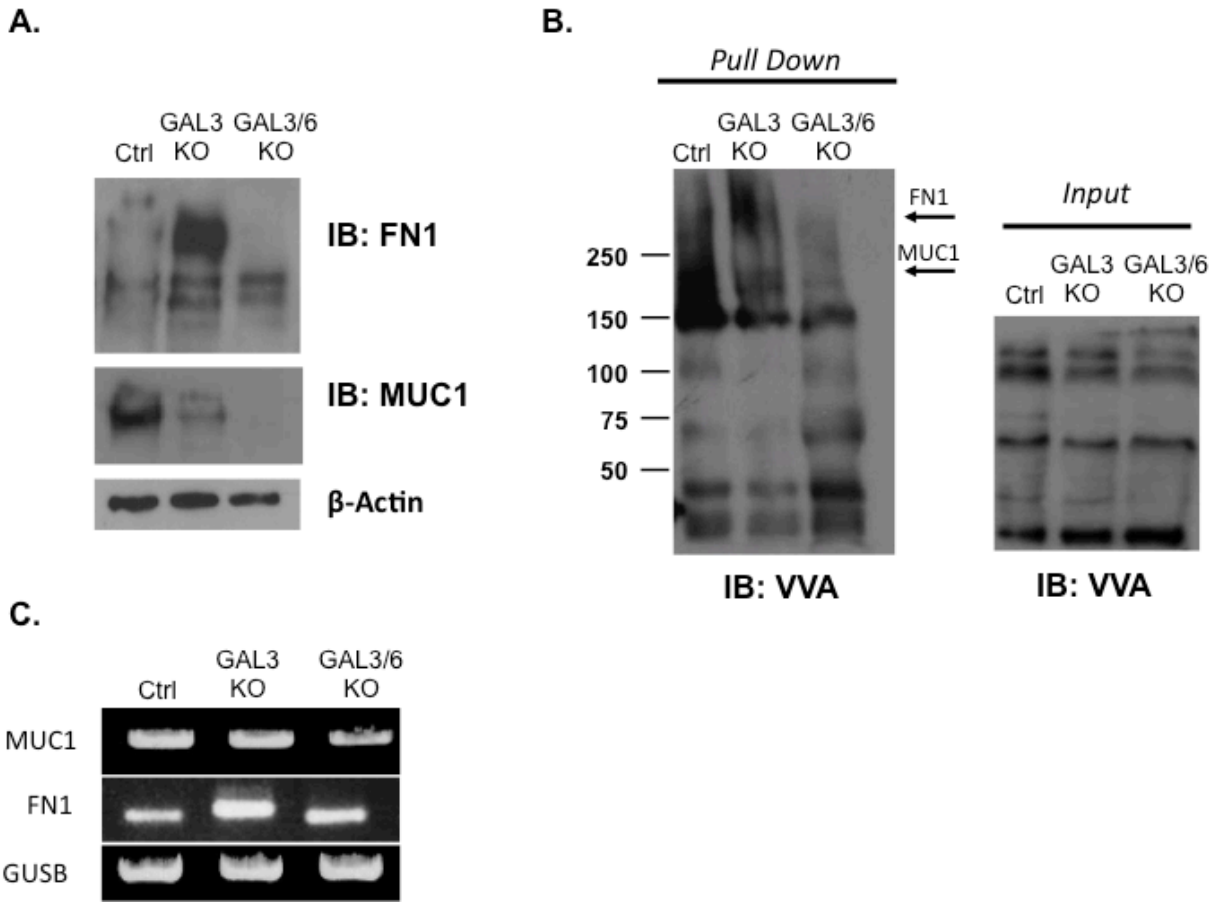


Figure 4.5. Western blot analysis of FN1 and MUC1 in the *GALNT3* KO and *GALNT3/T6* KO clones

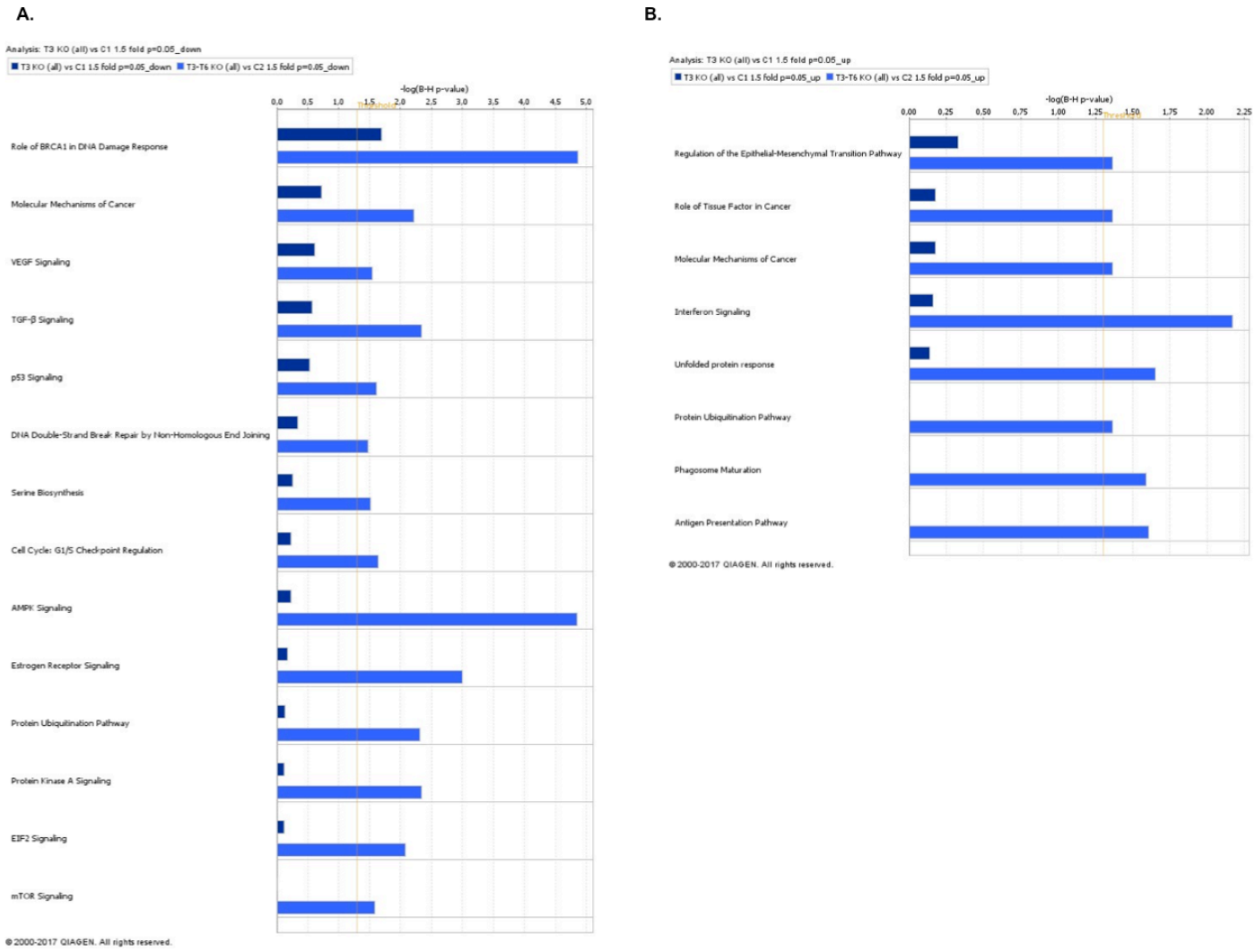


Figure 4.6. Canonical pathway analysis for a dataset of differentially expressed genes (≥ 1.5-fold) following *GALNT3* KO and *GALNT3/T6* KO in the A2780s cells

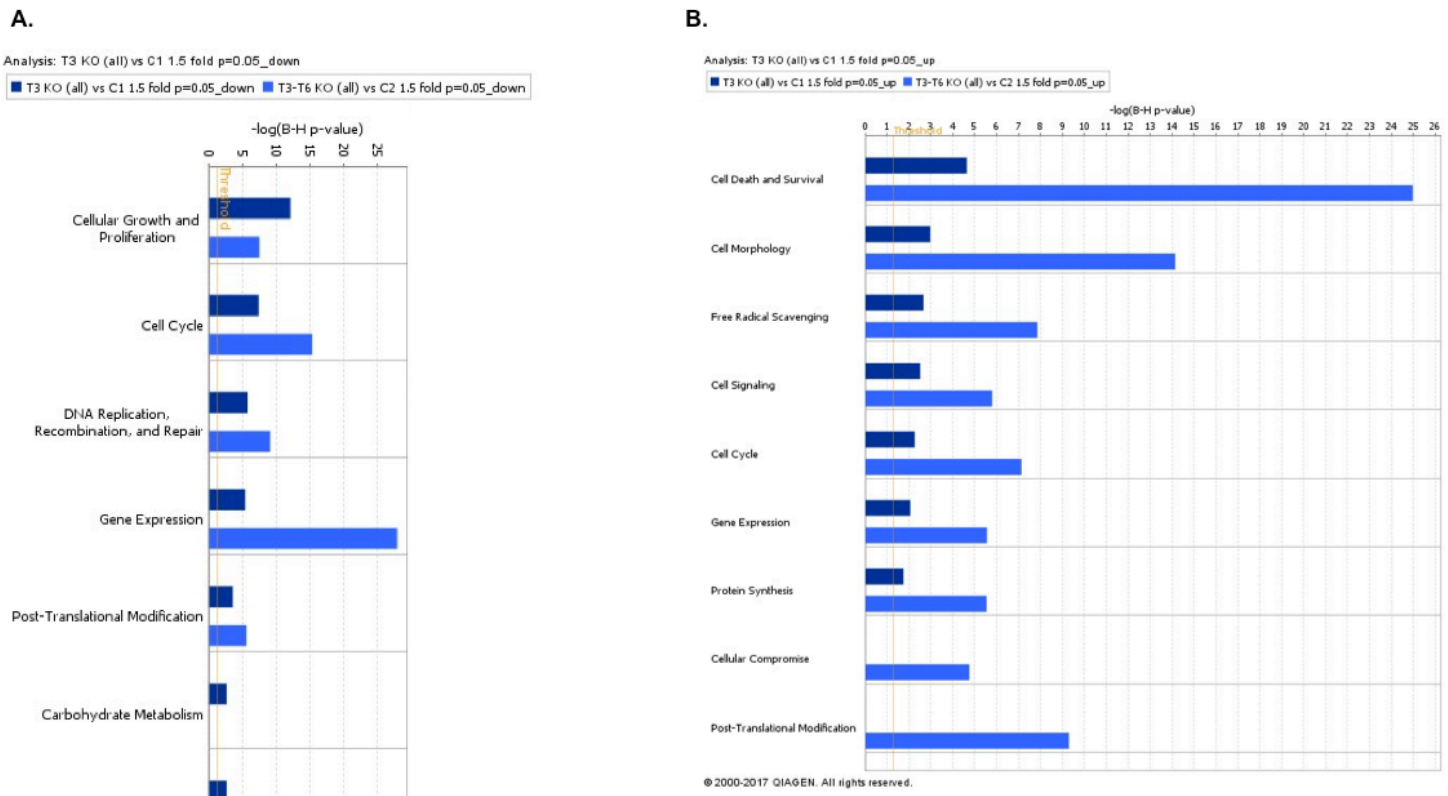


Figure 4.7. Functional pathway analysis for a dataset of differentially expressed genes (≥ 1.5-fold) following *GALNT3* KO and *GALNT3/T6* KO in the A2780s cells

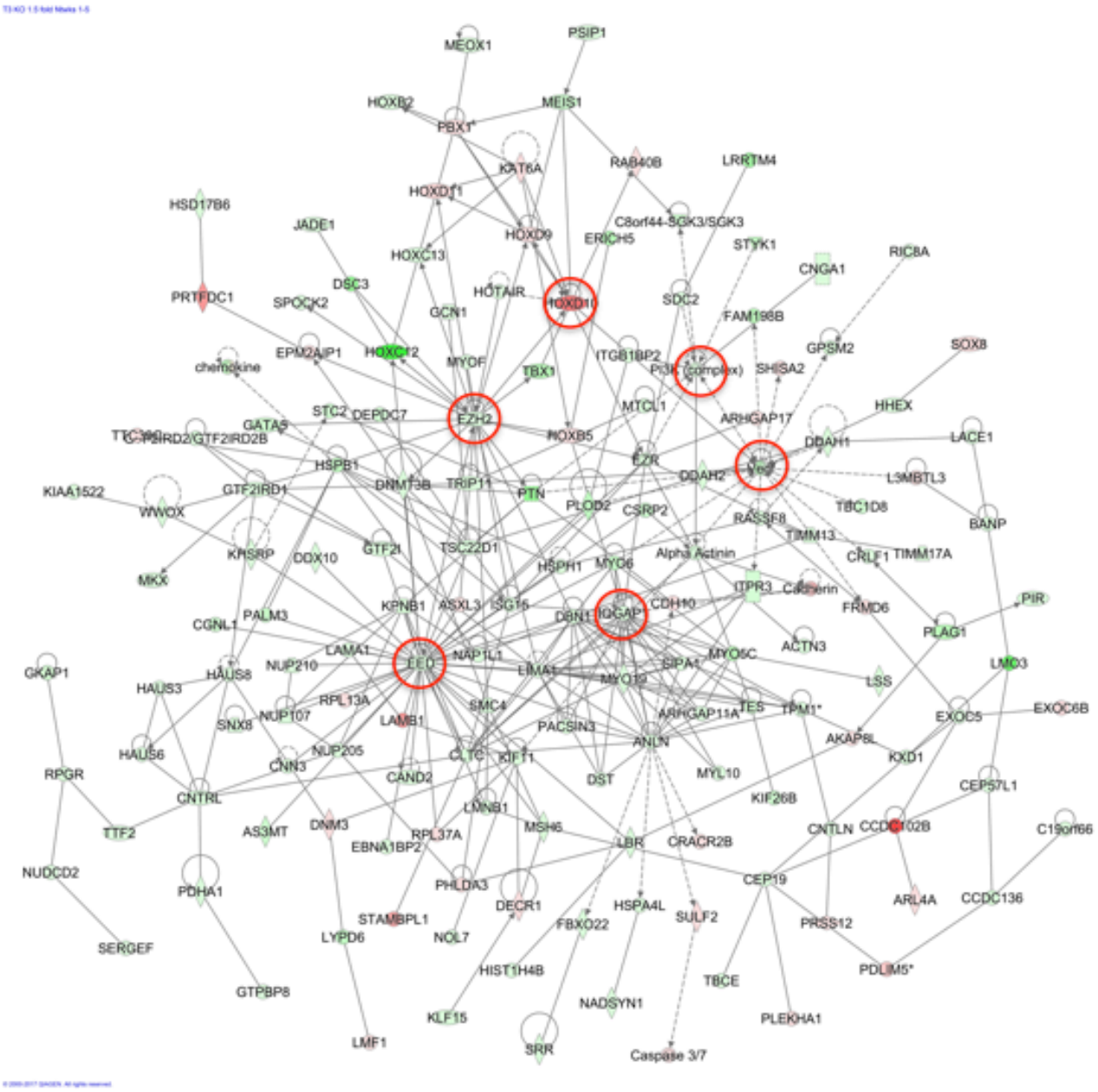


Figure 4.8. Network analysis of dynamic gene expression in A2780s cells based on the 1.5-fold gene expression list obtained following *GALNT3* KO

Figure 4.9. Network analysis of dynamic gene expression in A2780s cells based on the 1.5-fold gene expression list obtained following *GALNT3/T6* double KO

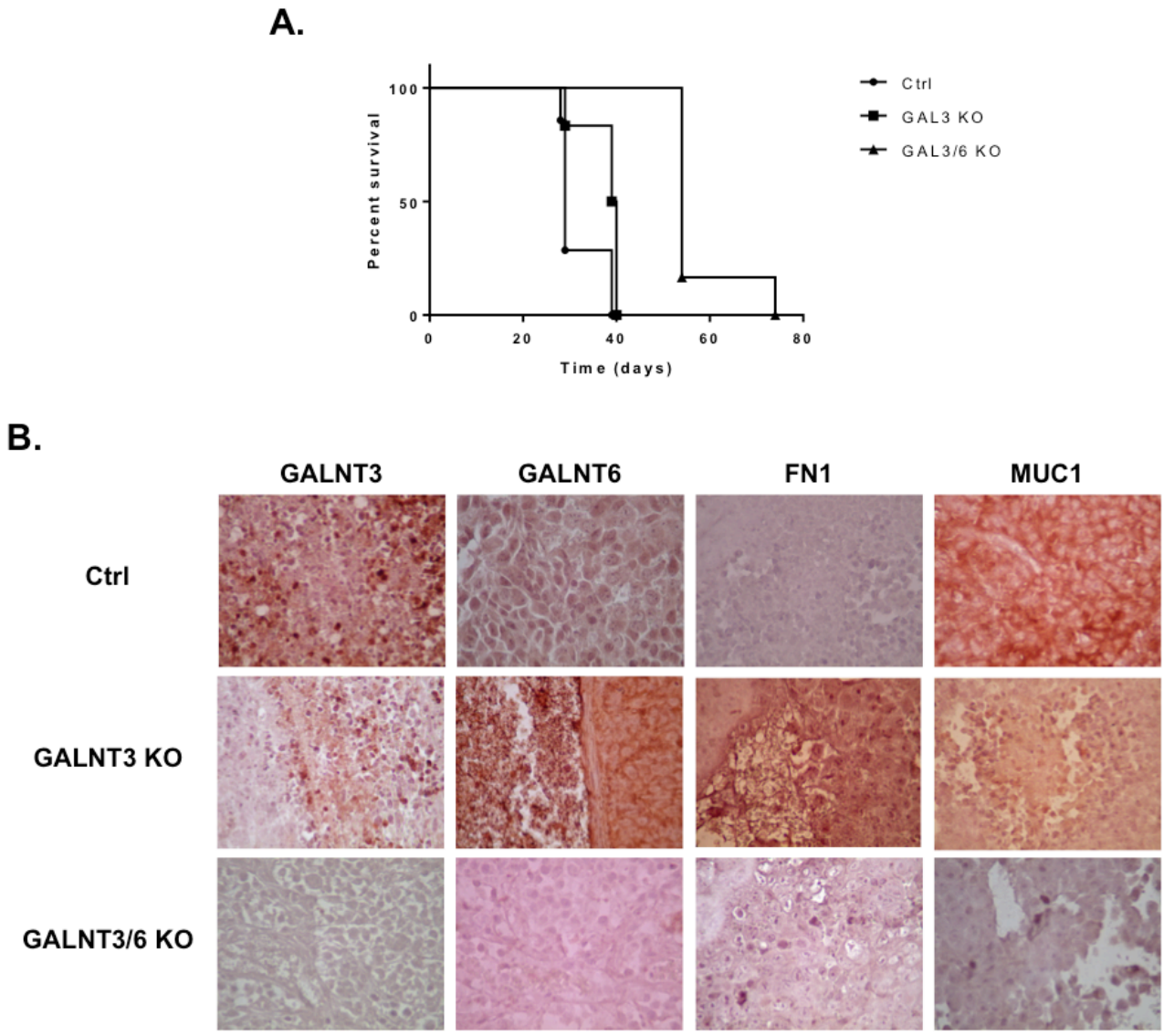
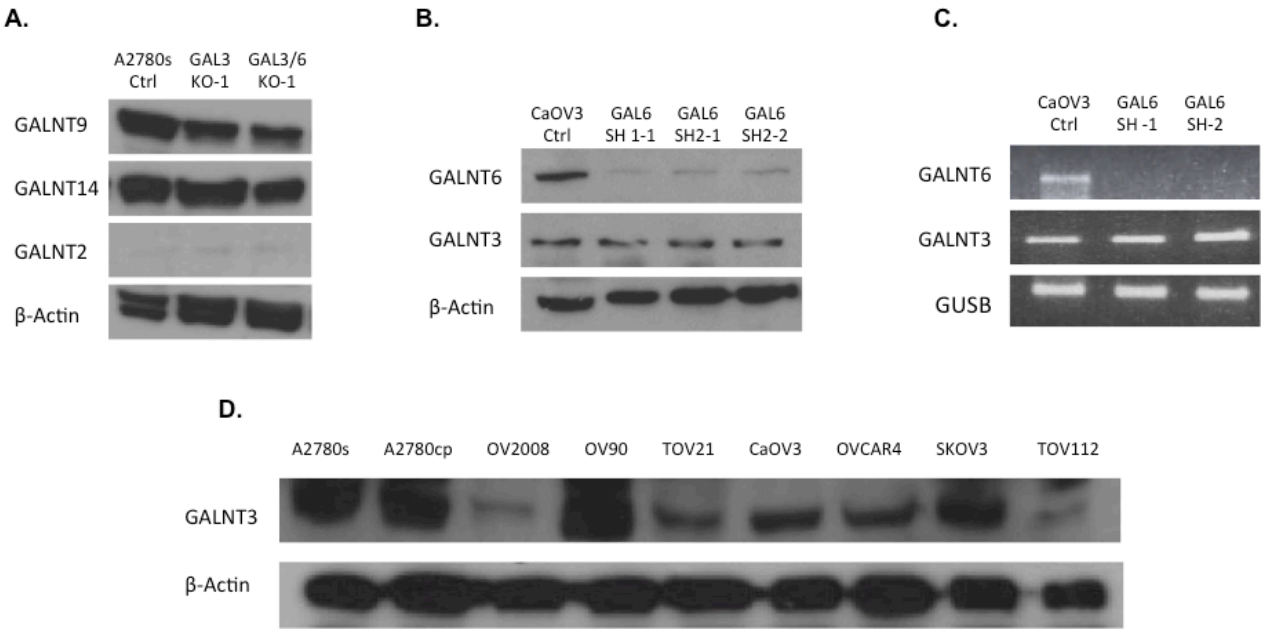
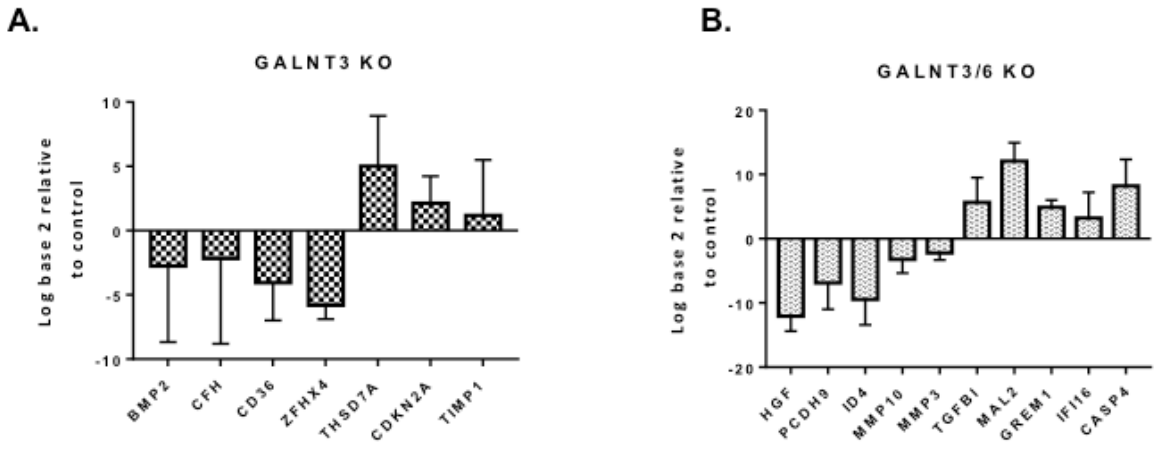


Figure 4.10. ***In vivo*** **examination of the effect of** ***GALNT3*** **and** ***GALNT3/T6*** **KO in tumor formation and survival in immunodeficient (SCID) mice**

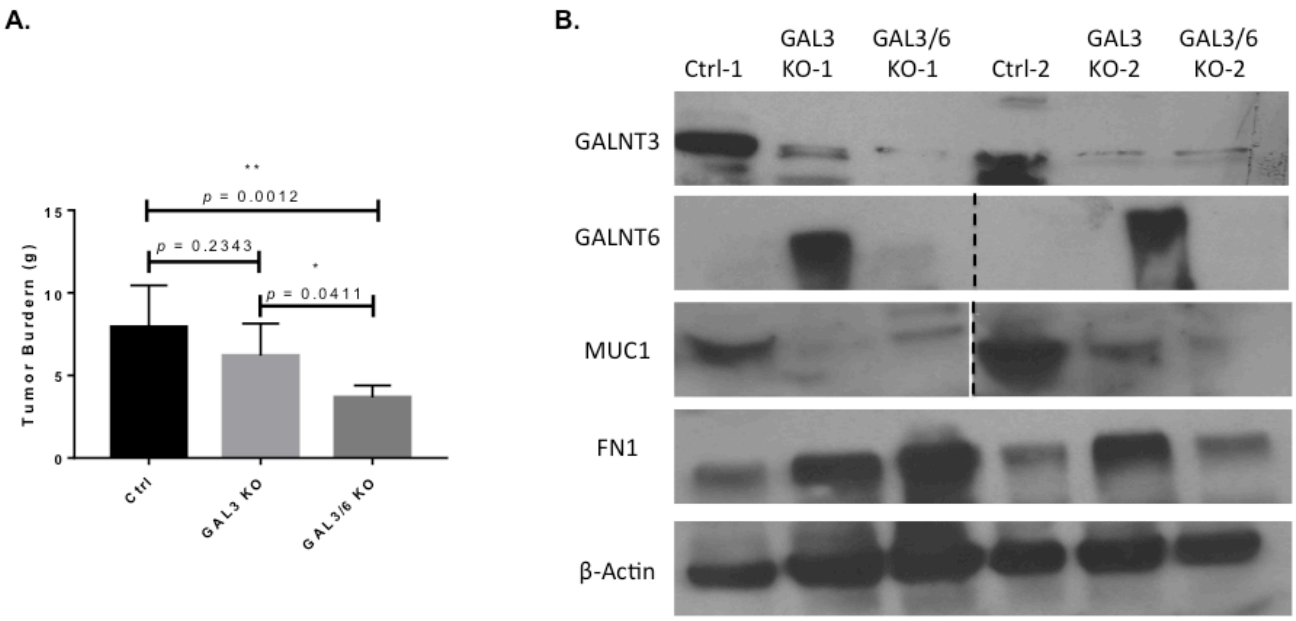
4.12 Supplemental figures



Supplemental Figure 4.1. Protein analysis of GALNT3 and GALNT6 expression in CaOV3 cell line



Supplemental Figure 4.2 Gene expression validation analysis of microarray data



Supplemental Figure 4.3. ***In vivo*** **analysis of** ***GALNT3*** **and** ***GALNT3/T6*** **double KO in SCID mice**

Chapter 5: General Discussion

5.1 Discussion

Epithelial Ovarian Cancer (EOC) is a disease that is responsible for more cancer deaths among women in the Western world than all other gynecologic malignancies (*400*). EOC lethality primarily stems from the inability to detect the disease at an early, organ-confined stage, and the lack of effective therapies for the advanced-stages of the disease. It is well established that cancer invasion and metastasis still represent the two major causes of the failure of cancer treatment (*400*). Thus, management of the metastatic disease becomes a crucial problem for the treatment of EOC. One possible way to resolve this problem is to target metastasis-specific pathways with novel therapies. Hence, focused identification of novel pro-metastatic mechanisms, target pathways and molecules could help identify new effective therapies.

