Stefan Mereiter

Characterization of Gastric Cancer Glycosylation Profile for the Discovery of Biomarkers.

Tese de Candidatura ao grau de Doutor em Ciências Biomédicas submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Orientador – Professor Doutor Celso Albuquerque Reis

Categoria – Professor Auxiliar Convidado

Afiliação – Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Coorientador – Doutora Ana Maria Rodrigues Leite de Magalhães

Categoria – Investigadora de pós-doutoramento do Instituto de Investigação e Inovação em Saúde (i3S) da Universidade do Porto

Afiliação – Instituto de Investigação e Inovação em Saúde (i3S) da Universidade do Porto

Funding

The work presented within this thesis was financially supported by the Marie Curie Actions and the Seventh Framework Programme of the European Union (Initial Training Networks: FP7-PEOPLE-2012-ITN; Project: GastricGlycoExplorer, 316929).

"Insanity is doing the same thing over and over again
and expecting different results."
- Albert Einstein

Einstein was obviously not working in life sciences.

Declaration

The author of this thesis declares that, in accordance with number 2, article 31 of "Decreto-lei nº 178/2009", he provided a major contribution to the design and technical execution of the work, interpretation of the results and manuscript preparation resulting in the following accepted articles and articles in preparation:

Scientific publications

- **Mereiter S**, Magalhães A, Adamczyk B, Jin C, Almeida A, Drici L, Ibáñez-Vea M, Gomes C, Ferreira JA, Afonso LP, Santos LL, Larsen MR, Kolarich D, Karlsson NG, Reis CA (2016). Glycomic analysis of gastric carcinoma cells discloses glycans as modulators of RON receptor tyrosine kinase activation in cancer. Biochim Biophys Acta. Aug;1860(8):1795-808.

- **Mereiter S**, Magalhães A, Adamczyk B, Jin C, Almeida A, Drici L, Ibáñez-Vea M, Larsen MR, Kolarich D, Karlsson NG, Reis CA (2016). Glycomic and sialoproteomic data of gastric carcinoma cells overexpressing ST3GAL4. Data Brief. Mar 14; 7:814-33.

- Magalhães A, **Mereiter S**, Reis C (2016) Reciprocal Modulation of Terminal Sialylation and Bisecting N-Glycans: A New Axis of Cancer-Cell Glycome Regulation? J Biol Chem. Apr; 8;291(15):8308.

- **Mereiter S**, Balmaña M, Gomes J, Magalhães A, Reis CA (2016). Glycomic Approaches for the Discovery of Targets in Gastrointestinal Cancer. Front Oncol. 2016 Mar; 9;6:55.

- **Mereiter S**, Gomes C, Barreira C, Macedo J, Drici L, Ibáñez-Vea M, Larsen MR, Magalhaes A, Reis CA. (*In preparation*). The Molecular and Functional Impact of *O*-Glycan Truncation on CD44 in Cancer.

- **Mereiter S**, Williams C, Persson N, Polonia A, Polom K, Roviello F, Blixt O, Kuras M, Magalhães A, Reis CA. (*In preparation*). Glycan Signature Acts as Predictive Marker in Gastric Cancer.

Acknowledgments

I have to start my acknowledgments by expressing my honest gratitude to Professor Celso A. Reis for the excellent supervision throughout my PhD. Thank you for having always been available for scientific consultations, even if it meant to squeeze in a mini-meeting on a busy day, and thank you for the trust in me and my scientific doing.

Thank you Ana Maria, for guiding me the past years through my experiments, projects and ultimately through this thesis. Thank you also for your patience, your understanding and for having always an open ear for my questions, doubts and worries. You never failed to surprise me with your profound knowledge to virtually all scientific subjects that we encountered. Without doubt, my time as PhD student would have been way harder without your continuous help. My sincere gratitude for that.

Talking of supervision, I obviously also have to thank my unofficial "co-co-supervisor", Txell. Can't write all the acknowledgments you'd deserve without making this part too cheesy. You were my bridge over turbulent waters. Thank you for always supporting me and for enduring the tough times we had.

Thanks to every current and former member of the glycobiology in cancer group that I had the pleasure to share the lab with so far. Every one of you guys has a special place in my heart: Ana Dias, Catarina, Cíntia, Cláudia, Daniela, Diana, Professora Fátima, Filipe, Henrique, Hugo, Inês, Joana Gomes, Joana Macedo, Juliana, Mariana, Rita, Salomé, Sandra, Sara, Telmo and Tiago Silva.

Also to the folks from the "other side", Lara, Lulu, Rita, Ricardo and Vânia; thank you.

I haven't forgotten about my buddies Bruno, Diogo, Márcia and Tiago O. You deserve to be highlighted; you turned the past three years, both in and outside the lab, into an awesome experience. I feel honored that I can call you guys my friends.

Thanks to the IPATIMUP football gang. My health and sanity were preserved because of this weekly timeout from all the worries and all the IPATIMUP-meninas.

I also would like to thank everyone involved in the consortium, in particular, Andreia, Barbara, Coralie, Davide, Felipe, Giulia, Josh, Karol, Lylia, Marcin, Nina, Sophie, Stefan, Yevgen and Yolanda, for being amazing colleagues. Thank you also Niclas and Masood for hosting me in Sweden.

Zu guter Letzt natürlich die Danksagung an meine Familie und Freunde daheim in Österreich. Euch ist meine Abschlussarbeit gewidmet. Der Gedanke euch stolz zu machen, oder zumindest euch einen Doktortitel unter die Nase zu reiben, gab mir Kraft und hat jeden Zweifel am eingeschlagenen Lebensweg im Keim erstickt. Mama und Papa, euch ist es vor allem zu verdanken, dass ich es so weit gebracht habe. Nach so vielen katastrophalen Schuljahren hoffe ich euch am Ende meines Bildungswegs schließlich stolz gemacht zu haben. Christian und Bernhard, wenn ihr das lest, wisst ihr, ich bin nun offiziell klüger als ihr - haha. Meinen besonderen Dank möchte ich auch an Martin S. 7B aussprechen - du hast mich via Skype immer begleitet. Danke liebe Familie und Freunde, dass, selbst wenn die Welt kopfsteht, ihr für mich da seid.

MUITO OBRIGADO A TODOS.

Resumo

O cancro gástrico é uma das principais causas de morte a nível mundial, afetando todas as regiões socioeconómicas e geográficas. O diagnóstico acontece geralmente em estados avançados e as tentativas de cura falham frequentemente. Por estes motivos, existe uma necessidade urgente de bons marcadores que permitam o diagnóstico precoce do cancro gástrico. Adicionalmente, são necessários novos marcadores preditivos para estratificar os pacientes e permitir estratégias clínicas mais personalizadas. Atualmente, algumas estruturas glicosiladas são importantes indicadores independentes de diagnóstico, constituindo também importantes ferramentas para monitorizar o tratamento e as recidivas. No entanto, a nossa compreensão das consequências fenotípicas das alterações de glicosilação é ainda muito reduzida, limitando o potencial da utilização das estruturas glicosiladas como biomarcadores na prática clínica.

Esta tese teve como objetivo geral caracterizar o papel das alterações da glicosilação no cancro gástrico. O trabalho desenvolvido focou-se na expressão de *O*-glicanos truncados e modificados com fucose e ácido siálico e procurou compreender o impacto destas alterações ao nível proteico, da célula tumoral e do tumor. Para alcançar este objetivo foram utilizadas três estratégias complementares: (i) utilização de linhas celulares geneticamente modificadas de forma a mimetizar diferentes glico-fenótipos malignos para a caracterização de alterações de glicosilação e identificação das glicoproteínas alvo modificadas; (ii) caracterização do impacto das alterações de glicosilação em glicoproteínas selecionadas com um papel crítico na carcinogénese e progressão tumoral; (iii) análise de associações entre alterações da glicosilação e características do tumor numa série de doentes com carcinoma gástrico.

Neste estudo, observamos que a sobre expressão da sialiltransferase ST3GAL4 em células de carcinoma gástrico resulta, para além do aumento da expressão da estrutura sialil Lewis X, em numerosas alterações do glicoma celular. Estas alterações incluem a redução da extensão de *O*-glicanos, a diminuição de *N*-glicanos bissectados e concomitante aumento de *N*-glicanos ramificados, bem como a transição de ácidos siálicos em ligação α2,6 para α2,3 em *N*-glicanos. Adicionalmente, identificamos glicoproteínas com glicosilação alterada nestes modelos celulares e verificamos que vários alvos relevantes na biologia tumoral eram afetados simultaneamente, incluindo o Recetor da Insulina, RON, Integrina αV, Integrina α3, Integrina β6, CEACAM1, CEACAM5 e as fosfatases de recetores de tirosina cinase DEP-1 and SAP-1. Verificamos ainda que a glicosilação alterada do recetor tirosina cinase oncogénico RON resulta na sua ativação e validamos a co-expressão de sialyl Lewis X e RON em carcinomas gástricos.

Demonstramos ainda nesta tese que a terminação precoce de *O*-glicanos, frequentemente observada nos carcinomas gástricos, induziu alterações significativas na glicoproteína CD44. Adicionalmente, demonstramos que o oncogene CD44 é abundantemente modificado com *O*-glicanos truncados e que estes afetam as suas características moleculares e a sua colocalização com o recetor tirosina cinase RON.

Finalmente, demonstramos novas associações entre estruturas glicosiladas bem caracterizadas, como o sialyl Lewis A or sialyl Tn, as características clinico-patológicas dos carcinomas e alterações moleculares das células tumorais. Estas análises permitiram identificar novas associações entre a expressão de sialyl Lewis A e infiltração perineural e mutações no gene *MLK3*, bem como entre a expressão de sialyl Tn e E-caderina.

Em resumo, esta tese permitiu estabelecer um conjunto de novos possíveis biomarcadores, bem como novas associações entre a expressão de estruturas

glicosiladas e características tumorais importantes. Adicionalmente, foram caracterizados mecanismos moleculares oncogénicos associados a um perfil de glicosilação aberrante que constituem a base para novas linhas de investigação com potencial para a utilização de marcadores da glicosilação na estratificação de doentes na prática clínica.

Abstract

Gastric cancer is among the leading causes of death worldwide, affecting every socioeconomic and geographic region. Patients are typically diagnosed in advanced stages in which curative attempts usually fail. Hence, there are urgent needs for good diagnostic biomarkers, that enable the early detection of gastric malignancies. In addition, novel predictive assays are required to stratify patients and tailor their clinical strategies. Cancer associated glycan epitopes already play a pivotal role as independent diagnostic indicators and tools to monitor treatment progress and cancer recurrence. However, our understanding of the phenotypic consequences of glycan alterations is fairly limited, impeding the use of glycan-based biomarkers to their full potential in the clinical setting.

In this thesis, we aimed to broadly tackle the role of glycan alterations in gastric cancer. We focused on the oncogenic expression of sialofucosylated and truncated *O*-glycan epitopes, and sought to understand the impact of these alterations at a tumor, cancer cell and protein level. Three complementary strategies were chosen to achieve this aim: (i) mimicking different malignant glycophentoypes by glycoengineered cell models to unravel concomitantly appearing glycan alterations and to identify altered glycoprotein targets; (ii) dissecting the impact of altered glycosylation on selected glycoproteins that are key players of carcinogenesis and cancer progression; (iii) analyzing associations between glycan alterations and tumor features in a cohort of patients' carcinomas.

We present that the overexpression of the sialyltransferase ST3GAL4 in gastric carcinoma cells led, besides the upregulation of the gastric cancer-associated epitope sialyl Lewis X, to numerous other glycomic changes. These alterations included the reduction of *O*-glycan extension, decrease in bisected and increase in branched *N*-glycans as well as the transition from α2-6 towards α2-3 linked sialic acids on *N*-glycans. We further determined glycoproteins with altered glycosylation in these cells

and revealed that numerous cancer relevant targets were affected simultaneously; among them known key players of malignancy such as Insulin receptor, RON, Integrin αV, Integrin α3, Integrin β6, CEACAM1, CEACAM5 and Receptor-type tyrosine-protein phosphatases DEP-1 and SAP-1. The oncogenic RON receptor tyrosine kinase was hyperactived upon this glycosylation alteration and the coexpression of sialyl Lewis X and RON was found in tumors of gastric cancer patients.

Moreover, we demonstrate in this thesis that the truncation of *O*-glycans, another common feature of gastric carcinomas, led to significant changes on the glycoprotein CD44. We disclose that the oncogene CD44 is a major carrier of truncated *O*-glycans, affecting its molecular characteristics and its colocalization with the RON receptor tyrosine kinase.

We finally show new associations of the well-studied glycan epitopes, such as sialyl Lewis A or sialyl Tn, with clinicopathological features of the tumor and molecular alterations of the cancer cells. These associations uncovered novel links between sialyl Lewis A expression and perineural infiltration and *MLK3* mutations, as well as between sialyl Tn and E-cadherin expression.

Overall, this thesis establishes a set of new putative biomarkers and novel associations between glycan epitopes and tumor features. We deliver an insight into oncogenic mechanisms of aberrant glycosylation and provide the basis for new lines of research that ultimately bare the potential to improve the utility of the glycan-based gastric cancer markers for patient stratification in the clinical setting.

Abbreviations

2-AB	2-aminobenzamide
ACN	Acetonitrile
AJCC	American Joint Committee on Cancer
CEA	Carcinoembryonic antigen
CEACAM1	Carcinoembryonic antigen-related cell adhesion molecule 1
ChABC	Chondroitinase ABC
CIN	Chromosomal instable tumors
CS	Chondroitin sulfate
DAB	3,3′-diaminobenzidine tetrahydrochloride
EBV	Epstein-Barr Virus
ECD	Extracellular domain
ECM	Extracellular matrix
EGFR	Epidermal growth factor receptor
ER	Endoplasmic reticulum
ESI	Electrospray ionization
FDA	Food and Drug Administration
FFPE	Formalin fixed paraffin embedded
FGFR2	Fibroblast growth factor receptor 2
FLD	Fluorescence Detector
Fuc	Fucose
GAGs	Glycosaminoglycans
Gal	Galactose
GalNAc	N-Acetylgalactosamine
GC	Gastric cancer
GlcA	Glucoronic acid
GlcN	Glucosamine
GlcNAc	N-Acetylglucosamine
GM	Gastric mucosa
GS	Genomic stable tumors
GU	Glucose units
H. pylori	Helicobacter pylori
HA	Hyaluronan
Hep	Heparinise
HER	Human epidermal growth factor receptor
Hex	Hexose
HexNAc	N-acetylhexoasmine
HILIC	Hydrophilic interaction liquid chromatography
HS	Heparan sulfate
ICD	Intracellular domain

IdoA	Iduronic acid
IF	Immunofluorescene
IHC	Immunohistochemistry
IM	Intestinal metaplasia
IP	Immunoprecipitation
IPT	Immunoglobulin-like-plexin-transcription
KEM®	Knowledge Extraction and Management
LacNAc	N-acetyllactosamine
M3 or M6	Mock transfected cells
Man	Mannose
MS/MS	Tandem mass spectrometry
MSI	Microstallite instability
PGC	Porous graphitized carbon
PLA	In situ Proximity Ligation Assay
PTM	Post-translational modification
RTK	Receptor tyrosine kinases
SC	SimpleCells
SEER	Surgically resected gastric adenocarcinomas of the Surveillance, Epidemiology, and End Results
Sia	Sialic acid
SLe^a	Sialyl Lewis A
SLe^x	Sialyl Lewis X
SNA	Sambucus nigra lectin
ST	Sialyl Thomsen-Friedenreich
ST3	ST3GAL4 transfected cells
ST6	ST6GALNAC1 transfected cells
STn	Sialyl Thomsen-nouvelle
T antigen	Thomsen-Friedenreich antigen
TCGA	The Cancer Genome Atlas
TM	Transmembrane
Tn antigen	Thomsen-nouvelle antigen
UPLC	Ultra Performance Liquid Chromatography
WB	Western blotting
WT	Wild type cells

TABLE OF CONTENTS

CHAPTER I

GENERAL INTRODUCTION

AIMS AND OBJECTIVES

1. GENERAL INTRODUCTION

1.1 CANCER AND GASTRIC CANCER

Cancer continues to be the most formidable challenge in modern medicine, with more than 8 million deaths worldwide each year [1] and even more than forty years after declaring "war" on cancer by Richard Nixon's national cancer act of 1971, still 15% of worldwide deaths can be attributed to this disease [2]. Cancer is the second leading cause of death today and knows no boundaries, affecting every region and socioeconomic level.

Among all cancers, gastric cancer is the fifth most common worldwide with 952,000 cases diagnosed in 2012 and ranks third as cause of cancer-related deaths due to its usually late diagnosis in advanced stages [1, 3].

1.1.1 GENERAL MECHANISMS OF CANCER

The term cancer (also malignant tumor or neoplasia) refers to a large group of diseases with the defining feature of abnormal cells that grow beyond their usual boundaries, and thereby invade adjacent tissue and other organs.

Carcinogenesis (the process of cancer development) is a mutation driven, multistep process (Figure 1) resulting in malignant cells that eventually acquire the following hallmark abilities: sustained proliferative signaling, avoiding cell death enabling replicative immortality, evasion of growth suppression and immune destruction, induction of angiogenesis, genomic instability, acquisition of tumor promoting inflammation, reprogramming of energy metabolism, invasion and ultimately, metastasis [4].

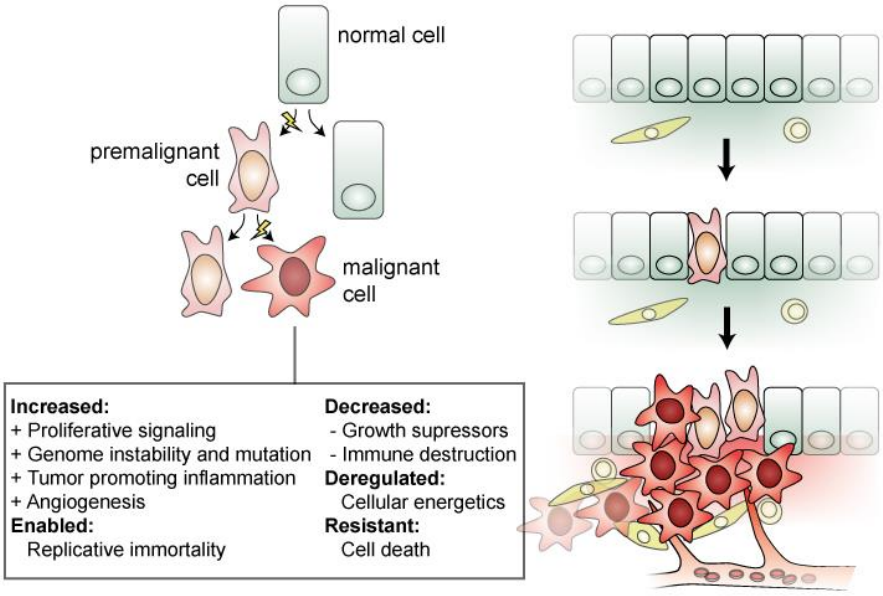


Figure 1. Schematic illustration of the mutation driven turn-over (transformation) from a normal cell into a malignant cell and the cancer hallmark abilities that are being acquired.

Normal epithelial cells that accumulate mutations may enter a premalignant state which is usually accompanied by disordered cellular morphology. These premalignant cells are poised to become malignant by the acquisition of additional mutations and ultimately attain a series of hallmark abilities (listed in the box). Mutation events are indicated by thunderbolts.

1.1.2 EPIDEMOLOGY OF GASTRIC CANCER

Despite being the fifth most common cancer worldwide, the geographic distribution of gastric cancer is exceptionally heterogeneous (Figure 2). In this regards, whereas among Asian countries such as China, Japan and Korea, gastric cancer is one of the most common cancers accounting for 70% of all gastric cancer incidences worldwide, it ranks only number 15 in cancer incidence rate in the US in the year 2016, being described as “relatively rare” by the NIH [5]. Between genders the incidence rate shows also a large discrepancy, with men being around twice as likely to develop this

malignancy as women [1, 3]. Less developed countries and especially rural areas carry a greater gastric cancer burden, probably related to environmental and nutritional factors. Risk factors to develop gastric cancer include diet which is rich in salted and smoked food and poor in fresh fruits and vegetables, and smoking [6-9]. The improved food preservation due to refrigeration and decreased tobacco consumption led to a decline in incidence during the past 80 years [3, 10, 11]. However, the strongest known risk factor to develop gastric cancer is a *Helicobacter pylori* infection. This gram-negative bacterium colonizes the human stomach of about half of the world's population and is the main cause of chronic gastritis. It is a class I carcinogen, and its eradication in endemic areas results in the decline of gastric cancer [12, 13]. Only a small subset of people infected with *H. pylori* develop gastric cancer as the infection outcome depends on multiple factors, such as host genetic susceptibility, strain virulence and other environmental factors [14, 15].

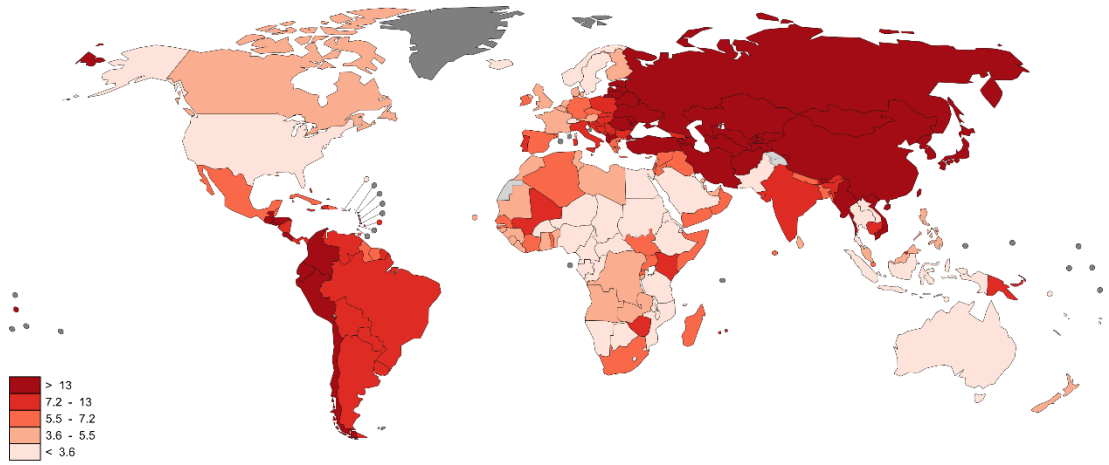


Figure 2. Estimated gastric cancer mortality worldwide in 2012 for men.

The mortality values are age-standardized and per 100,000 people. Hotspots for gastric cancer are the Middle East, Central Asia, East Asia, Eastern Europe and Latin America. Among Western European countries Portugal and Italy carry the highest cancer burden. Illustration from Globocan 2012 / Cancer Fact Sheet [1].

Epstein-Barr Virus (EBV), another ubiquitous pathogen, has also been associated to gastric cancer [16]. EBV is a DNA virus of the herpes family and about 95% of the world adult population has been infected and are lifelong carriers of the virus [17]. Around 10% of all gastric cancers are linked to EBV infection by expressing latent EBV genes and are thought to be derived from monoclonal proliferations of EBV-carrying epithelia cells. However, the mechanism by which EBV infects and subsequently contributes to the transformation of gastric epithelial cells remains uncertain [18].

1.1.3 CLASSIFICATION AND PATHOGENESIS

Several systems have been proposed throughout the years for the classification of gastric cancer. Today the Lauren classification is one of the most commonly used [19], which divides gastric cancer based on histological features into intestinal and diffuse type. Around 85% of all gastric cancers can be grouped into one of the two Lauren types. Intestinal type is in most geographic regions the more frequent subtype and is composed of well or moderately differentiated tumor cells that form recognizable glands. This subtype is often derived from preceding premalignant conditions triggered by *H. pylori* infection and is particularly prevalent in high risk regions and among elderly men [20]. Diffuse type tumors form the second major group and consist of poorly adhesive, infiltrative cells that do not form any glands. Malignant cells of the diffuse type show commonly a loss of expression for the cell-cell adhesion molecule E-cadherin [21, 22]. The remaining approximately 15% of the cases are mixed type gastric cancer, displaying features of both, intestinal and diffuse type or are indeterminate type, rare forms with histological features that resemble neither intestinal nor diffuse type [23].

Intestinal and diffuse gastric cancer display different carcinogenic pathways. Whereas the intestinal type progresses through a sequence of well-described histological stages often initiated by *H. pylori* infection [24], the diffuse type has no association

to premalignant lesions and is thought to arise spontaneous (Figure 3) [25]. While intestinal type gastric cancer preferentially metastasizes haematogenously to the liver, diffuse type gastric cancer frequently invades the duodenum and metastasize to peritoneal surfaces. Both types of tumors show an equal incidence of lymph node metastases and mixed type tumors exhibit the metastatic patterns of both intestinal and diffuse types [26, 27].

Genetic heterogeneity of the *H. pylori* genome plays an important role in determining the outcome of *H. pylori* infection [28]. In particular the presence of virulence factors CagA and VacA increase the risk for severe gastritis and subsequently gastric cancer [29, 30]. In addition the expression of carbohydrate adhesion molecules of *H. pylori* BabA and SabA were also associated with gastric disease severity [31, 32]. Moreover, genetic polymorphisms of the host influence the propensity to develop gastric cancer [28]. In particular, polymorphisms in genes involved in inflammatory response such as IL-1β, IL-8, IL-10 an TNF-α can play an important role in promoting gastric carcinogenesis [33].

Recently, the cancer genome atlas (TCGA) research network proposed a novel classification based on molecular markers for gastric cancer [34]. The TCGA classification divides gastric cancer in four major genomic subtypes: (i) Epstein-Barr virus positive tumors, which show increased DNA hypermethylation and amplification of PD-L1 and 2; (ii) microsatellite unstable tumors, which show very high mutation rates; (iii) tumors with chromosomal instability, which show increased aneuploidy and amplification of receptor tyrosine kinases (RTK); (iv) genomically stable tumors, which consist mainly of diffuse type cancer and show frequent alterations in Rho-GTPases.

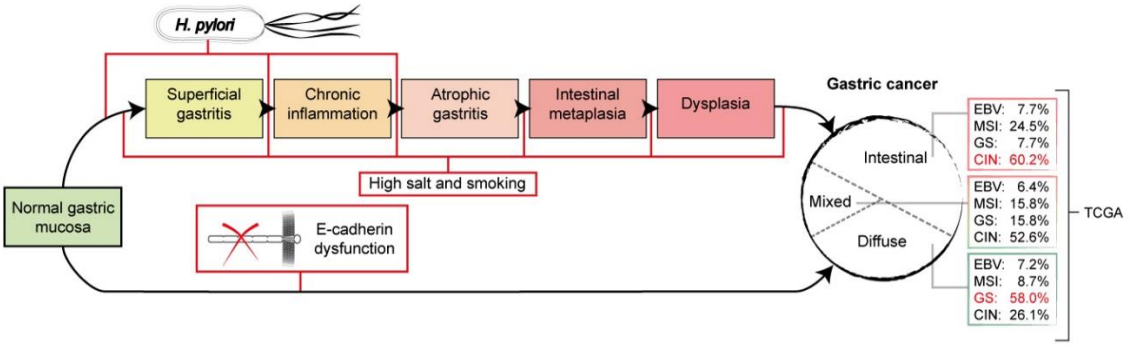


Figure 3. Carcinogenesis of intestinal and diffuse type gastric cancer and frequent drivers of progression.

Colonization of the stomach with certain virulent *H. pylori* strains, malnutrition and smoking can trigger gastritis and eventually chronic inflammation of the stomach lining. The continuous inflammation over a long period of time may first result in atrophic gastritis, a condition in which the stomach epithelium turns dysfunctional and then in intestinal metaplasia, where the gastric epithelial cells are replaced by intestinal epithelial cells. Intestinal metaplasia may eventually result in dysplasia, a premalignant condition characterized by disturbed differentiation and growth of epithelial cells. Genetic factors of the host may contribute to the progression, and tumors that arise from this cascade are often of intestinal type. The diffuse type gastric cancer arises spontaneous through mutations, and are often characterized by dysfunctional E-cadherin. The distribution of the cancer genome atlas (TCGA) classification [34] is shown to each histological subtype of the Lauren classification. EBV, Epstein-Barr virus positive tumor; MSI, microsatellite instable tumors; GS, genomic stable tumors; CIN, chromosomal instable tumors.

1.1.4 CURRENT DIAGNOSIS AND TREATMENT

In early stages gastric cancer remains often asymptomatic and therefore undiagnosed. The most common symptoms that are presented in advanced stages are weight loss, abdominal pain, nausea and loss of appetite. In the clinical practice, endoscopy remains the gold standard for the diagnosis, as no specific serological test for the screening exist. After endoscopic diagnosis with biopsy, staging (Figure 4) is further determined by endoscopic ultrasound and laparoscopy. Although very early stages (until T1a) may be removed by endoscopic resection, the majority of patients are subject to total or partial gastrectomy including the removal of regional lymph nodes [35].

In addition to surgery, chemotherapy or chemoradiation is usually applied. There is no global consensus about the optimum strategy. Depending on the country, these treatments are given either pre-, post- or perioperative with more than a dozen chemotherapy drugs being approved by national authorities that can be employed individually or in combination [35]. The exact therapy depends on many factors, such as patients' health conditions, location and stage of the tumor as well as country of treatment [35].

A new trend in cancer management is targeted therapy by applying monoclonal antibodies or small molecule inhibitors against aberrant molecular alterations of the given patients' cancer. In gastric cancer, Trastuzumab and Ramucirumab are two antibodies currently applied to target the oncogenic overexpression of human epidermal growth factor receptor 2 (HER2) and vascular endothelial growth factor receptor (VEGFR), respectively [35]. Numerous other antibodies and small molecule inhibitors are being evaluated for gastric cancer in ongoing clinical studies [36].

		Patients	5-year-survival
Stage 0 (*in situ*)			
Tis	N0	n/a	> 90%
Stage IA			
T1	N0	8,47%	70.8 %
Stage IB			
T2	N0	4,65%	57.4 %
T1	N1		
Stage IIA			
T3	N0	8,24%	45.5 %
T2	N1		
T1	N2		
Stage IIB			
T4a	N0	8,48%	32.8 %
T3	N1		
T2	N2		
T1	N3		
Stage IIIA			
T4a	N1	7,31%	19.8 %
T3	N2		
T2	N3		
Stage IIIB			
T4b	N0	11,78%	14.0 %
T4b	N1		
T4a	N2		
T3	N3		
Stage IIIC			
T4b	N2	7,47%	9.2 %
T4b	N3		
T4a	N3		
Stage IV			
M1, any T and N		43,61%	4.0 %

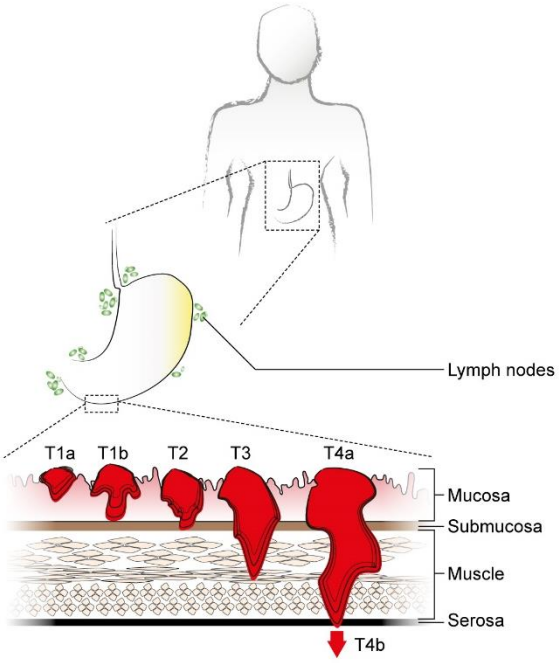


Figure 4. TNM staging system of gastric cancer according to the 7th edition of the American Joint Committee on Cancer (AJCC).

For each stage percentage of patients diagnosed and 5-year-survival rates are shown. The percentages were calculated from 10.601 surgically resected gastric adenocarcinomas of the Surveillance, Epidemiology, and End Results (SEER) database and extracted from [37]. T defines tumor invasion as shown at the right side. N categorizes the amount of lymph node metastases with N0: 0; N1: 1-2; N2: 3-6 and N3: >7 lymph nodes. M1 states the presence of distant metastases or tumor cells in the peritoneal fluid.

The three most used serological biomarker for gastric cancer so far are CA 19.9, CEA and CA 72.4 [38-40]. CA 19.9 detects the glycan structure sialyl Lewis A, which is commonly upregulated in gastric cancer. This marker cannot be used for diagnostic purposes, as benign diseases often overexpress sialyl Lewis A as well, but it is used in the clinical practice to monitor recurrence and response to therapy in gastric cancer patients [39, 40]. The CA 72.4 immunological assay detects elevated levels of sialyl Tn in the serum of gastric cancer patients. Similarly to CA 19.9, precancerous lesions

elevate STn serum levels and therefore cause a high false-positive rate of the CA 72.4 assay, preventing its application for gastric cancer diagnosis [41, 42]. The detection of this marker can be used though as an independent prognostic factor associated with aggressiveness, poor prognosis and tumor recurrence. The CEA serological assay detects the highly glycosylated CEA protein, which is commonly overexpressed in gastric cancer. However, sensitivity limitations of the CEA serological assay prevent early stage disease detection [43].

1.2 PROTEIN GLYCOSYLATION

Glycosylation is one of the most abundant and diverse forms of post-translational modifications (PTM) of proteins. In fact, approximately 60% of all naturally occurring proteins are believed to be glycosylated [44]. In human, around 2% of all genes are involved in the glycosylation process [45]. Glycosylation is non-template driven, solely controlled by the complex interplay of glycosyltransferases (monosaccharide adding enzymes) and glycosidases (glycan cleaving enzymes) and the availability of their substrates [46]. Ten monosaccharide building blocks are being used in human for the biosynthesis of glycans with various interlinkage possibilities, resulting in the vast diversity of glycans that we find (Figure 5, [47]).

We can distinguish between several types of protein glycosylation depending on the nature of the sugar-peptide bond and the oligosaccharide attached (Figure 5). The sugar-peptide bond can be *N*-, *O*- or *C*-linked; of which *N*-glycosylation (linked to the amine of asparagine) is the most studied. A variety of saccharides can be linked to *O*- (linked to the hydroxyl of usually Serine or Threonine), such as xylose, fucose, mannose, glucose, N-acetylgalactosamine (GalNAc) or N-acetylglucosamine (GlcNAc) [46]. The work presented in this thesis is focused on *N*-glycans, *O*-GalNAc glycans (also called "mucin type *O*-glycans" or simply "*O*-glycans") and to a lesser extent on glycosaminoglycans such as heparan sulfate and chondroitin sulfate.

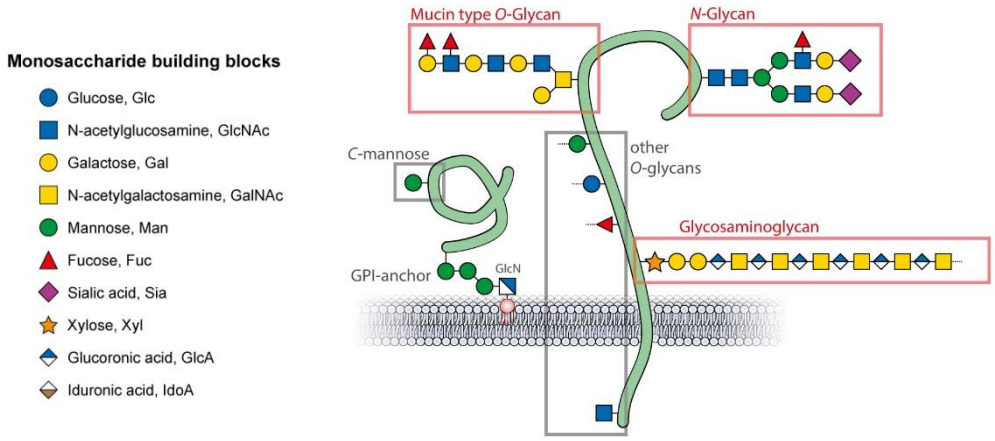


Figure 5. Types of protein glycosylation and their monosaccharide building blocks.

Most glycosylation types are *O*-linked to serine, threonine or tyrosine. So far 5 different types of *O*-linked glycosylation have been identified in humans: *O*-GalNAc glycosylation, *O*-mannosylation, *O*-glucosyation, *O*-fucosylation and *O*-GlcNAcylation. The glycosaminoglycan structures heparan sulfate, heparin, chondroitin sulfate and dermatan sulfate are also *O*-linked via xylose but are usually not considered *O*-glycans. *N*-glycosylation is linked to asparagine and represents the most studied type of glycosylation. Furthermore, proteins can be modified with *C*-linked mannose to tryptophan or C-terminal anchored to the glycan part of glycosylphosphatidylinositol (GPI). Framed with a red box are those protein glycosylation types that have been studied in this thesis.

Just as the structures, so are the functions of protein glycosylation manifold. In the secretory pathway glycosylation is important for protein folding and for the monitoring of protein folding as well as for protein sorting [48]. Glycans function as modulator of protein activity and interaction [49, 50]. It protects from proteolysis and may mask protein domains [51, 52]. Independent of its conjugate, glycans can act as binding sites for proteins, cells as well as pathogens, and may function as scaffold and structural element in the extra-cellular space [53-56].

1.2.1 *O*-GLYCOSYLATION

Mucin type *O*-glycosylation (in this thesis henceforth referred to as *O*-glycosylation) is initiated in the Golgi by the addition of *N*-Acetylgalactosamine (GalNAc) to serine or threonine in a non-consensus-sequence based process by one of 20 polypeptide GalNAc transferases (ppGalNAcT) [57]. The *O*-glycosylation pattern and site occupancy is defined by the tissue specific expression signature of ppGalNAcTs [57]. The term mucin type *O*-glycosylation is deceptive, as almost every protein that passes through the secretory pathway can be modified by *O*-GalNAc [58], however proteins with serine and threonine rich domains and/or local prolines are more efficiently *O*-glycosylated. Therefore, *O*-glycans often occur densely clustered as it is the case on mucins [58].

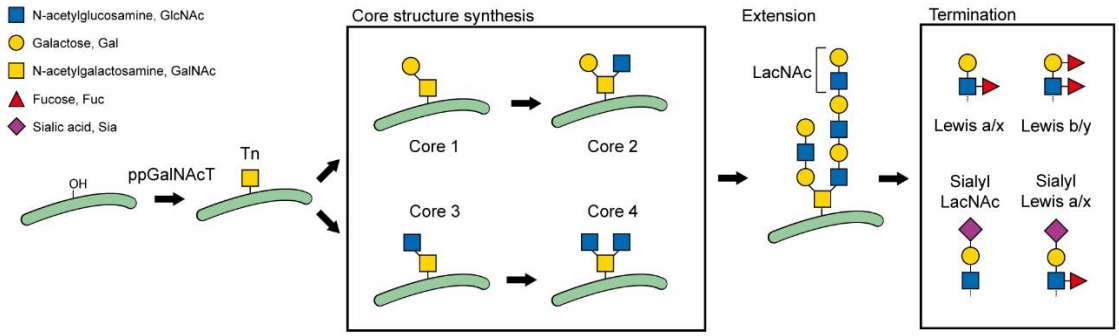


Figure 6. Major steps of the *O*-glycan biosynthesis.

Mucin type *O*-glycosylation is a multistep process initiated by the covalent addition of GalNAc to serine or threonine by a ppGalNAcT, followed by the formation of a core structure. These core structures are further extended and finally capped by a terminal structure. For image simplicity not all core and terminal structures are depicted.

The initial *O*-GalNAc (known as Tn antigen) is sequentially extended throughout the Golgi, forming at first a core-structure (Figure 6). There are four common *O*-glycan structures in human, numbered from 1 to 4 [59]. The core 1 structure is formed by the addition of β1-3 linked Gal through the glycosyltransferase C1GalT1 [59]. The

core 1 structure is also known as T antigen and is under healthy conditions further modified. This can be either by sialylation leading to sialyl-T and disialyl-T, which are abundant epitopes on serum proteins, by elongation or by the formation of the core 2 structure [60]. Core 2 is synthesized by the addition of β1-6 GlcNAc to the GalNAc of core 1 through C2GnT1, 2 or 3 [59]. In the gastric mucosa *O*-glycans are predominantly based on core 1 and 2 [61]. Alternatively to the core 1/2 synthesis, Tn antigen can be basis for the formation of core 3 and 4 by adding 1 or 2 GlcNAc residues to Tn respectively. Core 3 and 4 can be found in substantial amounts in the intestine [62], and their expression might therefore arise in the stomach during the process of intestinal metaplasia. Core structures can be further elongated by successive addition of GlcNAc and Gal, resulting in *N*-acetyllactosamine (LacNAc) units that are interlinked either β1-3 or β1-4 and referred to as type 1 or 2 extension, respectively. These *O*-glycans can be additionally decorated with a variety of functional groups such as fucoses, sialic acids and sulphates, which often precludes further chain elongation by forming a so called terminal epitope [63].

1.2.2 *N*-GLYCOSYLATION

The *N*-glycosylation process starts in the endoplasmic reticulum (ER) by a co-translational *en bloc* transfer of the oligosaccharide precursor to asparagine [64]. The consensus sequence for the *N*-glycosylation is Asn-Xxx-Ser/Thr, where Xxx can be any amino acid except proline. The precursor glycan (Figure 7) is subject to a series of trimming reactions in the ER and Golgi. The terminal 3 glucoses are essential for the protein quality control process in the ER and are fully removed as the protein backbone enters a correctly folded state [48]. Once in the Golgi, α-mannosidase 1 further removes several mannoses with variable efficiency, leaving typically 8 to 5 mannoses behind. The liberation of the lower *N*-glycan arm allows the addition of a GlcNAc through GnT-1, which is the prerequisite for α-mannosidase 2 to remove the remaining 2 mannoses that are not part of the core structure. As the upper *N*-glycan

arm is liberated as well the GlcNAc transferase GnT-2 can act to form a complex *N*-glycan. Incomplete removal of mannoses results in high-mannose and hybrid *N*-glycans (Figure 7) [64].

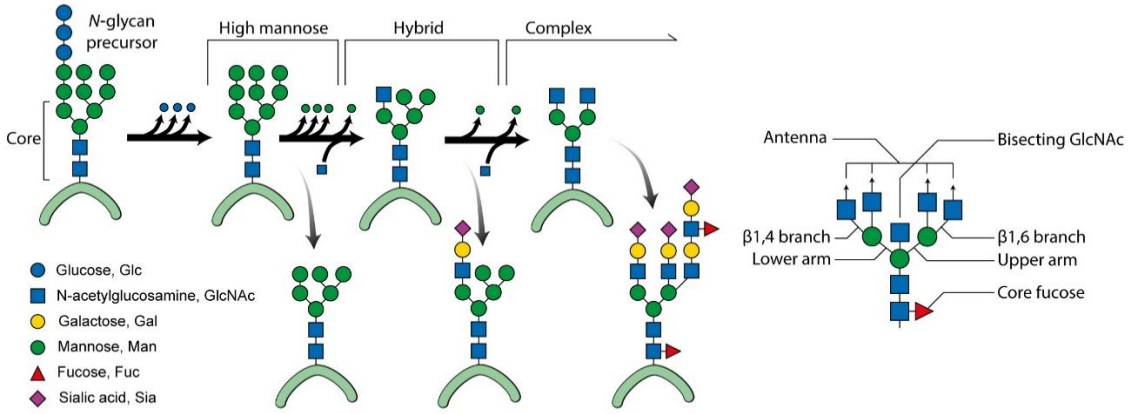


Figure 7. *N*-glycosylation processing.

The precursor *N*-glycan is transferred to the polypeptide backbone of the glycoprotein. The precursor is processed as the glycoprotein passes through the secretory pathway. At first the three glucoses are removed, followed by the removal of mannoses. Liberating the mannoses of the core structure provides the basis for the formation of so called antenna, and may result in hybrid or complex *N*-glycans. Further structural elements that can be added to hybrid and complex *N*-glycans are bisecting GlcNAc, core-fucoses and additional branches. Examples of typical high mannose, hybrid and complex structures are shown.

Complex *N*-glycans formed by GnT-1 and -2 are called bi-antennary. Additional branches can be initiated by GnT-4 and GnT-5 on the lower and upper arm respectively to form tri- or tetra-antennary *N*-glycans [64]. The GlcNAc antennas of hybrid and complex *N*-glycans are sequentially elongated forming LacNAc extensions and decorated with functional groups, such as sialic acids and fucoses, with the possible formation of terminal epitopes. Unique structural features that can be added

to *N*-glycans are the bisecting GlcNAc and the core-fucose (Figure 7) [65]. The bisecting GlcNAc is added in between the lower and upper N-glycan arm by GnT-3 and is not further modified. It prevents the addition of the upper arm branching by GnT-5. The core-fucose is added by FuT-8 [66].

N-glycans can modulate the activity and stability of proteins. The potential of *N*-glycan to tune protein functions is well exemplified by immunoglobulins, where different *N*-glycan moieties significantly alter pharmacokinetics and pharmacodynamics and modulate the triggered immune response [67].

1.2.3 GLYCOSAMINOGLYCANS

Glycosaminoglycans (GAGs) are long unbranched sugar chains consisting of disaccharide repeats [68]. These disaccharide building blocks contain one N-acetylhexosamine (such as GlcNAc or GalNAc) and one uronic acid (glucoronic acid or iduronic acid) or galactose and are frequently sulfated. Based on the composition of the repeat, GAGs can be grouped into 5 classes: Hyaluronan (HA), chondroitin sulfate (CS), dermatan sulfate (DS), heparan sulfate/heparin (HS) and keratansulfate (KS).

HA is distinguished from the other GAGs in that it is not covalently attached to any protein or peptide nor is it sulfated [69]. Further, HA is the only GAG that is not synthesized in the ER or Golgi, but at the inner surface of the plasma membrane where the GAG chain grows through the membrane into the extracellular space. It is a major component of the extracellular matrix (ECM) and anchors a wide range of proteins [69]. The other classes of GAGs are linked to a protein or peptide, forming a so called proteoglycan, and are synthesized in the ER and Golgi.

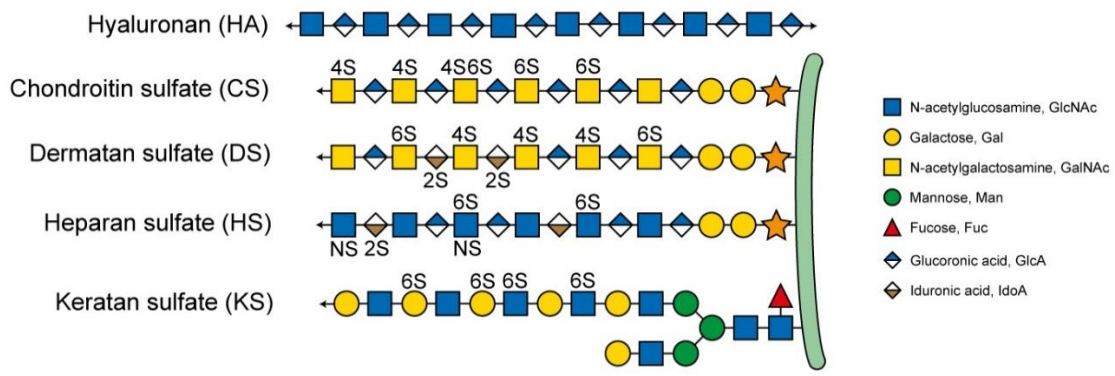


Figure 8. Glycosaminoglycan chains.

Hyaluronan is the only human GAG that is not protein linked. Its a faithful GlcNAc β1,4 GlcA β1,3 repeat that is not further modified and may reach up until 7 MDa. Keratan sulfate is a sulfated poly-LacNAc repeat on *N*-glycans (KS-1) and *O*-glycans (KS-2). The other GAG chains are attached to a protein or peptide via the linkage tetra saccharide GlcAβ1,3Galβ1,3Galβ1,4Xyl. Chondroitin sulfate is a poly sulfated GalNAc β1,4 GlcA β1,3 repeat. Dermatan sulfate is derived from chondroitin sulfate and has interspersed GalNAc β1,4 IdoA α1,3 dimers. Heparan sulfate is a combination of commonly sulfated GlcNAc α1,4 GlcA β1,4 dimers and GlcNAc α1,4 IdoA α 1,4 dimers. Sulphate groups and linkages are indicated by S and prefix, respectively.

Widely expressed in epithelial cells are CS and HS, which dominate mass and biochemical properties of the carrying protein due to their great length, often over 80 sugar residues, and strong negative charge [70]. CS and HS consist of repeating sulfate-substituted GalNAc-GlcA or GlcNAc-GlcA disaccharide respectively (Figure 8). Both are linked to serine residues in core proteins by way of xylose. A perfect consensus sequence does not exist, and the decision whether the chain elongates to heparan sulfate or chondroitin sulfate appears to be manifested at the level of enzyme recognition of the polypeptide substrate [68]. Heparan sulfate proteoglycans (HSPGs) are prevalently associated with the cell surface or found in the pericellular matrix [71]. The HSPGs function as major biological modifiers of growth factors such as FGF, VEGF

and PDGF among others in addition to their ability to interact with each other and with key constituents of the basement membrane, including various laminins, collagen type IV, and nidogen [70]. The presentation of growth factors to their receptors is a major function of cell surface HSPGs. As most Chondroitin-containing proteoglycans (CSPGs) are secreted, they predominate over HSPGs distal of the cells [71]. These proteoglycans function as structural constituents of the ECM and provide viscoelastic properties, retain water, keep osmotic pressure and dictate proper collagen organization.

1.2.4 ABERRANT GLYCOSYLATION IN GASTRIC CANCER

Protein glycosylation is predominantly found on membrane-bound glycoproteins, contributing to the glycocalyx, or on secreted glycoproteins, which can become integral parts of the ECM. Thus protein glycosylation are in the position to mediate cell adhesion and motility, as well as intracellular signaling events [72]. Changes in glycosylation associated with oncogenic transformation were first described over more than six decades ago [73]. Since then, the functional impact of cancer associated glycan epitopes has been associated with several cancer hallmark abilities such as tumor cell migration, invasion, proliferation, angiogenesis and metastasis [74]. In this regard, alterations of protein glycosylation have shown to be able to directly and indirectly activate cell receptors or facilitate the receptor-ligand interaction [75, 76]. Due to altered glycosylation tumor cells promote interaction with platelets, leukocytes and endothelial cells leading to increased haematogenous metastatic capacities [77]. Adhesion molecule functions are modulated by glycans. For instance, certain glycan structures on adhesion molecules decrease cell-cell and cell-matrix adhesion and promote cancer cell migration [78, 79].