Research in molecular biology is highly focused in understanding the paradigm of how information flows from DNA to RNA to protein. The major difficulty that lies in understanding this biological information flow is the ability to understand how small numbers of genes can develop large complexities during organisms' development. PTMs of proteins such as phosphorylation, ubiquitination and glycosylation are some of the major modifications that are involved in explaining the complexity of protein products. Glycosylation is the most complex form of PTM and is defined by the regulated process of adding lipids and carbohydrates to proteins, as this regulation depends on enzymes and proteases that allow for the diversification of protein function (*401*). Glycans are mostly found attached to the outer surface of cells or on secreted molecules, allowing them to mediate various events such as cell to cell and cell to molecule interactions, which are crucial for the development of multicellular organisms (*166*).

Glycosylation is defined by the regulated process of adding lipids and carbohydrates to proteins (*165*). Many theories have been developed to improve the understanding of the major roles of these glycans. However, no existing evidence can accurately explain glycan roles or functions, and this is mainly due to their enormous diversity. Increasing numbers of studies have investigated the implication of PTMs in cancer progression, notably the role of aberrant glycosylation in various cancer types (*402-404*). Cancer cells are characterized by rapid growth, adherence and invasion of neighboring cells; and these cellular changes are linked to alterations in the cell's glycosylation profiles (*164*). Tumors associated with different mucin families show complex alterations in protein glycosylation (such as premature termination, reductions in glycan number, increased branching, and terminal

modifications such as sialylation as fucosylation) when compared to normal cells (*164*). Different members of the GalNAc-Ts glycotransferases family have displayed aberrant expression in cancer, indicative for their potential implication in tumorigenesis (*305, 405-418*). Several mucins have been used as biomarkers for different cancers. (*419*). More glycoproteins are emerging as biomarkers for EOC, as changes associated with the progression of EOC from primary to metastatic disease have been shown to affect the expression of certain tumor-associated antigens (TAA) (*420*). Altered expression of specific glycosyltransferases leads to the formation of glycan structures that play a role in cancer growth and metastasis. Work from our lab has previously identified the *GALNT3* gene as a potential EOC oncogene, highly expressed in advanced disease, as *GALNT3* expression was significantly associated with poor outcome (*305*). This study also demonstrated that *GALNT3* might contribute to EOC dissemination through aberrant glycosylation of different O-glycoproteins, including the *MUC1* oncogene (*305*). Overall, these data provide evidence that the post-translational regulation of glycogens can contribute to cancer progression and metastasis in different tissue types and can potentially be used for identifying new EOC biomarkers. Work from this thesis project was focused to on characterizing the role of the GALNT3 enzyme and other members of the GalNAc-T family in aberrant O-glycosylation of mucin-like targets in EOC cells. We hypothesized that *GALNT3* overexpression contributes to EOC dissemination through aberrant mucin O-glycosylation. We addressed this hypothesis by analyzing altered expression of glycoproteins following *GALNT3* gene KD in EOC cells, using a metabolic labeling strategy for enrichment and glycoprotein analysis, coupled with MS characterization of enriched glycoproteins, as presented in chapter 2. This glycoproteomics approach led to the identification of numerous O-glycoproteins differentially expressed upon *GALNT3* KD, including O-glycoproteins implicated in mechanisms of cellular metabolism, PTMs and EOC dissemination and which could represent novel EOC biomarkers/therapeutic targets, as shown in chapter 2.

Moreover, we extended our studies to examine the role of other members of the GalNAc-T family in aberrant glycosylation in ovarian cancer. This objective was addressed, as discussed in chapter 1, evolutionary studies are not able to predict why the GalNAc-Ts family is so diverse, and if their diversity can be linked to some degree of redundancy

between the different GalNAc-Ts members. It is believed that their large number may reflect an essential functional requirement of the different GalNAc-Ts isoforms. This PhD project also investigated the expression pattern and prognostic value of 5 members of the GalNAc-T family in EOC, and explored the idea that there may be some redundancy observed between the GALNT3 transferase and other members of the GalNac-T family, as presented in chapter 3. Results from chapter 3 of this thesis, demonstrated that there is simultaneous overexpression of four GalNAc-Ts (*GALNT3, T6, T9* and *T14*) in advanced EOC, which raised the question about their specific and/or redundant functions in EOC. Although different GalNAc-Ts are differentially expressed within tissues, between cells, within a single tissue, and in different patterns at different stages in the development and differentiation, it is now becoming clear that a subset of the GalNAc-Ts display both distinct and overlapping substrate specificities (*316, 421*). The *GALNT3* and *GALNT6* gene members display 65% homology in their coding sequence (*422*). Interestingly, *GALNT6* exhibited quite similar oncogenic functions in breast cancer (modulating aberrant O-glycosylation and MUC1 stabilization) (*408*), as those found by us for *GALNT3* in EOC (*305*). Based on the above considerations we can suggest that different GalNAc-Ts may be performing redundant and/or overlapping functions in disease progression of women with HGSC.

5.1.1 The application of glycoproteomics studies in EOC

Glycosylation changes have become more important targets for therapeutic intervention (*423*); and thus, glycoproteomics is now a very promising field for glycoprotein-based biomarker discovery in EOC. Over the last 10-15 years, there were numerous publications focused on identifying novel EOC biomarkers and on examining their possible application for EOC diagnostics and prognosis. Table 5.1 summarizes the different studies focused on identifying and examining protein and glycoprotein biomarkers in EOC.

Table 5.1. PubMed search output of proteomics studies and ovarian cancer

Keyword search	**Number of publications**
Protein biomarkers and ovarian cancer	12,234
Glycoprotein and ovarian cancer	8,985

Prognostic glycoproteins and ovarian cancer	736
Predictive glycoproteins and ovarian cancer	685
Ovarian cancer and glycoproteomic	13

As demonstrated from Table 5.1, huge efforts have been put towards the identification of new EOC biomarkers, and recent efforts have been focused on the application of glycoproteomics in studying and understanding the EOC metastatic cascade (*363, 424-432*). Research in cancer biology has identified biomarkers as tools that can indicate disease progression, treatment response, or malignant transformation; all these biomarker discovery approaches have shown great success in multiple cancer types including EOC, and as data shows glycoproteomics has offered a hupe leap in the identification of these markers (*433-436*). The goal in biomarker discovery in cancer is the identification of a measurable biomolecule that shows significant level and expression changes in a specific type of cancer and during the cancer progression (*437*). Despite the major efforts put to identify potential biomarkers in EOC research, there are only a few studies using glycoproteomic approaches aiming at the identification of EOC glycoprotein markers (Table 5.1). When reviewing these studies, the data demonstrates that glycoproteomics could be an important tool for glycoprotein-based marker discovery in EOC. These studies apply a number of advanced proteomic techniques, in addition to the use of different types of samples, for identifying potential glycoprotein markers. Thus, Li *et al*. (2017) studied changes in glycoproteins expression and their glycosylation occupancy in both in HGSC and serous cystadenoma (*424*), using sophisticated glycoproteomic approaches for studying glycosylation changes in EOC tissue samples. Their work has shed light on the importance of glycosylation occupancy changes, and its importance in examining these differences between malignant (HGSC) and benign EOC tumors (*424*). One of the glycoproteins identified in their study included the folate receptor (FOLR), a receptor that has been detected in more than 78% of HGSC (*438*). Additionally, they identified changes in the protein expression levels of ADP-ribosyl cyclase/cyclic ADP-ribose hydrolase 1, an enzyme shown to be implicated in the

regulation of cancer cell invasion (*439*). Another research team by Y. Ji *et al*. (2017), examined glycosylation changes and their role in EOC acquired chemoresistance by applying another glycoproteomic technique referred to as hydrophilic interaction liquid chromatography (HILIC)-based enrichment, in combination with Stable isotope labeling with amino acids in cell culture (SILAC) labeling (*425*). Using this approach, they were able to compare N-glycosylation changes between EOC sensitive vs. EOC resistant cell lines (*425*). They identified significant changes in glycoproteins and their corresponding glycosites, which suggests for possible glycoproteins' implication in EOC multidrug resistance (*425*). For example, they found that ABCC5, which is an N-glycoprotein playing a role in multidrug resistance (also known as multidrug-resistance-associated protein 5), was significantly upregulated in EOC resistant cell lines (*425*). A study by Miyamoto *et al. (2016)*, has examined glycoprotein changes in ascites fluids using lectin enrichment followed by analysis using nanoHPLC-chip-time of flight (TOF) coupled with protein identification by Tandem Mass Spectrometry (LC–MS/MS) (*426*). This study has identified different N-glycan groups and proteins, quite enriched of large, fucosylated and sialylated glycan structures (*426*). Some of the most abundant glycoproteins identified in the ascites samples included haptoglobin, fibronectin, lumican, fibulin, hemopexin, ceruloplasmin, alpha-1-antitrypsin, and alpha-1-antichymotrypsin (*426*). They also performed a more extensive glycopeptide analysis of ascites fluids and identified N- and O-glycans in clusterin, hemopexin, and fibulin glycopeptides, which were suggested to be implicated in EOC tumorigenesis (*426*). Moreover, a study by Everest-Dass *et al*. (2016) applied a unique approach for investigating tissue-specific N-linked glycans using matrix-assisted laser desorption/ionization (MALDI) mass spectrometry imaging (MSI) on formalin-fixed paraffin-embedded (FFPE) tissue sections from EOC patients, as they were able to profile N-glycan compositions released from proteins from tissue samples (*440*). This technique can differentiate tumor from non-tumor tissue regions by using only N-glycan structure masses, as diverse and unique N-glycan structures can serve as signatures to differentiate between the tissues across biological samples and between necrotic, tumor, adipose and stromal tissue regions (*440*). Similarly, Kuzmanov *et al*. (2013) discovered 13 novel sialoglycopeptides in ovarian cyst and ascites fluids of EOC patients using electrospray ionization-LTQ Orbitrap tandem mass spectrometry (*441*). Abbott *et al*.

examined tumor-specific glycan changes between EOC tumor and normal ovarian tissue by performing lectin-binding assays followed by MS analysis (*442*). This study identified glycoprotein markers such as periostin and thrombospondin that show tumor-specific glycosylation changes in ovarian tissue and serum samples (*442*). A distinct set of N-linked sialylated glycopeptides were identified by Shetty *et al*. (2012), which displayed be significant upregulatation in serum samples of EOC patients (*443*). Another report by Li *et al*. (2008) also examined glycoprotein changes present on proteins from conditioned media of ovarian cancer cell lines, and in sera obtained from ovarian cancer patients and normal controls using high resolution Fourier transform ion cyclotron resonance MS (*444*), showing glycosylation changes in several proteins such as apolipoprotein B-100, fibronectin, and immunoglobulin A1, as observed between cancer and normal samples (*444*). Finally our work presented in chapter 2 represents the most recent data examining changes in O-glycosyaltion following GalNAc-Ts KD in ovarian cancer cell lines, as our study focused on the application of a metabolic labeling strategy for enrichment and mass spectrometry-based characterization of glycoproteins following *GALNT3* gene KD in A2780s EOC cells. We identified a total of 589 differentially expressed glycoproteins upon *GALNT3* KD, and interestingly most of these identified proteins were involved in mechanisms of cellular metabolic functions, post-translational modifications, and some implicated in EOC etiology (*427, 428*). To the best of our knowledge, our study is the first to make use of metabolic labeling to identify the impact of a GalNAc-T inhibition on altering the glycoproteins pattern in an EOC cell line. As reflected by the low percentage of glycoprotein entries annotated in SWISS-PROT (*168*), in this study we decided to use a developed prediction analysis system. Results from our data estimated that more than 88% of the identified proteins might be O-glycosylated. Additional prediction analysis were performed to compare our list of annotated proteins to previously published studies, and results indicated that more than 30% of the identified proteins are O-glycosylated Additionally, by using cellular fractionation, glycoprotein enrichment and high resolution LC/MS-based quantitative proteomics, we quantified over 2000 proteins. Among these proteins, we found several hundreds to be differentially expressed (≥ 2 fold, p-value ≤ 0.05) between the control and *GALNT3* KD A2780s EOC cells. GO cellular component classifications revealed that these differentially regulated proteins were predominantly

mapped to membrane, extracellular, cytoplasmic and intracellular organelle compartments. Furthermore, IPA functional pathway analyses of our proteomics data indicated that the *GALNT3* gene suppression resulted in reduced expression of essential pathways related to lipid, carbohydrate and amino acid metabolism, also demonstrating that changes in glycogen metabolism can provoke crucial alterations in various EOC metabolic pathways. This was further confirmed by IPA canonical pathway analyses, as the top five canonical pathways (PTM, cardiovascular and inflammatory disease, lipid metabolism, embryonic development and cell morphology, and cellular assembly) were significantly perturbed upon *GALNT3* KD. Our data also support others' findings for the implication of members of the GalNAc-Ts gene family in controlling cellular metabolism (*445*). Additionally, *GALNT3* KD was associated with the down-regulation of the metabolism related gene nodes *SOD2* and *MTHFD2*, previously shown to be overexpressed in different cancer types, including EOC (*446, 447*). Our findings suggest that metabolic pathways regulated via genes of the oxidative stress and/or mitochondrial enzymes might be controlled by PTMs such as glycosylation. Moreover, two oligosaccharyltransferases (OST) were found to be down-regulated upon *GALNT3* KD. The main function of OSTs is the transfer of lipid-linked oligosaccharides to selected asparagine residues within the consensus sequence Asn-Xaa-Ser/Thr on nascent polypeptides (*448*). The OST subunits examined from our network analysis include *DDOST* and *STT3B*, which were found to be significantly down-regulated upon *GALNT3* KD in EOC cells. Network analysis also demonstrated that *GALNT3* KD directs the down-regulation of several major gene nodes and corresponding pathways previously shown to be associated with EOC etiology, including *TGF-β*, *MAPK* and *NF-kappa-B*. Accordingly, some members of the *NF-kappa-B* pathway (*MTDH, ANXA4* and *RRM2*) were found to be significantly suppressed in *GALNT3* KD EOC cells. Moreover, several gene nodes implicated in EOC chemoresistance (including *TUBB3, MAP1B* and *ATG5*), displayed a significant decrease upon *GALNT3* KD. Indeed, augmented expression of *TUBB3* and *MAP1B* has been associated with worse prognosis and paclitaxel resistance in EOC patients (*449, 450*). Finally, the lipid kinases *PI3K* complex, known to be highly activated in human cancers (*451*), was up-regulated upon GALNT3 KD. Another gene node *B4GALT1*, previously reported as an oncogene, was also found to be up-regulated

following *GALNT3* KD. *B4GALT1* synthesizes Galβ1-4GlcNAc (N-acetyllactosamine) by transferring galactose from UDP-Gal to an acceptor sugar molecule, and thus represents a key enzyme in glycobiology (*452*). Several studies suggested that *B4GALT1* could play an essential role in the EOC dissemination (*453, 454*). The observed up-regulations of the *PI3K* complex and the *B4GALT1* gene may be due to the frequently observed redundancy associated with overlapping functions of other GalNAc-Ts enzymes (*175*), partially compensating for the KD of *GALNT3* in EOC cells. These findings reflect on the importance of the regulation of glycosylation on EOC-related gene expression. Suggesting that the detection of glycosylation-related genes could be of major importance in the characterization of the effect of GalNAc-Ts inhibition on the progression and invasivness of EOC cells.

This study has led to the identification of novel glycoproteins expressed in EOC and more importantly, sets of glycoproteins whose expression is altered by *GALNT3* KD, indicative for a potential role of *GALNT3* in modulating PTMs and metabolism pathways in EOC cells, which could significantly impact disease development. The concept of redundancy amongst the family of GalNAc-Ts has been well documented. Furthermore, changes in glycosylation patterns in cancer are associated not only with differential expression of glycoproteins (*455, 456*), but also with alterations in their glycan structures (*457*), as a common feature of tumors is the overexpression of truncated O-glycans. We have further expanded our analyses by applying a targeted O-glycoproteomic strategy that does not reduce O-glycan structures (see General conclusions and perspectives); thus information about the site and glycan structures can be studied simultaneously (*354, 365, 366, 399*). The application of such an innovative approach will allow for a better comprehension of the role of the GALNT3 enzyme and other members of the GalNAc-T isoforms in aberrant EOC glycosylation, associated with disease dissemination.

The above-mentioned studies, including ours, highlight the importance in analyzing glycoprotein signatures and defined changes in EOC cell lines, tissues and serum samples. Such gathered information at the glycoproteomic level can provide critical information for the discovery of novel tumor-associated EOC glycoproteins. More importantly, and as the above reports suggest, glycoproteins' alterations in EOC may be a result of changes in glycoprotein concentrations and/or protein glycosylation occupancy; thus, it seems

necessary to expand on protein profiling strategies and improve on current methods for examining glycoprotein changes in EOC.

5.1.2 Expression level analysis of different GalNAc-Ts in EOC and correlation with disease progression

The GalNAc-T family of enzymes provide cells with the functionality of differentially regulating the O-glycoproteome, thus allowing them to modulate multiple protein functions responsible for normal development in addition to the their role in cancerogenesis (*316, 344*). Multiple *in vitro* studies of the GalNAc-T isoenzymes suggest that this enzymes family has a wide variety of roles in the process of proteins' glycosylation, but their regulation and specified role in normal development and cancer biology have not been studied. Moreover, *in vivo* examination of deficiencies in the GalNAc-T genes appears to cause moderate or subtle phenotypic alterations in animal models when examining cancer development or metastasis (*316*). This is possibly related to the seemingly broad and partly overlapping roles of the different GalNAc-Ts in protein glycosylation, since many GalNAc-Ts can serve to fine-tune protein functions as observed in development and different pathologies, including cancer.

A frequently raised question is why evolution gave rise to this large family of isoenzymes: is due to **a)** an evolutionary requirement for multiple GalNAc-Ts which is important for efficiently regulating the glycosylation of a diverse subset of peptide and protein substrates; or **b)** a need for differential regulation of these multiple isoenzymes for the proper refinement of the distinct biological functions of their protein substrates. Several studies are indicative for non-redundant biological functions of these isoenzymes in human diseases such as cancer (*458*). Nonetheless, the challenge remains in properly identifying the individual contribution of each of these GalNAc-Ts in regulating protein function, and cellular processes in normal and pathological conditions.

Reports indicate for tumorigenic roles of the different GalNAc-Ts in cancer, suggesting a highly complex and specific O-glycosylation pattern of glycoproteins in the different cancer types. Work examining the implication of members of the GalNAc-Ts gene family in EOC dissemination is scarce. Indeed, only one study was suggestive for a possible role

of *GALNT14* in mediating the malignant behavior of EOC cells (*418*). In general, the most extensively studied GalNAc-Ts in cancer include *GALNT2*, *T4, T6, T9* and *T14* (*459*), which represent the transferases included in the study presented in chapter 3. In this study, we analyzed the expression levels of the selected five members of the polypeptide GalNAc-Ts family in EOC cell lines and EOC tumor samples (HGSCs). Similar, if not identical patterns of expression were observed for these GalNAc-Ts in both EOC cell lines and EOC tumors, and these protein expression analyses were also compared to normal ovarian cell line and tumors. Our data shows that *GALNT2* displays a very low or lack of expression in all EOC specimens studied, compared to control samples, while the expression levels of *GALNT6, T9* and *T14* were significantly higher in EOC cells and EOC tumors, as compared to very low/lack of expression in the corresponding controls. Only *GALNT4* showed very subtle, or absence of any expression in all EOC and control specimens analyzed, which is indicative for no implications of this enzyme in EOC progression. These observations were further confirmed by performing IHC examination of their expression levels in numerous EOC tumors and control tissues, using TMAs. Indeed, GALNT6, T9 and T14 exhibited very strong staining in HGSC tissues and very weak/no staining in the LMP tumors and control tissues, while GALNT2 displayed an inverse staining pattern indicative for a subtle or no expression in both HGSC and LMP tumors, compared to relatively strong expression in control tissues. Our data suggest for a possible correlation of *GALNT6, T9* and *T14* expression with EOC progression, while a putative protective role of *GALNT2* in EOC dissemination cannot be excluded. Kaplan–Meier survival curves and consecutive cox-regression analyses revealed that the expression levels of two of the GalNAc-Ts analyzed (*GALNT6* and *GALNT14*) were significantly associated with poor PFS in the studied cohort of HG serous EOC patients, suggesting the putative use of these two transferases as novel prognostic EOC markers. These results, as well as data from our previous study (*305*), are indicative for the significant inverse association of three members of the GalNAc-Ts gene family (*GALNT3, T6* and *T14*) with disease progression in serous ovarian adenocarcinoma patients.

Furthermore, the simultaneous overexpression of four GalNAc-Ts (*GALNT3, T6, T9* and *T14*) in advanced EOC raises the question about the specific and/or redundant functions of the members of this gene family in normal and pathological conditions, including EOC, and

also empahsizes on the importance of examining how these GalNAc-Ts are regulated in this cancer type. Our data shows an increase in the *GALNT3/T6* co-expression suggestive for *GALNT3/T6* possible complementary correlations, and thus for a possible functional overlap and redundancy of *GALNT3* and *GALNT6* in EOC. Based on the above considerations we presume that different GalNAc-Ts can be performing redundant and/or overlapping functions in disease progression of women with HGSC. We believe that these relationships need to be examined further, both *in vitro* and *in vivo*, to better understand how these transferases function together in initiating the biosynthesis of specific target glycoproteins, and if they do demonstrate compensatory functions that may enhance their implication in disease progression.

Data from chapter 3 demonstrate that four members of the GalNAc-Ts gene family are differentially expressed in EOC specimens. Importantly, *GALNT6* and *GALNT14* expression levels significantly correlated with poor prognosis of EOC patients with advanced disease. These results and our previously published data are indicative for a possible implication of different members of the GalNAc-Ts gene family (including *GALNT2, T3, T6, T9* and *T14*) in modulating EOC progression. Moreover, our results are suggestive for overlapping functions of some GalNAc-Ts, and especially *GALNT3* and *GALNT6*, in EOC, in conformity with *GALNT3/T6* functional redundancy described in other cancer types. Our studies highlight the neoplastic context of members of the GalNAc-Ts gene family in EOC, emphasizing the potential clinical importance in detecting their aberrant expression for use in diagnostic and prognostic purposes. Our data also presents the range of effects deficiencies in the proper glycosylation mechanism any of these GalNAc-Ts can have on proetin processing and thus enhancing neoplastic phenotypes observed in EOC. We believe that further studies focused on understanding the role of one or multiple GalNAc-Ts genes and the availability of biochemical characterization of these transferases will provide the possibility of assigning individual or multiple transferases to specific clinical contexts in EOC or other cancer types, and help in identifying a subset of transferases that could be therapeutically targeted.