In gastric cancer, major events of glycan alteration are the altered expression of ppGalNAcTs, truncation of *O*-glycans, changes in *N*-glycan branching, and increase in sialylation and fucosylation (Figure 9).

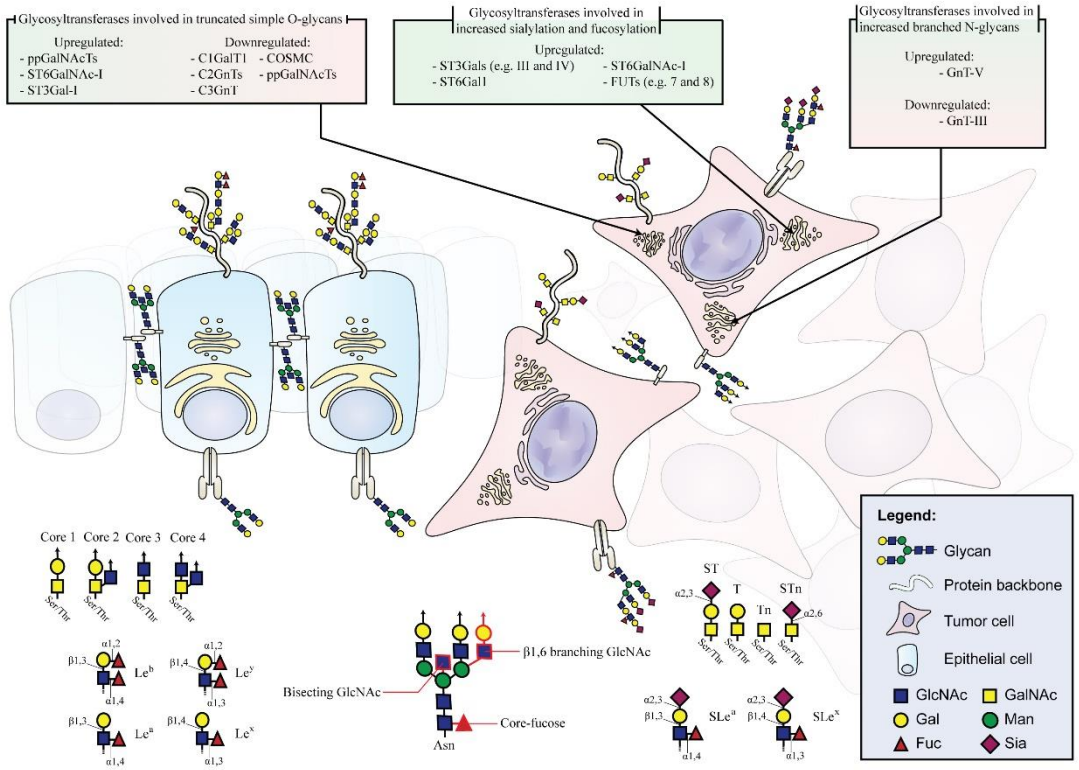


Figure 9. Gastric cancer associated glycan alterations.

Transformed cells show numerous alterations of the glycosylation machinery leading to the expression of aberrant glycan epitopes. Key glycosyltransferases of this process are shown on the top. Aberrantly expressed glycan epitopes are on the bottom right as opposed to glycan epitopes associated with healthy gastric epithelium left. (Extracted from Mereiter *et al.* 2016 [80])

The expression profile of ppGalNAcTs is tissue specific and controls the *O*-glycosylation site occupancy and density [57]. The ppGalNAcT expression profile and subcellular localization is altered in cancer [81]. The upregulation of specific members of this transferase family, such as ppGalNAcT-2, -3, -6 and -10, is associated with gastric cancer progression [82-85]. These alterations can

influence *O*-glycan structures through altering *O*-glycan density and steric accessibility [81].

O-GLYCAN TRUNCATION: The truncation of *O*-glycans results in increased Tn (GalNAcα1-Ser/Thr) and T antigen (also known as Thomsen-Friedenreich antigen or core 1 structure, Galβ1-3GalNAcα1-Ser/Thr) and their sialylated forms, STn (Neu5Acα2-6GalNAcα1-Ser/Thr) and ST (Neu5Acα2-3Galβ1-3GalNAcα1-Ser/Thr), respectively [86]. The underlying mechanisms are expression changes of glycosyltransferases or chaperones, glycosyltransferase relocation to the ER or aberration of structure and pH of the Golgi [87-91].

The Tn epitope is in the healthy gastric epithelium a biosynthetic intermediate located in the secretory pathway and is readily elongated before reaching the cell surface. An increase in Tn and its exposure on cell surface and on secreted proteins is associated with poor prognosis and survival. Several mechanisms are known that cause an accumulation of Tn: (i) prevention of core 1 elongation either by a direct down regulation of the core 1 synthetase (*C1GALT1*) or by the down regulation of its dedicated chaperone COSMC [92]; (ii) relocation of ppGalNAcTs from the Golgi to the ER [87]; (iii) deviation of the normal intra-Golgi pH [88].

The situation is similar for STn, which is absent from healthy gastric epithelium but occurs highly expressed in most gastric cancers, correlating with cancer cell adhesion, increased cancer cell invasion, and poor prognosis of the patients [91, 93-95]. The upregulation of Tn and STn often arise concomitant in cancer since mechanisms that lead to the accumulation of Tn imply an increase of substrate for the STn biosynthesis. A pathway that independently of the aforementioned mechanisms of Tn upregulation can lead to the oncogenic increase of STn is the overexpression of ST6GALNAC1 [91, 93]. The upregulation of ST6GALNAC1, the sialyltransferase that leads to STn

formation, can override the normal *O*-glycan elongation, leading to increased STn formation and inducing migration and invasion in gastric carcinoma cells [94].

INCREASED *N*-GLYCAN BRANCHING: Two structural features of complex *N*-glycans are the β1,6-branching, catalyzed by the glycosyltransferase GnT-V, and the bisecting-GlcNAc, added by the glycosyltransferase GnT-III. GnT-V is often upregulated in gastric carcinoma, leading to increased N-glycan branching and contributes to cancer cell invasion and the formation of metastases [96-98]. Another mechanism leading to increased branching is the downregulation of GnT-III and the decrease of bisecting-GlcNAc. Bisected *N*-glycans cannot be modified by GnT-V and preclude β1,6-branching. The interplay of GnT-III and GnT-V determines *N*-glycan branching, which has been shown to impact cell-cell and cell-matrix interactions in gastric cancer [99, 100]. Especially the alterations of branching on the *N*-glycans of E-cadherin and integrins have been linked to the migratory capacity of gastric cancer cells. [78, 79]

TERMINAL SIALYLATION AND SIALYL-LEWIS EPITOPES: The glycans of the healthy gastric mucosa are often terminated by fucosylated antigens such as Lewis A and Lewis B in the superficial glands and Lewis X and Y in the deeper glands. In gastric cancer on the other hand, glycans carry frequently sialylated epitopes and show a decline in neutral Lewis epitopes. As a terminal event, sialylation caps glycosylation chains resulting in exposed locations of its negative charge at the forefront of glycans and first encounter point for adjacent entities. Sialylation has, therefore, been shown to play important roles in modulating cellular recognition, cell adhesion, and cell signaling [101]. An increase in global sialylation owing to altered glycosyltransferases expression, has been closely associated with cancer and commonly described as one of the main modifications in gastric cancers [102, 103]. The major α2,3-sialylated antigens associated with cancer are sialyl Lewis A (SLe^a) and sialyl Lewis X (SLe^x). Although these structures can also be present in non-neoplastic cells, SLe^a and SLe^x

have been demonstrated to be highly expressed in many malignant tissues, including gastric tumors, correlating with poor survival for the patients [42, 104]. SLe^x functions as a ligand for selectins [105], cell adhesion molecules that mediate leukocyte-endothelial binding. In cancer, SLe^x expression of tumor cells increases the metastatic potential because of the favorable endothelial arrest of circulating tumor cells [77]. SLe^a, as previously described, is used as a cancer-associated marker in the clinical practice [39]. It is assessed by the serological assay CA 19.9 and elevated levels are associated with poor prognosis in gastric cancer [106].

1.3 KEY GLYCOPROTEINS IN GASTRIC CANCER

The majority of all secreted and membrane associated proteins is glycosylated. These glycoproteins and proteoglycans are key in fundamental cellular processes and contribute therefore often to carcinogenesis and cancer progression. Among the most frequently altered oncogenes are the receptor tyrosine kinases (RTKs), a family of glycosylated cell surface receptors whose upregulation and oncogenic activation are drivers of transformation in gastric cancer [107]. This class of glycoproteins are triggers of cancer cell migration, tumor growth, epithelial-mesenchymal transition and angiogenesis. RTKs involved in gastric cancer are human epidermal growth factor receptor 2 (HER2), epidermal growth factor receptor (EGFR), MET (from mesenchymal-epithelial transition factor), vascular endothelial growth factor receptor (VEGFR), fibroblast growth factor receptor 2 (FGFR2), platelet-derived growth factor receptors (PDGFR) and RON (from Recepteur d'Origine Nantais) [108-114].

Other frequently altered glycoproteins at the forefront of cells are adhesion molecules, such as integrins, syndecans, cadherins and CD44 [115]. Most of these adhesion proteins function also as co-receptors and thus have the ability to modulate signal transduction [116]. The glycoprotein CD44, for instance, incorporates adhesion and RTK co-receptor functions and finds itself at the intersection of multiple oncogenic processes [117].

In order to understand the role of glycan alterations in cancer it is of utmost importance to identify glycoproteins that are modified by aberrant glycosylation and to unravel their functional implications.

1.3.1 RON

RON (Recepteur d'Origine Nantais, also known as macrophage stimulating protein receptor, MSPR and MST1R) is a receptor tyrosine kinase (RTK) and constitutively expressed in many epithelia. RON has been shown to be commonly overexpressed and/or hyper-activated in gastric cancer leading to significantly worse survival rates of patients [118]. Oncogenic RON expression contributes in tumorigenesis, malignant progression, angiogenesis and chemoresistance [119-123]. In the last years numerous companies have been developing small molecules or antibodies to target RON with several being currently in clinical and pre-clinical trials [124-127].

The mature RON receptor is a single-pass type I transmembrane protein and consists of a 35 kDa α-chain and a 145 kDa β-chain linked by a disulphide bond. The β-chain comprises a large extracellular segment with 3 IPT (immunoglobulin-like-plexin-transcription) domains, a short transmembrane helix, and a cytoplasmic portion harboring a tyrosine kinase domain (Figure 10) [128]. RON is normally activated by a complex interaction with its ligand, the macrophage-stimulating protein (MSP) [129]. Several mechanisms have been identified that by-pass the MSP-dependent RON activation which are being especially exploited by cancer cells. In this regards, the IPT domains are of particular importance. Alteration such as deletion and insertions in the IPT domains are found in cancer and have been shown to cause constitutive activation of RON [130].

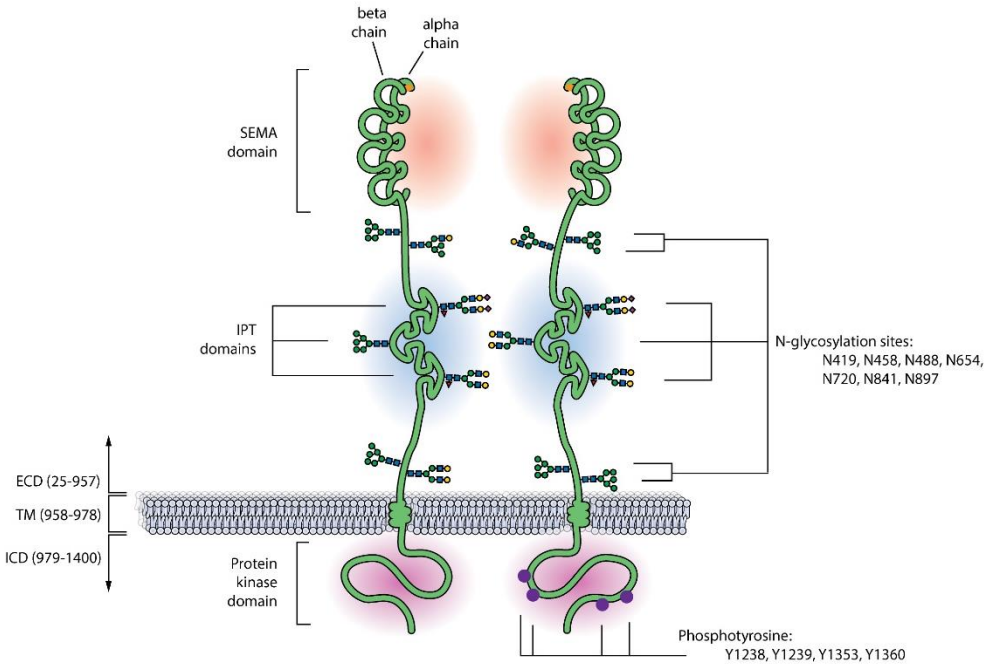


Figure 10. Illustration of RON receptors and protein domains.

Cleavage site in the SEMA domain leads to formation of α-chain and β-chain, which remain connected through a disulfide bond. This domain contains the MSP binding site (highlighted in red). MSP is the major ligand for RON activation and acts through dimerization of RON receptors. RON contains three IPT domains in the stem region (highlighted in blue), and a cytosolic protein kinase domain with 4 phosphotyrosines (highlighted in purple). Putative *N*-glycosylation sites are shown. ECD = Extracellular domain; TM = Transmembrane domain; ICD = Intracellular domain; IPT = immunoglobulin-like-plexin-transcription.

The RON receptor has 7 putative *N*-glycosylation sites of which each of the IPT domains is predicted to carry one. One proven *N*-glycan site (Asn488) has been shown to be essential for the proper formation of the MSP-binding groove [131].

1.3.2 CD44

CD44 is an exceptionally versatile glycoprotein and because of its diversity often described as a whole protein family encoded by a single gene [132, 133]. The molecular weight of CD44 varies in human between 80 and more than 250 kDa. This heterogeneity is generated by alternative splicing, its extensive glycosylation and proteolytic cleavages.

Alternative splicing allows the insertion of so-called variable exon products in the extracellular, membrane proximal region. Due to its 9 variable exons, named v2 to v10, hundreds of splice-variants of CD44 exist, each with a potentially different function [134-136]. A single cell can express different CD44 isoforms at the same time and many signaling mechanisms have been shown to alter this splicing signature [137], making CD44 a sentinel for cellular and environmental alterations. Not surprising, we therefore find numerous alterations of CD44 splicing in cancer. CD44 came under scrutiny as specific splice variants, such as CD44v3 and CD44v6, were shown to be highly upregulated in numerous cancers, including gastric cancer [138, 139].

Proteolytic cleavages of CD44 have been emerging as key regulatory events for the CD44 dependent cell-matrix interaction and signaling pathway. CD44 undergoes sequential proteolytic cleavages in the ectodomain and intramembranous domain, resulting in the release of a CD44 intracellular domain (CD44ICD) fragment [140]. The ectodomain cleavage of CD44 can be triggered by several proteases including MMP9 and ADAM10 and contributes to the regulation of cell attachment to and migration on HA matrix [141, 142]. The ectodomain cleavage subsequently induces the intramembranous cleavage, which is mediated by presenilin (PS)-dependent gamma-secretase. The intramembranous cleavage generates CD44ICD, which is translocated to the nucleus and activates transcription [143].

CD44 is highly glycosylated and a major part of its mass is derived from glycans [144, 145]. It has 9 putative *N*-glycosylation sites and more than 100 *O*-glycan sites predicted *in silico*. 41 *O*-glycan sites were recently confirmed [58]. Additionally, CD44 may carry heparan sulfate and chondroitin sulfate chains [146].

CD44 is the main receptor for hyaluronan (HA) [136]. The HA binding site is located at the N-terminal extracellular part and is constantly expressed in CD44 [147]. Allthough the HA binding domain is not subject to alternative splicing, the different CD44 splice variants have an impact on HA binding, with some variants leading to a loss of HA binding and others enhancing it [136, 148]. Glycosylation of CD44 was also shown to alter HA binding [136, 148, 149]. The CD44-HA interaction regulates cell-cell and cell-matrix adhesion and thus effects cell migration as well as cell growth and differentiation [147]. The second major function of CD44 is the modulation of receptors activation through several mechanisms: (i) Heparansulfate chains of CD44 bind growth factors such as FGF [149, 150]; (ii) CD44 can be platform for MMP7 binding and thereby facilitates proteolytic activation of receptor ligands such as HB-EGF [151]; (iii) Specific isoforms, such as CD44v6, act directly as co-receptor for Met, VEGFR-2 and RON [152, 153]; (iv) CD44 supports dimerization of RTKs such as HER2 and HER3 [154].

2. AIMS AND OBJECTIVES

2.1 RATIONALE AND GENERAL AIMS

Glycosylation alterations are a common feature in gastric cancer, and an increased body of evidence supports the role of glycan alterations in cancer cell biology, including the modulation of the process of cancer development and progression. This altered glycosylation can be found on specific proteins in cancer cells, modulating the functional role of these proteins and constituting an excellent source of biomarkers for screening and prognosis purposes.

The general aim of the present work is the characterization of glycosylation expression profile of gastric cancer cells and the identification of aberrantly glycosylated glycoprotein targets in order to unravel the role of these glycan alterations in carcinogenesis and cancer progression.

2.2 SPECIFIC OBJECTIVES

1. CHARACTERIZATION OF GASTRIC CANCER CELL GLYCOPHENOTYPE

The disruption of the glycosylation machinery leads to the synthesis of aberrant glycan epitopes in gastric cancer. In this process, alterations in expression levels of glycosyltransferases, such as ST3GAL4 or ST6GALNAC1, are commonly involved. In chapter 2 and 3 we aimed to assess how these glycosyltransferases modulate the glycome and glycoproteome of cancer cells. We performed *N*– and *O*–glycomic analyses based on mass spectrometry and fluorescent labeled glycans to demonstrate structural alterations induced by these two sialyltransferases.

2. UNRAVEL NEW MECHANISMS OF ONCOGENIC RECEPTOR ACTIVATION MEDIATED BY GLYCOSYLATION MODIFICATION

Glycan alterations are associated with cancer hallmarks, and have been proven to particularly impact cancer cell adhesion, migration, invasion and metastasis. Despite the large number of publications that report glycosylation association with the malignant phenotype, our understanding of the underlying molecular mechanisms is fairly limited. In chapter 2 and 3 we aimed to identify glycoproteins whose altered glycosylation contribute to more malignant phenotypes. We applied glycoproteomic analysis on gastric carcinoma cells with and without ST3GAL4 overexpression and identified protein targets that were modified with altered *N*-glycans. We assessed one of the targets, the RON receptor tyrosine kinase, for functional alterations. Furthermore, we evaluated the co-expression of RON and SLe^x in human gastric carcinomas with prospect of applicability as biomarker.

3. EVALUATE THE BIOLOGICAL IMPACT OF *O*-GLYCAN TRUNCATION ON CRITICAL CANCER CELL PROTEINS

The expression of short truncated carbohydrate antigens was one of the first evident glycan alteration described in cancer. Nowadays, we know that the *O*-glycan truncation can be achieved through various mechanisms with each of these mechanism being accompanied by the expression of a subset of truncated *O*-glycan epitopes. Therefore, we hypothesize that not only the expression of truncated glycan epitopes, but also the lack of *O*-glycan extension modulates fundamental molecular properties conveying advantageous cancer phenotypes. In chapter 4, we aimed to study the effect of truncated *O*-glycans on two major proteins expressed in gastrointestinal tissues, CD44 and Syndecan 1. We used three glyco-engineered gastric cancer cell line models: overexpressing ST3GAL4, overexpressing ST6GALNAC1 or COSMC knock-out for this evaluation. These model cell lines express

truncated *O*-glycans induced through different molecular mechanisms and are therefore of avail in search of fundamental molecular properties that *O*-glycan truncation may trigger.

4. DEFINE SPECIFIC GASTRIC GLYCAN SIGNATURE WITH POTENTIAL BIOMARKER APPLICATIONS

Good gastric cancer biomarkers for prognosis and patient stratification are urgently needed. Single molecules have not met the clinical requirements of specificity and sensitivity so far, leading to a paradigm shift in the field. Biomarkers are being combined to panels, increasing their specificity and enabling alteration signatures to more reliably predict clinical implications. Glycosylation plays already a pivotal role as gastric cancer marker, with CA 19.9 (assay for sialyl Lewis A) being an FDA approved biomarker to follow treatment response and to detect cancer recurrence and CA 72.3 (assay for sialyl Tn) being an assay with predictive potential for the outcome of gastric cancer patients. However, these biomarkers have limited screening application. To improve the applicability of glycan alterations as biomarker in gastric cancer we aimed to examine, in chapter 5, the potential of a panel of glycan alterations for biomarker application. For this reason we analyzed the expression of Tn, sialyl Tn, sialyl Lewis X and sialyl Lewis A in a cohort of 80 gastric carcinoma samples and associated the expression signature to clincopathological data of the patients and molecular alterations of the tumor cells.

3. REFERENCES

1 Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., . . . Bray, F. (2015) Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 136, E359-386

2 Global Burden of Disease Cancer, C., Fitzmaurice, C., Dicker, D., Pain, A., Hamavid, H., Moradi-Lakeh, M., . . . Naghavi, M. (2015) The Global Burden of Cancer 2013. JAMA Oncol. 1, 505-527

3 Ferro, A., Peleteiro, B., Malvezzi, M., Bosetti, C., Bertuccio, P., Levi, F., . . . Lunet, N. (2014) Worldwide trends in gastric cancer mortality (1980-2011), with predictions to 2015, and incidence by subtype. Eur J Cancer. 50, 1330-1344

4 Hanahan, D. and Weinberg, R. A. (2011) Hallmarks of cancer: the next generation. Cell. 144, 646-674

5 Howlader N, N. A., Krapcho M, Miller D, Bishop K, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA. (2016) SEER Cancer Statistics Factsheets: Stomach Cancer. ed.)^eds.), National Cancer Institute. Bethesda, MD

6 Gonzalez, C. A., Pera, G., Agudo, A., Palli, D., Krogh, V., Vineis, P., . . . Riboli, E. (2003) Smoking and the risk of gastric cancer in the European Prospective Investigation Into Cancer and Nutrition (EPIC). Int J Cancer. 107, 629-634

7 Ladeiras-Lopes, R., Pereira, A. K., Nogueira, A., Pinheiro-Torres, T., Pinto, I., Santos-Pereira, R. and Lunet, N. (2008) Smoking and gastric cancer: systematic review and meta-analysis of cohort studies. Cancer Causes Control. 19, 689-701

8 Gonzalez, C. A., Pera, G., Agudo, A., Bueno-de-Mesquita, H. B., Ceroti, M., Boeing, H., . . . Riboli, E. (2006) Fruit and vegetable intake and the risk of stomach and oesophagus adenocarcinoma in the European Prospective Investigation into Cancer and Nutrition (EPIC-EURGAST). Int J Cancer. 118, 2559-2566

9 Tsugane, S. (2005) Salt, salted food intake, and risk of gastric cancer: epidemiologic evidence. Cancer Sci. 96, 1-6

10 Howson, C. P., Hiyama, T. and Wynder, E. L. (1986) The decline in gastric cancer: epidemiology of an unplanned triumph. Epidemiol Rev. 8, 1-27

11 Levi, F., Lucchini, F., Gonzalez, J. R., Fernandez, E., Negri, E. and La Vecchia, C. (2004) Monitoring falls in gastric cancer mortality in Europe. Ann Oncol. 15, 338-345

12 Derakhshan, M. H. and Lee, Y. Y. (2012) Gastric cancer prevention through eradication of helicobacter pylori infection: feasibility and pitfalls. Arch Iran Med. 15, 662-663

13 Wang, C., Yuan, Y. and Hunt, R. H. (2007) The association between Helicobacter pylori infection and early gastric cancer: a meta-analysis. Am J Gastroenterol. 102, 1789-1798

14 Shiota, S., Suzuki, R. and Yamaoka, Y. (2013) The significance of virulence factors in Helicobacter pylori. J Dig Dis. 14, 341-349

15 David, S. and Meltzer, S. J. (2010) Stomach - Genetic and epigenetic alterations of preneoplastic and neoplastic lesions. Cancer Biomark. 9, 493-507

16 Shibata, D. and Weiss, L. M. (1992) Epstein-Barr virus-associated gastric adenocarcinoma. Am J Pathol. 140, 769-774

17 Young, L. S. and Rickinson, A. B. (2004) Epstein-Barr virus: 40 years on. Nat Rev Cancer. 4, 757-768

18 Iizasa, H., Nanbo, A., Nishikawa, J., Jinushi, M. and Yoshiyama, H. (2012) Epstein-Barr Virus (EBV)-associated gastric carcinoma. Viruses. 4, 3420-3439