5.1.3 Examining the concept of functional redundancy amongst members of the GalNAc-Ts in EOC

The concept of genetic and functional redundancy has been explored in many biological systems, and this phenomenon has been examined in genetic variants ranging from isoenzymes to transcription factors (*460*). Evolutionary studies have been helpful in understanding the need for such systems for the survival and fitness of organisms (*461*). The term genetic redundancy can also be referred to as functional redundancy and is defined as partial or complete overlap in function that occurs upon deleting a gene from a genome. This ablation shows minimal impact on phenotype due to compensation conferred by one or more genes with similar function (*461, 462*). Genetic studies of redundancy in S. cerevisiae (*463-465*) have shown that about 25% of gene duplicates are metabolic enzymes (*460*). Interestingly, the GalNAc-Ts family of isoenzymes are known for their genetic and functional redundancy, as examined through their high sequence similarity and their evolutionary identity (*316*). It has been suggested that these enzymes have been generated by genetic duplications, despite the fact that they display differences in their protein sequences (*316*). The twenty members of these GalNAc-Ts, all catalyze the same biochemical reaction defined by the addition of the first GalNAc sugar to a Ser/Thr residue on glycoproteins (*175*). The GalNAc-Ts functionality has been shown to vary both at the tissue and the cell type level, suggesting that these GalNAc-Ts isoenzymes differ in their spatial and temporal patterns of expression during development (*175, 316*). The GalNAc-Ts exhibit overall sequence similarity of around 45%, while the *GALNT3* and *GALNT6* display around 65% homology in their coding sequence (*316, 422*). Studies have shown that although the genomic structures of *GALNT3* and *GALNT6* are highly homologous, these genes are localized at different locus (2q31 and 12q13, respectively (*316, 422*)). Additionally, studies have shown that they have both similar and different substrate acceptor specificities (*422*). Initial *in vitro* studies examined in the lab provided evidence for a strong oncogenic capacity of the *GALNT3* gene in EOC, including its potential role in EOC cell proliferation, cell cycle control and cell migration/ invasion (*305*). In addition, gene expression profile analysis demonstrated a strong downregulation of major gene nodes implicated in EOC tumorigenesis (*305*). Several studies have focused on performing

genetic approaches to understand the roles of GalNac-Ts in mammalian O-glycosylation. Importantly, a number of synthetic deficiencies of the GalNac-Ts have been produced in animal model studies, and only few have shown signs of effects on animal survival or development (*316*). The lack of phenotypic differences is most likely not a result of subtle changes in the genome, but is the outcome of effective mechanisms that have allowed decreasing any major perturbations that may have been acquired from a single gene knockout or deletion. We thus hypothesized that the *GALNT6* gene may be providing biosynthetic backup in those cells ablated for the *GALNT3* gene, and we proposed that it may occur through a feedback mechanism. We suggested that when one isoenyzme (*GALNT3*) is knocked down or out, the functional pathway of GALNT3 gets reduced, which in turn triggers an increase in the expression of its closest homolog GALNT6 allowing for the reprograming of *GALNT3* pathway in these EOC cells. Indeed, our results show that when knocking out the *GALNT3* gene in the A2780s cells, an upregulation of the GALNT6 protein expression occurs. These data were suggestive for a mechanism of feedback regulation between these two genes. A report by Bennett *et al*. (1999), has demonstrated the existence of genetic redundancy of the GalNAc-transferases (*GALNT3* and *GALNT6*), but suggested that this genetic overlap does not provide full functional redundancy (*422*). This was observed when performing functional analyses comparing the effect of a single gene *GALNT3* KO versus a double gene KO for both *GALNT3* and *GALNT6*. Cell migration, invasion, proliferation and cell cycle analysis were significantly reduced in the double *GALNT3/T6* gene KO cells compared to the single *GALNT3* gene KO, which was suggestive that *GALNT6* expression provides backup genetic activity in the single *GALNT3* gene KO clones. This genetic activity was supported when examining the expression pattern of oncofetal fibronectin (FN1), an extracellular matrix (ECM) glycoprotein that is synthesized within the cell matrix surrounding tumors and fetal tissue (*466*). Earlier studies specified FN1 to be glycosylated by *GALNT3* (*467*), but more recent findings disproved the specificity of *GALNT3* to the glycosylation site of FN1 and showed FN1 as a candidate substrate of *GALNT6* (*422, 468*). The hypothesis was supported by several studies examining the role of *GALNT6* overexpression in mammary carcinogenesis (*408, 468*). Similarly, our protein expression analysis showed a significant increase in FN1 expression in the *GALNT3* KD and KO clones, properly supporting the stabilizing effect of

GALNT6 expression on FN1. This was further confirmed by VVA lectin pull down assays, confirming the activity of *GALNT6* in the *GALNT3* KO cells. Furthermore, MUC1 expression was downregulated in both *GALNT3* KO and *GALNT3/T6* KO cells when examined using protein expression analysis and VVA lectin pull down assays. More interestingly, *in vivo* studies very well supported the observed functional redundancy imposed by *GALNT6*, as animals injected with the double *GALNT3/T6* gene KO clones showed significant increase in survival rates compared to those injected with the single *GALNT3* gene KO or the control clone, respectively. Our study is the first to report a positive effect on mice survival upon knocking out two GalNAc-Ts genes, showing great promise in using both these genes as EOC markers/therapeutic targets. Collectively our data provide strong evidence of the possible genetic redundancy of the two polypeptide GalNAc-Ts (*GALNT3* and *T6*) in EOC. Nonetheless, the possible concept of functional redundancy between these two genes needs to be further validated to better confirm if these two genes carry cooperative or redundant functions in our model of EOC. We need to validate the possibility of functional complementation to confirm if these two genes are actually redundant in function, and as defined "Two genes are considered to be redundant if they can fully or partially substitute each others functions" (*469*). Further experiments examining the role of these two genes independently can help in confirming their individual roles, and thus understand if they do play redundant roles in different cancers such as EOC. To make a final conclusion on the extent of genetic redundancy between *GALNT3* and *GALNT6* in EOC requires a larger-scale analyses, but the observations we have acquired so so far have shed light on the possible role one or two of these genes play in the phenotypic complexity of EOC progression.

General Conclusions and Perspectives

Protein glycosylation is the most common PTM in mammalian cells, and it is involved in many biological pathways and molecular functions. Glycosylation is also suggested to be an excellent target for proteomics-based disease investigations, such as studies in cancer, neurodegenerative diseases, and metabolic related disorders. This is highly due to the fact that aberrant glycosylation has been associated with different pathologies including cancer. Hence, glycoforms of glycoproteins may serve as potential biomarkers for the early detection/treatment response of the disease and/or as putative therapeutic targets.

Mucin type O-glycosylation is a highly regulated process dependent on the enzymatic activity of GalNAc-Ts. The family of GalNAc-Ts plays a very important role in controlling sites of O-glycan attachments (*470*). This family of isoenzymes displays characteristics that are important in normal development and in cancer progression. These characteristics include: **a)** their sequence similarity, with some members showing higher similarities such as (*GALNT3* and *GALNT6*) (*471*), **b)** their genomic localization/organization with members sharing some of common genomic features (*471, 472*), and **c)** their substrate specificities with a certain level of overlap (*471, 473*). Collectively these features show that each GalNAc-T may have a different or unique function, without eliminating the possibility of functional overlap. Thus, as shown also by us, it appears to be necessary to gain more insight on the role of the different members of these GalNAc-Ts in EOC. We believe it would be important to better examine their specified expression pattern, their glycan structure alterations in tumor development, as well as their own potential changes in cancer (including EOC) progression. Thus far, studies on the specificities of the GalNAc-Ts have not properly defined their specified acceptor peptide sequence motifs, which would allow predicting the O-glycosylation capacity of each of these enzymes in different cancer types. This scientific delay is largely due to the experimental constraints/complexities that come in analyzing peptide and protein specificities of these transferases. We have applied an innovative approach focused on examining both the peptide and protein sequences affected by the aberrant expression of these GalNAc-Ts. This PhD project studied the effect of ablating the expression of one of this GalNAc-Ts (*GALNT3*) in EOC. We were able to show differential regulation of a plethora of glycoproteins affected by the KD of the

GALNT3 gene, some of which could represent new potential EOC biomarkers and/or therapeutic targets. This application demonstrates how advances in glycoproteomic analysis is an important part of proteomics, since as shown by us, it has helped reveal very important biological changes in terms of disease such as EOC, and it has also made it possible to study glycoprotein/glycan alterations in a more defined manner. In perspective, we are aiming to further benefit from the IsoTag technology in order to identify glycan structure perturbations/changes in terms of the level of elongation, site of attachment, and type of terminal modification acquired from the ablation of one or several of these GalNAc-Ts in EOC. Currently, this approach requires intense manual data processing for properly identifying and studying the glycan structure changes. Nonetheless, our lab and our collaborators are working on developing improved bioinformatics algorithms that can ease the way for a defined glycan identification software or database in order to help analyze glycoproteins and their attached glycan structures. Recently, our lab start using the Byonic software, which is a search engine that has been developed to identify peptides and proteins by tandem MS, and is now showing great promise in the field of glycoproteomics. Byonic has been also developed to identify glycosylated peptides, and with the use of predefined algorithms, it can identify glycan structures found on a specified peptide sequence. The application of the IsoTag platform in parallel with the use of the Byonic software has begun to show great promise in producing replicable results when examining glycan changes in different cancer cell lines as identified in studies presented by Dr. Woo (*354, 365, 366, 399*).

Moreover, we wanted to investigate for other members of the GalNAc-T family that may be implicated in aberrant O-glycosylation in EOC. Our focus was to examine the expression pattern of several GalNAc-Ts between normal and HG ovarian cancer patient samples. The objective of this part of the thesis was to provide a larger panel of biomarkers that can be potentially used for non-invasive screening of EOC patients. As previously mentioned, current screening methods for EOC have been based on serum tumor markers or ultrasound imaging of the ovaries and these approaches have shown to be not entirely specific for EOC. Therefore, there is a need for new biomarkers that are both sensitive and specific in for the different subtypes of EOC patients. In this PhD project we evaluated the expression of 5 members of the GalNAc-T family (*GALNT2, T4, T6, T9* and *T14*) in EOC

specimens and examined their correlation with the clinicopathologic features of these patients. Our results indicated that high GALNT6 and GALNT14 as negative prognostic predictors for patients with EOC, and we suggested their overexpression could directly contribute to aberrant mucin type O-glycosylation in EOC. More importantly, our findings support future studies that can focus on examining the implication of *GALNT6* and *GALNT14* in the molecular mechanisms of the abnormal glycosylation associated with EOC dissemination. Furthermore, this study has also led to the identification of possible redundant functions of two GalNAc-Ts (*GALNT3* and *T6*). Our preliminary analysis has points the need to more profoundly examine the hypothesis of synthetic redundancy of these two GalNAc-Ts in EOC progression and metastasis. We next provided experimental evidence that this may indeed be due to the compensatory role of its closest homolog *GALNT6*, which was strongly induced upon *GALNT3* KD both *in vitro*, and *in vivo*. Our data from this work strongly supported the perception of biosynthetic backup provided by these GalNAc-Ts. Our double KO experiments performed in this study strongly support the impact of the ablation of not one member, but two members of these GalNAc-Ts in the functional characteristics of EOC cells, as evaluated by their decreased level of migration, invasion, proliferation and cell cycle progression. More importantly, the role the double KO had an important implication on the glycosylation state of their proposed substrate proteins (MUC1 and FN1). Interestingly, we were also the first to show effects on the survival of animals injected with double KO EOC cells, where a significant increase in the overall survival of animals injected with cells ablated for the gene expression of the two members (*GALNT3* and *GALNT6*) was observed. Collectively, our data highlight the importance of integrating and examining multiple members of the GalNAc-T family into providing some prognostic information for EOC patients. Moreover, screening for *GALNT3* and *GALNT6* inhibitors could be valuable for the development of novel therapeutic modalities against EOC.

The research project is further focused on using the IsoTaG platform to analyze *GALNT3* KO and *GALNT3/T6* KO mediated alteration on the expression of O-glycoproteins. More importantly with the application of the IsoTaG approach, we aim to simultaneously identify altered expression of O-glycoproteins and/or alterations in their corresponding glycan structures in the EOC cell lines A2780s. It will be thus interesting to investigate the

implication of *GALNT3* and *GALNT6* in the aberrant O-glycosylation in EOC cells. Furthermore, the application of the glycoproteomic approach could better characterize both the synergistic and compensatory *GALNT3/T6* correlations, which could improve our understandings of their individual and/or combined role in EOC dissemination.

Bibliography

1. R. L. Siegel, K. D. Miller, A. Jemal, Cancer statistics, 2016. *CA: a cancer journal for clinicians* **66**, 7-30 (2016).
2. P. E. s. O. O. G. Group, A systematic review of patient values, preferences and expectations for the treatment of recurrent ovarian cancer. *Gynecologic oncology*, (2017).
3. C. f. f. D. Diaz-Gil *et al.*, Prediction of 5-year survival in advanced-stage ovarian cancer patients based on computed tomography peritoneal carcinomatosis index. *Abdominal radiology* **41**, 2196-2202 (2016).
4. P. Jasen, From the "silent killer" to the "whispering disease": ovarian cancer and the uses of metaphor. *Medical history* **53**, 489-512 (2009).
5. R. J. Kurman, M. Shih Ie, The origin and pathogenesis of epithelial ovarian cancer: a proposed unifying theory. *The American journal of surgical pathology* **34**, 433-443 (2010).
6. A. N. Vargas, Natural history of ovarian cancer. *Ecancermedicalscience* **8**, 465 (2014).
7. L. Dubeau, The cell of origin of ovarian epithelial tumors and the ovarian surface epithelium dogma: does the emperor have no clothes? *Gynecologic oncology* **72**, 437-442 (1999).
8. D. G. Rosen *et al.*, Ovarian cancer: pathology, biology, and disease models. *Frontiers in bioscience* **14**, 2089-2102 (2009).
9. A. Jemal *et al.*, Cancer statistics, 2009. *CA: a cancer journal for clinicians* **59**, 225-249 (2009).
10. S. Boussios *et al.*, Non-epithelial Ovarian Cancer: Elucidating Uncommon Gynaecological Malignancies. *Anticancer research* **36**, 5031-5042 (2016).
11. S. J. Lee *et al.*, Clinical characteristics of metastatic tumors to the ovaries. *Journal of Korean medical science* **24**, 114-119 (2009).
12. W. Li *et al.*, Ovarian metastases resection from extragenital primary sites: outcome and prognostic factor analysis of 147 patients. *BMC cancer* **12**, 278 (2012).
13. C. Straight, S. L. Zweizig, in *Diseases and Conditions*. (Wolters Kluwer, 5minuteconsult.com).
14. R. J. Kurman, Origin and molecular pathogenesis of ovarian high-grade serous carcinoma. *Annals of oncology : official journal of the European Society for Medical Oncology* **24 Suppl 10**, x16-21 (2013).
15. W. G. McCluggage, Morphological subtypes of ovarian carcinoma: a review with emphasis on new developments and pathogenesis. *Pathology* **43**, 420-432 (2011).
16. R. D. Alvarez, B. Y. Karlan, J. F. Strauss, "Ovarian cancers: Evolving paradigms in research and care": Report from the Institute of Medicine. *Gynecologic oncology* **141**, 413-415 (2016).
17. E. I. Braicu *et al.*, Role of histological type on surgical outcome and survival following radical primary tumour debulking of epithelial ovarian, fallopian tube and peritoneal cancers. *British journal of cancer* **105**, 1818-1824 (2011).
18. J. D. Seidman *et al.*, The histologic type and stage distribution of ovarian carcinomas of surface epithelial origin. *International journal of gynecological pathology : official journal of the International Society of Gynecological Pathologists* **23**, 41-44 (2004).
19. R. Vang, M. Shih Ie, R. J. Kurman, Ovarian low-grade and high-grade serous carcinoma: pathogenesis, clinicopathologic and molecular biologic features, and diagnostic problems. *Advances in anatomic pathology* **16**, 267-282 (2009).

20. C. B. Gilks *et al.*, Tumor cell type can be reproducibly diagnosed and is of independent prognostic significance in patients with maximally debulked ovarian carcinoma. *Human pathology* **39**, 1239-1251 (2008).
21. D. J. Storey *et al.*, Endometrioid epithelial ovarian cancer : 20 years of prospectively collected data from a single center. *Cancer* **112**, 2211-2220 (2008).
22. P. D. DePriest *et al.*, Endometrioid carcinoma of the ovary and endometriosis: the association in postmenopausal women. *Gynecologic oncology* **47**, 71-75 (1992).
23. M. Erzen, S. Rakar, B. Klancnik, K. Syrjanen, Endometriosis-associated ovarian carcinoma (EAOC): an entity distinct from other ovarian carcinomas as suggested by a nested case-control study. *Gynecologic oncology* **83**, 100-108 (2001).
24. H. Yoshikawa *et al.*, Prevalence of endometriosis in ovarian cancer. *Gynecologic and obstetric investigation* **50 Suppl 1**, 11-17 (2000).
25. P. T. Soliman *et al.*, Synchronous primary cancers of the endometrium and ovary: a single institution review of 84 cases. *Gynecologic oncology* **94**, 456-462 (2004).
26. R. Zaino *et al.*, Simultaneously detected endometrial and ovarian carcinomas--a prospective clinicopathologic study of 74 cases: a gynecologic oncology group study. *Gynecologic oncology* **83**, 355-362 (2001).
27. M. K. McConechy *et al.*, Ovarian and endometrial endometrioid carcinomas have distinct CTNNB1 and PTEN mutation profiles. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* **27**, 128-134 (2014).
28. S. Mabuchi, T. Sugiyama, T. Kimura, Clear cell carcinoma of the ovary: molecular insights and future therapeutic perspectives. *Journal of gynecologic oncology* **27**, e31 (2016).
29. M. S. Anglesio *et al.*, Clear cell carcinoma of the ovary: a report from the first Ovarian Clear Cell Symposium, June 24th, 2010. *Gynecologic oncology* **121**, 407-415 (2011).
30. E. L. Jenison *et al.*, Clear cell adenocarcinoma of the ovary: a clinical analysis and comparison with serous carcinoma. *Gynecologic oncology* **32**, 65-71 (1989).
31. J. Tammela, J. P. Geisler, P. N. Eskew, Jr., H. E. Geisler, Clear cell carcinoma of the ovary: poor prognosis compared to serous carcinoma. *European journal of gynaecological oncology* **19**, 438-440 (1998).
32. E. Veras *et al.*, Cystic and adenofibromatous clear cell carcinomas of the ovary: distinctive tumors that differ in their pathogenesis and behavior: a clinicopathologic analysis of 122 cases. *The American journal of surgical pathology* **33**, 844-853 (2009).
33. S. C. Jones *et al.*, Australian women's awareness of ovarian cancer symptoms, risk and protective factors, and estimates of own risk. *Cancer causes & control : CCC* **21**, 2231-2239 (2010).
34. T. Matsumoto *et al.*, Distinct beta-catenin and PIK3CA mutation profiles in endometriosis-associated ovarian endometrioid and clear cell carcinomas. *American journal of clinical pathology* **144**, 452-463 (2015).
35. K. R. Cho, M. Shih Ie, Ovarian cancer. *Annual review of pathology* **4**, 287-313 (2009).
36. M. K. McConechy *et al.*, Subtype-specific mutation of PPP2R1A in endometrial and ovarian carcinomas. *The Journal of pathology* **223**, 567-573 (2011).
37. D. A. Bell, Low-grade serous tumors of ovary. *International journal of gynecological pathology : official journal of the International Society of Gynecological Pathologists* **33**, 348-356 (2014).
38. A. Malpica *et al.*, Grading ovarian serous carcinoma using a two-tier system. *The American journal of surgical pathology* **28**, 496-504 (2004).
39. D. C. Bodurka *et al.*, Reclassification of serous ovarian carcinoma by a 2-tier system: a Gynecologic Oncology Group Study. *Cancer* **118**, 3087-3094 (2012).

40. C. Gourley *et al.*, Gynecologic Cancer InterGroup (GCIG) consensus review for ovarian and primary peritoneal low-grade serous carcinomas. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society* **24**, S9-13 (2014).
41. D. M. Gershenson *et al.*, Impact of Age and Primary Disease Site on Outcome in Women With Low-Grade Serous Carcinoma of the Ovary or Peritoneum: Results of a Large Single-Institution Registry of a Rare Tumor. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **33**, 2675-2682 (2015).
42. R. Dehari, R. J. Kurman, S. Logani, M. Shih Ie, The development of high-grade serous carcinoma from atypical proliferative (borderline) serous tumors and low-grade micropapillary serous carcinoma: a morphologic and molecular genetic analysis. *The American journal of surgical pathology* **31**, 1007-1012 (2007).
43. R. L. Parker, P. B. Clement, D. J. Chercover, T. Sornarajah, C. B. Gilks, Early recurrence of ovarian serous borderline tumor as high-grade carcinoma: a report of two cases. *International journal of gynecological pathology : official journal of the International Society of Gynecological Pathologists* **23**, 265-272 (2004).
44. E. G. Silva, C. S. Tornos, A. Malpica, D. M. Gershenson, Ovarian serous neoplasms of low malignant potential associated with focal areas of serous carcinoma. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* **10**, 663-667 (1997).
45. A. A. Tone *et al.*, Intratumoral heterogeneity in a minority of ovarian low-grade serous carcinomas. *BMC cancer* **14**, 982 (2014).
46. J. Brown, M. Frumovitz, Mucinous tumors of the ovary: current thoughts on diagnosis and management. *Current oncology reports* **16**, 389 (2014).
47. Y. Y. Chou *et al.*, Differentiation of ovarian mucinous carcinoma and metastatic colorectal adenocarcinoma by immunostaining with beta-catenin. *Histopathology* **43**, 151-156 (2003).
48. J. A. Ledermann *et al.*, Gynecologic Cancer InterGroup (GCIG) consensus review for mucinous ovarian carcinoma. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society* **24**, S14-19 (2014).
49. R. J. Zaino *et al.*, Advanced stage mucinous adenocarcinoma of the ovary is both rare and highly lethal: a Gynecologic Oncology Group study. *Cancer* **117**, 554-562 (2011).
50. S. Yasmeen, A. Hannan, F. Sheikh, A. A. Syed, N. Siddiqui, Borderline tumors of the ovary: A clinicopathological study. *Pakistan journal of medical sciences* **33**, 369-373 (2017).
51. M. Rechsteiner *et al.*, TP53 mutations are common in all subtypes of epithelial ovarian cancer and occur concomitantly with KRAS mutations in the mucinous type. *Experimental and molecular pathology* **95**, 235-241 (2013).
52. M. S. Anglesio *et al.*, Molecular characterization of mucinous ovarian tumours supports a stratified treatment approach with HER2 targeting in 19% of carcinomas. *The Journal of pathology* **229**, 111-120 (2013).
53. C. A. Doubeni, A. R. Doubeni, A. E. Myers, Diagnosis and Management of Ovarian Cancer. *American family physician* **93**, 937-944 (2016).
54. R. F. Ozols, R. C. Young, Ovarian cancer. *Current problems in cancer* **11**, 57-122 (1987).
55. M. Shih Ie, R. J. Kurman, Ovarian tumorigenesis: a proposed model based on morphological and molecular genetic analysis. *The American journal of pathology* **164**, 1511-1518 (2004).
56. K. P. Pennington *et al.*, Germline and somatic mutations in homologous recombination genes predict platinum response and survival in ovarian, fallopian tube, and

peritoneal carcinomas. *Clinical cancer research : an official journal of the American Association for Cancer Research* **20**, 764-775 (2014).
57. P. M. Jones, R. Drapkin, Modeling High-Grade Serous Carcinoma: How Converging Insights into Pathogenesis and Genetics are Driving Better Experimental Platforms. *Frontiers in oncology* **3**, 217 (2013).
58. T. Kaku *et al.*, Histological classification of ovarian cancer. *Medical electron microscopy : official journal of the Clinical Electron Microscopy Society of Japan* **36**, 9-17 (2003).
59. D. Hanahan, R. A. Weinberg, Hallmarks of cancer: the next generation. *Cell* **144**, 646-674 (2011).
60. L. A. Loeb, K. R. Loeb, J. P. Anderson, Multiple mutations and cancer. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 776-781 (2003).
61. R. L. Hollis, M. Churchman, C. Gourley, Distinct implications of different BRCA mutations: efficacy of cytotoxic chemotherapy, PARP inhibition and clinical outcome in ovarian cancer. *OncoTargets and therapy* **10**, 2539-2551 (2017).
62. T. Pal *et al.*, BRCA1 and BRCA2 mutations account for a large proportion of ovarian carcinoma cases. *Cancer* **104**, 2807-2816 (2005).
63. H. A. Risch *et al.*, Prevalence and penetrance of germline BRCA1 and BRCA2 mutations in a population series of 649 women with ovarian cancer. *American journal of human genetics* **68**, 700-710 (2001).
64. K. B. Kuchenbaecker *et al.*, Identification of six new susceptibility loci for invasive epithelial ovarian cancer. *Nature genetics* **47**, 164-171 (2015).
65. A. A. Ahmed *et al.*, Driver mutations in TP53 are ubiquitous in high grade serous carcinoma of the ovary. *The Journal of pathology* **221**, 49-56 (2010).
66. B. T. Hennessy *et al.*, Somatic mutations in BRCA1 and BRCA2 could expand the number of patients that benefit from poly (ADP ribose) polymerase inhibitors in ovarian cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **28**, 3570-3576 (2010).
67. F. C. Martins *et al.*, Combined image and genomic analysis of high-grade serous ovarian cancer reveals PTEN loss as a common driver event and prognostic classifier. *Genome biology* **15**, 526 (2014).
68. M. Verma, The Role of Epigenomics in the Study of Cancer Biomarkers and in the Development of Diagnostic Tools. *Advances in experimental medicine and biology* **867**, 59-80 (2015).
69. E. S. Hosseini *et al.*, Dysregulated expression of long noncoding RNAs in gynecologic cancers. *Molecular cancer* **16**, 107 (2017).
70. R. L. Huang *et al.*, Comprehensive methylome analysis of ovarian tumors reveals hedgehog signaling pathway regulators as prognostic DNA methylation biomarkers. *Epigenetics* **8**, 624-634 (2013).
71. K. Garg *et al.*, BRCA1 immunohistochemistry in a molecularly characterized cohort of ovarian high-grade serous carcinomas. *The American journal of surgical pathology* **37**, 138-146 (2013).
72. M. Keita *et al.*, Global methylation profiling in serous ovarian cancer is indicative for distinct aberrant DNA methylation signatures associated with tumor aggressiveness and disease progression. *Gynecologic oncology* **128**, 356-363 (2013).
73. M. Orsini, A. Travaglione, E. Capobianco, Warehousing re-annotated cancer genes for biomarker meta-analysis. *Computer methods and programs in biomedicine* **111**, 166-180 (2013).
74. A. Toss *et al.*, Ovarian cancer: can proteomics give new insights for therapy and diagnosis? *International journal of molecular sciences* **14**, 8271-8290 (2013).