19 Lauren, P. (1965) The Two Histological Main Types of Gastric Carcinoma: Diffuse and So-Called Intestinal-Type Carcinoma. An Attempt at a Histo-Clinical Classification. Acta Pathol Microbiol Scand. 64, 31-49

20 Lauren, P. A. and Nevalainen, T. J. (1993) Epidemiology of intestinal and diffuse types of gastric carcinoma. A time-trend study in Finland with comparison between studies from high- and low-risk areas. Cancer. 71, 2926-2933

21 Graziano, F., Humar, B. and Guilford, P. (2003) The role of the E-cadherin gene (CDH1) in diffuse gastric cancer susceptibility: from the laboratory to clinical practice. Ann Oncol. 14, 1705-1713

22 Carneiro, P., Fernandes, M. S., Figueiredo, J., Caldeira, J., Carvalho, J., Pinheiro, H., . . . Seruca, R. (2012) E-cadherin dysfunction in gastric cancer--cellular consequences, clinical applications and open questions. FEBS Lett. 586, 2981-2989

23 Qiu, M. Z., Cai, M. Y., Zhang, D. S., Wang, Z. Q., Wang, D. S., Li, Y. H. and Xu, R. H. (2013) Clinicopathological characteristics and prognostic analysis of Lauren classification in gastric adenocarcinoma in China. J Transl Med. 11, 58

24 Correa, P. (1992) Human gastric carcinogenesis: a multistep and multifactorial process--First American Cancer Society Award Lecture on Cancer Epidemiology and Prevention. Cancer Res. 52, 6735-6740

25 Correa, P. (2013) Gastric cancer: overview. Gastroenterol Clin North Am. 42, 211-217

26 Carneiro, F. and Sobrinho-Simoes, M. (1996) Metastatic pattern of gastric carcinoma. Hum Pathol. 27, 213-214

27 Mori, M., Sakaguchi, H., Akazawa, K., Tsuneyoshi, M., Sueishi, K. and Sugimachi, K. (1995) Correlation between metastatic site, histological type, and serum tumor markers of gastric carcinoma. Hum Pathol. 26, 504-508

28 Wroblewski, L. E., Peek, R. M., Jr. and Wilson, K. T. (2010) Helicobacter pylori and gastric cancer: factors that modulate disease risk. Clin Microbiol Rev. 23, 713-739

29 Parsonnet, J., Friedman, G. D., Orentreich, N. and Vogelman, H. (1997) Risk for gastric cancer in people with CagA positive or CagA negative Helicobacter pylori infection. Gut. 40, 297-301

30 Phadnis, S. H., Ilver, D., Janzon, L., Normark, S. and Westblom, T. U. (1994) Pathological significance and molecular characterization of the vacuolating toxin gene of Helicobacter pylori. Infect Immun. 62, 1557-1565

31 Ilver, D., Arnqvist, A., Ogren, J., Frick, I. M., Kersulyte, D., Incecik, E. T., . . . Boren, T. (1998) Helicobacter pylori adhesin binding fucosylated histo-blood group antigens revealed by retagging. Science. 279, 373-377

32 Mahdavi, J., Sonden, B., Hurtig, M., Olfat, F. O., Forsberg, L., Roche, N., . . . Boren, T. (2002) Helicobacter pylori SabA adhesin in persistent infection and chronic inflammation. Science. 297, 573-578

33 El-Omar, E. M., Rabkin, C. S., Gammon, M. D., Vaughan, T. L., Risch, H. A., Schoenberg, J. B., . . . Chow, W. H. (2003) Increased risk of noncardia gastric cancer associated with proinflammatory cytokine gene polymorphisms. Gastroenterol. 124, 1193-1201

34 Cancer Genome Atlas Research, N. (2014) Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 513, 202-209

35 Van Cutsem, E., Sagaert, X., Topal, B., Haustermans, K. and Prenen, H. (2016) Gastric cancer. Lancet

36 Wong, H. and Yau, T. (2013) Molecular targeted therapies in advanced gastric cancer: does tumor histology matter? Therap Adv Gastroenterol. 6, 15-31

37 Washington, K. (2010) 7th edition of the AJCC cancer staging manual: stomach. Ann Surg Oncol. 17, 3077-3079

38 Ferreira, J. A., Magalhaes, A., Gomes, J., Peixoto, A., Gaiteiro, C., Fernandes, E., . . . Reis, C. A. (2016) Protein glycosylation in gastric and colorectal cancers: Toward cancer detection and targeted therapeutics. Cancer Lett

39 Locker, G. Y., Hamilton, S., Harris, J., Jessup, J. M., Kemeny, N., Macdonald, J. S., . . . Asco. (2006) ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. J Clin Oncol. 24, 5313-5327

40 Shimada, H., Noie, T., Ohashi, M., Oba, K. and Takahashi, Y. (2014) Clinical significance of serum tumor markers for gastric cancer: a systematic review of literature by the Task Force of the Japanese Gastric Cancer Association. Gastric Cancer. 17, 26-33

41 Nakagoe, T., Sawai, T., Tsuji, T., Jibiki, M., Nanashima, A., Yamaguchi, H., . . . Ishikawa, H. (2001) Pre-operative serum levels of sialyl Tn antigen predict liver metastasis and poor prognosis in patients with gastric cancer. Eur J Surg Oncol. 27, 731-739

42 Nakagoe, T., Sawai, T., Tsuji, T., Jibiki, M. A., Nanashima, A., Yamaguchi, H., . . . Ishikawa, H. (2002) Predictive factors for preoperative serum levels of sialy Lewis(x), sialyl Lewis(a) and sialyl Tn antigens in gastric cancer patients. Anticancer Res. 22, 451-458

43 Duffy, M. J., Lamerz, R., Haglund, C., Nicolini, A., Kalousova, M., Holubec, L. and Sturgeon, C. (2014) Tumor markers in colorectal cancer, gastric cancer and gastrointestinal stromal cancers: European group on tumor markers 2014 guidelines update. Int J Cancer. 134, 2513-2522

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

45 Freeze, H. H., Chong, J. X., Bamshad, M. J. and Ng, B. G. (2014) Solving glycosylation disorders: fundamental approaches reveal complicated pathways. Am J Hum Genet. 94, 161-175

46 Rini, J., Esko, J. and Varki, A. (2009) Glycosyltransferases and Glycan-processing Enzymes. In Essentials of Glycobiology (Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W. and Etzler, M. E., eds.), Cold Spring Harbor (NY)

47 Rakus, J. F. and Mahal, L. K. (2011) New technologies for glycomic analysis: toward a systematic understanding of the glycome. Annu Rev Anal Chem (Palo Alto Calif). 4, 367-392

48 Helenius, A. and Aebi, M. (2001) Intracellular functions of N-linked glycans. Science. 291, 2364-2369

49 Skropeta, D. (2009) The effect of individual N-glycans on enzyme activity. Bioorg Med Chem. 17, 2645-2653

50 Contessa, J. N., Bhojani, M. S., Freeze, H. H., Rehemtulla, A. and Lawrence, T. S. (2008) Inhibition of N-linked glycosylation disrupts receptor tyrosine kinase signaling in tumor cells. Cancer Res. 68, 3803-3809

51 Sareneva, T., Pirhonen, J., Cantell, K. and Julkunen, I. (1995) N-glycosylation of human interferon-gamma: glycans at Asn-25 are critical for protease resistance. Biochem J. 308 (Pt 1), 9-14

52 Bachert, C. and Linstedt, A. D. (2013) A sensor of protein O-glycosylation based on sequential processing in the Golgi apparatus. Traffic. 14, 47-56

53 Laurent, T. C., Laurent, U. B. and Fraser, J. R. (1996) The structure and function of hyaluronan: An overview. Immunol Cell Biol. 74, A1-7

54 Ley, K. (2003) The role of selectins in inflammation and disease. Trends Mol Med. 9, 263-268

55 Steentoft, C., Vakhrushev, S. Y., Vester-Christensen, M. B., Schjoldager, K. T., Kong, Y., Bennett, E. P., . . . Clausen, H. (2011) Mining the O-glycoproteome using zinc-finger nuclease-glycoengineered SimpleCell lines. Nature methods. 8, 977-982

56 Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S. and Brunak, S. (2004) Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics. 4, 1633-1649

57 Bennett, E. P., Mandel, U., Clausen, H., Gerken, T. A., Fritz, T. A. and Tabak, L. A. (2012) Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. Glycobiology. 22, 736-756

58 Steentoft, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B., Schjoldager, K. T., . . . Clausen, H. (2013) Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J. 32, 1478-1488

59 Brockhausen, I. (1999) Pathways of O-glycan biosynthesis in cancer cells. Biochim Biophys Acta. 1473, 67-95

60 Brockhausen, I., Schachter, H. and Stanley, P. (2009) O-GalNAc Glycans. In Essentials of Glycobiology (Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W. and Etzler, M. E., eds.), Cold Spring Harbor (NY)

61 Rossez, Y., Maes, E., Lefebvre Darroman, T., Gosset, P., Ecobichon, C., Joncquel Chevalier Curt, M., . . . Robbe-Masselot, C. (2012) Almost all human gastric mucin O-glycans harbor blood group A, B or H antigens and are potential binding sites for Helicobacter pylori. Glycobiology. 22, 1193-1206

62 Bergstrom, K. S. and Xia, L. (2013) Mucin-type O-glycans and their roles in intestinal homeostasis. Glycobiology. 23, 1026-1037

63 Stanley P, Cummings RD: Structures Common to Different Glycans. In: Essentials of Glycobiology. 2nd edn. Edited by Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME. Cold Spring Harbor (NY); 2009.

64 Stanley, P., Schachter, H. and Taniguchi, N. (2009) N-Glycans. In Essentials of Glycobiology (Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W. and Etzler, M. E., eds.), Cold Spring Harbor (NY)

65 Takahashi, M., Kuroki, Y., Ohtsubo, K. and Taniguchi, N. (2009) Core fucose and bisecting GlcNAc, the direct modifiers of the N-glycan core: their functions and target proteins. Carbohydr Res. 344, 1387-1390

66 Wang, X., Gu, J., Ihara, H., Miyoshi, E., Honke, K. and Taniguchi, N. (2006) Core fucosylation regulates epidermal growth factor receptor-mediated intracellular signaling. J Biol Chem. 281, 2572-2577

67 Huhn, C., Selman, M. H., Ruhaak, L. R., Deelder, A. M. and Wuhrer, M. (2009) IgG glycosylation analysis. Proteomics. 9, 882-913

68 Esko, J. D., Kimata, K. and Lindahl, U. (2009) Proteoglycans and Sulfated Glycosaminoglycans. In Essentials of Glycobiology (Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W. and Etzler, M. E., eds.), Cold Spring Harbor (NY)

69 Toole, B. P. (2004) Hyaluronan: from extracellular glue to pericellular cue. Nat Rev Cancer. 4, 528-539

70 Bernfield, M., Gotte, M., Park, P. W., Reizes, O., Fitzgerald, M. L., Lincecum, J. and Zako, M. (1999) Functions of cell surface heparan sulfate proteoglycans. Annu Rev Biochem. 68, 729-777

71 Iozzo, R. V. and Schaefer, L. (2015) Proteoglycan form and function: A comprehensive nomenclature of proteoglycans. Matrix Biol. 42, 11-55

72 Fuster, M. M. and Esko, J. D. (2005) The sweet and sour of cancer: glycans as novel therapeutic targets. Nat Rev Cancer. 5, 526-542

73 Ladenson, R. P., Schwartz, S. O. and Ivy, A. C. (1949) Incidence of the blood groups and the secretor factor in patients with pernicious anemia and stomach carcinoma. Am J Med Sci. 217, 194-197

74 Pinho, S. S. and Reis, C. A. (2015) Glycosylation in cancer: mechanisms and clinical implications. Nat Rev Cancer. 15, 540-555

75 Girnita, L., Wang, M., Xie, Y., Nilsson, G., Dricu, A., Wejde, J. and Larsson, O. (2000) Inhibition of N-linked glycosylation down-regulates insulin-like growth factor-1 receptor at the cell surface and kills Ewing's sarcoma cells: therapeutic implications. Anticancer Drug Des. 15, 67-72

76 Lai, J., Chien, J., Staub, J., Avula, R., Greene, E. L., Matthews, T. A., . . . Shridhar, V. (2003) Loss of HSulf-1 up-regulates heparin-binding growth factor signaling in cancer. J Biol Chem. 278, 23107-23117

77 Nakamori, S., Kameyama, M., Imaoka, S., Furukawa, H., Ishikawa, O., Sasaki, Y., . . . Irimura, T. (1993) Increased expression of sialyl Lewisx antigen correlates with poor survival in patients with colorectal carcinoma: clinicopathological and immunohistochemical study. Cancer Res. 53, 3632-3637

78 Yoshimura, M., Ihara, Y., Matsuzawa, Y. and Taniguchi, N. (1996) Aberrant glycosylation of E-cadherin enhances cell-cell binding to suppress metastasis. J Biol Chem. 271, 13811-13815

79 Guo, H. B., Lee, I., Kamar, M., Akiyama, S. K. and Pierce, M. (2002) Aberrant N-glycosylation of beta1 integrin causes reduced alpha5beta1 integrin clustering and stimulates cell migration. Cancer Res. 62, 6837-6845

80 Mereiter, S., Balmana, M., Gomes, J., Magalhaes, A. and Reis, C. A. (2016) Glycomic Approaches for the Discovery of Targets in Gastrointestinal Cancer. Front Oncol. 6, 55

81 Gill DJ, Tham KM, Chia J, Wang SC, Steentoft C, Clausen H, Bard-Chapeau EA, Bard FA: Initiation of GalNAc-type O-glycosylation in the endoplasmic reticulum promotes cancer cell invasiveness. (2013) Proc Natl Acad Sci, 110(34):E3152-3161.

82 Gomes J, Marcos NT, Berois N, Osinaga E, Magalhaes A, …, Reis CA. (2009) Expression of UDP-N-acetyl-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase-6 in gastric mucosa, intestinal metaplasia, and gastric carcinoma. J Histochem Cytochem. 57(1):79-86.

83 Hua D, Shen L, Xu L, Jiang Z, …, Wu S. (2012) Polypeptide N-acetylgalactosaminyltransferase 2 regulates cellular metastasis-associated behavior in gastric cancer. Int J Mol Med. 30(6):1267-1274.

84 Onitsuka K, Shibao K, Nakayama Y, et al. (2003) Prognostic significance of UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase-3 (GalNAc-T3) expression in patients with gastric carcinoma. Cancer science.94:32-36.

85 Gao Y, Liu Z, Feng J, Sun Q, Zhang B, Zheng W, Ma W. (2013) Expression pattern of polypeptide N-acetylgalactosaminyltransferase-10 in gastric carcinoma. Oncology letters. 5(1):113-116.

86 Kudelka, M. R., Ju, T., Heimburg-Molinaro, J. and Cummings, R. D. (2015) Simple sugars to complex disease--mucin-type O-glycans in cancer. Adv Cancer Res. 126, 53-135

87 Gill, D. J., Chia, J., Senewiratne, J. and Bard, F. (2010) Regulation of O-glycosylation through Golgi-to-ER relocation of initiation enzymes. J Cell Biol. 189, 843-858

88 Kellokumpu, S., Sormunen, R. and Kellokumpu, I. (2002) Abnormal glycosylation and altered Golgi structure in colorectal cancer: dependence on intra-Golgi pH. FEBS Lett. 516, 217-224

89 Petrosyan, A., Ali, M. F. and Cheng, P. W. (2012) Glycosyltransferase-specific Golgi-targeting mechanisms. J Biol Chem. 287, 37621-37627

90 Hofmann, B. T., Schluter, L., Lange, P., Mercanoglu, B., Ewald, F., Folster, A., . . . Wolters-Eisfeld, G. (2015) COSMC knockdown mediated aberrant O-glycosylation promotes oncogenic properties in pancreatic cancer. Mol Cancer. 14, 109

91 Marcos, N. T., Bennett, E. P., Gomes, J., Magalhaes, A., Gomes, C., David, L., . . . Reis, C. A. (2011) ST6GalNAc-I controls expression of sialyl-Tn antigen in gastrointestinal tissues. Front Biosci (Elite Ed). 3, 1443-1455

92 Ju, T., Lanneau, G. S., Gautam, T., Wang, Y., Xia, B., Stowell, S. R., . . . Cummings, R. D. (2008) Human tumor antigens Tn and sialyl Tn arise from mutations in Cosmc. Cancer Res. 68, 1636-1646

93 Marcos, N. T., Pinho, S., Grandela, C., Cruz, A., Samyn-Petit, B., Harduin-Lepers, A., . . . Reis, C. A. (2004) Role of the human ST6GalNAc-I and ST6GalNAc-II in the synthesis of the cancer-associated sialyl-Tn antigen. Cancer Res. 64, 7050-7057

94 Pinho, S., Marcos, N. T., Ferreira, B., Carvalho, A. S., Oliveira, M. J., Santos-Silva, F., . . . Reis, C. A. (2007) Biological significance of cancer-associated sialyl-Tn antigen: modulation of malignant phenotype in gastric carcinoma cells. Cancer Lett. 249, 157-170

95 David, L., Nesland, J. M., Clausen, H., Carneiro, F. and Sobrinho-Simoes, M. (1992) Simple mucin-type carbohydrate antigens (Tn, sialosyl-Tn and T) in gastric mucosa, carcinomas and metastases. APMIS Suppl. 27, 162-172

96 Tian, H., Miyoshi, E., Kawaguchi, N., Shaker, M., Ito, Y., Taniguchi, N., . . . Matsuura, N. (2008) The implication of N-acetylglucosaminyltransferase V expression in gastric cancer. Pathobiology. 75, 288-294

97 Granovsky, M., Fata, J., Pawling, J., Muller, W. J., Khokha, R. and Dennis, J. W. (2000) Suppression of tumor growth and metastasis in Mgat5-deficient mice. Nat Med. 6, 306-312

98 Taniguchi, N. and Korekane, H. (2011) Branched N-glycans and their implications for cell adhesion, signaling and clinical applications for cancer biomarkers and in therapeutics. BMB Rep. 44, 772-781

99 Gu, J., Isaji, T., Xu, Q., Kariya, Y., Gu, W., Fukuda, T. and Du, Y. (2012) Potential roles of N-glycosylation in cell adhesion. Glycoconj J. 29, 599-607

100 Pinho, S. S., Reis, C. A., Paredes, J., Magalhaes, A. M., Ferreira, A. C., Figueiredo, J., . . . Seruca, R. (2009) The role of N-acetylglucosaminyltransferase III and V in the post-transcriptional modifications of E-cadherin. Hum Mol Genet. 18, 2599-2608

101 Dall'Olio, F. and Chiricolo, M. (2001) Sialyltransferases in cancer. Glycoconj J. 18, 841-850

102 Dall'Olio, F., Malagolini, N., Trinchera, M. and Chiricolo, M. (2014) Sialosignaling: sialyltransferases as engines of self-fueling loops in cancer progression. Biochim Biophys Acta. 1840, 2752-2764

103 Kim, Y. J. and Varki, A. (1997) Perspectives on the significance of altered glycosylation of glycoproteins in cancer. Glycoconj J. 14, 569-576

104 Baldus, S. E., Zirbes, T. K., Monig, S. P., Engel, S., Monaca, E., Rafiqpoor, K., . . . Dienes, H. P. (1998) Histopathological subtypes and prognosis of gastric cancer are correlated with the expression of mucin-associated sialylated antigens: Sialosyl-Lewis(a), Sialosyl-Lewis(x) and sialosyl-Tn. Tumour Biol. 19, 445-453

105 Rosen, S. D. and Bertozzi, C. R. (1994) The selectins and their ligands. Curr Opin Cell Biol. 6, 663-673

106 Marrelli, D., Pinto, E., De Stefano, A., de Manzoni, G., Farnetani, M., Garosi, L. and Roviello, F. (2001) Preoperative positivity of serum tumor markers is a strong predictor of hematogenous recurrence of gastric cancer. J Surg Oncol. 78, 253-258

107 Gschwind, A., Fischer, O. M. and Ullrich, A. (2004) The discovery of receptor tyrosine kinases: targets for cancer therapy. Nat Rev Cancer. 4, 361-370

108 De Vita, F., Giuliani, F., Silvestris, N., Catalano, G., Ciardiello, F. and Orditura, M. (2010) Human epidermal growth factor receptor 2 (HER2) in gastric cancer: a new therapeutic target. Cancer Treat Rev. 36 Suppl 3, S11-15

109 Moutinho, C., Mateus, A. R., Milanezi, F., Carneiro, F., Seruca, R. and Suriano, G. (2008) Epidermal growth factor receptor structural alterations in gastric cancer. BMC Cancer. 8, 10

110 Lee, H. E., Kim, M. A., Lee, H. S., Jung, E. J., Yang, H. K., Lee, B. L., . . . Kim, W. H. (2012) MET in gastric carcinomas: comparison between protein expression and gene copy number and impact on clinical outcome. Br J Cancer. 107, 325-333

111 Wang, X., Chen, X., Fang, J. and Yang, C. (2013) Overexpression of both VEGF-A and VEGF-C in gastric cancer correlates with prognosis, and silencing of both is effective to inhibit cancer growth. Int J Clin Exp Pathol. 6, 586-597

112 Turner, N. and Grose, R. (2010) Fibroblast growth factor signalling: from development to cancer. Nat Rev Cancer. 10, 116-129

113 Kodama, M., Kitadai, Y., Sumida, T., Ohnishi, M., Ohara, E., Tanaka, M., . . . Chayama, K. (2010) Expression of platelet-derived growth factor (PDGF)-B and PDGF-receptor beta is associated with lymphatic metastasis in human gastric carcinoma. Cancer Sci. 101, 1984-1989

114 Yao, H. P., Zhou, Y. Q., Zhang, R. and Wang, M. H. (2013) MSP-RON signalling in cancer: pathogenesis and therapeutic potential. Nat Rev Cancer. 13, 466-481

115 Schmidt, S. and Friedl, P. (2010) Interstitial cell migration: integrin-dependent and alternative adhesion mechanisms. Cell Tissue Res. 339, 83-92

116 Kim, S. H., Turnbull, J. and Guimond, S. (2011) Extracellular matrix and cell signalling: the dynamic cooperation of integrin, proteoglycan and growth factor receptor. J Endocrinol. 209, 139-151

117 Ponta, H., Sherman, L. and Herrlich, P. A. (2003) CD44: from adhesion molecules to signalling regulators. Nat Rev Mol Cell Biol. 4, 33-45

118 Lee, K. E., Park, J. S., Khoi, P. N., Joo, Y. E., Lee, Y. H. and Jung, Y. D. (2012) Upregulation of recepteur d'origine nantais tyrosine kinase and cell invasiveness via early growth response-1 in gastric cancer cells. J Cell Biochem. 113, 1217-1223

119 Wang, M. H., Lee, W., Luo, Y. L., Weis, M. T. and Yao, H. P. (2007) Altered expression of the RON receptor tyrosine kinase in various epithelial cancers and its

contribution to tumourigenic phenotypes in thyroid cancer cells. J Pathol. 213, 402-411

120 Catenacci, D. V., Cervantes, G., Yala, S., Nelson, E. A., El-Hashani, E., Kanteti, R., . . . Salgia, R. (2011) RON (MST1R) is a novel prognostic marker and therapeutic target for gastroesophageal adenocarcinoma. Cancer Biol Ther. 12, 9-46

121 McClaine, R. J., Marshall, A. M., Wagh, P. K. and Waltz, S. E. (2010) Ron receptor tyrosine kinase activation confers resistance to tamoxifen in breast cancer cell lines. Neoplasia. 12, 650-658

122 Thobe, M. N., Gurusamy, D., Pathrose, P. and Waltz, S. E. (2010) The Ron receptor tyrosine kinase positively regulates angiogenic chemokine production in prostate cancer cells. Oncogene. 29, 214-226

123 Logan-Collins, J., Thomas, R. M., Yu, P., Jaquish, D., Mose, E., French, R., . . . Lowy, A. M. (2010) Silencing of RON receptor signaling promotes apoptosis and gemcitabine sensitivity in pancreatic cancers. Cancer Res. 70, 1130-1140

124 Northrup, A. B., Katcher, M. H., Altman, M. D., Chenard, M., Daniels, M. H., Deshmukh, S. V., . . . Dinsmore, C. J. (2013) Discovery of 1-[3-(1-methyl-1H-pyrazol-4-yl)-5-oxo-5H-benzo[4,5]cyclohepta[1,2-b]pyridin-7-yl] -N-(pyridin-2-ylmethyl)methanesulfonamide (MK-8033): A Specific c-Met/Ron dual kinase inhibitor with preferential affinity for the activated state of c-Met. J Med Chem. 56, 2294-2310

125 Zhang, Y., Kaplan-Lefko, P. J., Rex, K., Yang, Y., Moriguchi, J., Osgood, T., . . . Dussault, I. (2008) Identification of a novel recepteur d'origine nantais/c-met small-molecule kinase inhibitor with antitumor activity in vivo. Cancer Res. 68, 6680-6687

126 Gunes, Z., Zucconi, A., Cioce, M., Meola, A., Pezzanera, M., Acali, S., . . . Vitelli, A. (2011) Isolation of Fully Human Antagonistic RON Antibodies Showing Efficient Block of Downstream Signaling and Cell Migration. Transl Oncol. 4, 38-46