75. M. M. Emori, R. Drapkin, The hormonal composition of follicular fluid and its implications for ovarian cancer pathogenesis. *Reproductive biology and endocrinology : RB&E* **12**, 60 (2014).
76. E. A. Braga, M. V. Fridman, N. E. Kushlinskii, Molecular Mechanisms of Ovarian Carcinoma Metastasis: Key Genes and Regulatory MicroRNAs. *Biochemistry. Biokhimiia* **82**, 529-541 (2017).
77. E. Lengyel, Ovarian cancer development and metastasis. *The American journal of pathology* **177**, 1053-1064 (2010).
78. O. Solheim *et al.*, Epithelial-mesenchymal transition markers in malignant ovarian germ cell tumors. *APMIS : acta pathologica, microbiologica, et immunologica Scandinavica*, (2017).
79. X. X. Jie, X. Y. Zhang, C. J. Xu, Epithelial-to-mesenchymal transition, circulating tumor cells and cancer metastasis: Mechanisms and clinical applications. *Oncotarget*, (2017).
80. V. Kumar, N. Jagadish, A. Suri, Role of A-Kinase anchor protein (AKAP4) in growth and survival of ovarian cancer cells. *Oncotarget*, (2017).
81. Y. Liu *et al.*, Annexin A2 inhibition suppresses ovarian cancer progression via regulating beta-catenin/EMT. *Oncology reports* **37**, 3643-3650 (2017).
82. Y. Zhou *et al.*, NID1, a new regulator of EMT required for metastasis and chemoresistance of ovarian cancer cells. *Oncotarget* **8**, 33110-33121 (2017).
83. J. Sun *et al.*, GOLPH3 induces epithelial-mesenchymal transition via Wnt/beta-catenin signaling pathway in epithelial ovarian cancer. *Cancer medicine* **6**, 834-844 (2017).
84. T. L. Yeung *et al.*, ELF3 is a negative regulator of epithelial-mesenchymal transition in ovarian cancer cells. *Oncotarget* **8**, 16951-16963 (2017).
85. R. Sheta *et al.*, Hic-5 regulates epithelial to mesenchymal transition in ovarian cancer cells in a TGFbeta1-independent manner. *Oncotarget* **8**, 82506-82530 (2017).
86. Y. Katsuno, S. Lamouille, R. Derynck, TGF-beta signaling and epithelial-mesenchymal transition in cancer progression. *Current opinion in oncology* **25**, 76-84 (2013).
87. A. Faddaoui *et al.*, The mannose receptor LY75 (DEC205/CD205) modulates cellular phenotype and metastatic potential of ovarian cancer cells. *Oncotarget* **7**, 14125-14142 (2016).
88. R. Catena *et al.*, Bone marrow-derived Gr1+ cells can generate a metastasis-resistant microenvironment via induced secretion of thrombospondin-1. *Cancer discovery* **3**, 578-589 (2013).
89. Z. Granot *et al.*, Tumor entrained neutrophils inhibit seeding in the premetastatic lung. *Cancer cell* **20**, 300-314 (2011).
90. D. F. Quail, J. A. Joyce, Microenvironmental regulation of tumor progression and metastasis. *Nature medicine* **19**, 1423-1437 (2013).
91. T. D. Tlsty, L. M. Coussens, Tumor stroma and regulation of cancer development. *Annual review of pathology* **1**, 119-150 (2006).
92. J. Permuth-Wey, T. A. Sellers, Epidemiology of ovarian cancer. *Methods in molecular biology* **472**, 413-437 (2009).
93. S. Jervis *et al.*, Ovarian cancer familial relative risks by tumour subtypes and by known ovarian cancer genetic susceptibility variants. *Journal of medical genetics* **51**, 108-113 (2014).
94. M. Soegaard *et al.*, Risk of ovarian cancer in women with first-degree relatives with cancer. *Acta obstetricia et gynecologica Scandinavica* **88**, 449-456 (2009).
95. J. F. Stratton, P. Pharoah, S. K. Smith, D. Easton, B. A. Ponder, A systematic review and meta-analysis of family history and risk of ovarian cancer. *British journal of obstetrics and gynaecology* **105**, 493-499 (1998).

96. B. A. Werness, G. H. Eltabbakh, Familial ovarian cancer and early ovarian cancer: biologic, pathologic, and clinical features. *International journal of gynecological pathology : official journal of the International Society of Gynecological Pathologists* **20**, 48-63 (2001).
97. H. T. Lynch *et al.*, Hereditary ovarian carcinoma: heterogeneity, molecular genetics, pathology, and management. *Molecular oncology* **3**, 97-137 (2009).
98. Y. Miki *et al.*, A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* **266**, 66-71 (1994).
99. T. R. Rebbeck *et al.*, Association of type and location of BRCA1 and BRCA2 mutations with risk of breast and ovarian cancer. *Jama* **313**, 1347-1361 (2015).
100. R. Wooster *et al.*, Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science* **265**, 2088-2090 (1994).
101. T. Walsh *et al.*, Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 18032-18037 (2011).
102. C. La Vecchia, Ovarian cancer: epidemiology and risk factors. *European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation* **26**, 55-62 (2017).
103. A. S. Whittemore, R. Harris, J. Itnyre, Characteristics relating to ovarian cancer risk: collaborative analysis of 12 US case-control studies. II. Invasive epithelial ovarian cancers in white women. Collaborative Ovarian Cancer Group. *American journal of epidemiology* **136**, 1184-1203 (1992).
104. N. N. Luan *et al.*, Breastfeeding and ovarian cancer risk: a meta-analysis of epidemiologic studies. *The American journal of clinical nutrition* **98**, 1020-1031 (2013).
105. S. S. Tworoger, K. M. Fairfield, G. A. Colditz, B. A. Rosner, S. E. Hankinson, Association of oral contraceptive use, other contraceptive methods, and infertility with ovarian cancer risk. *American journal of epidemiology* **166**, 894-901 (2007).
106. A. Lukanova, R. Kaaks, Endogenous hormones and ovarian cancer: epidemiology and current hypotheses. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **14**, 98-107 (2005).
107. C. Collaborative Group On Epidemiological Studies Of Ovarian *et al.*, Menopausal hormone use and ovarian cancer risk: individual participant meta-analysis of 52 epidemiological studies. *Lancet* **385**, 1835-1842 (2015).
108. B. M. Reid, J. B. Permuth, T. A. Sellers, Epidemiology of ovarian cancer: a review. *Cancer biology & medicine* **14**, 9-32 (2017).
109. U. Chandran *et al.*, Healthy eating index and ovarian cancer risk. *Cancer causes & control : CCC* **22**, 563-571 (2011).
110. J. Xie *et al.*, A prospective cohort study of dietary indices and incidence of epithelial ovarian cancer. *Journal of ovarian research* **7**, 112 (2014).
111. C. F. Gosvig *et al.*, Coffee, tea, and caffeine consumption and risk of epithelial ovarian cancer and borderline ovarian tumors: Results from a Danish case-control study. *Acta oncologica* **54**, 1144-1151 (2015).
112. Y. F. Zhang *et al.*, Tea consumption and the incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation* **24**, 353-362 (2015).

113. C. Collaborative Group on Epidemiological Studies of Ovarian *et al.*, Ovarian cancer and smoking: individual participant meta-analysis including 28,114 women with ovarian cancer from 51 epidemiological studies. *The Lancet. Oncology* **13**, 946-956 (2012).
114. in *PDQ Cancer Information Summaries*. (Bethesda (MD), 2002).
115. T. R. Rebbeck, N. D. Kauff, S. M. Domchek, Meta-analysis of risk reduction estimates associated with risk-reducing salpingo-oophorectomy in BRCA1 or BRCA2 mutation carriers. *Journal of the National Cancer Institute* **101**, 80-87 (2009).
116. J. B. Greer, F. Modugno, G. O. Allen, R. B. Ness, Androgenic progestins in oral contraceptives and the risk of epithelial ovarian cancer. *Obstetrics and gynecology* **105**, 731-740 (2005).
117. J. M. Schildkraut, B. Calingaert, P. A. Marchbanks, P. G. Moorman, G. C. Rodriguez, Impact of progestin and estrogen potency in oral contraceptives on ovarian cancer risk. *Journal of the National Cancer Institute* **94**, 32-38 (2002).
118. W. L. Yang, Z. Lu, R. C. Bast, Jr., The role of biomarkers in the management of epithelial ovarian cancer. *Expert review of molecular diagnostics* **17**, 577-591 (2017).
119. R. C. Bast, Jr. *et al.*, Reactivity of a monoclonal antibody with human ovarian carcinoma. *The Journal of clinical investigation* **68**, 1331-1337 (1981).
120. J. M. Niloff, R. C. Knapp, E. Schaetzl, C. Reynolds, R. C. Bast, Jr., CA125 antigen levels in obstetric and gynecologic patients. *Obstetrics and gynecology* **64**, 703-707 (1984).
121. C. Hising, I. M. Anjegard, N. Einhorn, Clinical relevance of the CA 125 assay in monitoring of ovarian cancer patients. *American journal of clinical oncology* **14**, 111-114 (1991).
122. E. Kobayashi *et al.*, Biomarkers for screening, diagnosis, and monitoring of ovarian cancer. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **21**, 1902-1912 (2012).
123. J. G. Cohen, M. White, A. Cruz, R. Farias-Eisner, In 2014, can we do better than CA125 in the early detection of ovarian cancer? *World journal of biological chemistry* **5**, 286-300 (2014).
124. R. G. Moore *et al.*, The use of multiple novel tumor biomarkers for the detection of ovarian carcinoma in patients with a pelvic mass. *Gynecologic oncology* **108**, 402-408 (2008).
125. M. Schummer *et al.*, Comparative hybridization of an array of 21,500 ovarian cDNAs for the discovery of genes overexpressed in ovarian carcinomas. *Gene* **238**, 375-385 (1999).
126. R. C. Bast, Jr. *et al.*, New tumor markers: CA125 and beyond. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society* **15 Suppl 3**, 274-281 (2005).
127. R. Drapkin *et al.*, Human epididymis protein 4 (HE4) is a secreted glycoprotein that is overexpressed by serous and endometrioid ovarian carcinomas. *Cancer research* **65**, 2162-2169 (2005).
128. N. S. Karlsen, M. A. Karlsen, C. K. Hogdall, E. V. Hogdall, HE4 tissue expression and serum HE4 levels in healthy individuals and patients with benign or malignant tumors: a systematic review. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **23**, 2285-2295 (2014).
129. F. A. Eggink, C. M. Koopmans, H. W. Nijman, Surgery for patients with newly diagnosed advanced ovarian cancer: which patient, when and extent? *Current opinion in oncology*, (2017).

130. T. B. Stage, T. K. Bergmann, D. L. Kroetz, Clinical Pharmacokinetics of Paclitaxel Monotherapy: An Updated Literature Review. *Clinical pharmacokinetics*, (2017).
131. in *PDQ Cancer Information Summaries*. (Bethesda (MD), 2002).
132. B. A. Winter-Roach, H. C. Kitchener, T. A. Lawrie, Adjuvant (post-surgery) chemotherapy for early stage epithelial ovarian cancer. *The Cochrane database of systematic reviews*, CD004706 (2012).
133. NIH consensus conference. Ovarian cancer. Screening, treatment, and follow-up. NIH Consensus Development Panel on Ovarian Cancer. *Jama* **273**, 491-497 (1995).
134. E. L. Trimble *et al.*, Neoadjuvant therapy in cancer treatment. *Cancer* **72**, 3515-3524 (1993).
135. J. Morrison, K. Haldar, S. Kehoe, T. A. Lawrie, Chemotherapy versus surgery for initial treatment in advanced ovarian epithelial cancer. *The Cochrane database of systematic reviews*, CD005343 (2012).
136. J. Morrison, A. Swanton, S. Collins, S. Kehoe, Chemotherapy versus surgery for initial treatment in advanced ovarian epithelial cancer. *The Cochrane database of systematic reviews*, CD005343 (2007).
137. J. P. Neijt *et al.*, Randomised trial comparing two combination chemotherapy regimens (Hexa-CAF vs CHAP-5) in advanced ovarian carcinoma. *Lancet* **2**, 594-600 (1984).
138. G. Omura *et al.*, A randomized trial of cyclophosphamide and doxorubicin with or without cisplatin in advanced ovarian carcinoma. A Gynecologic Oncology Group Study. *Cancer* **57**, 1725-1730 (1986).
139. K. Bertelsen *et al.*, A randomized study of cyclophosphamide and cis-platinum with or without doxorubicin in advanced ovarian carcinoma. *Gynecologic oncology* **28**, 161-169 (1987).
140. M. A. Jordan, L. Wilson, Microtubules as a target for anticancer drugs. *Nature reviews. Cancer* **4**, 253-265 (2004).
141. J. Baselga, Targeting tyrosine kinases in cancer: the second wave. *Science* **312**, 1175-1178 (2006).
142. C. H. Yeon, M. D. Pegram, Anti-erbB-2 antibody trastuzumab in the treatment of HER2-amplified breast cancer. *Investigational new drugs* **23**, 391-409 (2005).
143. S. B. Kaye *et al.*, Phase II, open-label, randomized, multicenter study comparing the efficacy and safety of olaparib, a poly (ADP-ribose) polymerase inhibitor, and pegylated liposomal doxorubicin in patients with BRCA1 or BRCA2 mutations and recurrent ovarian cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **30**, 372-379 (2012).
144. G. Kim *et al.*, FDA Approval Summary: Olaparib Monotherapy in Patients with Deleterious Germline BRCA-Mutated Advanced Ovarian Cancer Treated with Three or More Lines of Chemotherapy. *Clinical cancer research : an official journal of the American Association for Cancer Research* **21**, 4257-4261 (2015).
145. R. L. Coleman, B. J. Monk, A. K. Sood, T. J. Herzog, Latest research and treatment of advanced-stage epithelial ovarian cancer. *Nature reviews. Clinical oncology* **10**, 211-224 (2013).
146. P. Bhosale *et al.*, Clinical utility of positron emission tomography/computed tomography in the evaluation of suspected recurrent ovarian cancer in the setting of normal CA-125 levels. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society* **20**, 936-944 (2010).
147. A. Prat *et al.*, Risk of recurrence during follow-up for optimally treated advanced epithelial ovarian cancer (EOC) with a low-level increase of serum CA-125 levels. *Annals of oncology : official journal of the European Society for Medical Oncology* **20**, 294-297 (2009).

148. E. A. White, H. A. Kenny, E. Lengyel, Three-dimensional modeling of ovarian cancer. *Advanced drug delivery reviews* **79-80**, 184-192 (2014).
149. A. A. Nemudryi, K. R. Valetdinova, S. P. Medvedev, S. M. Zakian, TALEN and CRISPR/Cas Genome Editing Systems: Tools of Discovery. *Acta naturae* **6**, 19-40 (2014).
150. A. J. Song, R. D. Palmiter, Detecting and Avoiding Problems When Using the Cre-lox System. *Trends Genet*, (2018).
151. P. Chang *et al.*, The Application of NHEJ-CRISPR/Cas9 and Cre-Lox System in the Generation of Bivalent Duck Enteritis Virus Vaccine against Avian Influenza Virus. *Viruses* **10**, (2018).
152. Z. Zou *et al.*, Efficient strategy for constructing duck enteritis virus-based live attenuated vaccine against homologous and heterologous H5N1 avian influenza virus and duck enteritis virus infection. *Vet Res* **46**, 42 (2015).
153. N. Sternberg, D. Hamilton, Bacteriophage P1 site-specific recombination. I. Recombination between loxP sites. *J Mol Biol* **150**, 467-486 (1981).
154. A. Richmond, Y. Su, Mouse xenograft models vs GEM models for human cancer therapeutics. *Disease models & mechanisms* **1**, 78-82 (2008).
155. N. Hasan, A. W. Ohman, D. M. Dinulescu, The promise and challenge of ovarian cancer models. *Translational cancer research* **4**, 14-28 (2015).
156. T. J. Shaw, M. K. Senterman, K. Dawson, C. A. Crane, B. C. Vanderhyden, Characterization of intraperitoneal, orthotopic, and metastatic xenograft models of human ovarian cancer. *Molecular therapy : the journal of the American Society of Gene Therapy* **10**, 1032-1042 (2004).
157. S. Orsulic *et al.*, Induction of ovarian cancer by defined multiple genetic changes in a mouse model system. *Cancer cell* **1**, 53-62 (2002).
158. X. Fu, R. M. Hoffman, Human ovarian carcinoma metastatic models constructed in nude mice by orthotopic transplantation of histologically-intact patient specimens. *Anticancer research* **13**, 283-286 (1993).
159. J. Li *et al.*, [Monitoring impact of measurement of urinary beta-human chorionic gonadotropin on orthotopic implantation model of human ovarian carcinoma in athymic nude mice]. *Ai zheng = Aizheng = Chinese journal of cancer* **21**, 949-953 (2002).
160. D. Amadori, E. Sansoni, A. Amadori, Ovarian cancer: natural history and metastatic pattern. *Frontiers in bioscience : a journal and virtual library* **2**, g8-10 (1997).
161. A. Feki *et al.*, Dissemination of intraperitoneal ovarian cancer: Discussion of mechanisms and demonstration of lymphatic spreading in ovarian cancer model. *Critical reviews in oncology/hematology* **72**, 1-9 (2009).
162. V. M. Howell, Genetically engineered mouse models for epithelial ovarian cancer: are we there yet? *Seminars in cell & developmental biology* **27**, 106-117 (2014).
163. K. Garson, T. J. Shaw, K. V. Clark, D. S. Yao, B. C. Vanderhyden, Models of ovarian cancer--are we there yet? *Molecular and cellular endocrinology* **239**, 15-26 (2005).
164. D. L. Meany, D. W. Chan, Aberrant glycosylation associated with enzymes as cancer biomarkers. *Clinical proteomics* **8**, 7 (2011).
165. R. G. Spiro, Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **12**, 43R-56R (2002).
166. A. Varki, Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **3**, 97-130 (1993).
167. S. A. Brooks *et al.*, Altered glycosylation of proteins in cancer: what is the potential for new anti-tumour strategies. *Anti-cancer agents in medicinal chemistry* **8**, 2-21 (2008).

168. R. Apweiler, H. Hermjakob, N. Sharon, On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochimica et biophysica acta* **1473**, 4-8 (1999).
169. A. Furmanek, J. Hofsteenge, Protein C-mannosylation: facts and questions. *Acta biochimica Polonica* **47**, 781-789 (2000).
170. F. Schwarz, M. Aebi, Mechanisms and principles of N-linked protein glycosylation. *Current opinion in structural biology* **21**, 576-582 (2011).
171. B. Schenk, F. Fernandez, C. J. Waechter, The ins(ide) and out(side) of dolichyl phosphate biosynthesis and recycling in the endoplasmic reticulum. *Glycobiology* **11**, 61R-70R (2001).
172. P. Hu, S. Shimoji, G. W. Hart, Site-specific interplay between O-GlcNAcylation and phosphorylation in cellular regulation. *FEBS letters* **584**, 2526-2538 (2010).
173. P. Stanley, Golgi glycosylation. *Cold Spring Harbor perspectives in biology* **3**, (2011).
174. A. Halim *et al.*, Site-specific characterization of threonine, serine, and tyrosine glycosylations of amyloid precursor protein/amyloid beta-peptides in human cerebrospinal fluid. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 11848-11853 (2011).
175. K. G. Ten Hagen, T. A. Fritz, L. A. Tabak, All in the family: the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases. *Glycobiology* **13**, 1R-16R (2003).
176. H. C. Hang, C. R. Bertozzi, The chemistry and biology of mucin-type O-linked glycosylation. *Bioorganic & medicinal chemistry* **13**, 5021-5034 (2005).
177. C. B. Madsen *et al.*, Cancer associated aberrant protein O-glycosylation can modify antigen processing and immune response. *PloS one* **7**, e50139 (2012).
178. P. Radhakrishnan *et al.*, Immature truncated O-glycophenotype of cancer directly induces oncogenic features. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E4066-4075 (2014).
179. J. Munkley, I. G. Mills, D. J. Elliott, The role of glycans in the development and progression of prostate cancer. *Nature reviews. Urology* **13**, 324-333 (2016).
180. I. Brockhausen, H. Schachter, P. Stanley, in *Essentials of Glycobiology,* A. Varki *et al.*, Eds. (Cold Spring Harbor (NY), 2009).
181. I. Brockhausen, Pathways of O-glycan biosynthesis in cancer cells. *Biochimica et biophysica acta* **1473**, 67-95 (1999).
182. I. O. Potapenko *et al.*, Glycan gene expression signatures in normal and malignant breast tissue; possible role in diagnosis and progression. *Molecular oncology* **4**, 98-118 (2010).
183. J. C. Yeh, E. Ong, M. Fukuda, Molecular cloning and expression of a novel beta-1, 6-N-acetylglucosaminyltransferase that forms core 2, core 4, and I branches. *The Journal of biological chemistry* **274**, 3215-3221 (1999).
184. in *Essentials of Glycobiology,* A. Varki *et al.*, Eds. (Cold Spring Harbor (NY), 2009).
185. A. Vasconcelos-Dos-Santos *et al.*, Biosynthetic Machinery Involved in Aberrant Glycosylation: Promising Targets for Developing of Drugs Against Cancer. *Frontiers in oncology* **5**, 138 (2015).
186. S. M. Muthana, C. T. Campbell, J. C. Gildersleeve, Modifications of glycans: biological significance and therapeutic opportunities. *ACS chemical biology* **7**, 31-43 (2012).
187. M. M. Fuster, J. D. Esko, The sweet and sour of cancer: glycans as novel therapeutic targets. *Nature reviews. Cancer* **5**, 526-542 (2005).
188. L. C. de Mattos, Structural diversity and biological importance of ABO, H, Lewis and secretor histo-blood group carbohydrates. *Revista brasileira de hematologia e hemoterapia* **38**, 331-340 (2016).