127 O'Toole, J. M., Rabenau, K. E., Burns, K., Lu, D., Mangalampalli, V., Balderes, P., . . . Pereira, D. S. (2006) Therapeutic implications of a human neutralizing antibody to the macrophage-stimulating protein receptor tyrosine kinase (RON), a c-MET family member. Cancer Res. 66, 9162-9170

128 Ronsin, C., Muscatelli, F., Mattei, M. G. and Breathnach, R. (1993) A novel putative receptor protein tyrosine kinase of the met family. Oncogene. 8, 1195-1202

129 Danilkovitch, A., Miller, M. and Leonard, E. J. (1999) Interaction of macrophage-stimulating protein with its receptor. Residues critical for beta chain binding and evidence for independent alpha chain binding. J Biol Chem. 274, 29937-29943

130 Ma, Q., Zhang, K., Guin, S., Zhou, Y. Q. and Wang, M. H. (2010) Deletion or insertion in the first immunoglobulin-plexin-transcription (IPT) domain differentially regulates expression and tumorigenic activities of RON receptor Tyrosine Kinase. Mol Cancer. 9, 307

131 Chao, K. L., Tsai, I. W., Chen, C. and Herzberg, O. (2012) Crystal structure of the Sema-PSI extracellular domain of human RON receptor tyrosine kinase. PLoS One. 7, e41912

132 Screaton, G. R., Bell, M. V., Jackson, D. G., Cornelis, F. B., Gerth, U. and Bell, J. I. (1992) Genomic structure of DNA encoding the lymphocyte homing receptor CD44 reveals at least 12 alternatively spliced exons. Proc Natl Acad Sci USA. 89, 12160-12164

133 Marhaba, R. and Zoller, M. (2004) CD44 in cancer progression: adhesion, migration and growth regulation. Journal of molecular histology. 35, 211-231

134 Goodison, S., Urquidi, V. and Tarin, D. (1999) CD44 cell adhesion molecules. Mol Pathol. 52, 189-196

135 Sneath, R. J. and Mangham, D. C. (1998) The normal structure and function of CD44 and its role in neoplasia. Mol Pathol. 51, 191-200

136 Naor, D., Sionov, R. V. and Ish-Shalom, D. (1997) CD44: structure, function, and association with the malignant process. Adv Cancer Res. 71, 241-319

137 Banky, B., Raso-Barnett, L., Barbai, T., Timar, J., Becsagh, P. and Raso, E. (2012) Characteristics of CD44 alternative splice pattern in the course of human colorectal adenocarcinoma progression. Mol Cancer. 11, 83

138 Bourguignon, L. Y., Gunja-Smith, Z., Iida, N., Zhu, H. B., Young, L. J., Muller, W. J. and Cardiff, R. D. (1998) CD44v(3,8-10) is involved in cytoskeleton-mediated tumor cell migration and matrix metalloproteinase (MMP-9) association in metastatic breast cancer cells. J Cell Physiol. 176, 206-215

139 Mulder, J. W., Kruyt, P. M., Sewnath, M., Oosting, J., Seldenrijk, C. A., Weidema, W. F., . . . Pals, S. T. (1994) Colorectal cancer prognosis and expression of exon-v6-containing CD44 proteins. Lancet. 344, 1470-1472

140 Okamoto, I., Kawano, Y., Murakami, D., Sasayama, T., Araki, N., Miki, T., . . . Saya, H. (2001) Proteolytic release of CD44 intracellular domain and its role in the CD44 signaling pathway. J Cell Biol. 155, 755-762

141 Chetty, C., Vanamala, S. K., Gondi, C. S., Dinh, D. H., Gujrati, M. and Rao, J. S. (2012) MMP-9 induces CD44 cleavage and CD44 mediated cell migration in glioblastoma xenograft cells. Cell Signal. 24, 549-559

142 Anderegg, U., Eichenberg, T., Parthaune, T., Haiduk, C., Saalbach, A., Milkova, L., . . . Simon, J. C. (2009) ADAM10 is the constitutive functional sheddase of CD44 in human melanoma cells. J Invest Dermatol. 129, 1471-1482

143 Murakami, D., Okamoto, I., Nagano, O., Kawano, Y., Tomita, T., Iwatsubo, T., . . . Saya, H. (2003) Presenilin-dependent gamma-secretase activity mediates the intramembranous cleavage of CD44. Oncogene. 22, 1511-1516

144 Lokeshwar, V. B. and Bourguignon, L. Y. (1991) Post-translational protein modification and expression of ankyrin-binding site(s) in GP85 (Pgp-1/CD44) and its biosynthetic precursors during T-lymphoma membrane biosynthesis. J Biol Chem. 266, 17983-17989

145 Camp, R. L., Kraus, T. A., Birkeland, M. L. and Pure, E. (1991) High levels of CD44 expression distinguish virgin from antigen-primed B cells. J Exp Med. 173, 763-766

146 Lesley, J., Hyman, R. and Kincade, P. W. (1993) CD44 and its interaction with extracellular matrix. Adv Immunol. 54, 271-335

147 Laurent, T. C. and Fraser, J. R. (1992) Hyaluronan. FASEB J. 6, 2397-2404

148 Orian-Rousseau, V. and Sleeman, J. (2014) CD44 is a multidomain signaling platform that integrates extracellular matrix cues with growth factor and cytokine signals. Adv Cancer Res. 123, 231-254

149 Bennett, K. L., Modrell, B., Greenfield, B., Bartolazzi, A., Stamenkovic, I., Peach, R., . . . Aruffo, A. (1995) Regulation of CD44 binding to hyaluronan by glycosylation of variably spliced exons. J Cell Biol. 131, 1623-1633

150 Jackson, D. G., Bell, J. I., Dickinson, R., Timans, J., Shields, J. and Whittle, N. (1995) Proteoglycan forms of the lymphocyte homing receptor CD44 are alternatively spliced variants containing the v3 exon. J Cell Biol. 128, 673-685

151 Yu, W. H., Woessner, J. F., Jr., McNeish, J. D. and Stamenkovic, I. (2002) CD44 anchors the assembly of matrilysin/MMP-7 with heparin-binding epidermal growth factor precursor and ErbB4 and regulates female reproductive organ remodeling. Genes Dev. 16, 307-323

152 Tremmel, M., Matzke, A., Albrecht, I., Laib, A. M., Olaku, V., Ballmer-Hofer, K., . . . Orian-Rousseau, V. (2009) A CD44v6 peptide reveals a role of CD44 in VEGFR-2 signaling and angiogenesis. Blood. 114, 5236-5244

153 Matzke, A., Herrlich, P., Ponta, H. and Orian-Rousseau, V. (2005) A five-amino-acid peptide blocks Met- and Ron-dependent cell migration. Cancer Res. 65, 6105-6110

154 Sherman, L. S., Rizvi, T. A., Karyala, S. and Ratner, N. (2000) CD44 enhances neuregulin signaling by Schwann cells. J Cell Biol. 150, 1071-1084

CHAPTER II

GLYCOMIC AND SIALOPROTEOMIC CHARACTERIZATION OF GLYCO-ENGINEERED GASTRIC CANCER MODEL CELL LINES

Data in Brief 7 (2016) 814–833

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Glycomic and sialoproteomic data of gastric carcinoma cells overexpressing ST3GAL4

Stefan Mereiter [a,b,c], Ana Magalhães [a,b], Barbara Adamczyk [d], Chunsheng Jin [d], Andreia Almeida [e,f], Lylia Drici [g], Maria Ibáñez-Vea [g], Martin R. Larsen [g], Daniel Kolarich [e], Niclas G. Karlsson [d], Celso A. Reis [a,b,c,h,*]

[a] *I3S - Instituto de Investigação e Inovação em Saúde, University of Porto, Portugal*
[b] *Institute of Molecular Pathology and Immunology of the University of Porto - IPATIMUP, Porto, Portugal*
[c] *Institute of Biomedical Sciences of Abel Salazar - ICBAS, University of Porto, Portugal*
[d] *Department of Medical Biochemistry and Cell Biology, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Sweden*
[e] *Department of Biomolecular Systems, Max Planck Institute of Colloids and Interfaces, 14424 Potsdam, Germany*
[f] *Free University Berlin, Berlin, Germany*
[g] *Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark*
[h] *Medical Faculty, University of Porto, Portugal*

ARTICLE INFO

Article history:
Received 22 December 2015
Received in revised form
22 February 2016
Accepted 4 March 2016
Available online 14 March 2016

Keywords:
N-glycome
O-glycome
Gastric cancer
Sialyltransferase
Sialoproteome

ABSTRACT

Gastric carcinoma MKN45 cells stably transfected with the full-length *ST3GAL4* gene were characterised by glycomic and sialoproteomic analysis. Complementary strategies were applied to assess the glycomic alterations induced by ST3GAL4 overexpression. The *N*- and *O*-glycome data were generated in two parallel structural analyzes, based on PGC-ESI-MS/MS. Data on glycan structure identification and relative abundance in ST3GAL4 overexpressing cells and respective mock control are presented. The sialoproteomic analysis based on titanium-dioxide enrichment of sialopeptides with subsequent LC-MS/MS identification was performed. This analysis identified 47 proteins with significantly increased sialylation. The data in this article is associated with the research article published in Biochim Biophys Acta "Glycomic analysis of gastric carcinoma cells discloses

DOI of original article: http://dx.doi.org/10.1016/j.bbagen.2015.12.016
* Corresponding author at: I3S - Instituto de Investigação e Inovação em Saúde, University of Porto, Portugal.
E-mail address: celsor@ipatimup.pt (C.A. Reis).

http://dx.doi.org/10.1016/j.dib.2016.03.022

glycans as modulators of RON receptor tyrosine kinase activation in cancer" [1].

Specifications Table

Subject area	*Biology*
More specific subject area	*Glycobiology in cancer*
Type of data	*Tables*
How data was acquired	*N-and O-glycome analysed by PGC-LC ESI MS/MS performed on an LTQ ion trap mass spectrometer and PGC-nanoLC ESI MS/MS performed on an amaZon ETD Speed ion trap. Sialoproteome performed using Easy-nLC II system and Orbitrap Fusion Tribrid system.*
Data format	*Analyzed*
Experimental factors	*MKN45 cells stably transfected with the full-length ST3GAL4 gene or empty vector (mock).*
Experimental features	*MKN45 cells grown in vitro in RPMI medium with 10% FBS.*
Data source location	*Not applicable*
Data accessibility	*Data is within this article.*

Value of the data

- Data shows the *N*- and *O*-glycome of the MKN45 gastric carcinoma cells and ST3GAL4 sialyltransferase overexpressing cells.
- Data provides a list of glycoproteins with altered sialylated *N*-glycans in ST3GAL4 overexpressing cells, including several cancer associated proteins.
- These data are valuable as a source for novel biomarkers for gastric cancer.

1. Data

This data article includes the *N*- and *O*-glycome of MKN45 cells and assesses the glycosylation alterations induced by ST3GAL4 overexpression. In addition, we provide the data on the proteins with significant increased sialylated *N*-glycans upon ST3GAL4 overexpression.

2. Experimental design, materials and methods

Glycomic and sialoproteomic analyses were performed as described below comparing MKN45 cells stably transfected with the full-length *ST3GAL4* gene or empty vector (mock). Glycome data and statistical evaluation are shown in Tables 1, 2, 3 and 4. Sialoproteome data and statistical evaluation are shown in Tables 5 and 6.

3. *N*-glycomic strategy I: sample preparation and PGC LC-ESI-MS/MS

Frozen cell pellets (10^7 cells) of mock or *ST3GAL4* transfected MKN45 cells [2] were directly resuspended in 7 M urea, 2 M thiourea, 40 mM Tris, 2% CHAPS, 10 mM DTT and 1% protease inhibitor (Sigma-Aldrich, St. Louis, MO). The cell membranes were disrupted by 10 times 10 s sonication with 16 amplitudes and 1 min on ice in between, and subsequent shaking at 4 °C overnight. To reduce the viscosity of the lysates, the DNA was degraded by adding 1 µl benzonase® nuclease (250 units, Sigma-Aldrich) and 30 min incubation at 37 °C. In order to impair refolding of proteins, 25 mM iodoacetamide were added for alkylation during 1 h in the dark. The lysates were centrifuged for 30 min at 14,000 rpm and the supernatants were transferred to a fresh tube.

Then, solubilized proteins were concentrated by adding 150 µl of supernatant on a 10 kDa cut-off spinfilter (PALL, Port Washington, NY), spinning down for 5 min with 12,000xg and washing 3 times with 100 µl 50 mM NH_4HCO_3, pH 8.4. *N*-linked oligosaccharides were released in the spinfilter using 20 µl 50 mM NH_4HCO_3 and PNGase F (5 mU, Prozyme, Hayward, CA) with incubation at 37 °C overnight. Subsequently, the *N*-glycans were collected by washing 3 times with 20 µl H_2O and dried in Speedvac. Reactions were quenched with 1 µl of glacial acetic acid and *N*-glycan samples were desalted and dried as previously described [3]. *N*-glycan samples were subjected to LC-ESI-MS/MS analysis using a 10 cmx250 µm I.D. column, prepared in-house, containing 5 µm porous graphitized carbon (PGC) particles (Thermo Scientific, Waltham, MA). Glycans were eluted using a linear gradient from 0% to 40% acetonitrile in 10 mM NH_4HCO_3 over 40 min at a flow rate of 10 µl/min. The eluted *N*-glycans were detected using a LTQ ion trap mass spectrometer (Thermo Scientific) in negative-ion mode with an electrospray voltage of 3.5 kV, capillary voltage of −33.0 V and capillary temperature of 300 °C. Air was used as a sheath gas and mass ranges were defined dependent on the specific structure to be analyzed. The data were processed using the Xcalibur software (version 2.0.7, Thermo Scientific) and manually interpreted from their MS/MS spectra.

Optionally prior to analysis *N*-glycan were digested in 50 mM sodium phosphate, pH 6.0 at 37 °C overnight using sialidase S (4 mU, ProZyme) that releases α2-3 linked non-reducing terminal sialic acids

Table 1
Identified *N*-glycan structures I.

m/z		Composition	Proposed structure	Mock replicates (n=3)					ST3GAL4 replicates (n=3)					
z=1	z=2			%	%	%	Avg	SD	%	%	%	Avg	SD	p-value
708	/	HexNAc1Hex3		0.04	0.02	0.2	0.09	0.10	0.08	0.08	0.03	0.06	0.03	0.3636
749	/	HexNAc2Hex2		0.36	0.3	0.3	0.32	0.03	0.3	0.29	0.23	0.27	0.04	0.0955
895	/	HexNAc2Hex2Fuc1		1.18	0.95	1.58	1.24	0.32	0.75	0.69	1.38	0.94	0.38	0.181
911	/	HexNAc2Hex3		0.17	0.27	1.27	0.57	0.61	0.32	0.39	0.37	0.36	0.04	0.3054
911	/	HexNAc2Hex3		0.02	0.03	0.45	0.17	0.25	0.07	0.09	0.13	0.10	0.03	0.3356
911	/	HexNAc2Hex3		0.1	0.09	0.04	0.08	0.03	0.09	0.09	0.06	0.08	0.02	0.4421
1057	/	HexNAc2Hex3Fuc1		0.96	0.84	2.28	1.36	0.80	0.75	0.65	0.98	0.79	0.17	0.1719
1073	/	HexNAc2Hex4		0.06	0.12	0.68	0.29	0.34	0.11	0.13	0.12	0.12	0.01	0.2437
1073	/	HexNAc2Hex4		0.3	0.25	0.18	0.24	0.06	0.23	0.3	0.13	0.22	0.09	0.3604
1202	/	2HexNAc3Hex1Sia		2.29	2.24	6.23	3.59	2.29	1.94	3.09	3.34	2.79	0.75	0.3077
1202	/	HexNAc2Hex3Sia1		0.29	0.21	0.69	0.40	0.26	0.18	0.28	0.48	0.31	0.15	0.33
1235	/	HexNAc2Hex5		0.41	0.45	0.24	0.37	0.11	0.41	0.55	0.24	0.40	0.16	0.3896

Table 1 (*continued*)

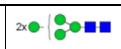
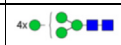
1235	/	HexNAc2Hex5		0.2	0.17	0.09	0.15	0.06	0.17	0.21	0.07	0.15	0.07	0.4765
1235	/	HexNAc2Hex5		1.53	1.25	2.26	1.68	0.52	1.81	1.8	2.55	2.05	0.43	0.1974
1364	681	HexNAc2Hex4Sia1		1.2	1.28	2.29	1.59	0.61	0.85	1.28	1.25	1.13	0.24	0.1592
1397	698	HexNAc2Hex6		0.26	0.24	0.28	0.26	0.02	0.3	0.33	0.25	0.29	0.04	0.1463
1397	698	HexNAc2Hex6		9.32	8.64	7.54	8.50	0.90	7.52	8.27	6.52	7.44	0.88	0.1082
1422	711	HexNAc3Hex4Fuc1		na	na	1.29	1.29	/	na	na	0.49	0.49	/	n/a
1422	711	HexNAc3Hex4Fuc1		na	na	0.41	0.41	/	na	na	0.22	0.22	/	n/a
1438	719	HexNAc3Hex5		na	na	0.31	0.31	/	na	na	0.2	0.20	/	n/a
1559	779	HexNAc2Hex7		0.25	0.24	0.17	0.22	0.04	0.32	0.37	0.19	0.29	0.09	0.1542
1559	779	HexNAc2Hex7		6.14	6.24	4.61	5.66	0.91	4.72	5.47	3.41	4.53	1.04	0.116
1559	779	HexNAc2Hex7		3.71	3.39	3.28	3.46	0.22	3.15	3.45	2.79	3.13	0.33	0.1173
1567	783	HexNAc3Hex3Sia1		na	na	0.6	0.60	/	0	0	0	0.00	0.00	n/a
1567	783	HexNAc3Hex4Sia1		na	na	0.7	0.70	/	na	na	0.63	0.63	/	n/a
1584	792	HexNAc3Hex5Fuc1		na	na	0.63	0.63	/	na	na	0.57	0.57	/	n/a
1600	800	HexNAc3Hex6		na	na	0.52	0.52	/	na	na	0.59	0.59	/	n/a
1713	856	HexNAc3Hex4Sia1Fuc1		7.34	5.42	8.25	7.00	1.44	4.71	3.79	6.94	5.15	1.62	0.1067
1721	860	HexNAc2Hex8		9.65	9.88	6.89	8.81	1.66	8.53	10.15	5.82	8.17	2.19	0.3544
1729	864	HexNAc3Hex5Sia1		0.95	1.59	3.19	1.91	1.15	1.19	1.48	1.77	1.48	0.29	0.2944
1729	864	HexNAc3Hex5Sia1		0.96	1	1.36	1.11	0.22	1.02	0.89	1.15	1.02	0.13	0.2978
1729	864	HexNAc3Hex5Sia1		na	na	0.8	0.80	#DIV/0!	na	na	1.09	1.09	/	n/a
1729	864	HexNAc3Hex5Sia1		0	0	0	0.00	0.00	na	na	0.58	0.58	/	n/a
1746	873	HexNAc3Hex6Fuc1		na	na	0.41	0.41	/	na	na	0.32	0.32	/	n/a
1787	893	HexNAc4Hex5Fuc1		0.39	0.41	1.7	0.83	0.75	0.38	0.27	0.63	0.43	0.18	0.2247

Table 1 (*continued*)

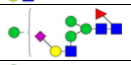
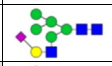
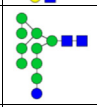
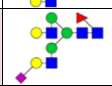
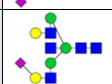
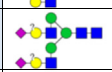
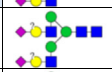
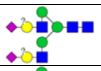
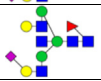
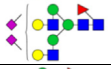
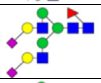
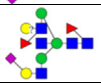
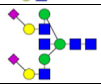
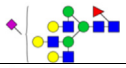
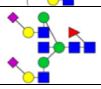
/	922	HexNAc5Hex5		0.06	0.1	0.09	0.08	0.02	0	0.01	0	0.00	0.01	0.0081 **
/	937	HexNAc3Hex5Sia1Fuc1		1.74	1.45	1.48	1.56	0.16	1.83	1.3	2.04	1.72	0.38	0.2703
/	941	HexNAc2Hex9		3.31	3.36	2.63	3.10	0.41	3.32	3.74	2.36	3.14	0.71	0.4687
/	945	HexNAc3Hex6Sia1		1.44	1.31	1.12	1.29	0.16	1.51	1.32	1.62	1.48	0.15	0.1024
/	945	HexNAc3Hex6Sia1		0.04	0.03	0	0.02	0.02	0.3	0.27	0.35	0.31	0.04	0.0009 ***
/	966	HexNAc4Hex5Sia1		0.87	0.78	0.71	0.79	0.08	0.98	0.96	0.73	0.89	0.14	0.1706
/	995	HexNAc5Hex5Fuc1		na	na	0.74	0.74	/	na	na	0.15	0.15	/	n/a
/	1010	HexNAc3Hex5Sia2		7.58	9.96	7.37	8.30	1.44	5.23	7.11	6.89	6.41	1.03	0.0723
/	1010	HexNAc3Hex5Sia2		3.17	3.43	1.89	2.83	0.82	5.67	6.23	6.69	6.20	0.51	0.0034 **
/	1010	HexNAc3Hex5Sia2		1.24	1.1	0.25	0.86	0.54	3.07	2.89	4.24	3.40	0.73	0.0052 **
/	1018	HexNAc3Hex6Sia1Fuc1		1.04	0.82	0.86	0.91	0.12	0.96	0.98	1.07	1.00	0.06	0.1464
/	1022	HexNAc2Hex10		0.37	0.38	0.38	0.38	0.01	0.39	0.37	0.39	0.38	0.01	0.2191
/	1039	HexNAc4Hex5Sia1Fuc1		3.73	3.32	5.22	4.09	1.00	3.86	3.45	5.57	4.29	1.12	0.4133
/	1039	HexNAc4Hex5Sia1Fuc1		0.19	0.06	0.55	0.27	0.25	0.97	0.62	1.51	1.03	0.45	0.0389 *
/	1067	HexNAc5Hex5Sia1		0.45	0.56	na	0.51	0.08	0.1	0.04	na	0.07	0.04	0.0189 *
/	1111	HexNAc4Hex5Sia2		1.78	2.37	na	2.08	0.42	0.09	0.15	na	0.12	0.04	0.0463 *
/	1111	HexNAc4Hex5Sia2		5.39	5.23	na	5.31	0.11	5.49	7.31	na	6.40	1.29	0.2209
/	1111	HexNAc4Hex5Sia2		1.74	1.92	na	1.83	0.13	2.18	1.4	na	1.79	0.55	0.4677

Table 1 (*continued*)

/	1111	HexNAc4Hex5Sia2		0.18	0.2	na	0.19	0.01	1.41	0.98	na	1.20	0.30	0.0668
/	1140	HexNAc5Hex5Sia1Fuc1		2.01	2.54	na	2.28	0.37	0.33	0.24	na	0.29	0.06	0.0385 *
/	1184	HexNAc4Hex5Sia2Fuc1		9.37	9	7.91	8.76	0.76	7.41	5.95	7.95	7.10	1.03	0.0475 *
/	1184	HexNAc4Hex5Sia2Fuc1		2.52	2.81	2.25	2.53	0.28	7.47	4.87	7.03	6.46	1.39	0.0175 *
/	1184	HexNAc4Hex5Sia2Fuc1		0.38	0.54	0.23	0.38	0.16	6.22	4.33	5.88	5.48	1.01	0.0057 **
/	1213	HexNAc5Hex5Sia1Fuc2		1.27	1.06	0.85	1.06	0.21	0.12	0.1	0	0.07	0.06	0.0048 **
/	1213	HexNAc5Hex5Sia2		1.08	1.13	0.98	1.06	0.08	0.17	0.18	0	0.12	0.10	0.0002 ***
/	1221	HexNAc5Hex6Sia1Fuc1		0.96	1.02	na	0.99	0.04	0.99	0.75	na	0.87	0.17	0.2474
/	1286	HexNAc5Hex5Sia2Fuc1		na	na	2.76	2.76	/	0	0	0	0.00	0.00	n/a
			Total:	100	100	100			100	100	100			

Structures are represented by the mass to charge ratio (m/z) in which they were identified and quantified, by monosaccharide composition and proposed structure based on MS/MS analyses. The relative quantities were determined by base-peak intensity of extracted ion chromatograms. The average value (Avg) and standard deviation (SD) of triplicates are shown, as well as the p-value (t-test; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$). Increased or decreased relative abundance are shown in red or blue, respectively. Structures marked as not analyzed (na) were either not detected or overlapped with other structures in given sample, precluding their quantification. Unknown linkage is represented by "?".

(recombinant sialidase from *Streptococcus pneumoniae*, expressed in *Escherichia coli*) and sialidase A (5 mU, ProZyme) that releases α2-3/6/8 linked non-reducing terminal sialic acid (recombinant gene from *Arthrobacter ureafaciens*, expressed in *Escherichia coli*) to confirm sialic acid linkage.

All analyzes were performed in three independent replicates and results were subjected to statistical analyses (Average, standard deviation and unpaired t-test) Table 1.

4. *N*-glycomic strategy II: sample preparation and PGC nanoLC-ESI MS/MS

Frozen cell pellets (10^7 cells) of mock or *ST3GAL4* transfected MKN45 cells were directly resuspended in 2 mL of lysis buffer (50 mM Tris–HCl, 100 mM NaCl, 1 mM EDTA and protease inhibitor at pH 7.4) and stored on ice for 20 min. The cells were lysed using a Polytron homogenizer for at least three times for 10 s in a cold room. Cellular debris and unlysed cells were sedimented by centrifugation at 2000*g* for 20 min at 4 °C. The supernatant was collected and the pellets resuspended in 1 mL of lysis buffer and centrifuged again at 2000*g* for 20 min at 4 °C. All the supernatants were combined, diluted in 20 mM Tris–HCl (pH 7.4) and ultracentrifuged at 120,000*g* for 90 min at 4 °C. The supernatant was separated from the pellet containing the cell membrane proteins. The membrane proteins were resuspended with 150 μL of 100 mM ammonium bicarbonate buffer and lyophilized overnight. The dried samples were solubilized in 50 μL of 8 M urea and 10 μL aliquots were dot-blotted onto PVDF membranes as described previously [4]. *N*- and *O*-glycan release as well as

Table 2
Identified *N*-glycan structures II.

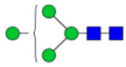
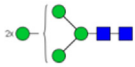
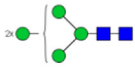
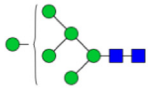
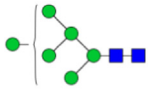
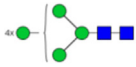
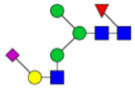
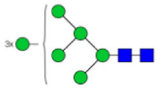
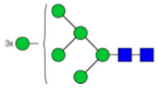
[M-H]-	z	Composition	Proposed structure	mock	ST3GAL4	ST3GAL4/mock
1073.44	1	HexNAc2Hex4		0.26	0.23	0.91
1235.48	1	HexNAc2Hex5		0.35	0.31	0.87
1235.5	1	HexNAc2Hex5		0.47	0.59	1.25
1235.54	1	HexNAc2Hex5		1.21	1.30	1.08
1397.52	1	HexNAc2Hex6		0.37	0.31	0.83
1397.56	1	HexNAc2Hex6		4.06	2.67	0.66
1559.64	2	HexNAc2Hex7		2.18	1.72	0.79
1559.69	2	HexNAc2Hex7		4.66	3.36	0.72
1713.82	2	HexNAc3Hex4Sia1Fuc1		1.61	1.52	0.94
1721.69	2	HexNAc2Hex8		0.06	0.11	1.84
1721.72	2	HexNAc2Hex8		0.21	0.16	0.80

Table 2 (*continued*)

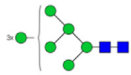
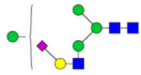
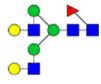
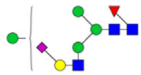
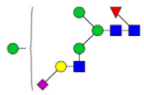
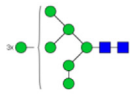
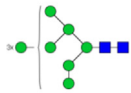
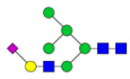
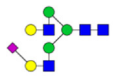
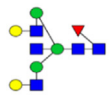
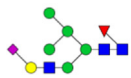
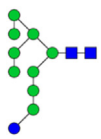
1721.78	2	HexNAc2Hex8		14.36	12.98	0.90
1729.76	2	HexNAc3Hex5Sia1		0.43	0.41	0.95
1787.79	2	HexNAc4Hex5Fuc1		0.87	0.69	0.80
1875.78	2	HexNAc3Hex5Sia1Fuc1		2.99	2.51	0.84
1875.89	2	HexNAc3Hex5Sia1Fuc1		0.00	0.63	/
1883.74	2	HexNAc2Hex9		11.94	11.97	1.00
1883.75	2	HexNAc2Hex9		2.40	2.19	0.91
1891.78	2	HexNAc3Hex6Sia1		2.26	2.11	0.93
1932.80	2	HexNAc4Hex5Sia1		0.97	0.65	0.67
1990.84	2	HexNAc5Hex5Fuc1		0.85	0.00	0.00
2037.83	2	HexNAc3Hex6Sia1Fuc1		1.23	0.90	0.73
2045.73	2	HexNAc1Hex10		2.05	2.16	1.05
2053.70	2	HexNAc3Hex7Sia1Fuc1		0.49	0.28	0.57

Table 2 (*continued*)

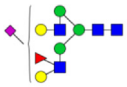
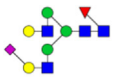
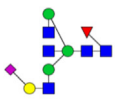
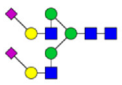
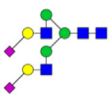
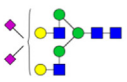
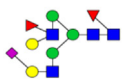
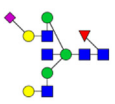
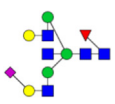
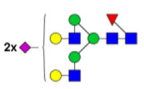
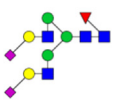
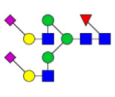
2078.80	2	HexNAc4Hex5Sia1Fuc1		0.91	1.20	1.32
2078.85	2	HexNAc4Hex5Sia1Fuc1		1.48	0.58	0.39
2078.90	2	HexNAc4Hex5Sia1Fuc1		11.24	7.82	0.70
2119.83	2	HexNAc5Hex4Sia1Fuc1		1.07	0.00	0.00
2223.85	2	HexNAc4Hex5Sia2		1.38	1.62	1.17
2223.87	2	HexNAc4Hex5Sia2		0.00	0.45	/
2223.89	2	HexNAc4Hex5Sia2		0.34	0.72	2.12
2224.92	2	HexNAc4Hex5Sia1Fuc2		5.64	3.56	0.63
2281.82	2	HexNAc5Hex5Sia1Fuc1		0.28	0.00	0.00
2281.89	2	HexNAc5Hex5Sia1Fuc1		1.97	0.32	0.16
2369.88	2	HexNAc4Hex5Sia2Fuc1		1.31	6.45	4.93
2369.91	2	HexNAc4Hex5Sia2Fuc1		0.00	5.26	/
2369.97	2	HexNAc4Hex5Sia2Fuc1		8.23	8.53	1.04

Table 2 (*continued*)

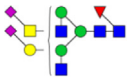
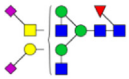
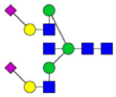
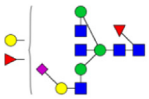
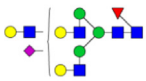
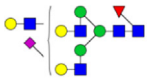
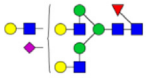
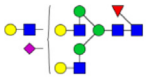
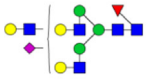
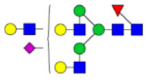
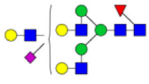
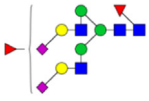
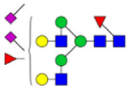
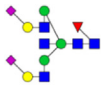
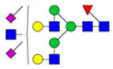
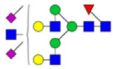
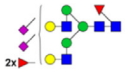
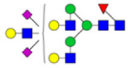
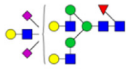
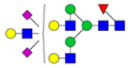
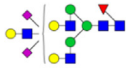
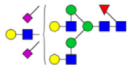
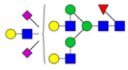
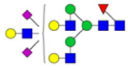
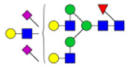
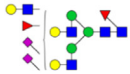
2410.85	2	HexNAc5Hex4Sia2Fuc1		0.16	0.14	0.90
2410.91	2	HexNAc5Hex4Sia2Fuc1		0.00	0.49	/
2426.89	2	HexNAc5Hex5Sia2		0.35	0.00	0.00
2427.87	2	HexNAc5Hex5Sia1Fuc2		1.29	0.10	0.07
2443.86	2	HexNAc5Hex6Sia1Fuc1		0.17	0.35	2.05
2443.87	2	HexNAc5Hex6Sia1Fuc1		0.00	0.10	/
2443.89	2	HexNAc5Hex6Sia1Fuc1		0.41	0.15	0.38
2443.90	2	HexNAc5Hex6Sia1Fuc1		0.54	0.46	0.84
2443.91	2	HexNAc5Hex6Sia1Fuc1		1.21	1.58	1.31
2443.95	2	HexNAc5Hex6Sia1Fuc1		1.12	0.00	0.00
2444.17	2	HexNAc5Hex6Sia1Fuc1		0.00	0.84	/
2515.91	2	HexNAc4Hex5Sia2Fuc2		0.00	0.55	/
2516.01	2	HexNAc4Hex5Sia2Fuc2		0.00	0.91	/

Table 2 (*continued*)

2572.95	2	HexNAc5Hex5Sia2Fuc1		1.81	0.25	0.14
2572.97	2	HexNAc5Hex5Sia2Fuc1		0.00	1.16	/
2572.97	2	HexNAc5Hex5Sia2Fuc1		0.00	1.37	/
2662.01	2	HexNAc4Hex5Sia2Fuc3		0.00	0.33	/
2734.91	2	HexNAc5Hex6Sia2Fuc1		0.72	0.72	0.99
2734.96	2	HexNAc5Hex6Sia2Fuc1		0.33	0.40	1.21
2734.96	2	HexNAc5Hex6Sia2Fuc1		0.38	0.67	1.74
2734.99	2	HexNAc5Hex6Sia2Fuc1		0.44	0.18	0.40
2735.00	2	HexNAc5Hex6Sia2Fuc1		0.00	0.45	/
2735.03	2	HexNAc5Hex6Sia2Fuc1		0.51	0.40	0.79
2735.04	2	HexNAc5Hex6Sia2Fuc1		0.00	0.60	/
2735.08	2	HexNAc5Hex6Sia2Fuc1		0.31	0.00	0.00
2880.98	2	HexNAc5Hex6Sia2Fuc2		0.00	0.48	/
3025.85	2	HexNAc5Hex6Sia3Fuc1		0.00	0.78	/
3026.03	2	HexNAc5Hex6Sia3Fuc1		0.14	0.27	1.97

Structures are represented by their [M–H] value and charge state in which they were identified and quantified, by monosaccharide composition and proposed structure based on MS/MS analyzes. The relative quantities were determined by base-peak intensity of extracted ion chromatograms. The ratio of glycan abundance in ST3GAL4 transfected cells relative to mock transfected is shown on the right column. Increases or decreases larger than 2 fold in glycan abundance are highlighted red or blue, respectively.

PGC-nanoLC ESI MS/MS analysis on an amaZon ETD Speed ion trap (Bruker, Bremen, Germany) were performed as described in detail previously [4,5] (Table 2).

5. *O*-glycomic strategy I: sample preparation and LC-ESI-MS/MS

After the removal of *N*-glycan, as previously described in "1. *N*-glycomic strategy I: Sample preparation and LC-ESI-MS/MS", the *O*-linked glycans were released from retained glycoproteins in spinfilter using reductive β-elimination (0.5 M $NaBH_4$, 50 mM NaOH at 50 °C, 16 h). Reactions were quenched with 1 μl of glacial acetic acid and glycan samples were desalted and dried as previously described [3]. Glycans were subjected to LC-ESI-MS/MS analysis using a 10 cmx250 μm I.D. column, prepared in-house, containing 5 μm porous graphitized carbon (PGC) particles (Thermo Scientific. Waltham. MA). Glycans were eluted using a linear gradient from 0% to 40% acetonitrile in 10 mM NH_4HCO_3 over 40 min at a flow rate of 10 μl/min. The eluted *O*-glycans were detected using a LTQ ion trap mass spectrometer (Thermo Scientific) in negative-ion mode with an electrospray voltage of 3.5 kV, capillary voltage of −33.0 V and capillary temperature of 300 °C. Air was used as a sheath gas and mass ranges were defined dependent on the specific structure to be analyzed. The data were processed using the Xcalibur software (version 2.0.7. Thermo Scientific) and manually interpreted from their MS/MS spectra (Table 3).

6. *O*-glycomic strategy II: sample preparation and PGC nanoLC-ESI MS/MS

Sample preparation and analysis were performed as previously described in Section 2. "*N*-glycomic strategy II: Sample preparation and PGC nanoLC-ESI MS/MS" (Table 4).

7. Sialoproteomic analysis

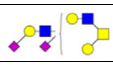
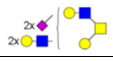
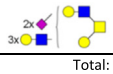
7.1. Cell lysis, protein digestion and iTRAQ labeling

Cell pellets were redissolved in ice-cold Na_2CO_3 buffer (0.1 M, pH 11) supplemented with protease inhibitor (Roche complete EDTA free), PhosSTOP phosphatase inhibitor cocktail (Roche) and 10 mM sodium pervanadate on ice. The suspensions were tip probe sonicated for 20 s (amplitude = 50%) twice and incubated at 4 °C for 1 h. The lysates were then centrifuged at 100,000 × *g* for 90 min at 4 °C to enrich membrane proteins (pellet). The pellets were washed with 50 mM triethylammonium

Table 3
Identified *O*-glycan structures I.

m/z		Composition	Proposed structure	Mock replicates (n=3)					ST3GAL4 replicates (n=3)					p-value
z=1	z=2			%	%	%	Avg	SD	%	%	%	Avg	SD	
384	/	HexNAc1Hex1		3.52	0.91	0.95	1.79	1.50	2.65	1.65	1.8	2.03	0.54	0.407
513	/	HexNAc1Sia1		1.52	0.83	0.93	1.09	0.37	0.74	1.39	0.73	0.95	0.38	0.336
587	/	HexNAc2Hex1		1.62	0	0	0.54	0.94	1.34	0	0	0.45	0.77	0.450
587	/	HexNAc2Hex1		0.94	0	0	0.31	0.54	0.6	0	0	0.20	0.35	0.390
675	/	HexNAc1Hex1Sia1		4.79	2.22	1.79	2.93	1.62	3.77	2.13	2.41	2.77	0.88	0.444
675	/	HexNAc1Hex1Sia1		6.13	8.1	5.79	6.67	1.25	9.5	8.99	9.9	9.46	0.46	0.024 *

Table 3 (*continued*)

749	/	HexNAc2Hex2		3.59	0.84	0.6	1.68	1.66	2.9	0.38	0.84	1.37	1.34	0.409
749	/	HexNAc2Hex2		1.65	0.88	0.34	0.96	0.66	0.65	0.5	0.57	0.57	0.08	0.210
878	/	HexNAc2Hex1Sia1		6.38	3.46	2.16	4.00	2.16	6.05	4.44	3.43	4.64	1.32	0.344
895	/	HexNAc2Hex2Fuc1		0.74	0	0	0.25	0.43	0.35	0	0	0.12	0.20	0.334
966	/	HexNAc1Hex1Sia2		3.77	11.25	9.05	8.02	3.84	6.32	9.04	9.39	8.25	1.68	0.466
1040	/	HexNAc2Hex2Sia1		8.09	2.82	1.2	4.04	3.60	9.11	3.04	4.93	5.69	3.11	0.290
1040	/	HexNAc2Hex2Sia1		11.86	3.85	2.05	5.92	5.22	11.98	5.77	6.99	8.25	3.29	0.278
1114	/	HexNAc3Hex3		1.77	0	0	0.59	1.02	0.27	0	0	0.09	0.16	0.244
1114	/	HexNAc3Hex3		0.87	0	0	0.29	0.50	0.15	0	0	0.05	0.09	0.248
1186	/	HexNAc2Hex2Sia1Fuc1		0.27	0	0	0.09	0.16	0.37	0	0	0.12	0.21	0.419
1186	/	HexNAc2Hex2Sia1Fuc1		2.26	0	0	0.75	1.30	1.14	0	0	0.38	0.66	0.344
1331	665	HexNAc2Hex2Sia2		22.75	15.58	15.81	18.05	4.07	41.62	41.5	42.3	41.81	0.43	0.005 **
1477	738	HexNAc2Hex2Sia2Fuc1		0	0	0	0.00	0.00	0.48	0	0	0.16	0.28	0.211
1696	848	HexNAc3Hex3Sia2		0	8.78	11.34	6.71	5.95	0	3.99	4.64	2.88	2.51	0.061
1696	848	HexNAc3Hex3Sia2		0	7.09	10.78	5.96	5.48	0	8.65	6.74	5.13	4.54	0.314
/	1030	HexNAc4Hex4Sia2		9.38	18.42	22.07	16.62	6.53	0	8.51	5.34	4.62	4.30	0.033 *
/	1213	HexNAc5Hex5Sia2		8.09	14.96	15.15	12.73	4.02	0	0	0	0.00	0.00	0.016 *
			Total:	100	100	100			100	100	100			

Structures are represented by the mass to charge ratio (m/z) in which they were identified and quantified, by monosaccharide composition and proposed structure based on MS/MS analyzes. The relative quantities were determined by base-peak intensity of extracted ion chromatograms. The average value (Avg) and standard deviation (SD) of triplicates are shown, as well as the p-value (t-test; $^{*}p < 0.05$; $^{**}p < 0.01$). Increased or decreased relative abundance are shown in red or blue, respectively. Unknown linkage is represented by "?".

bicarbonate (TEAB) to remove any remaining soluble protein. Membrane fraction was resuspended directly in 6 M urea and 2 M thiourea, reduced in 10 mM DTT for 30 min and then alkylated in 20 mM IAA for 30 min at room temperature in the dark.

Samples were incubated with endoproteinase Lys-C (Wako, Osaka, Japan) for 2 h (1:100 w/w). Following the incubation, the samples were diluted 8 times with 50 mM TEAB (pH 8) and trypsin was added at a ratio of 1:50 (w/w) and left overnight at room temperature. Trypsin digestion was stopped by the addition of 2% formic acid and then the samples were centrifuged at 14,000 × g for 10 min to

precipitate any lipids present in the sample. The supernatant was purified using in-house packed staged tips with a mixture of Poros R2 and Oligo R3 reversed phase resins (Applied Biosystem, Foster City, CA, USA). Briefly, a small plug of C18 material (3 M Empore) was inserted in the end of a P200 tips, followed by packing of the stage tip with the resins (resuspended in 100% ACN) by applying gentle air pressure. The acidified samples were loaded onto the micro-column after equilibration of the column with 0.1% trifluoroacetic acid (TFA), washed twice with 0.1% TFA and peptides were eluted with 60% ACN/0.1% TFA. A small amount of purified peptides (1 μl) from each sample was subjected to Qubit assay to determine the concentration, while the remaining samples were dried by vacuum centrifugation. Afterwards, peptides were redissolved in dissolution buffer and a total of 150 μg for each condition was labeled with 4-plex iTRAQ™ (Applied Biosystems, Foster City, CA) as described by the manufacturer. After labeling, the samples were mixed 1:1:1:1 and lyophilized by vacuum centrifugation.

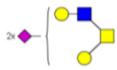
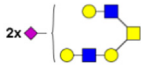
7.2. *Sialic acid containing glycopeptide enrichment by TiSH protocol*

The method used for sialylated glycopeptides enrichment is a modification of the TiSH protocol [6] described in [7,8]. Briefly, samples were resuspended in loading buffer (1 M glycolic acid, 80% ACN, 5%

Table 4
Identified *O*-glycan structures II.

[M-H]-	z	Composition	Type	Assigned structures	Mock	ST3GAL4	ST3GAL4/mock
675.30	1	HexNAc1Hex1Sia1	core 1		0.00	1.48	/
675.31	1	HexNAc1Hex1Sia1	core 1		0.00	1.17	/
749.40	1	HexNAc2Hex2	core 2		0.00	0.00	/
895.39	1	HexNAc2Hex2Fuc1	core 1 / 3	/	0.00	0.36	/
966.40	1	HexNAc1Hex1Sia2	core 1		0.00	0.09	/
1040.43	1	HexNAc2Hex2Sia1	core 1	/	18.81	17.40	0.93
1040.43	1	HexNAc2Hex2Sia1	core 2		2.33	1.44	0.62
1040.44	1	HexNAc2Hex2Sia1	core 2		6.65	7.91	1.19
1186.49	1	HexNAc2Hex2Sia1Fuc1	core 2 / 4	/	3.05	4.29	1.40
1186.52	1	HexNAc2Hex2Sia1Fuc1	core 2 / 4	/	0.51	0.15	0.30
1227.53	1	HexNAc3Hex1Sia1Fuc1	core 2 / 4	/	0.00	2.47	/
1331.49	1	HexNAc2Hex2Sia2	core 2		0.00	1.32	/
1331.52	1	HexNAc2Hex2Sia2	core 2	/	0.99	0.20	0.20

Table 4 (*continued*)

1331.57	1	HexNAc2Hex2Sia2	core 2		2.90	1.01	0.35
1389.56	1	HexNAc3Hex2Sia1Fuc1	core 2 / 4	/	26.90	45.91	1.71
1389.58	1	HexNAc3Hex2Sia1Fuc1	core 2 / 4	/	0.00	0.00	/
1405.55	1	HexNAc3Hex3Sia1	core 2 / 4	/	0.00	0.10	/
1405.55	1	HexNAc3Hex3Sia1	core 2 / 4	/	0.00	1.10	/
1405.58	1	HexNAc3Hex3Sia1	core 2 / 4	/	0.00	0.00	/
1405.59	1	HexNAc3Hex3Sia1	core 2 / 4	/	1.84	0.31	0.17
1477.56	1	HexNAc2Hex2Sia2Fuc1	core 2 / 4	/	3.01	0.57	0.19
1477.58	1	HexNAc2Hex2Sia2Fuc1	core 2 / 4	/	1.28	0.00	0.00
1477.60	1	HexNAc2Hex2Sia2Fuc1	core 2 / 4	/	1.49	0.17	0.11
1696.62	2	HexNAc3Hex3Sia2	core 2 / 4	/	0.00	2.17	/
1696.63	2	HexNAc3Hex3Sia2	core 2		1.81	0.00	0.00
1696.66	2	HexNAc3Hex3Sia2	core 2 / 4	/	0.60	0.00	0.00
2061.76	2	HexNAc4Hex4Sia2	core 2 / 4	/	1.46	0.00	0.00
2061.79	2	HexNAc4Hex4Sia2	core 2 / 4	/	1.61	0.44	0.27
2061.80	2	HexNAc4Hex4Sia2	core 2 / 4	/	10.19	4.86	0.48
				Total:	100	100	

Structures are represented by their [M–H] value and charge state in which they were identified and quantified, by monosaccharide composition, type of core and proposed structure based on MS/MS analyzes. The relative quantities were determined by base-peak intensity of extracted ion chromatograms. The ratio of glycan abundance in *ST3GAL4* transfected cells relative to mock transfected is shown on the right column. Increases or decreases greater than 1.5 fold in glycan abundance are highlighted red or blue, respectively.

TFA) and incubated with TiO_2 beads (GL Sciences, Japan, 10 μm; using a total of 0.6 mg TiO_2 beads per 100 μg of peptides). The supernatant containing the un-modified peptides was carefully separated. The TiO_2 beads were sequentially washed with loading buffer, washing buffer 1 (80% ACN, 1% TFA) and washing buffer 2 (20% ACN, 0.1% TFA), saving the washings with the previous supernatant. The bound peptides were eluted with 1.5% ammonium hydroxide by shaking for 15 min. The eluted fraction containing the phosphopeptides and sialylated glycopeptides was dried by vacuum centrifugation and subjected to an enzymatic deglycosylation in 20 mM TEAB buffer using 500 U of PNGase F (New England Biolabs, Ipswich, MA) and 0.1 U of Sialidase A (Prozyme, Hayward, CA) overnight at 37 °C.

Table 5
Proteins with increased *N*-glycan sialylation.