189. V. Ravn, E. Dabelsteen, Tissue distribution of histo-blood group antigens. *APMIS : acta pathologica, microbiologica, et immunologica Scandinavica* **108**, 1-28 (2000).
190. J. Le Pendu *et al.*, ABH and Lewis histo-blood group antigens in cancer. *APMIS : acta pathologica, microbiologica, et immunologica Scandinavica* **109**, 9-31 (2001).
191. S. Hakomori, Antigen structure and genetic basis of histo-blood groups A, B and O: their changes associated with human cancer. *Biochimica et biophysica acta* **1473**, 247-266 (1999).
192. P. Leon-Atance *et al.*, Prognostic influence of loss of blood group A antigen expression in pathologic stage I non-small-cell lung cancer. *Archivos de bronconeumologia* **48**, 49-54 (2012).
193. K. Fukushima, T. Satoh, S. Baba, K. Yamashita, alpha1,2-Fucosylated and beta-N-acetylgalactosaminylated prostate-specific antigen as an efficient marker of prostatic cancer. *Glycobiology* **20**, 452-460 (2010).
194. V. Vitiazeva *et al.*, The O-Linked Glycome and Blood Group Antigens ABO on Mucin-Type Glycoproteins in Mucinous and Serous Epithelial Ovarian Tumors. *PloS one* **10**, e0130197 (2015).
195. R. Mollicone *et al.*, Five specificity patterns of (1----3)-alpha-L-fucosyltransferase activity defined by use of synthetic oligosaccharide acceptors. Differential expression of the enzymes during human embryonic development and in adult tissues. *Carbohydrate research* **228**, 265-276 (1992).
196. L. Yan *et al.*, Lewis (y) antigen overexpression increases the expression of MMP-2 and MMP-9 and invasion of human ovarian cancer cells. *International journal of molecular sciences* **11**, 4441-4452 (2010).
197. M. Isla Larrain *et al.*, Breast cancer humoral immune response: involvement of Lewis y through the detection of circulating immune complexes and association with Mucin 1 (MUC1). *Journal of experimental & clinical cancer research : CR* **28**, 121 (2009).
198. J. A. Westwood *et al.*, The Lewis-Y carbohydrate antigen is expressed by many human tumors and can serve as a target for genetically redirected T cells despite the presence of soluble antigen in serum. *Journal of immunotherapy* **32**, 292-301 (2009).
199. P. J. Sabbatini *et al.*, Immunization of ovarian cancer patients with a synthetic Lewis(y)-protein conjugate vaccine: a phase 1 trial. *International journal of cancer* **87**, 79-85 (2000).
200. P. J. Sabbatini *et al.*, Pilot study of a heptavalent vaccine-keyhole limpet hemocyanin conjugate plus QS21 in patients with epithelial ovarian, fallopian tube, or peritoneal cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **13**, 4170-4177 (2007).
201. M. N. Saleh *et al.*, Phase I trial of the anti-Lewis Y drug immunoconjugate BR96-doxorubicin in patients with lewis Y-expressing epithelial tumors. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **18**, 2282-2292 (2000).
202. Y. J. Kim, A. Varki, Perspectives on the significance of altered glycosylation of glycoproteins in cancer. *Glycoconjugate journal* **14**, 569-576 (1997).
203. T. Nakayama, M. Watanabe, T. Katsumata, T. Teramoto, M. Kitajima, Expression of sialyl Lewis(a) as a new prognostic factor for patients with advanced colorectal carcinoma. *Cancer* **75**, 2051-2056 (1995).
204. K. Iwai *et al.*, Importance of E-selectin (ELAM-1) and sialyl Lewis(a) in the adhesion of pancreatic carcinoma cells to activated endothelium. *International journal of cancer* **54**, 972-977 (1993).

205. J. Ogawa, T. Tsurumi, S. Yamada, S. Koide, A. Shohtsu, Blood vessel invasion and expression of sialyl Lewisx and proliferating cell nuclear antigen in stage I non-small cell lung cancer. Relation to postoperative recurrence. *Cancer* **73**, 1177-1183 (1994).
206. T. Jorgensen *et al.*, Up-regulation of the oligosaccharide sialyl LewisX: a new prognostic parameter in metastatic prostate cancer. *Cancer research* **55**, 1817-1819 (1995).
207. J. Renkonen, T. Paavonen, R. Renkonen, Endothelial and epithelial expression of sialyl Lewis(x) and sialyl Lewis(a) in lesions of breast carcinoma. *International journal of cancer* **74**, 296-300 (1997).
208. R. Kannagi, Carbohydrate antigen sialyl Lewis a--its pathophysiological significance and induction mechanism in cancer progression. *Chang Gung medical journal* **30**, 189-209 (2007).
209. A. Harduin-Lepers, M. A. Recchi, P. Delannoy, 1994, the year of sialyltransferases. *Glycobiology* **5**, 741-758 (1995).
210. J. Chia, G. Goh, F. Bard, Short O-GalNAc glycans: regulation and role in tumor development and clinical perspectives. *Biochimica et biophysica acta* **1860**, 1623-1639 (2016).
211. S. Chugh *et al.*, Pathobiological implications of mucin glycans in cancer: Sweet poison and novel targets. *Biochimica et biophysica acta* **1856**, 211-225 (2015).
212. S. Julien, P. A. Videira, P. Delannoy, Sialyl-tn in cancer: (how) did we miss the target? *Biomolecules* **2**, 435-466 (2012).
213. H. Kobayashi, T. Terao, Y. Kawashima, Sialyl Tn as a prognostic marker in epithelial ovarian cancer. *British journal of cancer* **66**, 984-985 (1992).
214. H. Kobayashi, T. Terao, Y. Kawashima, Serum sialyl Tn as an independent predictor of poor prognosis in patients with epithelial ovarian cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **10**, 95-101 (1992).
215. H. Kobayashi, T. Terao, Y. Kawashima, [Circulating serum STN antigen as a prognostic marker in epithelial ovarian cancer]. *Gan to kagaku ryoho. Cancer & chemotherapy* **18**, 1651-1655 (1991).
216. R. Soares, A. Marinho, F. Schmitt, Expression of sialyl-Tn in breast cancer. Correlation with prognostic parameters. *Pathology, research and practice* **192**, 1181-1186 (1996).
217. J. Munkley, The Role of Sialyl-Tn in Cancer. *International journal of molecular sciences* **17**, 275 (2016).
218. K. Ishikawa, S. H. Medina, J. P. Schneider, A. J. Klar, Glycan Alteration Imparts Cellular Resistance to a Membrane-Lytic Anticancer Peptide. *Cell chemical biology* **24**, 149-158 (2017).
219. S. Ricardo *et al.*, Mucins MUC16 and MUC1 are major carriers of SLe(a) and SLe(x) in borderline and malignant serous ovarian tumors. *Virchows Archiv : an international journal of pathology* **468**, 715-722 (2016).
220. D. J. Moloney, A. I. Lin, R. S. Haltiwanger, The O-linked fucose glycosylation pathway. Evidence for protein-specific elongation of o-linked fucose in Chinese hamster ovary cells. *The Journal of biological chemistry* **272**, 19046-19050 (1997).
221. A. Sarrats *et al.*, Differential percentage of serum prostate-specific antigen subforms suggests a new way to improve prostate cancer diagnosis. *The Prostate* **70**, 1-9 (2010).
222. C. S. Timpte, A. E. Eckhardt, J. L. Abernethy, R. L. Hill, Porcine submaxillary gland apomucin contains tandemly repeated, identical sequences of 81 residues. *The Journal of biological chemistry* **263**, 1081-1088 (1988).
223. R. Gupta, N. Jentoft, Subunit structure of porcine submaxillary mucin. *Biochemistry* **28**, 6114-6121 (1989).

224. N. Moniaux, F. Escande, N. Porchet, J. P. Aubert, S. K. Batra, Structural organization and classification of the human mucin genes. *Frontiers in bioscience : a journal and virtual library* **6**, D1192-1206 (2001).
225. S. K. Linden, P. Sutton, N. G. Karlsson, V. Korolik, M. A. McGuckin, Mucins in the mucosal barrier to infection. *Mucosal immunology* **1**, 183-197 (2008).
226. W. Gevers, Mucus and mucins. *South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde* **72**, 39-42 (1987).
227. S. M. Kreda, C. W. Davis, M. C. Rose, CFTR, mucins, and mucus obstruction in cystic fibrosis. *Cold Spring Harbor perspectives in medicine* **2**, a009589 (2012).
228. T. D. Green, A. L. Crews, J. Park, S. Fang, K. B. Adler, Regulation of mucin secretion and inflammation in asthma: a role for MARCKS protein? *Biochimica et biophysica acta* **1810**, 1110-1113 (2011).
229. M. A. Hollingsworth, B. J. Swanson, Mucins in cancer: protection and control of the cell surface. *Nature reviews. Cancer* **4**, 45-60 (2004).
230. C. L. Hattrup, S. J. Gendler, Structure and function of the cell surface (tethered) mucins. *Annual review of physiology* **70**, 431-457 (2008).
231. K. M. McDermott *et al.*, Overexpression of MUC1 reconfigures the binding properties of tumor cells. *International journal of cancer* **94**, 783-791 (2001).
232. D. D. Carson, The cytoplasmic tail of MUC1: a very busy place. *Science signaling* **1**, pe35 (2008).
233. S. Senapati, S. Das, S. K. Batra, Mucin-interacting proteins: from function to therapeutics. *Trends in biochemical sciences* **35**, 236-245 (2010).
234. P. Thai, A. Loukoianov, S. Wachi, R. Wu, Regulation of airway mucin gene expression. *Annual review of physiology* **70**, 405-429 (2008).
235. M. Andrianifahanana, N. Moniaux, S. K. Batra, Regulation of mucin expression: mechanistic aspects and implications for cancer and inflammatory diseases. *Biochimica et biophysica acta* **1765**, 189-222 (2006).
236. D. J. Thornton, K. Rousseau, M. A. McGuckin, Structure and function of the polymeric mucins in airways mucus. *Annual review of physiology* **70**, 459-486 (2008).
237. D. W. Kufe, Mucins in cancer: function, prognosis and therapy. *Nature reviews. Cancer* **9**, 874-885 (2009).
238. V. Apostolopoulos, G. A. Pietersz, I. F. McKenzie, MUC1 and breast cancer. *Current opinion in molecular therapeutics* **1**, 98-103 (1999).
239. J. M. Hanson *et al.*, MUC1 expression in primary breast cancer: the effect of tamoxifen treatment. *Breast cancer research and treatment* **67**, 215-222 (2001).
240. J. Z. Zaretsky *et al.*, MUC1 gene overexpressed in breast cancer: structure and transcriptional activity of the MUC1 promoter and role of estrogen receptor alpha (ERalpha) in regulation of the MUC1 gene expression. *Molecular cancer* **5**, 57 (2006).
241. R. Beatson *et al.*, The Breast Cancer-Associated Glycoforms of MUC1, MUC1-Tn and sialyl-Tn, Are Expressed in COSMC Wild-Type Cells and Bind the C-Type Lectin MGL. *PloS one* **10**, e0125994 (2015).
242. H. Rajabi *et al.*, MUCIN 1 ONCOPROTEIN EXPRESSION IS SUPPRESSED BY THE miR-125b ONCOMIR. *Genes & cancer* **1**, 62-68 (2010).
243. P. Mukhopadhyay *et al.*, Mucins in the pathogenesis of breast cancer: implications in diagnosis, prognosis and therapy. *Biochimica et biophysica acta* **1815**, 224-240 (2011).
244. D. W. Kufe, MUC1-C oncoprotein as a target in breast cancer: activation of signaling pathways and therapeutic approaches. *Oncogene* **32**, 1073-1081 (2013).
245. J. Deng *et al.*, The role of tumour-associated MUC1 in epithelial ovarian cancer metastasis and progression. *Cancer metastasis reviews* **32**, 535-551 (2013).

246. Y. Hinoda *et al.*, Increased expression of MUC1 in advanced pancreatic cancer. *Journal of gastroenterology* **38**, 1162-1166 (2003).
247. R. Q. Wang, D. C. Fang, Alterations of MUC1 and MUC3 expression in gastric carcinoma: relevance to patient clinicopathological features. *Journal of clinical pathology* **56**, 378-384 (2003).
248. N. Scholler, N. Urban, CA125 in ovarian cancer. *Biomarkers in medicine* **1**, 513-523 (2007).
249. O. Dorigo, J. S. Berek, Personalizing CA125 levels for ovarian cancer screening. *Cancer Prev Res (Phila)* **4**, 1356-1359 (2011).
250. J. A. Gubbels *et al.*, MUC16 provides immune protection by inhibiting synapse formation between NK and ovarian tumor cells. *Molecular cancer* **9**, 11 (2010).
251. C. Theriault *et al.*, MUC16 (CA125) regulates epithelial ovarian cancer cell growth, tumorigenesis and metastasis. *Gynecologic oncology* **121**, 434-443 (2011).
252. Z. N. Dolgun *et al.*, The Use of Human Epididymis 4 and Cancer Antigen 125 Tumor Markers in the Benign or Malignant Differential Diagnosis of Pelvic or Adnexal Masses. *Balkan medical journal* **34**, 156-162 (2017).
253. A. Remo *et al.*, Clear cell colorectal carcinoma: Time to clarify diagnosis. *Pathology, research and practice*, (2017).
254. N. Abdalla *et al.*, Analysis of serum level of HE4 and CA125 considering selected risk factors among patients with endometrioid endometrial cancer. *Contemp Oncol (Pozn)* **20**, 463-467 (2016).
255. X. Ma *et al.*, Study of lung cancer regulatory network that involves erbB4 and tumor marker gene. *Saudi journal of biological sciences* **24**, 649-657 (2017).
256. D. Haridas *et al.*, Pathobiological implications of MUC16 expression in pancreatic cancer. *PloS one* **6**, e26839 (2011).
257. S. H. Chen, W. C. Hung, P. Wang, C. Paul, K. Konstantopoulos, Mesothelin binding to CA125/MUC16 promotes pancreatic cancer cell motility and invasion via MMP-7 activation. *Scientific reports* **3**, 1870 (2013).
258. I. Lakshmanan *et al.*, MUC16 induced rapid G2/M transition via interactions with JAK2 for increased proliferation and anti-apoptosis in breast cancer cells. *Oncogene* **31**, 805-817 (2012).
259. D. M. Maher, B. K. Gupta, S. Nagata, M. Jaggi, S. C. Chauhan, Mucin 13: structure, function, and potential roles in cancer pathogenesis. *Molecular cancer research : MCR* **9**, 531-537 (2011).
260. S. Khan *et al.*, MUC13 interaction with receptor tyrosine kinase HER2 drives pancreatic ductal adenocarcinoma progression. *Oncogene* **36**, 491-500 (2017).
261. Y. Nishii *et al.*, A newly developed anti-Mucin 13 monoclonal antibody targets pancreatic ductal adenocarcinoma cells. *International journal of oncology* **46**, 1781-1787 (2015).
262. P. Xia *et al.*, Cell membrane-anchored MUC4 promotes tumorigenicity in epithelial carcinomas. *Oncotarget* **8**, 14147-14157 (2017).
263. J. Xuan *et al.*, The diagnostic performance of serum MUC5AC for cholangiocarcinoma: A systematic review and meta-analysis. *Medicine* **95**, e3513 (2016).
264. J. Y. Jeong *et al.*, IL-1ra Secreted by ATP-Induced P2Y2 Negatively Regulates MUC5AC Overproduction via PLCbeta3 during Airway Inflammation. *Mediators of inflammation* **2016**, 7984853 (2016).
265. M. Khaidakov *et al.*, Gastric Proteins MUC5AC and TFF1 as Potential Diagnostic Markers of Colonic Sessile Serrated Adenomas/Polyps. *American journal of clinical pathology* **146**, 530-537 (2016).

266. H. Zhang *et al.*, High mucin 5AC expression predicts adverse postoperative recurrence and survival of patients with clear-cell renal cell carcinoma. *Oncotarget*, (2017).
267. M. Guo *et al.*, Gene signature driving invasive mucinous adenocarcinoma of the lung. *EMBO molecular medicine* **9**, 462-481 (2017).
268. H. Hoshi *et al.*, Tumor-associated MUC5AC stimulates in vivo tumorigenicity of human pancreatic cancer. *International journal of oncology* **38**, 619-627 (2011).
269. X. Zhu *et al.*, Abrogation of MUC5AC Expression Contributes to the Apoptosis and Cell Cycle Arrest of Colon Cancer Cells. *Cancer biotherapy & radiopharmaceuticals* **31**, 261-267 (2016).
270. J. Yang *et al.*, The Pulmonary Fibrosis Associated MUC5B Promoter Polymorphism Is Prognostic of the Overall Survival in Patients with Non-Small Cell Lung Cancer (NSCLC) Receiving Definitive Radiotherapy. *Translational oncology* **10**, 197-202 (2017).
271. J. Lee, J. H. Yun, D. G. Jeong, J. H. Kim, DUSP28 links regulation of Mucin 5B and Mucin 16 to migration and survival of AsPC-1 human pancreatic cancer cells. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine* **37**, 12193-12202 (2016).
272. E. P. Garcia *et al.*, MUC5B silencing reduces chemo-resistance of MCF-7 breast tumor cells and impairs maturation of dendritic cells. *International journal of oncology* **48**, 2113-2123 (2016).
273. R. Nagashio *et al.*, Diagnostic and prognostic significances of MUC5B and TTF-1 expressions in resected non-small cell lung cancer. *Scientific reports* **5**, 8649 (2015).
274. M. D. Walsh *et al.*, Expression of MUC2, MUC5AC, MUC5B, and MUC6 mucins in colorectal cancers and their association with the CpG island methylator phenotype. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* **26**, 1642-1656 (2013).
275. F. Pantano *et al.*, MUC2 but not MUC5 expression correlates with prognosis in radically resected pancreatic cancer patients. *Cancer biology & therapy* **8**, 996-999 (2009).
276. H. Kawashima, Roles of the gel-forming MUC2 mucin and its O-glycosylation in the protection against colitis and colorectal cancer. *Biological & pharmaceutical bulletin* **35**, 1637-1641 (2012).
277. T. Utsunomiya *et al.*, Expression of MUC1 and MUC2 mucins in gastric carcinomas: its relationship with the prognosis of the patients. *Clinical cancer research : an official journal of the American Association for Cancer Research* **4**, 2605-2614 (1998).
278. S. S. Pinho, C. A. Reis, Glycosylation in cancer: mechanisms and clinical implications. *Nature reviews. Cancer* **15**, 540-555 (2015).
279. T. M. Horm, J. A. Schroeder, MUC1 and metastatic cancer: expression, function and therapeutic targeting. *Cell adhesion & migration* **7**, 187-198 (2013).
280. J. Wang *et al.*, Mucin1 promotes the migration and invasion of hepatocellular carcinoma cells via JNK-mediated phosphorylation of Smad2 at the C-terminal and linker regions. *Oncotarget* **6**, 19264-19278 (2015).
281. P. Ciborowski, O. J. Finn, Non-glycosylated tandem repeats of MUC1 facilitate attachment of breast tumor cells to normal human lung tissue and immobilized extracellular matrix proteins (ECM) in vitro: potential role in metastasis. *Clinical & experimental metastasis* **19**, 339-345 (2002).
282. Q. Shen *et al.*, MUC1 initiates Src-CrkL-Rac1/Cdc42-mediated actin cytoskeletal protrusive motility after ligating intercellular adhesion molecule-1. *Molecular cancer research : MCR* **6**, 555-567 (2008).
283. A. Aleshin, R. S. Finn, SRC: a century of science brought to the clinic. *Neoplasia* **12**, 599-607 (2010).

284. C. B. Madsen *et al.*, Glycan elongation beyond the mucin associated Tn antigen protects tumor cells from immune-mediated killing. *PloS one* **8**, e72413 (2013).
285. Y. Wang *et al.*, Cosmc is an essential chaperone for correct protein O-glycosylation. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 9228-9233 (2010).
286. D. Nath *et al.*, Macrophage-tumour cell interactions: identification of MUC1 on breast cancer cells as a potential counter-receptor for the macrophage-restricted receptor, sialoadhesin. *Immunology* **98**, 213-219 (1999).
287. W. J. Janssen, A. L. Stefanski, B. S. Bochner, C. M. Evans, Control of lung defence by mucins and macrophages: ancient defence mechanisms with modern functions. *The European respiratory journal* **48**, 1201-1214 (2016).
288. M. Inoue, H. Fujii, H. Kaseyama, I. Yamashina, H. Nakada, Stimulation of macrophages by mucins through a macrophage scavenger receptor. *Biochemical and biophysical research communications* **264**, 276-280 (1999).
289. Y. J. Kim, L. Borsig, H. L. Han, N. M. Varki, A. Varki, Distinct selectin ligands on colon carcinoma mucins can mediate pathological interactions among platelets, leukocytes, and endothelium. *The American journal of pathology* **155**, 461-472 (1999).
290. S. H. Chen, M. R. Dallas, E. M. Balzer, K. Konstantopoulos, Mucin 16 is a functional selectin ligand on pancreatic cancer cells. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **26**, 1349-1359 (2012).
291. J. A. Belisle *et al.*, Identification of Siglec-9 as the receptor for MUC16 on human NK cells, B cells, and monocytes. *Molecular cancer* **9**, 118 (2010).
292. D. H. Dube, C. R. Bertozzi, Glycans in cancer and inflammation--potential for therapeutics and diagnostics. *Nature reviews. Drug discovery* **4**, 477-488 (2005).
293. J. Renkonen, O. Tynninen, P. Hayry, T. Paavonen, R. Renkonen, Glycosylation might provide endothelial zip codes for organ-specific leukocyte traffic into inflammatory sites. *The American journal of pathology* **161**, 543-550 (2002).
294. Y. Dong *et al.*, Expression of MUC1 and MUC2 mucins in epithelial ovarian tumours. *The Journal of pathology* **183**, 311-317 (1997).
295. C. H. Van Elssen *et al.*, Expression of aberrantly glycosylated Mucin-1 in ovarian cancer. *Histopathology* **57**, 597-606 (2010).
296. S. Ricardo *et al.*, Detection of glyco-mucin profiles improves specificity of MUC16 and MUC1 biomarkers in ovarian serous tumours. *Molecular oncology* **9**, 503-512 (2015).
297. K. A. Williams *et al.*, Polymorphisms of MUC16 (CA125) and MUC1 (CA15.3) in relation to ovarian cancer risk and survival. *PloS one* **9**, e88334 (2014).
298. J. Wang, M. A. El-Bahrawy, Expression profile of mucins in ovarian mucinous tumors: distinguishing primary ovarian from metastatic tumors. *International journal of gynecological pathology : official journal of the International Society of Gynecological Pathologists* **33**, 166-175 (2014).
299. D. Shahbazi-Gahrouei, M. Abdolahi, Detection of MUC1-expressing ovarian cancer by C595 monoclonal antibody-conjugated SPIONs using MR imaging. *TheScientificWorldJournal* **2013**, 609151 (2013).
300. J. W. Pedersen *et al.*, Cancer-associated autoantibodies to MUC1 and MUC4--a blinded case-control study of colorectal cancer in UK collaborative trial of ovarian cancer screening. *International journal of cancer* **134**, 2180-2188 (2014).
301. K. Kannan, G. K. Kordestani, A. Galagoda, C. Coarfa, L. Yen, Aberrant MUC1-TRIM46-KRTCAP2 Chimeric RNAs in High-Grade Serous Ovarian Carcinoma. *Cancers* **7**, 2083-2093 (2015).

302. R. Falahat *et al.*, A Cell ELISA for the quantification of MUC1 mucin (CD227) expressed by cancer cells of epithelial and neuroectodermal origin. *Cellular immunology* **298**, 96-103 (2015).
303. S. Cascio, O. J. Finn, Intra- and Extra-Cellular Events Related to Altered Glycosylation of MUC1 Promote Chronic Inflammation, Tumor Progression, Invasion, and Metastasis. *Biomolecules* **6**, (2016).
304. T. Ju, R. D. Cummings, 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**, 16613-16618 (2002).
305. Z. Q. Wang *et al.*, Role of the polypeptide N-acetylgalactosaminyltransferase 3 in ovarian cancer progression: possible implications in abnormal mucin O-glycosylation. *Oncotarget* **5**, 544-560 (2014).
306. W. Fiedler *et al.*, A phase I study of PankoMab-GEX, a humanised glyco-optimised monoclonal antibody to a novel tumour-specific MUC1 glycopeptide epitope in patients with advanced carcinomas. *Eur J Cancer* **63**, 55-63 (2016).
307. B. N. Vajaria, P. S. Patel, Glycosylation: a hallmark of cancer? *Glycoconjugate journal* **34**, 147-156 (2017).
308. J. Munkley, D. J. Elliott, Hallmarks of glycosylation in cancer. *Oncotarget* **7**, 35478-35489 (2016).
309. G. F. Springer, T and Tn, general carcinoma autoantigens. *Science* **224**, 1198-1206 (1984).
310. S. Itzkowitz *et al.*, Expression of Tn, sialosyl Tn, and T antigens in human pancreas. *Gastroenterology* **100**, 1691-1700 (1991).
311. T. Ju *et al.*, Tn and sialyl-Tn antigens, aberrant O-glycomics as human disease markers. *Proteomics. Clinical applications* **7**, 618-631 (2013).
312. T. Ju *et al.*, Human tumor antigens Tn and sialyl Tn arise from mutations in Cosmc. *Cancer research* **68**, 1636-1646 (2008).
313. R. Sewell *et al.*, 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**, 3586-3594 (2006).
314. S. Cotton *et al.*, Targeted O-glycoproteomics explored increased sialylation and identified MUC16 as a poor prognosis biomarker in advanced-stage bladder tumours. *Molecular oncology*, (2017).
315. S. Hakomori, Glycosylation defining cancer malignancy: new wine in an old bottle. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 10231-10233 (2002).
316. E. P. Bennett *et al.*, Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. *Glycobiology* **22**, 736-756 (2012).
317. T. Lubke *et al.*, Complementation cloning identifies CDG-IIc, a new type of congenital disorders of glycosylation, as a GDP-fucose transporter deficiency. *Nature genetics* **28**, 73-76 (2001).
318. M. Schwarzkopf *et al.*, Sialylation is essential for early development in mice. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 5267-5270 (2002).
319. K. Kumamoto *et al.*, Increased expression of UDP-galactose transporter messenger RNA in human colon cancer tissues and its implication in synthesis of Thomsen-Friedenreich antigen and sialyl Lewis A/X determinants. *Cancer research* **61**, 4620-4627 (2001).
320. B. Hadley *et al.*, Structure and function of nucleotide sugar transporters: Current progress. *Computational and structural biotechnology journal* **10**, 23-32 (2014).