Accession number	Protein name	Sequence	*N*-glycan site	ST3GAL4 /Mock	Limma&Rank (*p*-value)	Func. Grp.
O15031	Plexin-B2	ISVAGRNdeCSFQPER	N844	2.89	0.0342	RS
		ALSNdeISLR	N127	2.08	0.0358	
O60637	Tetraspanin-3	TYNdeGTNdePDAASR	N127	2.96	0.0358	RS
		TYNdeGTNPDAASR	N127	2.98	0.0433	
O75882	Attractin	GICNdeSSDVR	N300	2.35	0.0342	IR
P02786	Transferrin receptor protein 1	KDFEDLYTPVNdeGSIVIVR	N251	1.72	0.0433	TT
P05026	Sodium/potassium-transporting ATPase subunit beta-1	FKLEWLGNdeCSGLNDETYGYK	N158	2.79	0.0479	TT
P06213	Insulin receptor	HNdeLTITQGK	N445	2.45	0.0479	RS
P06731	Carcinoembryonic antigen-related cell adhesion molecule 5	TLTLFNdeVTR	N204	2.67	0.0342	AM
		TLTLFNVTRNdeDTASYK	N208	3.11	0.0342	
P06756	Integrin alpha-V	ISSLQTTEKNdeDTVAGQGER	N874	4.02	0.0342	AM
		NdeMTISR	N554	4.02	0.0342	
P07602	Prosaposin	TNdeSTFVQALVEHVKEECDR	N215	2.59	0.0479	O
P08962	CD63 antigen	CCGAANdeYTDWEK	N150	1.94	0.0394	RS
P10909	Clusterin	EIRHNdeSTGCLR	N291	2.12	0.0342	O
P11117	Lysosomal acid phosphatase	QTPEYQNdeESSR	N177	3.05	0.0383	E
P12821	Angiotensin-converting enzyme	IGLLDRVTNdeDTESDINYLLK	N445	2.62	0.0456	P
P13473	Lysosome-associated membrane glycoprotein 2	LNdeSSTIK	N275	3.09	0.0342	AM
		VASVINdeINPNdeTTHSTGSCR	N253	1.96	0.0358	
		VASVININPNdeTTHSTGSCR	N257	2.06	0.0456	
P13688	Carcinoembryonic antigen-related cell adhesion molecule 1	NdeQSLPSSER	N363	2.76	0.0479	AM
P13726	Tissue factor	RNdeNTFLSLR	N169	5.52	0.0342	P
		NdeNTFLSLR	N169	5.76	0.0358	
P18564	Integrin beta-6	EVEVNdeSSK	N463	3.71	0.0479	AM
P21589	5'-nucleotidase	LDNdeYSTQELGK	N333	1.67	0.0479	E
P26006	Integrin alpha-3	ELAVPDGYTNdeRTGAVYLCPLTAHK	N86	2.68	0.0433	AM
P30825	High affinity cationic amino acid transporter 1	LCLNNdeDTK	N235	3.48	0.0297	TT
		LCLNdeNdeDTK	N234	3.01	0.0342	
P35613	Basigin	ALMNdeGSESR	N268	2.91	0.0342	P
P42892	Endothelin-converting enzyme 1	NdeSSVEAFK	N632	1.89	0.0394	P
		DYYLNdeKTENEK	N270	5.64	0.0479	
P43007	Neutral amino acid transporter A	VVTQNdeSSSGNdeVTHEK	N201	2.58	0.0477	TT
P46059	Solute carrier family 15 member 1	NdeDSCPEVK	N562	4.66	0.0358	TT
P48960	CD97 antigen	WCPQNSSCVNdeATACR	N38	1.87	0.0479	AM
P54760	Ephrin type-B receptor 4	CAQLTVNdeLTRFPETVPR	N203	2.34	0.0342	RS
Q04912	Macrophage-stimulating protein receptor	LPEYVVRDPQGWVAGNdeLSAR	N841	3.26	0.0358	RS

Table 5 (*continued*)

Accession number	Protein name	Sequence	*N*-glycan site	ST3GAL4 /Mock	Limma&Rank (*p*-value)	Func. Grp.
Q08380	Galectin-3-binding protein	DAGVVCTNdeETR	N125	2.43	0.0358	AM
Q08722	Leukocyte surface antigen CD47	SDAVSHTGNdeYTCEVTELTREGETIIELK	N111	6.51	0.0056	AM
Q11206	ST3GAL4	LFGNdeYSR	N61	5.77	0.0088	O
Q12913	Receptor-type tyrosine-protein phosphatase eta	HGSNdeHTSTYDK	N582	3.15	0.0342	RS
		VSDNdeESSSNdeYTYK	N391	2.94	0.0342	
		VSDNESSSNdeYTYK	N396	2.74	0.0383	
		GPNdeGTEGASR	N525	2.76	0.0427	
		IHVAGETDSSNdeLNdeVSEPR	N411	1.69	0.0479	
		SNdeDTAASEYK	N142	2.77	0.0479	
Q13641	Trophoblast glycoprotein	CVNRNdeLTEVPTDLPAYVR	N81	1.85	0.0429	RS
Q14108	Lysosome membrane protein 2	ANIQFGDNdeGTTISAVSNdeK	N105	2.03	0.0394	T
		ANdeIQFGDNdeGTTISAVSNK	N99	1.78	0.0479	
Q15043	Zinc transporter ZIP14	ALLNHLDVGVGRGNdeVTQHVQGHR	N77	2.42	0.0479	TT
Q15758	Neutral amino acid transporter B(0)	NdeITGTR	N212	2.93	0.0479	TT
Q7Z7H5	Transmembrane emp24 domain-containing protein 4	QYGSEGRFTFTSHTPGDHQICLHSNdeSTR	N117	2.50	0.0479	T
Q8N271	Prominin-2	ILRNdeVSECFLAR	N725	3.11	0.0433	O
Q92542	Nicastrin	RPNdeQSQPLPPSSLQR	N417	1.66	0.0483	P
Q92673	Sortilin-related receptor	LTIVNdeSSVLDRPR	N871	1.92	0.0479	T
		SRNdeSTVEYTLNK	N1986	1.67	0.0479	
Q9BXB1	Leucine-rich repeat-containing G-protein coupled receptor 4	TLDLSYNNIRDLPSFNdeGCHALEEISLQR	N362	4.00	0.0358	RS
Q9BXS4	Transmembrane protein 59	LFSICQFVDDGIDLNdeRTK	N90	4.09	0.0342	T
Q9H330	Transmembrane protein 245	ILGDKVNdeNTAVIEK	N551	3.47	0.0433	O
Q9H5V8	CUB domain-containing protein 1	ASVSFLNFNdeLSNCERK	N270	3.33	0.0342	AM
		IGTFCSNdeGTVSR	N180	2.93	0.0358	
		NdeVSGFSIANR	N205	2.33	0.0358	
Q9HD43	Receptor-type tyrosine-protein phosphatase H	NdeATTAHNPVR	N203	6.86	0.0342	RS
		ETRNdeATTAPNPVR	N381	2.93	0.0479	
		NdeTTNTSVTAER	N434	2.72	0.0479	
Q9P2B2	Prostaglandin F2 receptor negative regulator	QRNdeNSWVK	N525	3.60	0.0429	RS
Q9UN76	Sodium- and chloride-dependent neutral and basic amino acid transporter B(0+)	SPIVTHCNdeVSTVNdeK	N174	2.44	0.0420	TT
Q9Y639	Neuroplastin	ANdeATIEVK	N229	2.59	0.0429	AM

List of significantly increased sialylated *N*-glycan modified peptides in the ST3GAL4 overexpressing cells compared to mock control shown with accession number, protein name, peptide sequence and the identified *N*-glycan site. The fold in increase, p-value and the protein's functional group are also presented. AM: Adhesion and migration; E: Metabolic enzyme; RS: Receptor and signaling; T: Trafficking; TT: Transmembrane transporter; P: Protease; O: Others.

Table 6
Proteins with decreased *N*-glycan sialylation.

Accession number	Protein name	Sequence	N-glycan site	ST3GalIV/Mock	Limma & rank (*p*-value)
O14672	Disintegrin and metalloproteinase domain-containing protein 10	NdeISQVLEK	N439	0.57	0.0149
P07602	Proactivator polypeptide	NdeSTKQEILAALEK	N426	0.57	0.0474

List of significantly decreased sialylated *N*-glycan modified peptides in the ST3GAL4 overexpressing cells compared to mock control shown with accession number, protein name, peptide sequence and the identified *N*-glycan site, fold in increase and the *p*-value.

To separate phosphorylated peptides and formerly glycosylated peptides, the samples were subjected to a second TiO_2 enrichment procedure to separate phosphorylated from deglycosylated peptides. The supernatant containing the deglycosylated peptides was saved and the beads were washed with 50% ACN, 0.1% TFA. The washing was added to the supernatant. The deglycosylated fraction was desalted on Oligo R3 staged tip column and dried prior to the HILIC fractionation [7]. All fractions were dried by vacuum centrifugation prior nLC-MS/MS analysis.

7.3. Sialic acid containing glycopeptide analysis by nLC-MS/MS

Samples were resuspended in 6 μL of 0.1% TFA for analysis. Peptides were loaded on an in-house packed Reprosil-Pur C18-AQ (2 cmx100 μm, 5 μm; Dr. Maisch GmbH, Germany) pre-column and separated on an in-house packed Reprosil-Pur C18-AQ (17 cmx75 μm, 3 μm; Dr. Maisch GmbH, Germany) column using an Easy-nLC II system (Thermo Scientific, Bremen, Germany) and eluted at a flow of 250 nL/min. Mobile phase was 95% acetonitrile (B) and water (A) both containing 0.1% formic acid. Depending on the samples, gradient was from 1% to 30% solvent B in 80 or 110 min, 30–50% B in 10 min, 50–100% B in 5 min and 8 min at 100% B. Mass spectrometric analyses were performed in an Orbitrap Fusion Tribrid system (Thermo Scientific, Bremen, Germany). MS scans (400–1200 m/z) were acquired in the orbitrap at a resolution of 120000 at 200 m/z for a AGC target of 5×10^5 ions and a maximum injection time of 60 ms. Data-dependent HCD MS/MS analysis at top speed of the most intense ions were performed at a resolution of 30000 at 200 m/z for a AGC target of 5×10^4 and a maximum injection time of 150 ms using the quadrupole to isolate the ions and an isolation window of 1.2 m/z, a NCE of 38% and a dynamic exclusion of 20 s.

The raw data were processed and quantified by Proteome Discoverer (version 1.4.1.14, Thermo Scientific) against SwissProt and Uniprot human reference databases by using Mascot (v2.3.02, Matrix Science Ltd, London, UK) and Sequest HT, respectively. Database searches were performed using the following parameters: precursor mass tolerance of 10 ppm, product ion mass tolerance of 0.02 Da, 1 missed cleavages for trypsin, carbamidomethylation of Cys and iTRAQ labeling on protein *N*-terminal and Lys as fixed modifications, and phosphorylation on S/T/Y and deamidation of Asn as dynamic modifications. The iTRAQ datasets were quantified using the centroid peak intensity with the “reporter ions quantifier” node. Only peptides with up to a *q*-value of 0.01 (Percolator), Mascot and Sequest HT rank 1, Sequest HT ΔCn of 0.1, cut-off value of Mascot score $\geq$ 18 and a cut-off value of XCorr score for charge states of +1, +2, +3, and +4 higher than 1.5, 2, 2.25 and 2.5, respectively, were considered for further analysis.

7.4. Data normalization and significance analysis

Three biological replicates were analyzed and submitted to the statistical analysis. The log2 values of the measured intensities were normalized by the median. Modified peptides were merged with the R Rollup function (http://www.omics.pnl.gov) allowing for one-hit-wonders and using the mean of the normalized intensities for each peptide. Quantification of proteins was obtained by merging the

un-modified peptides with the R Rollup function considering at least 2 unique peptides not allowing for one-hit-wonders and using the mean of the intensities. Then the mean over the experimental conditions for each peptide in each replicate was subtracted in order to merge data from different iTRAQ runs. Formerly sialylated glycopeptides containing the consensus motif for *N*-linked glycosylation (NXS/T/C; where X # P) were normalized based on the protein expression in each of the replicates. Significant up/down-regulations between experimental conditions were calculated allowing a false discovery rate of 0.05. Therefore, we applied combined limma and rank product tests [9], subsequently corrected for multiple testing according to Storey.

Since spontaneous deamidation is frequently observed for asparagine residues, especially when the C-terminal amino acid is glycine (NG), the sites with NGS/T/C are considered as only potential glycosylation. However, in order to reduce the contribution from spontaneous deamidation in the final list, we sort first for the *N*-linked consensus site (NXS/T/C) and then we filter for proteins that are membrane-associated in order to exclude intracellular proteins that are not *N*-linked glycosylated (Tables 5 and 6).

Acknowledgments

We acknowledge the support from the European Union, Seventh Framework Programme, Gastric Glyco Explorer Initial Training Network: Grant number 316929. IPATIMUP integrates the i3S Research Unit, which is partially supported by FCT, the Portuguese Foundation for Science and Technology. This work is funded by FEDER funds through the Operational Programme for Competitiveness Factors-COMPETE (FCOMP-01-0124-FEDER028188) and National Funds through the FCT-Foundation for Science and Technology, under the projects: PEst-C/SAU/LA0003/2013, PTDC/BBB-EBI/0786/2012, PTDC/BBB-EBI/0567/2014 (CR). This work was also supported by "Glycoproteomics" project Grant number PCIG09-GA-2011-293847 (to DK) and the Danish Natural Science Research Council and a generous Grant from the VILLUM Foundation to the VILLUM Center for Bioanalytical Sciences at the University of Southern Denmark (to MRL). AM acknowledges FCT, POPH (Programa Operacional Potencial Humano) and FSE (Fundo Social Europeu) (SFRH/BPD/75871/2011). The UPLC instrument was obtained with a grant from the Ingabritt and Arne Lundbergs Research Foundation. C.J. was supported by the Knut and Alice Wallenberg Foundation. The mass spectrometer (LTQ) was obtained by a grant from the Swedish Research Council (342-2004-4434).

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2016.03.022.

References

[1] S. Mereiter, A. Magalhães, B. Adamczyk, C. Jin, A. Almeida, L. Drici, M. Ibáñez-Vea, C. Gomes, J.A. Ferreira, L.P. Afonso, L.L. Santos, M.R. Larsen, D. Kolarich, N.G. Karlsson, C.A. Reis, Glycomic analysis of gastric carcinoma cells discloses glycans as modulators of RON receptor tyrosine kinase activation in cancer, Biochim Biophys Acta, (in press).

[2] A.S. Carvalho, A. Harduin-Lepers, A. Magalhaes, E. Machado, N. Mendes, L.T. Costa, R. Matthiesen, R. Almeida, J. Costa, C. A. Reis, Differential expression of alpha-2,3-sialyltransferases and alpha-1,3/4-fucosyltransferases regulates the levels of sialyl lewis a and sialyl lewis x in gastrointestinal carcinoma cells, Int. J. Biochem. Cell Biol. 42 (2010) 80–89.

[3] B.L. Schulz, N.H. Packer, N.G. Karlsson, Small-scale analysis of O-linked oligosaccharides from glycoproteins and mucins separated by gel electrophoresis, Anal. Chem. 74 (2002) 6088–6097.

[4] P.H. Jensen, N.G. Karlsson, D. Kolarich, N.H. Packer, Structural analysis of N-and O-glycans released from glycoproteins, Nat. Protoc. 7 (2012) 1299–1310.

[5] D. Kolarich, M. Windwarder, K. Alagesan, F. Altmann, Isomer-specific analysis of released N-glycans by LC-ESI MS/MS with porous graphitized carbon, Methods Mol. Biol. 1321 (2015) 427–435.

[6] K. Engholm-Keller, P. Birck, J. Storling, F. Pociot, T. Mandrup-Poulsen, M.R. Larsen, TiSH – a robust and sensitive global phosphoproteomics strategy employing a combination of TiO_2, SIMAC, and HILIC, J. Proteom. 75 (2012) 5749–5761.

[7] M.N. Melo-Braga, M. Ibanez-Vea, M.R. Larsen, K. Kulej, Comprehensive protocol to simultaneously study protein phosphorylation, acetylation, and N-linked sialylated glycosylation, Methods Mol. Biol. 1295 (2015) 275–292.
[8] M.R. Larsen, S.S. Jensen, L.A. Jakobsen, N.H. Heegaard, Exploring the sialiome using titanium dioxide chromatography and mass spectrometry, Mol. Cell. Proteom.: MCP 6 (2007) 1778–1787.
[9] V. Schwammle, I.R. Leon, O.N. Jensen, Assessment and improvement of statistical tools for comparative proteomics analysis of sparse data sets with few experimental replicates, Journal. Proteom. Res. 12 (2013) 3874–3883.

ADDITIONAL RESULTS: *N*- AND *O*-GLYCOMIC ANALYSIS OF MKN45 GASTRIC CARCINOMA CELLS OVEREXPRESSING THE SIALYLTRANSFERASE ST6GALNAC1

1. Material and methods

1.1 Cell culture

The gastric carcinoma cell line MKN45 was obtained from the Japanese Cancer Research Bank (Tsukuba, Japan) and was stably transfected with the full length human *ST6GALNAC1* gene and the corresponding empty vector pcDNA3.1 (mock) as previously shown [1]. The cells were grown in monolayer culture in uncoated cell culture flasks. Cells were maintained at 37°C in an atmosphere of 5% CO2, in RPMI 1640 GlutaMAX, HEPES medium supplemented with 10% fetal bovine serum (FBS), 1% penicillin-streptomycin and in the presence of 0.5 mg/mL G418 (all from Invitrogen, Waltham, MA). Cell culture medium was replaced every two days.

1.2 Sample preparation and analysis

Samples were prepared and analyzed as described previously in this chapter. Briefly, frozen cell pellets (10^7 cells) of *ST6GALNAC1* transfected cells or mock were directly resuspended in 7 M urea, 2 M thiourea, 40 mM Tris, 2% CHAPS, 10 mM DTT and 1% protease inhibitor (Sigma-Aldrich, St. Louis, MO). The cell membranes were disrupted by sonication and the viscosity of the lysates was reduced by benzonase® nuclease (250 units, Sigma-Aldrich). Iodoacetamide was added and supernatants collected. *N*-linked oligosaccharides of the supernatant glycoproteins were released on 10 kDa cut-off spinfilter (PALL, Port Washington, NY) by PNGase F (Prozyme, Hayward, CA). The released *N*-glycans were collected, dried in Speedvac and overnight reduced with 0.5 M $NaBH_4$, 10 mM NaOH. *O*-linked oligosaccharides were released from retained glycoproteins in spinfilter by reductive β-elimination (0.5 M NaBH4, 50 mM NaOH). Reactions were quenched with 1 µl of glacial acetic acid and *N*- and *O*-glycans were

desalted and dried. Released glycans were analyzed by LC-ESI-MS/MS using a column containing 5 µm porous graphitized carbon (PGC) particles (Thermo Scientific, Waltham, MA). Glycans were eluted using an acetonitrile gradient in 10 mM NH_4HCO_3. The eluted *N*- and *O*-glycans were detected using a LTQ ion trap mass spectrometer (Thermo Scientific) in negative-ion mode of electrospray ionization. The data were processed using the Xcalibur software (version 2.0.7, Thermo Scientific) and manually interpreted from their MS/MS spectra. Structural information on sialic acid linkage were inferred from the in depth structural elucidation on MKN45 cells (chapter 2.1). The identified glycan structures were quantified by the area under the peak at the extracted ion chromatogram.

2. Results

We identified 18 *N*-glycan structures in both ST6GALNAC1 transfected cells and mock control, covering pauci-mannose, oligo-mannosidic, hybrid and complex *N*-glycans (table 1). The same *N*-glycan structures were identified in both cell lines, and only limited quantitative differences were detected, indicating no effect of ST6GALNAC1 overexpression on the *N*-glycome.

The *O*-glycomic analysis revealed 19 *O*-glycan structures, including STn, core 1, core 2 and core 3 structures (table 2). Most structures identified had terminal sialic acids with core 2 structures being the most elongated. ST6GALNAC1 overexpressing cells showed in accordance with previous results a significant increase in STn. In addition, as a result of ST6GALNAC1 overexpression the α2,6-ST structure seemed to be slightly increased. Surprisingly, we also found a significant increase in unsialylated and monosialylated core 2 structures.

Table 1: Identified *N*-linked glycans of ST6GALNAC1 and mock transfected cells. For each structure theoretical and detected mass of singly charged (sc) and doubly charged (dc) ion, retention time (rt) of elution, area under the peak (intensity) and relative abundance (%) are displayed.

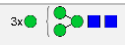
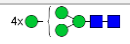
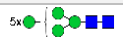
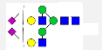
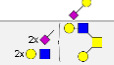
Theoretical mass		Composition	Proposed structure	Detected mass		Mock			ST6GALNAC1			
sc	dc			sc	dc	rt (min)	Intensity	%	rt (min)	Intensity	%	Ratio
911,34	/	HexNAc2Hex3		911,03	/	17,52-17,76	10391,2	0,14	17,51-17,76	7167,6	0,12	**0,89**
1057,4	/	HexNAc2Hex3Fuc1		1057,09	/	23,30-23,64	67457,5	0,88	23,03-23,29	74470	1,25	**1,42**
1235,44	/	HexNAc2Hex5		1235,07	/	21,57-21,91	146487,5	1,91	21,34-21,68	100866	1,69	**0,88**
1397,5	698,24	HexNAc2Hex6		1397,08	698,31	19,29-19,60	1226141	16,01	19,30-19,52	904836,4	15,18	**0,95**
1559,55	779,27	HexNAc2Hex7		1559,03	779,17	18,78-19,07	1314644	17,17	18,78-19,07	1047276	17,57	**1,02**
1713,62	856,31	HexNAc3Hex4Sia1Fuc1		1713,12	856,83	23,72-24,13	341831,4	4,46	23,29-23,70	232156,2	3,89	**0,87**
1721,6	860,3	HexNAc2Hex8		1721,02	860,3	18,93-19,21	1644482	21,47	18,86-19,15	1457053	24,44	**1,14**
/	864,3	HexNAc3Hex5Sia1		/	864,12	21,16-21,49	97387	1,27	20,93-21,26	55601,5	0,93	**0,73**
/	937,34	HexNAc3Hex5Sia1Fuc1		/	937,19	23,05-23,39	83674	1,09	22,69-23,03	68687	1,15	**1,05**
/	941,33	HexNAc2Hex9		/	941,09	19,14-19,44	645903	8,43	19,07-19,38	526569,5	8,83	**1,05**
/	945,33	HexNAc3Hex6Sia1		/	945,05	22,46-22,71	164440	2,15	22,01-22,35	114456	1,92	**0,89**
/	965,84	HexNAc4Hex5Sia1		/	965,61	22,54-23,05	153594	2,01	22,10-22,69	105002,4	1,76	**0,88**
/	1022,35	HexNAc2Hex10		/	1022,1	19,93-20,18	112636,8	1,47	19,77-20,02	87682	1,47	**1,00**
/	1038,87	HexNAc4Hex5Sia1Fuc1		/	1038,52	24,38-24,79	315053,4	4,11	23,96-24,37	273725,4	4,59	**1,12**
/	1111,39	HexNAc4Hex5Sia2		/	1111,07	24,88-25,29	281691,6	3,68	24,46-25,14	125333,1	2,10	**0,57**
/	1111,39	HexNAc4Hex5Sia2		/	1111,12	28,14-28,73	183112,8	2,39	27,85-28,87	121747,6	2,04	**0,85**
/	1184,42	HexNAc4Hex5Sia2Fuc1		/	1184,08	26,47-26,88	630633,6	8,24	26,07-26,75	380829,6	6,39	**0,78**
/	1184,42	HexNAc4Hex5Sia2Fuc1		/	1184,06	29,56-30,15	238402,4	3,11	29,38-30,31	278018,4	4,66	**1,50**
Total:							7657963	100		5961478	100	

Table 2: Identified *O*-linked glycans of ST6GALNAC1 and mock transfected cells. For each structure theoretical and detected mass of singly charged (sc) and doubly charged (dc) ion, retention time (rt) of elution, area under the peak (intensity) and relative abundance (%) are displayed.

Theoretical mass		Composition	Proposed structure	Detected mass		Mock			ST6GALNAC1			
sc	dc			sc	dc	rt (min)	Intensity	%	rt (min)	Intensity	%	Ratio
384,15	/	HexNAc1Hex1		384,09	/	5,64-6,41	33872,4	1,3%	5,02-5,55	57664	2,3%	**1,74**
513,19	/	HexNAc1Sia1		513,09	/	11,95-12,18	5238,8	0,2%	14,27-14,82	61897,6	2,5%	**12,09**
587,23	/	HexNAc2Hex1		587,11	/			0,0%	18,64-19,11	19113,5	0,8%	
675,25	/	HexNAc1Hex1Sia1		675,11	/	13,59-14,07	80399,9	3,2%	16,06-16,61	168817,6	6,8%	**2,15**
675,25	/	HexNAc1Hex1Sia1		675,12	/	16,25-16,79	288041,6	11,3%	18,72-19,27	285753,6	11,5%	**1,01**
733,29	/	HexNAc2Hex1Fuc1		733,07	/	16,95-17,11	7194	0,3%	20,12-20,27	10173,6	0,4%	**1,45**
749,28	/	HexNAc2Hex2		749,09	/	15,63-15,79	6359,4	0,2%	18,80-19,11	63956	2,6%	**10,29**
749,28	/	HexNAc2Hex2		749,09	/	16,48-16,64	5402,7	0,2%	19,65-19,81	15597,3	0,6%	**2,95**
749,28	/	HexNAc2Hex2		749,09	/	16,95-17,11	15933,3	0,6%	20,12-20,35	23124,4	0,9%	**1,48**
878,36	/	HexNAc2Hex1Sia1		878,11	/	24,81-25,38	27052,8	1,1%	26,47-26,80	49573	2,0%	**1,87**
895,34	/	HexNAc2Hex2Fuc1		895,08	/	13,28-13,59	3933	0,2%	16,53-16,85	14475	0,6%	**3,77**
966,34	/	HexNAc1Hex1Sia2		966	/	20,08-21,28	375865,6	14,8%	21,28-22,23	191331,4	7,7%	**0,52**
1040,38	/	HexNAc2Hex2Sia1		1040,08	/	21,51-22,00	60053,7	2,4%	23,51-23,99	168144,9	6,8%	**2,86**
1040,38	/	HexNAc2Hex2Sia1		1040,05	/	22,48-23,22	132493	5,2%	24,32-24,72	56116,8	2,3%	**0,43**
1040,38	/	HexNAc2Hex2Sia1		1040	/	24,56-24,65	34399,6	1,4%	25,99-26,39	156181,8	6,3%	**4,64**
1114,42	/	HexNAc3Hex3		1114,09	/	20,24-20,55	5890,5	0,2%	22,96-23,28	30929	1,2%	**5,37**
1331,47	665,23	HexNAc2Hex2Sia2		/	665,13	28,57-29,28	154088	6,1%	29,77-30,50	160701	6,5%	**1,07**
2061,74	1030,37	HexNAc4Hex4Sia2		/	1030,13	33,41-36,42	700233,6	27,5%	33,56-35,22	471209,2	18,9%	**0,69**
2426,87	1212,93	HexNAc5Hex5Sia2		/	1212,45	34,79-37,72	610155,9	24,0%	34,44-38,07	484536	19,5%	**0,81**
Total:							2546608	100%		2489296	100%	

3. Discussion

The sialyltransferase ST6GALNAC1 catalyzes the addition of α2,6 linked sialic acid to the Tn antigen, leading to the formation of sialyl Tn (STn). We demonstrate here, in accordance to previous results, that the overexpression of ST6GALNAC1 leads to increased levels of STn. We also found an increase in ST with α2,6-linked sialic acid in cells with ST6GALNAC1 overexpression. This is according to expectation as ST6GALNAC1 has slightly overlapping substrate specificities with ST6GALNAC2, the major enzyme responsible for the formation of α2,6-ST [1, 2]. Reciprocally, we found disialyl-T slightly decreased as α2,6-ST precludes the addition of α2,3-linked sialic acid [2].