321. S. Goto *et al.*, UDP-sugar transporter implicated in glycosylation and processing of Notch. *Nature cell biology* **3**, 816-822 (2001).
322. S. Kellokumpu, R. Sormunen, I. Kellokumpu, Abnormal glycosylation and altered Golgi structure in colorectal cancer: dependence on intra-Golgi pH. *FEBS letters* **516**, 217-224 (2002).
323. M. Udono *et al.*, Impaired ATP6V0A2 expression contributes to Golgi dispersion and glycosylation changes in senescent cells. *Scientific reports* **5**, 17342 (2015).
324. C. Rosnoblet, R. Peanne, D. Legrand, F. Foulquier, Glycosylation disorders of membrane trafficking. *Glycoconjugate journal* **30**, 23-31 (2013).
325. J. Raman, Y. Guan, C. L. Perrine, T. A. Gerken, L. A. Tabak, Erratum: UDP-N-acetyl-alpha-D-galactosamine: polypeptide N-acetylgalactosaminyltransferases: completion of the family tree. *Glycobiology* **25**, 465 (2015).
326. M. R. Pratt *et al.*, Deconvoluting the functions of polypeptide N-alpha-acetylgalactosaminyltransferase family members by glycopeptide substrate profiling. *Chemistry & biology* **11**, 1009-1016 (2004).
327. D. Kapitonov, R. K. Yu, Conserved domains of glycosyltransferases. *Glycobiology* **9**, 961-978 (1999).
328. A. Imberty, V. Piller, F. Piller, C. Breton, Fold recognition and molecular modeling of a lectin-like domain in UDP-GalNac:polypeptide N-acetylgalactosaminyltransferases. *Protein engineering* **10**, 1353-1356 (1997).
329. P. D. Kingsley, K. G. Hagen, K. M. Maltby, J. Zara, L. A. Tabak, Diverse spatial expression patterns of UDP-GalNAc:polypeptide N-acetylgalactosaminyl-transferase family member mRNAs during mouse development. *Glycobiology* **10**, 1317-1323 (2000).
330. G. Talavera, J. Castresana, Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol* **56**, 564-577 (2007).
331. A. Dell, A. Galadari, F. Sastre, P. Hitchen, Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes. *International journal of microbiology* **2010**, 148178 (2010).
332. J. E. Hansen *et al.*, NetOglyc: prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility. *Glycoconjugate journal* **15**, 115-130 (1998).
333. L. Kirnarsky *et al.*, Structural analysis of peptide substrates for mucin-type O-glycosylation. *Biochemistry* **37**, 12811-12817 (1998).
334. T. A. Gerken, C. L. Owens, M. Pasumarthy, Determination of the site-specific O-glycosylation pattern of the porcine submaxillary mucin tandem repeat glycopeptide. Model proposed for the polypeptide:galnac transferase peptide binding site. *The Journal of biological chemistry* **272**, 9709-9719 (1997).
335. F. G. Hanisch *et al.*, Dynamic epigenetic regulation of initial O-glycosylation by UDP-N-Acetylgalactosamine:Peptide N-acetylgalactosaminyltransferases. site-specific glycosylation of MUC1 repeat peptide influences the substrate qualities at adjacent or distant Ser/Thr positions. *The Journal of biological chemistry* **274**, 9946-9954 (1999).
336. T. A. Gerken, C. Tep, J. Rarick, Role of peptide sequence and neighboring residue glycosylation on the substrate specificity of the uridine 5'-diphosphate-alpha-N-acetylgalactosamine:polypeptide N-acetylgalactosaminyl transferases T1 and T2: kinetic modeling of the porcine and canine submaxillary gland mucin tandem repeats. *Biochemistry* **43**, 9888-9900 (2004).
337. K. G. Ten Hagen *et al.*, Characterization of a UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase that displays glycopeptide N-

acetylgalactosaminyltransferase activity. *The Journal of biological chemistry* **274**, 27867-27874 (1999).
338. K. G. Ten Hagen *et al.*, Cloning and characterization of a ninth member of the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase family, ppGaNTase-T9. *The Journal of biological chemistry* **276**, 17395-17404 (2001).
339. C. Peng *et al.*, Identification of a novel human UDP-GalNAc transferase with unique catalytic activity and expression profile. *Biochemical and biophysical research communications* **402**, 680-686 (2010).
340. J. Raman, Y. Guan, C. L. Perrine, T. A. Gerken, L. A. Tabak, UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferases: completion of the family tree. *Glycobiology* **22**, 768-777 (2012).
341. D. K. Banfield, Mechanisms of protein retention in the Golgi. *Cold Spring Harbor perspectives in biology* **3**, a005264 (2011).
342. S. Rottger *et al.*, Localization of three human polypeptide GalNAc-transferases in HeLa cells suggests initiation of O-linked glycosylation throughout the Golgi apparatus. *Journal of cell science* **111 (Pt 1)**, 45-60 (1998).
343. J. W. Pedersen *et al.*, Lectin domains of polypeptide GalNAc transferases exhibit glycopeptide binding specificity. *The Journal of biological chemistry* **286**, 32684-32696 (2011).
344. H. H. Wandall *et al.*, 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**, 374-387 (2007).
345. M. R. Hussain, D. C. Hoessli, M. Fang, N-acetylgalactosaminyltransferases in cancer. *Oncotarget* **7**, 54067-54081 (2016).
346. W. L. Ho, W. M. Hsu, M. C. Huang, K. Kadomatsu, A. Nakagawara, Protein glycosylation in cancers and its potential therapeutic applications in neuroblastoma. *Journal of hematology & oncology* **9**, 100 (2016).
347. T. Hennet, F. K. Hagen, L. A. Tabak, J. D. Marth, 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**, 12070-12074 (1995).
348. M. Tenno *et al.*, Initiation of protein O glycosylation by the polypeptide GalNAcT-1 in vascular biology and humoral immunity. *Molecular and cellular biology* **27**, 8783-8796 (2007).
349. L. A. Tabak, The role of mucin-type O-glycans in eukaryotic development. *Seminars in cell & developmental biology* **21**, 616-621 (2010).
350. A. E. Manzi *et al.*, Exploring the glycan repertoire of genetically modified mice by isolation and profiling of the major glycan classes and nano-NMR analysis of glycan mixtures. *Glycobiology* **10**, 669-689 (2000).
351. J. Brookfield, Can genes be truly redundant? *Current biology : CB* **2**, 553-554 (1992).
352. H. Pearson, Surviving a knockout blow. *Nature* **415**, 8-9 (2002).
353. M. Gentzsch, W. Tanner, Protein-O-glycosylation in yeast: protein-specific mannosyltransferases. *Glycobiology* **7**, 481-486 (1997).
354. C. M. Woo, A. T. Iavarone, D. R. Spiciarich, K. K. Palaniappan, C. R. Bertozzi, Isotope-targeted glycoproteomics (IsoTaG): a mass-independent platform for intact N- and O-glycopeptide discovery and analysis. *Nature methods* **12**, 561-567 (2015).
355. R. Raman, S. Raguram, G. Venkataraman, J. C. Paulson, R. Sasisekharan, Glycomics: an integrated systems approach to structure-function relationships of glycans. *Nature methods* **2**, 817-824 (2005).

356. Y. Yang, G. Wang, T. Song, C. B. Lebrilla, A. J. R. Heck, Resolving the micro-heterogeneity and structural integrity of monoclonal antibodies by hybrid mass spectrometric approaches. *mAbs* **9**, 638-645 (2017).
357. G. Alvarez-Manilla *et al.*, Tools for glycoproteomic analysis: size exclusion chromatography facilitates identification of tryptic glycopeptides with N-linked glycosylation sites. *Journal of proteome research* **5**, 701-708 (2006).
358. B. A. Budnik, R. S. Lee, J. A. Steen, Global methods for protein glycosylation analysis by mass spectrometry. *Biochimica et biophysica acta* **1764**, 1870-1880 (2006).
359. Z. Roth, G. Yehezkel, I. Khalaila, Identification and Quantification of Protein Glycosylation. *International Journal of Carbohydrate Chemistry* **2012**, 10 (2012).
360. C. C. Chen *et al.*, Interaction modes and approaches to glycopeptide and glycoprotein enrichment. *The Analyst* **139**, 688-704 (2014).
361. H. Kaji, T. Isobe, Liquid Chromatography/Mass Spectrometry (LC/MS)-Based Glycoproteomics Technologies for Cancer Biomarker Discovery. *Clinical proteomics* **4**, 14-24 (2008).
362. H. Zhang, X. J. Li, D. B. Martin, R. Aebersold, Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nature biotechnology* **21**, 660-666 (2003).
363. B. Sun *et al.*, Shotgun glycopeptide capture approach coupled with mass spectrometry for comprehensive glycoproteomics. *Molecular & cellular proteomics : MCP* **6**, 141-149 (2007).
364. H. Kaji *et al.*, Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins. *Nature biotechnology* **21**, 667-672 (2003).
365. C. M. Woo, C. R. Bertozzi, Isotope Targeted Glycoproteomics (IsoTaG) to Characterize Intact, Metabolically Labeled Glycopeptides from Complex Proteomes. *Current protocols in chemical biology* **8**, 59-82 (2016).
366. C. M. Woo *et al.*, Development of IsoTaG, a Chemical Glycoproteomics Technique for Profiling Intact N- and O-Glycopeptides from Whole Cell Proteomes. *Journal of proteome research* **16**, 1706-1718 (2017).
367. Q. K. Li, E. Gabrielson, H. Zhang, Application of glycoproteomics for the discovery of biomarkers in lung cancer. *Proteomics. Clinical applications* **6**, 244-256 (2012).
368. M. V. Novotny, Y. Mechref, New hyphenated methodologies in high-sensitivity glycoprotein analysis. *Journal of separation science* **28**, 1956-1968 (2005).
369. H. J. An, T. R. Peavy, J. L. Hedrick, C. B. Lebrilla, Determination of N-glycosylation sites and site heterogeneity in glycoproteins. *Analytical chemistry* **75**, 5628-5637 (2003).
370. P. Hagglund, J. Bunkenborg, F. Elortza, O. N. Jensen, P. Roepstorff, A new strategy for identification of N-glycosylated proteins and unambiguous assignment of their glycosylation sites using HILIC enrichment and partial deglycosylation. *Journal of proteome research* **3**, 556-566 (2004).
371. Y. Wada, M. Tajiri, S. Yoshida, Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics. *Analytical chemistry* **76**, 6560-6565 (2004).
372. C. F. Teo *et al.*, Glycopeptide-specific monoclonal antibodies suggest new roles for O-GlcNAc. *Nature chemical biology* **6**, 338-343 (2010).
373. J. M. Prendergast *et al.*, Novel anti-Sialyl-Tn monoclonal antibodies and antibody-drug conjugates demonstrate tumor specificity and anti-tumor activity. *mAbs* **9**, 615-627 (2017).
374. H. Nakada *et al.*, Elucidation of an essential structure recognized by an anti-GalNAc alpha-Ser(Thr) monoclonal antibody (MLS 128). *The Journal of biological chemistry* **266**, 12402-12405 (1991).

375. Z. Wang, A. Pandey, G. W. Hart, Dynamic interplay between O-linked N-acetylglucosaminylation and glycogen synthase kinase-3-dependent phosphorylation. *Molecular & cellular proteomics : MCP* **6**, 1365-1379 (2007).
376. L. Kasturi, H. Chen, S. H. Shakin-Eshleman, Regulation of N-linked core glycosylation: use of a site-directed mutagenesis approach to identify Asn-Xaa-Ser/Thr sequons that are poor oligosaccharide acceptors. *The Biochemical journal* **323 (Pt 2)**, 415-419 (1997).
377. S. J. Qu, H. Z. Fan, F. Blanco-Vaca, H. J. Pownall, Effects of site-directed mutagenesis on the N-glycosylation sites of human lecithin:cholesterol acyltransferase. *Biochemistry* **32**, 8732-8736 (1993).
378. C. Steentoft *et al.*, Mining the O-glycoproteome using zinc-finger nuclease-glycoengineered SimpleCell lines. *Nature methods* **8**, 977-982 (2011).
379. E. Saxon, C. R. Bertozzi, Cell surface engineering by a modified Staudinger reaction. *Science* **287**, 2007-2010 (2000).
380. S. Mereiter, M. Balmana, J. Gomes, A. Magalhaes, C. A. Reis, Glycomic Approaches for the Discovery of Targets in Gastrointestinal Cancer. *Frontiers in oncology* **6**, 55 (2016).
381. J. A. Prescher, C. R. Bertozzi, Chemistry in living systems. *Nature chemical biology* **1**, 13-21 (2005).
382. M. Zheng, L. Zheng, P. Zhang, J. Li, Y. Zhang, Development of bioorthogonal reactions and their applications in bioconjugation. *Molecules* **20**, 3190-3205 (2015).
383. J. A. Prescher, D. H. Dube, C. R. Bertozzi, Chemical remodelling of cell surfaces in living animals. *Nature* **430**, 873-877 (2004).
384. C. S. McKay, M. G. Finn, Click chemistry in complex mixtures: bioorthogonal bioconjugation. *Chemistry & biology* **21**, 1075-1101 (2014).
385. S. T. Laughlin, C. R. Bertozzi, Imaging the glycome. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 12-17 (2009).
386. H. C. Hang, C. Yu, D. L. Kato, C. R. Bertozzi, A metabolic labeling approach toward proteomic analysis of mucin-type O-linked glycosylation. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 14846-14851 (2003).
387. V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angewandte Chemie* **41**, 2596-2599 (2002).
388. A. E. Speers, B. F. Cravatt, Chemical strategies for activity-based proteomics. *Chembiochem : a European journal of chemical biology* **5**, 41-47 (2004).
389. S. T. Laughlin, J. M. Baskin, S. L. Amacher, C. R. Bertozzi, In vivo imaging of membrane-associated glycans in developing zebrafish. *Science* **320**, 664-667 (2008).
390. Y. Mechref, M. V. Novotny, Mass spectrometric mapping and sequencing of N-linked oligosaccharides derived from submicrogram amounts of glycoproteins. *Analytical chemistry* **70**, 455-463 (1998).
391. C. Steentoft *et al.*, Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *The EMBO journal* **32**, 1478-1488 (2013).
392. M. Boyce *et al.*, Metabolic cross-talk allows labeling of O-linked beta-N-acetylglucosamine-modified proteins via the N-acetylgalactosamine salvage pathway. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 3141-3146 (2011).
393. N. Daude *et al.*, Molecular cloning, characterization, and mapping of a full-length cDNA encoding human UDP-galactose 4'-epimerase. *Biochemical and molecular medicine* **56**, 1-7 (1995).

394. D. J. Vocadlo, H. C. Hang, E. J. Kim, J. A. Hanover, C. R. Bertozzi, A chemical approach for identifying O-GlcNAc-modified proteins in cells. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 9116-9121 (2003).
395. M. Yang, J. Li, P. R. Chen, Transition metal-mediated bioorthogonal protein chemistry in living cells. *Chemical Society reviews* **43**, 6511-6526 (2014).
396. J. Dubuisson, C. M. Rice, Hepatitis C virus glycoprotein folding: disulfide bond formation and association with calnexin. *Journal of virology* **70**, 778-786 (1996).
397. J. V. Olsen, S. E. Ong, M. Mann, Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Molecular & cellular proteomics : MCP* **3**, 608-614 (2004).
398. K. K. Palaniappan *et al.*, Isotopic signature transfer and mass pattern prediction (IsoStamp): an enabling technique for chemically-directed proteomics. *ACS chemical biology* **6**, 829-836 (2011).
399. C. M. Woo, A. Felix, L. Zhang, J. E. Elias, C. R. Bertozzi, Isotope-targeted glycoproteomics (IsoTaG) analysis of sialylated N- and O-glycopeptides on an Orbitrap Fusion Tribrid using azido and alkynyl sugars. *Analytical and bioanalytical chemistry* **409**, 579-588 (2017).
400. R. Siegel, E. Ward, O. Brawley, A. Jemal, Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *CA: a cancer journal for clinicians* **61**, 212-236 (2011).
401. A. Varki, N. Sharon, in *Essentials of Glycobiology,* A. Varki *et al.*, Eds. (Cold Spring Harbor (NY), 2009).
402. A. M. Bode, Z. Dong, Post-translational modification of p53 in tumorigenesis. *Nature reviews. Cancer* **4**, 793-805 (2004).
403. T. M. Karve, A. K. Cheema, Small changes huge impact: the role of protein posttranslational modifications in cellular homeostasis and disease. *Journal of amino acids* **2011**, 207691 (2011).
404. K. M. Sheehan *et al.*, Use of reverse phase protein microarrays and reference standard development for molecular network analysis of metastatic ovarian carcinoma. *Molecular & cellular proteomics : MCP* **4**, 346-355 (2005).
405. M. X. Ding *et al.*, ppGalNAc T1 as a potential novel marker for human bladder cancer. *Asian Pacific journal of cancer prevention : APJCP* **13**, 5653-5657 (2012).
406. M. C. Lin, M. J. Huang, C. H. Liu, T. L. Yang, M. C. Huang, GALNT2 enhances migration and invasion of oral squamous cell carcinoma by regulating EGFR glycosylation and activity. *Oral oncology* **50**, 478-484 (2014).
407. K. Taniuchi *et al.*, Overexpression of GalNAc-transferase GalNAc-T3 promotes pancreatic cancer cell growth. *Oncogene* **30**, 4843-4854 (2011).
408. J. H. Park *et al.*, Critical roles of mucin 1 glycosylation by transactivated polypeptide N-acetylgalactosaminyltransferase 6 in mammary carcinogenesis. *Cancer research* **70**, 2759-2769 (2010).
409. J. Gomes *et al.*, Expression of UDP-N-acetyl-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase-6 in gastric mucosa, intestinal metaplasia, and gastric carcinoma. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society* **57**, 79-86 (2009).
410. T. C. Lin *et al.*, GALNT6 expression enhances aggressive phenotypes of ovarian cancer cells by regulating EGFR activity. *Oncotarget*, (2017).
411. R. Q. Peng *et al.*, MicroRNA-214 suppresses growth and invasiveness of cervical cancer cells by targeting UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 7. *The Journal of biological chemistry* **287**, 14301-14309 (2012).

412. A. Gaziel-Sovran *et al.*, miR-30b/30d regulation of GalNAc transferases enhances invasion and immunosuppression during metastasis. *Cancer cell* **20**, 104-118 (2011).
413. N. Berois *et al.*, GALNT9 gene expression is a prognostic marker in neuroblastoma patients. *Clinical chemistry* **59**, 225-233 (2013).
414. Y. Gao *et al.*, Expression pattern of polypeptide N-acetylgalactosaminyltransferase-10 in gastric carcinoma. *Oncology letters* **5**, 113-116 (2013).
415. M. G. Libisch *et al.*, GALNT11 as a new molecular marker in chronic lymphocytic leukemia. *Gene* **533**, 270-279 (2014).
416. K. Guda *et al.*, Inactivating germ-line and somatic mutations in polypeptide N-acetylgalactosaminyltransferase 12 in human colon cancers. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 12921-12925 (2009).
417. N. Berois *et al.*, ppGalNAc-T13: a new molecular marker of bone marrow involvement in neuroblastoma. *Clinical chemistry* **52**, 1701-1712 (2006).
418. R. Wang, C. Yu, D. Zhao, M. Wu, Z. Yang, The mucin-type glycosylating enzyme polypeptide N-acetylgalactosaminyltransferase 14 promotes the migration of ovarian cancer by modifying mucin 13. *Oncology reports* **30**, 667-676 (2013).
419. R. Singh, D. Bandyopadhyay, MUC1: a target molecule for cancer therapy. *Cancer biology & therapy* **6**, 481-486 (2007).
420. J. Taylor-Papadimitriou, J. Burchell, D. W. Miles, M. Dalziel, MUC1 and cancer. *Biochimica et biophysica acta* **1455**, 301-313 (1999).
421. E. M. Beaman, S. A. Brooks, The extended ppGalNAc-T family and their functional involvement in the metastatic cascade. *Histology and histopathology* **29**, 293-304 (2014).
422. E. P. Bennett *et al.*, Cloning and characterization of a close homologue of human UDP-N-acetyl-alpha-D-galactosamine:Polypeptide N-acetylgalactosaminyltransferase-T3, designated GalNAc-T6. Evidence for genetic but not functional redundancy. *The Journal of biological chemistry* **274**, 25362-25370 (1999).
423. R. Schiess, B. Wollscheid, R. Aebersold, Targeted proteomic strategy for clinical biomarker discovery. *Molecular oncology* **3**, 33-44 (2009).
424. Q. K. Li *et al.*, An integrated proteomic and glycoproteomic approach uncovers differences in glycosylation occupancy from benign and malignant epithelial ovarian tumors. *Clinical proteomics* **14**, 16 (2017).
425. Y. Ji *et al.*, Integrated proteomic and N-glycoproteomic analyses of doxorubicin sensitive and resistant ovarian cancer cells reveal glycoprotein alteration in protein abundance and glycosylation. *Oncotarget* **8**, 13413-13427 (2017).
426. S. Miyamoto *et al.*, Glycoproteomic Analysis of Malignant Ovarian Cancer Ascites Fluid Identifies Unusual Glycopeptides. *Journal of proteome research* **15**, 3358-3376 (2016).
427. R. Sheta *et al.*, Proteomic dataset for altered glycoprotein expression upon GALNT3 knockdown in ovarian cancer cells. *Data in brief* **8**, 342-349 (2016).
428. R. Sheta *et al.*, A metabolic labeling approach for glycoproteomic analysis reveals altered glycoprotein expression upon GALNT3 knockdown in ovarian cancer cells. *Journal of proteomics* **145**, 91-102 (2016).
429. J. Wu, X. Xie, S. Nie, R. J. Buckanovich, D. M. Lubman, Altered expression of sialylated glycoproteins in ovarian cancer sera using lectin-based ELISA assay and quantitative glycoproteomics analysis. *Journal of proteome research* **12**, 3342-3352 (2013).
430. J. Wu *et al.*, Identification and confirmation of differentially expressed fucosylated glycoproteins in the serum of ovarian cancer patients using a lectin array and LC-MS/MS. *Journal of proteome research* **11**, 4541-4552 (2012).

431. Y. Tian, Z. Yao, R. B. Roden, H. Zhang, Identification of glycoproteins associated with different histological subtypes of ovarian tumors using quantitative glycoproteomics. *Proteomics* **11**, 4677-4687 (2011).
432. M. Di Michele *et al.*, Glycoproteomics of paclitaxel resistance in human epithelial ovarian cancer cell lines: towards the identification of putative biomarkers. *Journal of proteomics* **73**, 879-898 (2010).
433. G. J. Kelloff *et al.*, Progress in chemoprevention drug development: the promise of molecular biomarkers for prevention of intraepithelial neoplasia and cancer--a plan to move forward. *Clinical cancer research : an official journal of the American Association for Cancer Research* **12**, 3661-3697 (2006).
434. N. Rauniyar *et al.*, Data-Independent Acquisition and Parallel Reaction Monitoring Mass Spectrometry Identification of Serum Biomarkers for Ovarian Cancer. *Biomarker insights* **12**, 1177271917710948 (2017).
435. J. Gallego, C. Lopez, R. Pazo-Cid, F. Lopez-Rios, A. Carrato, Biomarkers in pancreatic ductal adenocarcinoma. *Clinical & translational oncology : official publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico*, (2017).
436. S. P. Kar *et al.*, Common Genetic Variation and Susceptibility to Ovarian Cancer: Current Insights and Future Directions. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, (2017).
437. E. Boja *et al.*, Evolution of clinical proteomics and its role in medicine. *Journal of proteome research* **10**, 66-84 (2011).
438. M. Kobel *et al.*, Evidence for a time-dependent association between FOLR1 expression and survival from ovarian carcinoma: implications for clinical testing. An Ovarian Tumour Tissue Analysis consortium study. *British journal of cancer* **111**, 2297-2307 (2014).
439. S. Morone *et al.*, Overexpression of CD157 contributes to epithelial ovarian cancer progression by promoting mesenchymal differentiation. *PloS one* **7**, e43649 (2012).
440. A. V. Everest-Dass *et al.*, N-glycan MALDI Imaging Mass Spectrometry on Formalin-Fixed Paraffin-Embedded Tissue Enables the Delineation of Ovarian Cancer Tissues. *Molecular & cellular proteomics : MCP* **15**, 3003-3016 (2016).
441. U. Kuzmanov *et al.*, Glycoproteomic identification of potential glycoprotein biomarkers in ovarian cancer proximal fluids. *Clinical chemistry and laboratory medicine* **51**, 1467-1476 (2013).
442. K. L. Abbott *et al.*, Identification of candidate biomarkers with cancer-specific glycosylation in the tissue and serum of endometrioid ovarian cancer patients by glycoproteomic analysis. *Proteomics* **10**, 470-481 (2010).
443. V. Shetty, J. Hafner, P. Shah, Z. Nickens, R. Philip, Investigation of ovarian cancer associated sialylation changes in N-linked glycopeptides by quantitative proteomics. *Clinical proteomics* **9**, 10 (2012).
444. B. Li *et al.*, Glycoproteomic analyses of ovarian cancer cell lines and sera from ovarian cancer patients show distinct glycosylation changes in individual proteins. *Journal of proteome research* **7**, 3776-3788 (2008).
445. K. T. Schjoldager *et al.*, Probing isoform-specific functions of polypeptide GalNAc-transferases using zinc finger nuclease glycoengineered SimpleCells. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 9893-9898 (2012).
446. R. Nilsson *et al.*, Metabolic enzyme expression highlights a key role for MTHFD2 and the mitochondrial folate pathway in cancer. *Nature communications* **5**, 3128 (2014).

447. S. H. Olson *et al.*, Genetic variants in SOD2, MPO, and NQO1, and risk of ovarian cancer. *Gynecologic oncology* **93**, 615-620 (2004).
448. S. Pan *et al.*, Quantitative glycoproteomics analysis reveals changes in N-glycosylation level associated with pancreatic ductal adenocarcinoma. *Journal of proteome research* **13**, 1293-1306 (2014).
449. G. Ferrandina *et al.*, Class III beta-tubulin overexpression is a marker of poor clinical outcome in advanced ovarian cancer patients. *Clinical cancer research : an official journal of the American Association for Cancer Research* **12**, 2774-2779 (2006).
450. Y. Yu *et al.*, Inhibition of Spleen Tyrosine Kinase Potentiates Paclitaxel-Induced Cytotoxicity in Ovarian Cancer Cells by Stabilizing Microtubules. *Cancer cell* **28**, 82-96 (2015).
451. P. Liu, H. Cheng, T. M. Roberts, J. J. Zhao, Targeting the phosphoinositide 3-kinase pathway in cancer. *Nature reviews. Drug discovery* **8**, 627-644 (2009).
452. P. L. Lee, J. J. Kohler, S. R. Pfeffer, Association of beta-1,3-N-acetylglucosaminyltransferase 1 and beta-1,4-galactosyltransferase 1, trans-Golgi enzymes involved in coupled poly-N-acetyllactosamine synthesis. *Glycobiology* **19**, 655-664 (2009).
453. H. Yamashita *et al.*, Alteration in the metastatic potential of ovarian cancer cells by transfection of the antisense gene of beta-1,4-galactosyltransferase. *Oncology reports* **10**, 1857-1862 (2003).
454. J. Ma *et al.*, Expression of human beta 1,4-galactosyltransferase in gynecological cancer cell lines. *International journal of oncology* **11**, 117-122 (1997).
455. H. J. An, C. B. Lebrilla, A glycomics approach to the discovery of potential cancer biomarkers. *Methods in molecular biology* **600**, 199-213 (2010).
456. S. Hua, H. J. An, Glycoscience aids in biomarker discovery. *BMB reports* **45**, 323-330 (2012).
457. D. Kolarich, P. H. Jensen, F. Altmann, N. H. Packer, Determination of site-specific glycan heterogeneity on glycoproteins. *Nature protocols* **7**, 1285-1298 (2012).
458. K. T. Schjoldager *et al.*, Deconstruction of O-glycosylation--GalNAc-T isoforms direct distinct subsets of the O-glycoproteome. *EMBO reports* **16**, 1713-1722 (2015).
459. R. Sheta, D. Bachvarov, Role of aberrant glycosylation in ovarian cancer dissemination. *Biomdecial Reviews* **25**, 83-92 (2014).
460. R. Kafri, M. Springer, Y. Pilpel, Genetic redundancy: new tricks for old genes. *Cell* **136**, 389-392 (2009).
461. J. Zhang, Genetic redundancies and their evolutionary maintenance. *Advances in experimental medicine and biology* **751**, 279-300 (2012).
462. F. B. Pickett, D. R. Meeks-Wagner, Seeing double: appreciating genetic redundancy. *The Plant cell* **7**, 1347-1356 (1995).
463. E. J. Dean, J. C. Davis, R. W. Davis, D. A. Petrov, Pervasive and persistent redundancy among duplicated genes in yeast. *PLoS genetics* **4**, e1000113 (2008).
464. A. DeLuna *et al.*, Exposing the fitness contribution of duplicated genes. *Nature genetics* **40**, 676-681 (2008).
465. R. Kafri, O. Dahan, J. Levy, Y. Pilpel, Preferential protection of protein interaction network hubs in yeast: evolved functionality of genetic redundancy. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 1243-1248 (2008).
466. H. Matsuura *et al.*, The oncofetal structure of human fibronectin defined by monoclonal antibody FDC-6. Unique structural requirement for the antigenic specificity provided by a glycosylhexapeptide. *The Journal of biological chemistry* **263**, 3314-3322 (1988).

467. U. Mandel *et al.*, Expression of polypeptide GalNAc-transferases in stratified epithelia and squamous cell carcinomas: immunohistological evaluation using monoclonal antibodies to three members of the GalNAc-transferase family. *Glycobiology* **9**, 43-52 (1999).
468. J. H. Park, T. Katagiri, S. Chung, K. Kijima, Y. Nakamura, Polypeptide N-acetylgalactosaminyltransferase 6 disrupts mammary acinar morphogenesis through O-glycosylation of fibronectin. *Neoplasia* **13**, 320-326 (2011).
469. I. Barbaric, G. Miller, T. N. Dear, Appearances can be deceiving: phenotypes of knockout mice. *Brief Funct Genomic Proteomic* **6**, 91-103 (2007).
470. H. H. Wandall *et al.*, Substrate specificities of three members of the human UDP-N-acetyl-alpha-D-galactosamine:Polypeptide N-acetylgalactosaminyltransferase family, GalNAc-T1, -T2, and -T3. *The Journal of biological chemistry* **272**, 23503-23514 (1997).
471. E. P. Bennett, H. Hassan, H. Clausen, cDNA cloning and expression of a novel human UDP-N-acetyl-alpha-D-galactosamine. Polypeptide N-acetylgalactosaminyltransferase, GalNAc-t3. *The Journal of biological chemistry* **271**, 17006-17012 (1996).
472. E. P. Bennett *et al.*, Cloning of a human UDP-N-acetyl-alpha-D-Galactosamine:polypeptide N-acetylgalactosaminyltransferase that complements other GalNAc-transferases in complete O-glycosylation of the MUC1 tandem repeat. *The Journal of biological chemistry* **273**, 30472-30481 (1998).
473. F. K. Hagen *et al.*, cDNA cloning and expression of a novel UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase. *The Journal of biological chemistry* **272**, 13843-13848 (1997).

Annex

Role of aberrant glycosylation in ovarian cancer dissemination

Razan Sheta and Dimcho Bachvarov

Department of Molecular Medicine, Laval University, Québec PQ, Canada Centre de recherche du CHU de Québec, L'Hôtel-Dieu de Québec, Québec PQ, Canada

(Biomedical Reviews, accepted 23 December 2014).

Abstract

Epithelial ovarian cancer (EOC) is the most lethal gynecologic malignancy, and understanding the molecular changes associated with EOC etiology could lead to the identification of novel targets for more effective therapeutic interventions. Glycosylation represents a post-translational modification (PTM) of proteins playing a major role in various cellular functions. Moreover, glycosylation participates in major pathobiological events during tumor progression, as aberrant expression of glycan structures has been shown to contribute in alterations of specific cellular onco-phenotypes, including tumor cell proliferation, migration and invasion. This review aims to describe what is currently known about aberrant glycosylation in EOC, and more specifically, the contribution of aberrant O-linked glycosylation in EOC progression. We also discuss our findings about the altered *GALNT3* overexpression in EOC and its involvement in disease dissemination through aberrant mucin O-glycosylation, as well as the potential to exploit the role of *GALNT3* in understanding the general mechanisms of abnormal glycosylation implicated in EOC spreading. Further analyses in cancer glycobiology could significantly enhance our understanding of the molecular mechanisms of cancer progression, including EOC dissemination, and could lead to the identification of novel biomarkers/therapeutic targets for better management of this deadly disease. Biomed Rev 2014; 25: 83-92 Keywords: post-translational modification, N-linked glycosylation, O-linked glycosylation, GalNAc-transferases, ovarian cancer metastasis, mucins

Introduction

Etiology of epithelial ovarian cancer Epithelial ovarian cancer (EOC) is a disease that is responsible for more cancer deaths among women in the Western world than all other gynecologic malignancies (1). EOC lethality primarily stems from the inability to detect the disease at an early, organ-confined stage, and the lack of effective therapies for advanced-stage disease. So there is urgent need for new therapeutic targets and a better understanding of the mechanisms involved in the spread of ovarian carcinoma. EOC is histologically classified into five different carcinoma subtypes, including low-grade serous, high-grade serous, endometrioid, mucinous and clear cell (2); among these, serous carcinomas

represent the most frequent type, comprising approximately 80 % of all advanced EOC (2). Recent findings indicate that EOC subtypes are different diseases and suggest that future biomarker identification and clinical research studies should investigate each EOC subtype separately (3). Treatment of EOC includes the combination of surgery and platinum based chemotherapy (4). These approaches, in addition to hormonal therapy, show up to 80% responsiveness in patients, but unfortunately more than 50% of EOC patients eventually relapse and will need additional treatments (4). Currently, there is no effective adjuvant treatment for EOC patients after surgery or chemotherapy. It is well established that cancer invasion and metastasis represent the major cause of treatment failure. Approximately 70% of women with advanced-stage EOC have widespread intraperitoneal metastases, including the formation of malignant serous effusions within the peritoneal cavity (1, 2). Pleural effusions constitute the most frequent site of distant metastasis (FIGO stage IV disease). Unlike the majority of solid tumors, particularly at the primary site, cancer cells in effusions are not amenable to surgical removal, and failure in their eradication is one of the main causes of treatment failure. Indeed, despite advances in cytotoxic therapies (5, 6), only 30% of women with advanced-stage EOC survive 5 years after initial diagnosis (1). Thus, management of the metastatic disease becomes a crucial problem for the treatment of EOC. One possible way to resolve this problem is to target metastasis-specific pathways with novel therapies. Hence, focused identification of novel pro-metastatic mechanisms, target pathways and molecules could enhance our chances of discovering new and effective therapies. Post-translational modifications (PTM) alter the three dimensional structures of proteins by covalently binding small molecules to them and therefore represent a major diversification mechanism for altering protein function. The main PTM types include protein covalent modifications such as phosphorylation, acetylation, glycosylation, methylation, and ubiquitination, and can be classified according to the type of the amino acid side chain modified, the category of the modifying enzyme, and the extent of reversibility (7). There are many documented studies investigating different PTMs and their association with cancer progression; among these, aberrant glycosylation has displayed important role in cancer progression (8, 9), including EOC dissemination (10).

Protein glycosylation

Glycosylation represents the most complex form of all PTM, and is defined by the regulated process of adding lipids and carbohydrates to proteins, as this regulation depends on various building enzymes and proteases that allow for the diversification of protein function. Many studies have advanced our understanding of the major roles of glycans, but no evidence has yet comprehensively explained their role or function under different normal and pathological conditions, which is mainly due to their enormous diversity. Glycans have been shown to play crucial roles in development, growth and the functionality of organisms (11). Cell surface glycans are mainly implicated in recognizing molecules that aid in the process of communication and adhesion (12). Analysis of the SWISS-PROT database showed that the majority of proteins of the sequon are glycosylated, as more than 50% of all proteins are characterized as glycoproteins (13). Glycosylation of proteins is relatively rare in prokaryotes and quite common in eukaryotes. Three types of glycosylation of proteins are currently known (including C-linked glycosylation, N-linked glycosylation, and O-linked glycosylation), but C-glycosylation, representing mannosylation of C2 of the indole ring of tryptophan residues is quite rare (14), as the major glycosylation types represent N-linked and O-linked (mucin) glycosylation.

N-linked glycosylation

N-linked-(or asparagine-type) glycosylation represents a fundamental and extensive PTM that results in the covalent attachment of an oligosaccharide onto asparagine residues of the polypeptide chains (15). The N-linked glycosylation of glycoproteins takes place in both prokaryotic and eukaryotic cells (15). Through the attachment of glycans at asparagine residues, N-linked glycosylation promotes protein folding by enhancing solubility and mediating interactions between nascent proteins and cellular proteins, such as chaperones. Three key processes in N-linked glycosylation have been identified: (i) the lipid-mediated assembly of monosaccharides into glycans, which is performed by various enzymes in the endoplasmic reticulum (ER) - Golgi complex; (ii) the acceptance of a glycan by the consensus sequence AsnX-Ser/Thr (X for any amino acid except proline), and (iii) the

oligosaccharyltransferase-catalyzed attachment of the glycan to the side chain of the asparagine residue (15). The initiation step of N-glycan synthesis starts with the transfer of GlcNAc-P from UDP-GlcNAc to membrane-bound Dol-P to form GlcNAc-P-P-Dol, and this process is catalyzed by the enzyme GlcNAc-1-phosphotransferase (16). See Figure 1 for illustration of the principles of the N-glycosylation.

O-linked (mucin) glycosylation

O-glycosylation of proteins represents the most diverse PTM form. There are several types of O-glycosylation, and they have been characterized and differentiated based on their biosynthesis and linkage processes. These include: O-GalNAc or mucin-type, O-mannose, O-xylose, O-fucose, O-glucose, O-galactose and O-GlcNAc (12). The initiation process of O-glycosylation, except for O-GalNAc, O-Xyl and O-GlcNAc, occurs in the ER. O-GalNAc initiation of proteins takes place in the Golgi (12) and the O-GlcNAc glycosylation takes place in the cytosol and the nucleus (17). O-linked glycosylation mostly happens on serine and threonine residues (18), although as recently demonstrated, O-glycosylation can also be found on tyrosine residues (19). The initiation step of O-glycosylation is more complex than N-glycosylation and involves the transfer of different monosaccharides to each of the six O-linked glycans. O-GalNAc glycosylation is more distinct as it involves two steps including the initiation step which is directed by 20 GalNAc-transferases, a family of enzymes known as the UDPN-acetylgalactosamine: polypeptide N-acetyl galactosaminyl transferases (GalNAcTs). Thus, the O-GalNAc initiation step involves the transfer of the monosaccharide GalNAc from UDP-GalNAc to the hydroxyl group of the serine, threonine or tyrosine residues found in the target protein substrate, which is followed by a processing step, where 30 or more glycosyltransferases create distinct glycan structures (20, 21). See Figure 2 for illustration of the principles of GalNAc O-glycosylation. As described above, cellular glycosylation mechanisms and their biosynthetic pathways are very complex and have shown to be fundamental to the changes in glycan processing and divergence. Numerous studies have shown that alterations in surface glycans are linked to disease progression and most importantly in cancer, as these alterations play pivotal roles in cancer initiation and progression (22, 23).

Aberrant glycosylation in cancer

Cancer cells have been characterized to undergo activation, rapid growth, adherence and invasion of neighbouring cells/issues, and these processes are also linked to changes in the cell's glycosylation profiles (24). Studies have provided strong evidence for the existence of important link between abnormal glycosylation and tumorigenesis. Glycosylation changes characterized in cancer cells follow a variety of forms, as glycan alterations can be associated with loss or gain of expression, depending on the cell type and the specific glycan's structure. Aberrant glycosylation in cancer cells could affect certain ligand-receptor interactions and more importantly, could favour cancer cell proliferation, migration and invasion/metastasis. Altered N-glycosylation of proteins has been reported in different cancer types, which has led to the identification of biomarkers for cancer prediction and diagnosis (25). Notably, the branching of N-glycans is involved in various biological functions including signal transduction, cell adhesion, proliferation, differentiation, and programmed cell death (26). The alteration of the branching alters the activity of relevant glycosyltransferases, thus affecting the regulation of various malignant processes, including cancer cells invasion and metastasis (26). A number of N-glycoproteins has been shown to be differentially expressed in different cancer types. Notably, differential glycosylation and increased expression of some N-linked glycopeptides (including apolipoprotein B-100 and alpha-1-antichymotrypsin) has been identified in lung cancer serum pools compared to control pools (27). Some N-glycans, including ceruloplasmin, alpha-1 antichymotrypsin, and multimerin-1 were identified as biomarkers associated with hepatocellular carcinoma (HCC) (28). In addition, a study by Chen *et al*. (2011) found substantial concentration changes in N-linked glycoproteins between normal and HCC serum samples, as some of these glycoproteins included galectin-3 binding protein, insulinlike growth factor binding protein 3, and thrombospondin 1, previously found to be associated with HCC development (28). A strong overexpression of N-glycoproteins, including cathepsin L and periostin (proteins, involved in extracellular matrix remodeling during metastasis) was shown to be associated with aggressive prostate cancer (30). Similarly, O-glycans can play important role in cancerogenesis, as they could be in part responsible for the attachment, invasion and survival of cancer cells. Studies have demonstrated that O-glycan structures are truncated in many cancer types (31). Tumors

associated with different mucin families show complex alterations in glycosylation profiles compared to normal cells. Recent studies showed that the downregulation of mucins contributes to the alterations observed in various cancer cell lines, which is associated with disease progression. MUC1 is a glycoprotein integrated in the plasma membranes of epithelial cells as MUC1 has been repeatedly characterized as a major oncogene in different cancer types (32). Indeed, MUC1 is aberrantly glycosylated and overexpressed in >90% of breast carcinomas and in other cancer types of epithelial origin, including ovarian, lung, colon, and pancreatic carcinomas (33, 34). MUC4 has also been shown to be responsible for tumor progression (35). MUC4 expression has been associated with pancreatic cancer, as MUC4 is not expressed in normal pancreatic cells or tissue (35). MUC16 (or CA-125) is now considered to be the most effective marker for early EOC detection and the monitoring of EOC reoccurrence after therapy (36). The different members of the glycotransferase (GALNACTs) family were shown to be differentially expressed in malignant tumors compared to normal tissue. The deregulation in the expression of the different GalNAc-Ts allows them to play diverse roles in cancerogenesis. Thus, *GALNT1* is highly expressed in bladder cancer tissues, making it a potential prognostic marker in bladder cancer (37); *GALNT2* is overexpressed in oral squamous cell carcinoma (38) and was also found to play a role in modifying EGFR glycosylation in HCC, which contributes to its malignant phenotype (39). *GALNT3* is overexpressed in pancreatic cancer, and its suppression significantly correlates with pancreatic cancer cell growth inhibition (40). We have recently demonstrated that *GALNT3* is strongly overexpressed in high-grade serous EOC tumors, as compared to normal ovarian tissue (41). Upregulation of *GALNT6* stabilizes MUC1 protein by aberrant glycosylation in breast cancer cells, which allows the accumulation of the MUC1 protein in these cancer cells (42). Further, *GALNT6* is expressed in gastric carcinomas and its expression was shown to associate with invasive disease (43). A role of *GALNT7* in progression of cervical cancer (44) and human melanoma (45) has been demonstrated. *GALNT9* expression was coupled with increased survival of neuroblastoma cancer cells making this enzyme a potential prognostic marker for this cancer type (46). *GALNT10* was found to be overexpressed in gastric cancer, as its expression significantly correlated with the histological type and degree of this cancer type (47). Overexpression of *GALNT11* was demonstrated in chronic lymphocytic leukemia (48). Mutations in the

GALNT12 gene were linked to the development of colon cancer (49), and *GALNT13* was specifically expressed in high-risk neuroblastoma (stage 4) patients, while no *GALNT13* expression was detected in healthy individuals (50). *GALNT14* contributes to the glycosylation of MUC13, which was observed to be significantly higher in EOC cells compared to their normal counterparts, suggesting a role of *GALNT14* in EOC carcinogenesis, by directing aberrant MUC13 glycosylation (51). Moreover, a positive correlation was demonstrated between *GALNT14* expression and the pro-apoptotic factor TRAIL in pancreatic cancer, indicative for a protective role of the *GALNT14* enzyme in this cancer type (52).

Altered glycosylation implicated in EOC etiology

N-linked glycosylation in ovarian cancer

N-glycans play crucial roles that influence some of the glycoproteins properties and induce effects on the conformation, solubility, and recognition capacity of the proteins by glycanbinding proteins (16). Deficiencies in N-glycan synthesis has lead to a variety of human diseases. Aberrant N-linked glycosylation has been also described and investigated in ovarian cancer. Several studies have examined this association; thus a number of N-glycans displayed aberrant expression in different EOC subtypes, as mesothelin was shown to be overexpressed in high-grade serous, low-grade serous, and transitional-cell carcinomas, while versican and periostin were overexpressed in most subtypes of ovarian tumors (29). Gubbels *et al*. (2006) demonstrated that mesothelin-MUC16 binding is N-glycan dependent, as this interaction was shown to facilitate EOC peritoneal dissemination (53). Another study by Zhang *et al*. (2014) investigated the association between N-glycosylation alterations in glycoproteins secreted by EOC cells and metastasis of the disease. By using metabolic stable isotopic labeling, they have revealed a notable decrease in bisecting GlcNAc structure, as this decrease was found to be associated with a higher EOC metastatic potential (54). Interestingly, Anugraham *et al.* have shown an expression of bisecting GlcNAc on N-linked glycoproteins, and this association was found to highly correlate with MGAT3 gene expression, a gene responsible for expressing the enzyme

implicated in the synthesis of the core of the N-glycans (55). Their data demonstrated that definite N-glycan substructures and their complexes are associated with specific epigenetic programming of synthetic enzymes in EOC (55). These studies suggest the involvement of N-glycosylation in EOC spreading and metastasis, which further suggests the potential use of N-glycans for biomarker development and targeted EOC therapy.

O-linked (mucin-type) glycosylation in ovarian cancer

Aberrant mucin-type O-glycosylation is one of the most abundant cancer-related PTMs, causing major pathological changes that in turn influence growth and survival of cancer cells and their ability for invasion and metastasis (31). Changes associated with the progression of EOC from primary to metastatic disease can affect the expression of certain tumorassociated antigens (TAA), subject to altered O-glycosylation. Thus, MUC1 has been identified as a novel TAA, and several studies have demonstrated that MUC1 is overexpressed and aberrantly O-glycosylated in most adenocarcinomas (34), in addition to being overexpressed on the cell surface of 90% of EOC tumors, including metastatic lesions (56). The major structural difference observed in tumor-associated MUC1 compared to a normal MUC1 is that tumor MUC1 contains shorter and less dense O-glycan chains, thus exposing more regions of the MUC1 protein core (57). This feature of the tumorassociated MUC1 makes it a possible EOC therapeutic target, which could delineate novel approaches for more effective EOC therapy. MUC13 represents another mucin associated with EOC dissemination. Indeed, MUC13 has been overexpressed in EOC compared to normal tissue, and was shown to modulate various cellular functions such as cell adhesion and proliferation (58). A study by Wang *et al*. (2013) has provided evidence that *GALNT14* contributes to the aberrant glycosylation of MUC13 in ovarian carcinogenesis (51). As mentioned above, *MUC16* (CA-125) is strongly associated with EOC pathobiology, as it currently represents the only clinically recognized marker for EOC progression (59). A study by Robert *et al.* has characterized the expression of several mucin genes in differnet EOC histological subtypes (60). They have shown that *MUC1, MUC2* and *MUC5AC* were overexpressed in EOC, while advanced disease was associated with decreased *MUC3* and *MUC4* levels (60). Moreover, other studies have revealed that ovarian cyst fluid can be a valuable source of mucins. *MUC6* is a major component in

ovarian cyst mucins (61) making it a good candidate to further elucidate the role of cyst mucins in tumorigenesis. In addition, *MUC9* was also found to be highly present in the serum of EOC patients (62). Using an epigenomics approach, we have previously shown that DNA hypermethylation occurs in all (including less invasive/early) stages of ovarian tumorigenesis, while advanced disease was exclusively associated with DNA hypomethylation of a number of oncogenes, implicated in EOC progression, including invasion/metastasis (63). These experiments identified the *GALNT3* gene as hypomethylated and overexpressed in high-grade EOC tumors, when compared to low-malignant potential EOC tumors and normal ovarian tissue (41). Additionally, the *GALNT3* expression significantly correlated with shorter progressionfree survival intervals in women with advanced EOC (41). Consecutive short-hairpin (shRNA)-mediated *GALNT3* KD experiments were strongly indicative for *GALNT3* implication in EOC cell proliferation, migration, invasion and cell cycle control. Gene expression profiling and successive network and pathway analyses confirmed these findings, as numerous genes and pathways known previously to be implicated in EOC tumor invasion and metastasis were found to be downregulated upon *GALNT3* suppression, while some tumor suppressor genes were induced (41). Moreover, *GALNT3* KD resulted in considerable reduction of MUC1 protein in EOC cells, although the transcriptional level of *MUC1* remained unchanged (41). A significantly diminished protein expression of other glycosylated proteins in EOC cells upon *GALNT3* suppression was also observed, suggesting that *GALNT3* may influence the posttranslational modification and stabilization of MUC1 and other O-glycosylated mucin-like targets in EOC cells (41). Also, the *GALNT3* KD-mediated MUC1 suppression promoted the expression of the cell adhesion molecules E-cadherin and β-catenin, suggesting that the involvement of the *GALNT3-MUC1* pathway in EOC invasion could include the destabilization of the E-cadherin/ β-catenin complex formation (41). Our data specify some of the mechanisms of abnormal O-glycosylation in ovarian carcinogenesis and the identification of the *GALNT3* enzyme as EOC oncogene and novel EOC biomarker/possible molecular target for EOC therapy. These results warrant further and more profound studies of aberrant O-glycosylation in EOC.

Conclusion and perspectives

Glycosylation mediates various events associated with important biological expansion of the proteome. Aberrant glycosylation of glycan structures is now a new hallmark common of neoplastic transformation. Determining glycan structures is an essential step towards establishing a more comprehensive understanding of their roles in normal and pathological conditions, including cancer. In this review we focused on discussing the role of aberrant glycosylation in EOC. EOC is a highly metastatic disease and the leading cause of death from gynecologic malignancy. Hence, understanding the molecular changes associated with ovarian cancer metastasis could lead to the identification of targets for novel therapeutic interventions. Studies of disease-related alterations in glycan structures of glycoproteins can also lead to the discovery of new biomarkers for early EOC diagnostics and more efficient follow-up. Global analysis of proteins and their modifications has been majorly dependent on the recent advances in mass spectrometry tools for sequencing N- and O-glycans from cells and tissues. Progress in the field has developed strategies in performing both glycomic and glycoproteomic analysis. With these high-throughput technologies the dynamic area of N- type and mucin-type O-glycosylation can be further explored and their role in tumorigenesis, including that of EOC, better understood. Moreover, the discovery of new and efficient inhibitors that can target glycosylation biosynthetic pathways could play important role in the understanding of specific/molecular mechanisms of normal and aberrant (disease-related) glycosylation, which undoubtedly will lead to the development of new and improved cancer treatment strategies, including more efficient EOC therapy.

Conflict of interest

The authors declare no conflict of interest.

References

1. Siegel R, Ward E, Brawley O, Jemal, A. Cancer statistics. Cancer J Clinicians 2011; 61: 212–236. DOI:10.3322/ caac.20121.

2. Cho KR, Shih IM. Ovarian cancer. Annu Rev Pathol Mech Dis 2009; 4: 287–313. DOI:10.1146/annurev. pathol.4.110807.092246.

3. Brooks SA, Carter TM, Royle L, Harvey DJ, Fry SA, Kinch C, et al. Ovarian carcinoma subtypes are different diseases: implications for biomarker studies. PLos Med 2008; 5; e232. DOI:10.1371/journal.pmed.

4. Bhoola S, Hoskins WJ. Diagnosis and management of epithelial ovarian cancer. Obst Gynecol 2006; 107: 1399–1410.

5. Fruscio R, Corso S, Ceppi L, Garavaglia D, Garbi A, Floriani I, et al. Conservative management of early-stage epithelial ovarian cancer: results of a large retrospective series. Ann Oncol 2012; 24: 138–144. DOI:10.1093/ annonc/mds241.

6.Armstrong DK. Relapsed ovarian cancer: Challenges and management strategies for a chronic disease. Oncologist 2002; 7: 20–28. DOI:10.1634/theoncologist.7-suppl_5-20.

7. Walsh CT, Garneau-Tsodikova S, Gatto GJ. Protein posttranslational modifications: The chemistry of proteome diversifications. Angew Chem Int Ed 2005; 44: 7342–7372. DOI:10.1002/anie.200501023.

8. Bode AM, Dong Z. Post-translational modification of p53 in tumorigenesis. Natl Rev Cancer 2004; 4: 793–805. DOI:10.1038/nrc1455.

9. Karve TM, Cheema AK. Small changes huge impact: The role of protein posttranslational modifications in cellular homeostasis and disease. J Amino Acids 2011; 1–13. DOI:10.1042/BST0380137.

10. Sheehan KM, Calvert VS, Kay EW, Lu Y, Fishman D, Espina V, et al. Use of reverse phase protein microarrays and reference standard development for molecular network analysis of metastatic ovarian carcinoma. Mol Cell Proteomics 2005; 4: 346–355. DOI:10.1074/mcp. T500003-MCP200.

11. Varki A. Biological roles of oligosaccharides: all of the theories are correct. Glycobiology 1993; 3: 97–130. DOI:10.1093/glycob/3.2.97.

12. Brooks SA, Carter TM, Royle L, Harvey DJ, Fry SA, Kinch C, et al. Altered glycosylation of proteins in cancer: what is the potential for new anti-tumor strategies. Anticancer Agents Med Chem 2008; 8: 2–21. DOI:10.2174/187152008783330860.

13. Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochim Biophys Acta General Subjects 1999; 1473: 4–8. DOI:10.1016/S0304- 4165(99)00165-8.

14. Furmanek A, Hofsteenge J. Protein C-mannosylation: Facts and questions. Acta Biochim Pol 2000; 47: 781–789.

15. Schwarz F, Aebi M. Mechanisms and principles of N-linked protein glycosylation. Curr Opin Struct Biol 2011; 21: 576–582. DOI:10.1016/j.sbi.2011.08.005.

16. Schenk B, Fernandez F, Waechter CJ. The ins(ide) and outs(ide) of dolichyl phosphate biosynthesis and recycling in the endoplasmic reticulum. Glycobiology 2001; 11: 61R–70R. DOI:10.1093/glycob/11.5.61R.

17. Hu P, Shimoji S, Hart GW. Site-specific interplay between O-GlcNAcylation and phosphorylation in cellular regulation. FEBS Lett 2010; 584: 2526–2538. DOI:10.1016/j.febslet.2010.04.044.

18. Stanley P. Golgi glycosylation. Cold Spring Harb Persp Biol 2011; 3: a005199–a005199. DOI:10.1101/ cshperspect.a005199.

19. Halim A, Brinkmalm G, Rüetschi U, WestmanBrinkmalm A, Portelius E, Zetterberg H, et al. Sitespecific characterization of threonine, serine, and tyrosine glycosylations of amyloid precursor protein/amyloid β-peptides in human cerebrospinal fluid. Proc Natl Acad Sci USA 2011; 108: 11848–11853.

20. Hagen, Ten KG. All in the family: the UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferases. Glycobiology 2002; 13: 1R–16. DOI:10.1093/glycob/cwg007.

21. Hang HC, Bertozzi CR. The chemistry and biology of mucin-type O-linked glycosylation. Bioorg Med Chem 2005; 13: 5021–5034. DOI:10.1016/j.bmc.2005.04.085.

22. Madsen CB, Petersen C, Lavrsen K, Harndahl M, Buus S, Clausen H, et al. Cancer associated aberrant protein o-glycosylation can modify antigen processing and immune response. PLoS One 2012; 7: e50139. DOI:10.1371/journal.pone.0050139.s002.

23. Radhakrishnan P, Dabelsteen S, Madsen FB, Francavilla C, Kopp KL, Steentoft C, et al. Immature truncated O-glycophenotype of cancer directly induces oncogenic features. Proc Natl Acad Sci USA 2014; 111: E4066– E4075. DOI:10.1073/pnas.1406619111.

24. Meany DL, Chan DW. Aberrant glycosylation associated with enzymes as cancer biomarkers. Clin Proteomics 2011; 8: 7. DOI:10.1186/1559-0275-8-7.

25. Tian Y, Zhang H. Characterization of disease-associated N-linked glycoproteins. Proteomics 2013; 13: 504–511. DOI: 10.1002/pmic.201200333.

26. Zhao Y, Sato Y, Isaji T, Fukuda T, Matsumoto A, Miyoshi E, et al. Branched N-glycans regulate the biological functions of integrins and cadherins. FEBS J 2008; 275: 1939–1948. DOI:10.1111/j.1742-4658.2008.06346.x.

27. Zeng X, Hood BL, Sun M, Conrads TP, Day RS, Weissfeld JL, et al. Lung cancer serum biomarker discovery using glycoprotein capture and liquid chromatography mass spectrometry. J Proteome Res 2010; 9: 6440–6449. DOI:10.1021/pr100696n.

28. Chen R, Tan Y, Wang M, Wang F, Yao Z, Dong L, et al. Development of glycoprotein capture-based label-free method for the high-throughput screening of differential glycoproteins in hepatocellular carcinoma. Mol Cell Proteomics 2011; 10: M110.006445–M110.006445. DOI:10.1074/mcp.M110.006445.

29. Tian Y, Yao Z, Roden RBS, Zhang H. Identification of glycoproteins associated with different histological subtypes of ovarian tumors using quantitative glycoproteomics. Proteomics 2011; 11: 4677–4687. DOI: 10.1002/pmic.201000811.

30. Tian Y, Bova GS, Zhang H. Quantitative glycoproteomic analysis of optimal cutting temperature-embedded frozen tissues identifying glycoproteins associated with aggressive prostate cancer. Anal Chem 2011; 83: 7013–7019. DOI: 10.1021/ac200815q.

31. Brockhausen I. Pathways of O-glycan biosynthesis in cancer cells. Biochim Biophys Acta - General Subjects 1999; 1473: 67–95. DOI:10.1016/S0304-4165(99)00170-1.

32. Patton S, Gendler SJ, Spicer AP. The epithelial mucin, MUC 1, of milk, mammary gland and other tissues. Biochim Biophys Acta - General Subjects 1995; 1241: 407–424. DOI:10.1016/0304-4157(95)00014-3.

33. Singh R, Bandyopadhyay D. MUC1: A target molecule for cancer therapy. Cancer Biol Ther 2014; 6: 481–486. DOI:10.4161/cbt.6.4.4201.

34. Taylor-Papadimitriou J, Burchell J, Miles DW, Dalziel M. MUC1 and cancer. Biochim BiophysActa - General Subjects 1999; 1455: 301–313. DOI:10.1016/S0925-4439(99)00055-1.

35. Swartz MJ, Batra SK, Varshney GC, Hollingsworth MA, Yeo CJ, Cameron JL, et al. MUC4 expression increases progressively in pancreatic intraepithelial neoplasia. Am J Clin Pathol 2002; 117: 791–796. DOI:10.1309/7y7nm1wm-r0yk-m2va.

36. Gupta D, Lis CG. Role of CA-125 in predicting ovarian cancer survival - a review of the epidemiological literature. J Ovarian Res 2009; 2: 13. DOI:10.1186/1757- 2215-2-13.

37. Ding M-X, Wang H-F, Wang J-S, Zhan H, Zuo Y-G., Yang D-L, et al. ppGalNAc T1 as a potential novel marker for human bladder cancer. Asian Pacific J Cancer Preven 2012; 13: 5653–5657. DOI:10.7314/ APJCP.2012.13.11.5653.

38. Lin M-C, Huang M-J, Liu C-H, Yang T-L, Huang M-C. GALNT2 enhances migration and invasion of oral squamous cell carcinoma by regulating EGFR glycosylation and activity. Oral Oncol 2014; 50: 478– 484. DOI:10.1016/j.oraloncology.2014.02.003.

39. Wu YM, Liu CH, Hu RH, Huang MJ, Lee JJ, Chen CH, et al. Mucin glycosylating enzyme GALNT2 regulates the malignant character of hepatocellular carcinoma by modifying the EGF receptor. Cancer Res 2011; 71: 7270–7279. DOI:10.1158/0008-5472.CAN-11-1161.

40. Taniuchi K, Cerny RL, Tanouchi A, Kohno K, Kotani N, Honke K, et al. Overexpression of GalNActransferase GalNAc-T3 promotes pancreatic cancer cell growth. Oncogene 2011; 30: 4843–4854. DOI:10.1038/ onc.2011.194.

41. Wang Z-Q, Bachvarova M, Morin C, Plante M, Gregoire J, Renaud M-C, et al. Role of the polypeptide N-acetylgalactosaminyltransferase 3 in ovarian cancer progression: possible implications in abnormal mucin O-glycosylation. Oncotarget 2014; 5: 544–560.

42. Park JH, Nishidate T, Kijima K, Ohashi T, Takegawa K, Fujikane, T, et al. Critical roles of mucin 1 glycosylation by transactivated polypeptide N-acetylgalactosaminyltransferase 6 in mammary carcinogenesis. Cancer Res 2010; 70: 2759–2769. DOI:10.1158/0008-5472.CAN-09-3911.

43. Gomes J, Marcos NT, Berois N, Osinaga E, Magalhaes A, Pinto-de-Sousa J, et al. Expression of UDP-N-acetyl-D-galactosamine: Polypeptide N-acetylgalactosaminyltransferase-6 in gastric mucosa, intestinal metaplasia, and gastric carcinoma. J Histochem Cytochem 2008; 57: 79–86. DOI:10.1369/ jhc.2008.952283.

44. Peng RQ, Wan HY, Li HF, Liu M, Li X, Tang H. MicroRNA-214 suppresses growth and invasiveness of cervical cancer cells by targeting UDP-N-acetyl- -D-galactosamine: Polypeptide N-Acetylgalactosaminyltransferase 7. J Biol Chem 2012; 287: 14301–14309. DOI:10.1074/jbc.M111.337642.

45. Gaziel-Sovran A, Segura MF, Di Micco R, Collins MK, Hanniford D, Vega-Saenz de Miera E, et al. miR-30b/30d regulation of GalNAc transferases enhances invasion and immunosuppression during metastasis. Cancer Cell 2011; 20: 104–118. DOI:10.1016/j.ccr.2011.05.027.

46. Berois N, Gattolliat CH, Barrios E, Capandeguy L, Douc-Rasy S, Valteau-Couanet D, et al. GALNT9 gene expression is a prognostic marker in neuroblastoma patients. Clin Chemistry 2013; 59: 225–233. DOI:10.1373/ clinchem.2012.192328.

47. Gao Y, Liu Z, Feng J, Sun Q, Zhang B, Zheng W, et al. Expression pattern of polypeptide N-acetylgalactosaminyltransferase-10 in gastric carcinoma. Oncol Lett 2012; 5: 113-116. DOI:10.3892/ ol.2012.980.

48. Libisch MG, Casás M, Chiribao ML, Moreno P, Cayota A, Osinaga E, et al GALNT11 as a new molecular marker in chronic lymphocytic leukemia. Gene 2014; 533: 270–279. DOI:10.1016/j.gene.2013.09.052.

49. Gudaa K, Moinova H, He J, Jamison O, Ravi L, Natale L, et al Inactivating germ-line and somatic mutationsin polypeptide N-acetylgalactosaminyltransferase 12 in human colon cancers. Proc Natl Acad Sci USA 2009; 106: 12921–12925. DOI: 10.1073/pnas.0901454106.

50. Berois N. ppGalNAc-T13: A new molecular marker of bone marrow involvement in neuroblastoma. Clin Chem 2006; 52: 1701–1712. DOI:10.1373/ clinchem.2006.067975.

51. Wang R, Yu C, Zhao D, Wu M, Yang Z. The mucin-type glycosylating enzyme polypeptide N-acetylgalactosaminyltransferase 14 promotes the migration of ovarian cancer by modifying mucin 13. Oncol Rep 2013; 30: 667-676. DOI: 10.3892/or.2013.2493.

52. Holoch PA, Griffith TS. TNF-related apoptosis-inducing ligand (TRAIL): A new path to anti-cancer therapies. Eur J Pharmacol 2009: 625: 63–72. DOI: 10.1016/j. ejphar.2009.06.066.

53. Gubbels JAA, Belisle J, Onda M, Rancourt C, Migneault M, Ho M, et al. Mesothelin-MUC16 binding is a high affinity, N-glycan dependent interaction that facilitates peritoneal metastasis of ovarian tumors. Mol Cancer 2006; 5: 50. DOI:10.1186/1476-4598-5-50.

54. Zhang X, Wang Y, Qian Y, Wu X, Zhang Z, Liu X, et al. Discovery of specific metastasis-related n-glycan alterations in epithelial ovarian cancer based on quantitative glycomics. PLoS One 2014; 9: e87978. DOI:10.1371/journal.pone.0087978.s007.

55. Anugraham M, Jacob F, Nixdorf S, Everest-Dass AV, Heinzelmann-Schwarz V, Packer NH. Specific glycosylation of membrane proteins in epithelial ovarian cancer cell lines: Glycan structures reflect gene expression and DNA methylation status. Mol Cell Proteomics 2014; 13: 2213–2232. DOI:10.1074/mcp.M113.037085.

56. Wang L, Ma J, Liu F, Yu Q, Chu G, Perkins AC, et al. Expression of MUC1 in primary and metastatic human epithelial ovarian cancer and its therapeutic significance. Gynecol Oncol 2007; 105: 695–702. DOI:10.1016/j. ygyno.2007.02.004.

57. Price MR, Hudecz F, O'Sullivan C, Baldwin RW, Edwards PM, Tendler SJ. Immunological and structural features of the protein core of human polymorphic epithelial mucin. Mol Immunol 1990; 27: 795–802. DOI:10.1016/0161- 5890(90)90089-I.

58. Chauhan SC, Vannatta K, Ebeling MC, Vinayek N, Watanabe A, Pandey KK, et al. Expression and functions of transmembrane mucin muc13 in ovarian cancer. Cancer Res 2009; 69: 765–774. DOI:10.1158/0008-5472. CAN-08-0587.

59. Guppy AE, Rustin GJ. CA-125 response: Can it replace the traditional response criteria in ovarian cancer? Oncologist 2002; 7: 437–443. DOI:10.1634/theoncologist.7-5-437.

60. Giuntoli RL 2nd, Rodriguez GC, Whitaker RS, Dodge R, Voynow JA. Mucin gene expression in ovarian cancers. Cancer Res 1998; 58: 5546–5550.

61. Lloyd KO, Yin BW, Tempst P, Erdjument-Bromage H. MUC-6 mucin is a major component of `blood group substance' from human ovarian cyst fluid. Bioch Bioph Acta - General Subjects 2000; 1474: 410–414.

62. Maines-Bandiera S, Woo MM, Borugian M, Molday LL, Hii T, Gilks B, et al. Oviductal glycoprotein (OVGP1, MUC9), a differentiation-based mucin present in serum of women with ovarian cancer. Int J Gynecol Cancer 2010; 20: 16–22. DOI:10.1111/ IGC.0b013e3181bcc96d.

63. Keita M, Wang ZQ, Pelletier JF, Bachvarova M, Plante M, Gregoire J, et al. Global methylation profiling in serous ovarian cancer is indicative for distinct aberrant DNA methylation signatures associated with tumor aggressiveness and disease progression. Gynecol Oncol 2013; 128: 356-363. DOI:10.1016/j. ygyno.2012.11.036.

Figure legends

Figure 1. Schematic depiction of N-linked glycosylation. N-linked glycosylation is mainly linked to asparagine residues of proteins, specifically the Asn-X-Ser/Thr motif. N-linked glycosylation is initiated by en bloc transfer of an assembled complex glycan to the Asn residue. N-glycosylation biosynthesis starts with the high-mannose structure, followed by the branching of glycans and the addition of a variety of monosaccharides, this process is mainly dependent on the glycosyltransferases present, here we show the role of GlcNAc-1-phosphotransferase involved in the initiation process found in the ER-Golgi complex. The arrow indicates the transferase involved in the initiation process. The linkage of the first oligosaccharide is depicted with a (-).

Figure 2. Schematic depiction of O-linked-(mucin type) glycosylation. O-linked glycosylation is mainly linked to Ser/Thr and Tyr residues. O-glycosylation is initiated by transfer of one of six monosaccharides; among these, the GalNAc-type-O-linked glycosylation pathway as depicted in the figure. GalNAc-type-O-linked glycosylation is uniquely controlled by the family of 20 polypeptide GalNAc-transferases located in the Golgi complex. The arrow indicates the genes involved in the initiation process. The linkage of the first oligosaccharide is depicted with a (-).

Figures

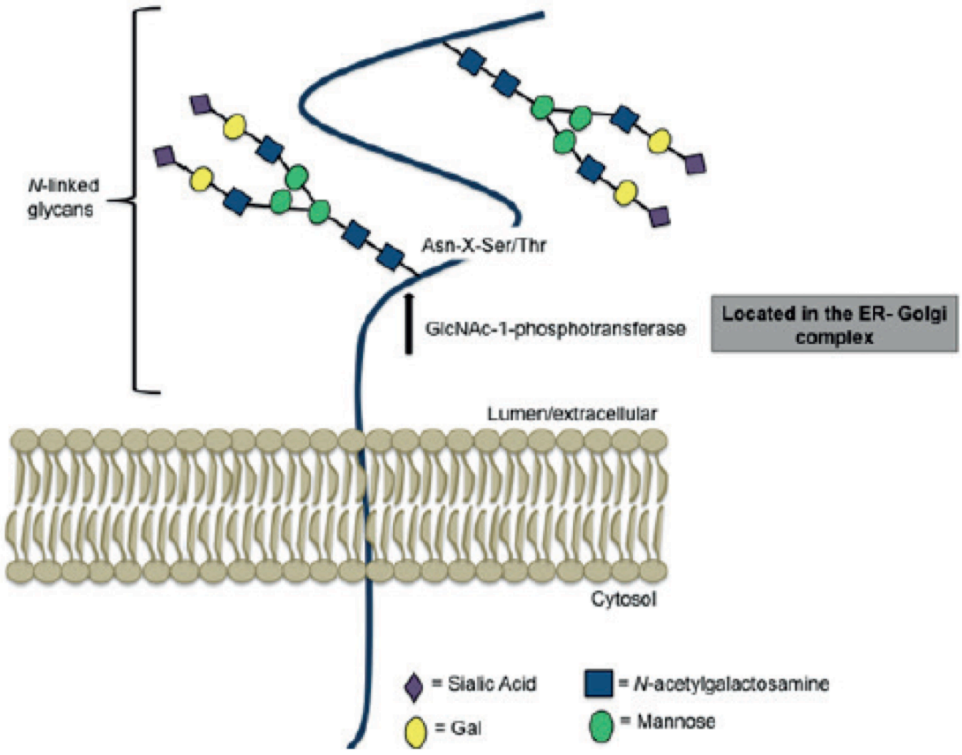


Figure 1. Schematic depiction of N-linked glycosylation

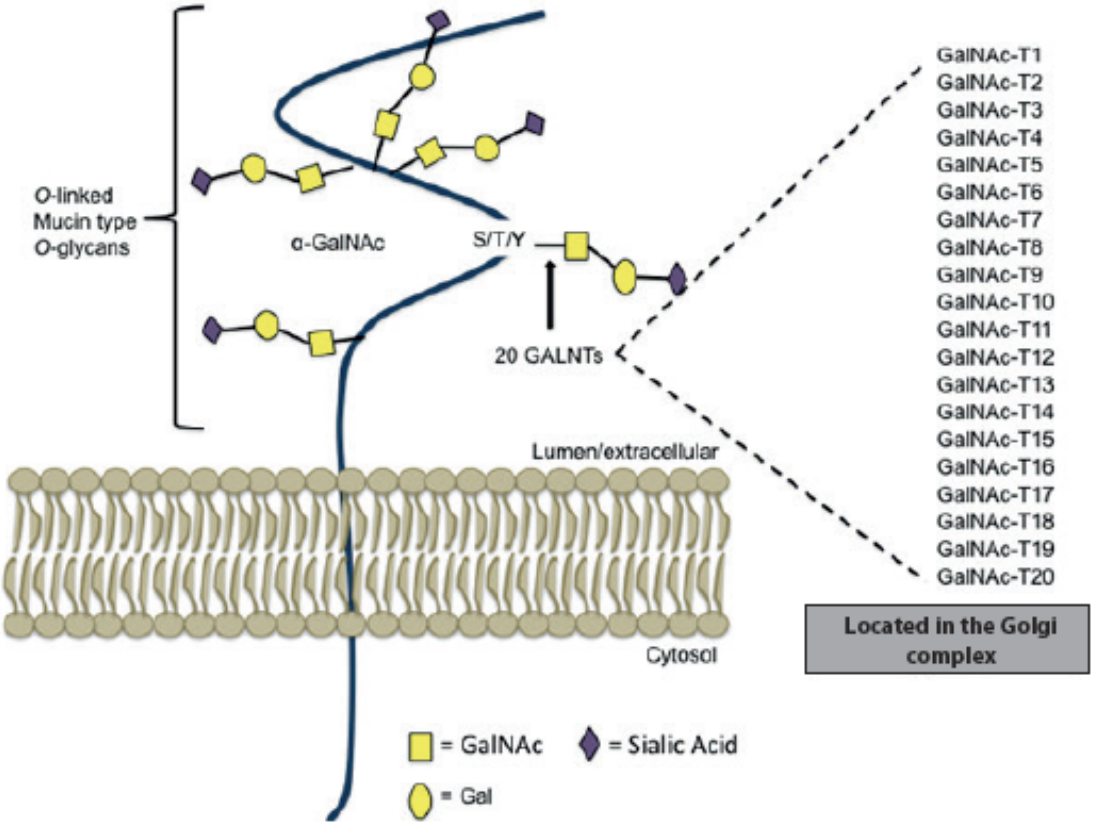


Figure 2. Schematic depiction of O-linked-(mucin type) glycosylation