Furthermore, unsialylated and monosialylated core 2 structures were increased and disialylated core 2 structures decreased. Direct interaction of ST6GALNAC1 and core-2 synthases have never been reported to our knowledge. We hypothesize that the increased formation of STn reduced the amount and availability of the sialyl-CMP donor in the Golgi apparatus. The reduction of available sialyl-CMP could explain the decline of other completely sialylated structures and the increase of unsialylated and incomplete sialylated structures. Minor alterations in amounts of sialylated *N*-glycans were detected, with most of them being slightly reduced.

Increase of ST6GALNAC1 expression increased the formation of STn and changed indirectly the abundance of several other *O*-glycan structures likewise. This underlines that numerous glycan structures are affected when single enzymes of the glycosylation machinery are changed. It also highlights the imperative of glycomic evaluation when biological conclusions are to be drawn from glycoengineered cell lines.

4. References

1 Marcos, N. T., Pinho, S., Grandela, C., Cruz, A., Samyn-Petit, B., Harduin-Lepers, A., . . . Reis, C. A. (2004) Role of the human ST6GalNAc-I and ST6GalNAc-II in the synthesis of the cancer-associated sialyl-Tn antigen. Cancer Res. 64, 7050-7057

2 Harduin-Lepers, A., Vallejo-Ruiz, V., Krzewinski-Recchi, M. A., Samyn-Petit, B., Julien, S. and Delannoy, P. (2001) The human sialyltransferase family. Biochimie. 83, 727-737

CHAPTER III

GLYCOMIC ANALYSIS OF GASTRIC CARCINOMA CELLS DISCLOSES GLYCANS AS MODULATORS OF RON RECEPTOR TYROSINE KINASE ACTIVATION IN CANCER

Biochimica et Biophysica Acta 1860 (2016) 1795–1808

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagen

Glycomic analysis of gastric carcinoma cells discloses glycans as modulators of RON receptor tyrosine kinase activation in cancer☆

Stefan Mereiter [a,b,c], Ana Magalhães [a,b], Barbara Adamczyk [d], Chunsheng Jin [d], Andreia Almeida [e,f], Lylia Drici [g], Maria Ibáñez-Vea [g], Catarina Gomes [a,b], José A. Ferreira [a,b,h], Luis P. Afonso [i], Lúcio L. Santos [h,j], Martin R. Larsen [g], Daniel Kolarich [e], Niclas G. Karlsson [d], Celso A. Reis [a,b,c,k,*]

[a] I3S — Instituto de Investigação e Inovação em Saúde, University of Porto, Portugal
[b] Institute of Molecular Pathology and Immunology of the University of Porto — IPATIMUP, Porto, Portugal
[c] Institute of Biomedical Sciences of Abel Salazar — ICBAS, University of Porto, Portugal
[d] Department of Medical Biochemistry and Cell Biology, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Sweden
[e] Department of Biomolecular Systems, Max Planck Institute of Colloids and Interfaces, 14424 Potsdam, Germany
[f] Free University Berlin, Berlin, Germany
[g] Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark
[h] Experimental Pathology and Therapeutics Group, Portuguese Institute of Oncology of Porto, Portugal
[i] Department of Pathology, Portuguese Institute of Oncology of Porto, Portugal
[j] Department of Surgical Oncology, Portuguese Institute of Oncology of Porto, Portugal
[k] Medical Faculty, University of Porto, Portugal

ARTICLE INFO

Article history:
Received 30 October 2015
Received in revised form 18 December 2015
Accepted 19 December 2015
Available online 22 December 2015

Keywords:
ST3GAL4
Sialyl Lewis X (SLeX)
RON
Gastric cancer
Glycome
Sialome

ABSTRACT

Background: Terminal α2-3 and α2-6 sialylation of glycans precludes further chain elongation, leading to the biosynthesis of cancer relevant epitopes such as sialyl-Lewis X (SLe^X). SLe^X overexpression is associated with tumor aggressive phenotype and patients' poor prognosis.
Methods: MKN45 gastric carcinoma cells transfected with the sialyltransferase *ST3GAL4* were established as a model overexpressing sialylated terminal glycans. We have evaluated at the structural level the glycome and the sialoproteome of this gastric cancer cell line applying liquid chromatography and mass spectrometry. We further validated an identified target expression by proximity ligation assay in gastric tumors.
Results: Our results showed that ST3GAL4 overexpression leads to several glycosylation alterations, including reduced *O*-glycan extension and decreased bisected and increased branched *N*-glycans. A shift from α2-6 towards α2-3 linked sialylated *N*-glycans was also observed. Sialoproteomic analysis further identified 47 proteins with significantly increased sialylated *N*-glycans. These included integrins, insulin receptor, carcinoembryonic antigens and RON receptor tyrosine kinase, which are proteins known to be key players in malignancy. Further analysis of RON confirmed its modification with SLe^X and the concomitant activation. SLe^X and RON co-expression was validated in gastric tumors.
Conclusion: The overexpression of ST3GAL4 interferes with the overall glycophenotype of cancer cells affecting a multitude of key proteins involved in malignancy. Aberrant glycosylation of the RON receptor was shown as an alternative mechanism of oncogenic activation.
General significance: This study provides novel targets and points to an integrative tumor glycomic/proteomic-profiling for gastric cancer patients' stratification. This article is part of a Special Issue entitled "Glycans in personalised medicine" Guest Editor: Professor Gordan Lauc.

1. Introduction

Even more than four decades after declaring "war" on cancer at least 7 million patients die annually from the consequences of this disease, which represents 15% of all deaths worldwide [1,2]. Gastric cancer is globally the fifth most common cancer with near 1 million cases diagnosed in 2012 [3] and represents the third most common cause of cancer-related death. Due to its usually late diagnosis in already advanced stages it remains difficult to treat even in developed countries. Nowadays, there is

Abbreviations: SLe^a, Sialyl Lewis A; SLe^X, Sialyl Lewis X; PTM, Post-translational modifications; RTK, Receptor tyrosine kinase; IHC, Immunohistochemistry; PLA, Proximity ligation assay; GC, gastric cancer; HILIC-FLD-UPLC, hydrophilic interaction liquid chromatography fluorescence detection ultra-performance liquid chromatography.

☆ This article is part of a Special Issue entitled "Glycans in personalised medicine" Guest Editor: Professor Gordan Lauc.

* Corresponding author at: Institute of Molecular Pathology and Immunology of the University of Porto — IPATIMUP, Porto, Portugal.
E-mail address: celsor@ipatimup.pt (C.A. Reis).

http://dx.doi.org/10.1016/j.bbagen.2015.12.016
0304-4165/

no specific serological assay for the screening and diagnosis of gastric cancer, and endoscopy remains the gold standard in the clinical practice. However, the SLe^a glycan antigen CA19-9 is currently used for monitoring gastric cancer patients' response to treatment and recurrence [4,5].

One of the most abundant forms of posttranslational modifications (PTM) of proteins is glycosylation, which is a complex process coordinated by the interplay of numerous glycosyltransferases and glycosidases, resulting in a vast diversity of structures. In cancer the disruption of the glycosylation machinery leads to aberrant expression of short truncated carbohydrate chains, known as cancer-associated simple carbohydrate antigens, and to altered expression of terminal sialylated chains [6,7]. These cancer-associated antigens are detected in different types of carcinomas and are associated with disease prognosis, constituting a pool of potential cancer biomarkers, especially when combined with information on the carrying proteins.

It has been long recognized that each cancer patient is different regarding prognosis, clinical presentation, tumor response and tolerance to treatment. With the constant improvement in sequencing and other large-scale analytical technologies, molecular tumor profiling has become unprecedentedly feasible [8]. This enables the identification of the unique combination of alterations, especially that of well-established cancer-related genes, for each patient and personalize their treatment. This customized attempt of tackling cancer has been on the focus of researchers for a long time, but for many cancer-associated alterations the understanding of the biological consequences and therapy-response implications lags behind [9].

One typical target of the personalized therapy strategies is the inhibition of activated signaling pathways, either by interfering directly with the receptor activation or interrupting the downstream signaling cascade by blocking specific GTPases [10]. Receptors such as EGFR, c-MET and RON (Recepteur d'Origine Nantais) can be targeted by small molecules or monoclonal antibodies that act as inhibitors by blocking ligand binding and signaling or by antibodies that tag tumor cells for immune response [10,11]. For instance, the RON receptor tyrosine kinase (also known as macrophage-stimulating protein receptor (MSPR) or MST1R) is constitutively transcribed in many epithelial cells but commonly shows aberrant activation in various tumors due to overexpression and generation of oncogenic variants. Currently several therapeutic agents and monoclonal antibodies targeting RON are tested in clinical and preclinical phases [12].

Therefore, two main goals in gastric cancer research are the discovery of new biomarkers for early diagnosis [5] and the elucidation of altered mechanisms that can be targeted for the development of novel directed and personalized treatments.

In the present work, we have performed a comprehensive analysis of cellular clones derived from MKN45 gastric carcinoma cells overexpressing the human α2,3-sialyltransferase *ST3GAL4*. The *N*- and *O*-glycomic, sialoproteomic and gene expression analysis were performed using a combination of high-throughput methods which revealed several cancer-associated glycomic alterations and generated a list of 47 putative glycoprotein targets with altered glycosylation in gastric cancer cells. Further characterization of one of the targets, the RON receptor tyrosine kinase, confirmed the altered glycosylation and revealed an increased activation of this oncogene in the gastric cancer cells. Finally, we demonstrate that this altered glycosylation of RON receptor tyrosine kinase occurs in human gastric tumors, showing that glycosylation alterations are an alternative pathway for cancer cells to activate receptors that lead to malignant growth, and which can be targets for personalized therapy in gastric cancer patients.

2. Material and Methods

2.1. Cell culture

The gastric carcinoma cell line MKN45 was obtained from the Japanese Cancer Research Bank (Tsukuba, Japan) and was stably transfected with the full length human *ST3GAL4* gene and the corresponding empty vector pcDNA3.1 (Mock) as previously shown [13]. The cells were grown in monolayer culture in uncoated cell culture flasks or cell culture flasks coated with fibronectin or collagen IV (BD BioCoat, BD Biosciences, Franklin Lakes, NJ). Cells were maintained at 37 °C in an atmosphere of 5% CO_2, in RPMI 1640 GlutaMAX, HEPES medium supplemented with 10% fetal bovine serum (FBS), 1% penicillin-streptomycin and in the presence of 0.5 mg/mL G418 (all from Invitrogen, Waltham, MA). Cell culture medium was replaced every two days.

2.2. Quantitative real-time PCR (RT-PCR)

Total RNA extracts from mock and *ST3GAL4* transfected cell lysates were isolated with TRI Reagent (Sigma-Aldrich, St. Louis, MO) and converted into cDNA using the SuperScript® II Reverse Transcriptase (Invitrogen) according to the manufacturer's protocol. The following primers for ST3GAL4 were used applying the protocol previously described [13]: for5′-cctggtagctttcaaggcaatg-3′; rev5′-cctttcgcacccgcttct-3′. Expression level of 18S rRNA was used for mRNA expression normalization (for5′-cgccgctagaggtgaaattc-3′; rev5′-cattcttggcaaatgctttcg-3′).

2.3. RNA next generation sequencing

Total RNA was extracted as previously described in Section 2.2. The mRNAs of over 20,000 primed targets were sequenced by using Ion AmpliSeq Transcriptome Human Gene Expression Kit. The Ion Chef system was used for templating and the loaded chips sequenced using the Ion Proton System (both from Life Technologies). Sequencing data was automatically transfered to the dedicated Ion Torrent server to generate sequencing reads. Reads quality and trimming was performed using Torrent Server v4.2 before read alignment using TMAP 4.2. The TS plugin CoverageAnalysis v4.2 was used to generate reads count. The sequencing was performed in duplicates and sequence reads were normalized to the total read count.

2.4. Immunofluorescence (IF)

Cells were grown in 15 μ-Chamber 12 well glass slides (IBIDI, Martinsried, Germany) and either fixed with 4% paraformaldehyde (PFA) or cold acetone for 15 and 5 min, respectively. The IF protocol performed was as previously described [13] using the antibodies and dilutions noted in Table 1. Samples were examined under a Zeiss Imager.Z1 Axio fluorescence microscope (Zeiss, Welwyn Garden City, UK). Images were acquired using a Zeiss Axio cam MRm and the AxioVision Release 4.8.1 software.

2.5. Protein and phosphoprotein arrays

Confluent cells were lysed in lysis buffer 17 or 6 (R&D Systems, McKinley Place, MN) supplemented with 1 mM sodium orthovanadate, 1 mM phenylmethanesulfonylfluoride (PMSF) and protease inhibitor cocktail (Roche, Basel, Switzerland). The protein concentrations of lysates were determined by the DC protein assay (BioRad, Hercules, CA) and the recommended total protein amounts were used for the human non-haematopoietic soluble receptor array, human phospho-RTK array kit and human phospho-kinase array (R&D Systems). The array protocols were performed according to manufacturer's instructions.

2.6. Western (WB) and Lectin blotting (LB)

Proteins were obtained from total cell lysates as previously described in Section 2.5. Total lysates were denatured and charged using Laemmli buffer, separated by SDS-PAGE and blotted onto a nitrocellulose membrane (GE Healthcare, Chalfont, UK) in a semi-dry system (BioRad). Membranes were incubated with primary antibodies overnight at 4 °C or with biotinylated lectins for 2 h at room temperature (Table 1). Secondary antibodies or avidin (ABC Standard Kit, Vector

Laboratories, Burlingame, CA) conjugated with peroxidase were incubated for 1 h at room temperature. Target proteins were revealed by chemiluminescence using the ECL Western blotting detection reagent and films (both from GE Healthcare).

2.7. In situ proximity ligation assay (PLA)

In situ Proximity Ligation Assay (PLA) was performed in acetone fixed cells as described in 2.4 or in paraffin sections from human gastric carcinoma tissues for the detection of co-expression in proximity of SLe^X and RON. PLA was performed adapting the procedure previously described [14]. Duolink II reagents (Olink Bioscience, Uppsala, Sweden) were used according to the manufacture instructions. Paraffin sections were dewaxed, rehydrated, antigen retrieval using sodium citrate buffer (10 mM, pH 6.0) was performed before sections were incubated with blocking solution (Olink Bioscience) for 30 min at 37 °C. Primary antibodies against SLe^X and RON were used in equal concentrations as for IF (Table 1) and incubated overnight at 4 °C. Antibodies conjugated with oligonucleotides were utilized as secondary probes (DUO92001 and DUO92005 from Olink Bioscience). Ligation and amplification were performed at 37 °C for 30 min or 120 min respectively and cell nuclei were visualized by DAPI (Sigma-Aldrich, 0.4 mg/ml). Samples were examined under a Zeiss Imager.Z1 Axio fluorescence microscope (Zeiss, Welwyn Garden City, UK) equipped with DAPI, FITC and Texas Red filters. Proximity ligation assays products are seen as bright red fluorescent dots. Images were acquired using a Zeiss Axio cam MRm and the AxioVision Rel. 4.8 software. The resulting images were modified using ImageJ as follows: background with radius 5 was subtracted from the red channel of the RGB images and a multiply filter of 10 was applied. The result was intensity-scaled to suit printing demands.

2.8. Immunohistochemistry (IHC)

Tissue samples from gastric carcinoma patients were obtained from the archives of the Portuguese Institute of Oncology (IPO), Porto, Portugal. All procedures were performed after patient's written informed consent and approved by the local Ethical committee. All clinicopathological information was obtained from patients' clinical records. Expression of SLe^X and RON were evaluated in 15 cases of human gastric carcinomas (10 cases of intestinal subtype and 5 cases of diffuse subtype according to Lauren's classification [15]). Paraffin sections were dewaxed, rehydrated and antigen retrieval was carried out by microwave treatment in sodium citrate buffer (10 mM, pH 6.0) for 20 min. Endogenous peroxidases were blocked with 3% hydrogen peroxide (H_2O_2) in methanol. Tissue sections were further blocked for 30 min with normal goat serum in PBS with 10% BSA, followed by incubation with primary antibodies against SLe^X and RON (Table 1) overnight at 4 °C. Biotin-labeled secondary antibodies (Dako, Glostrup, Denmark) were applied for 30 min and the ABC kit (Vector Labs, Burlingame, CA) for 30 min. Finally, sections were stained by 3,3′-diaminobenzidine tetrahydrochloride (DAB) and counterstained with Mayers' hematoxylin solution. Slides were examined using a Zeiss Optical Microscope.

2.9. Sample preparation for LC-ESI-MS/MS and HILIC-FLD-UPLC analyses

Frozen cell pellets (10^7 cells) were directly resuspended in 7 M urea, 2 M thiourea, 40 mM Tris, 2% CHAPS, 10 mM DTT and 1% protease inhibitor (Sigma-Aldrich, St. Louis, MO). The cell membranes were disrupted by 10 times 10 sec sonication with 16 amplitudes and 1 minute on ice in between, and subsequent shaking at 4 °C overnight. To reduce the viscosity of the lysates, the DNA was degraded by adding 1 μl benzonase® nuclease (250 units, Sigma-Aldrich) and 30 min incubation at 37 °C. In order to impair refolding of proteins, 25 mM iodoacetamide were added for alkylation during 1 h in the dark. The lysates were centrifuged for 30 min with 14,000 rpm and the supernatants transferred to a fresh tube.

Then, solubilized proteins were concentrated by adding 150 μl of supernatant on a 10 kDa cut-off spinfilter (PALL, Port Washington, NY), spinning down for 5 min with 12,000 x g and washing 3 times with 100 μl 50 mM NH_4HCO_3, pH 8.4. *N*-linked oligosaccharides were released in the spinfilter using 20 μl 50 mM NH_4HCO_3 and PNGase F (5 mU, Prozyme, Hayward, CA) with incubation at 37 °C overnight. Subsequently, the *N*-glycans were collected by washing 3 times with 20 μl H_2O and dried in Speedvac. For UPLC analysis samples were further processed as described below and for LC-ESI-MS/MS analysis samples were reduced with 0.5 M $NaBH_4$, 10 mM NaOH at 50 °C overnight. The *O*-linked glycans were released from retained glycoproteins in spinfilter using reductive β-elimination (0.5 M $NaBH_4$, 50 mM NaOH at 50 °C, 16 h). Reactions were quenched with 1 μl of glacial acetic acid and glycan (both *N*-glycans and *O*-glycans) samples were desalted and dried as previously described [16] and subjected to LC-ESI-MS/MS analysis.

2.10. LC-ESI-MS/MS for N- and O-glycomic analysis

Released glycans were analyzed by LC-ESI-MS/MS using a 10 cm x 250 μm I.D. column, prepared in-house, containing 5 μm porous graphitized carbon (PGC) particles (Thermo Scientific, Waltham, MA). Glycans were eluted using a linear gradient from 0 to 40% acetonitrile in 10 mM NH_4HCO_3 over 40 min at a flow rate of 10 μl/min. The eluted *N*- and *O*-glycans were detected using a LTQ ion trap mass spectrometer (Thermo Scientific) in negative-ion mode with an electrospray voltage of 3.5 kV, capillary voltage of −33.0 V and capillary temperature of 300 °C. Air was used as a sheath gas and mass ranges were defined dependent on the specific structure to be analyzed. The data were processed using the Xcalibur software (version 2.0.7, Thermo Scientific) and manually interpreted from their MS/MS spectra.

2.11. Ultra performance liquid chromatography (UPLC)

Released *N*-glycans were labeled by reductive amination with the fluorophore 2-aminobenzamide (2-AB) (Sigma-Aldrich, St.Louis, MO) with sodium cyanoborohydride in 30% v/v acetic acid in DMSO at 65 °C for 2 h. Excess of 2-AB reagent was removed on Glycoworks HILIC cartridges according to the manufacturer's instructions (Waters, Milford, MA) and then concentrated to dryness in speed-vac.

Hydrophilic interaction liquid chromatography (HILIC) of fluorescently labeled *N*-glycans was carried out on a 1.7 μm BEH glycan column (2.1 mm x 15 mm, Waters, Milford, MA) and analyzed using

Table 1
Antibodies and lectins.

Antibody/lectin	Manufacturer	Type	Dil. for WB	Dil. for IHC/IF	Ref.
KM93; Anti-SLe^X	Millipore, Billerica, MA	Mouse IgM	1:500	1:60	[22]
C-20, Anti-RON	Santa Cruz Biotechnology, Dallas, TX	Rabbit IgG	1:1000	1:60	
Anti-p-RON	Santa Cruz Biotechnology	Rabbit IgG	1:1000	1:60	
SNA, biotinylated	Vector Laboratories, Burlingame, CA	Lectin	1:600		[80]

Waters ACQUITY UPLC® I-class with fluorescence detection. The column temperature was kept at 40 °C and the flow rate set to 0.561 mL/min using a linear gradient of 50 mM ammonium formate (pH 4.4) against acetonitrile with ammonium formate increasing from 30% to 47% over a 25 min period. An injection volume of 25 μL sample prepared in 70% v/v acetonitrile was used throughout. Fluorescence detection was achieved using excitation and emission wavelengths of 330 nm and 420 nm, respectively.

The 2-AB labeled glycans were digested in 10 μL of 50 mM sodium phosphate, pH 6.0 at 37 °C overnight using sialidase S (4 mU, ProZyme) that releases α2-3 linked non-reducing terminal sialic acids (recombinant sialidase from *Streptococcus pneumoniae*, expressed in *Escherichia coli*) and sialidase A (5 mU, ProZyme) that releases α2-3/6/8 linked non-reducing terminal sialic acid (recombinant gene from *Arthrobacter ureafaciens*, expressed in *Escherichia coli*) to confirm sialic acid linkage. After incubation, enzymes were removed by filtration through a 10 kDa cut-off spinfilter (PALL, Port Washington, NY) and the *N*-glycans were analysed by HILIC-FLD-UPLC. The system was calibrated by running an external standard of 2AB–dextran ladder (2AB–glucose homopolymer, Ludger, Oxfordshire, UK) alongside the sample runs. A fifth-order polynomial distribution curve was fitted to the dextran ladder and used to allocate GU values from retention times (using Empower 3 software from Waters) [17].

2.12. Cell lysis, protein digestion and iTRAQ labeling

Cell pellets were redissolved in ice-cold Na_2CO_3 buffer (0.1 M, pH 11) supplemented with protease inhibitor (Roche complete EDTA free), PhosSTOP phosphatase inhibitor cocktail (Roche) and 10 mM sodium pervanadate on ice. The suspensions were tip probe sonicated for 20 s (amplitude = 50%) twice and incubated at 4 °C for 1 h. The lysates were then centrifuged at 100,000× g for 90 min at 4 °C to separate soluble proteins from membrane proteins (pellet). The pellets were washed with 50 mM triethylammonium bicarbonate (TEAB) to remove any remaining soluble protein. The supernatants containing soluble proteins were concentrated using 10 kDa cutoff Amicon ultra centrifugal filters units (Millipore, Billerica, MA. USA) while membrane fractions were resuspended directly in 6 M urea and 2 M thiourea. Soluble and membrane fractions were both reduced in 10 mM DTT for 30 min and then alkylated in 20 mM IAA for 30 min at room temperature in the dark.

Samples were incubated with endoproteinase Lys-C (Wako, Osaka, Japan) for 2 h (1:100 w/w). Following the incubation, the samples were diluted 8 times with 50 mM TEAB (pH 8) and trypsin was added at a ratio of 1:50 (w/w) and left overnight at room temperature. Trypsin digestion was stopped by the addition of 2% formic acid and then the samples were centrifuged at 14,000× g for 10 min to precipitate any lipids present in the sample. The supernatant was purified using in-house packed staged tips with a mixture of Poros R2 and Oligo R3 reversed phase resins (Applied Biosystem, Foster City, CA, USA). Briefly, a small plug of C18 material (3 M Empore) was inserted in the end of a P200 tip, followed by packing of the stage tip with the resins (resuspended in 100% ACN) by applying gentle air pressure. The acidified samples were loaded onto the micro-column after equilibration of the column with 0.1% trifluoroacetic acid (TFA), washed twice with 0.1% TFA and peptides were eluted with 60% ACN/0.1% TFA. A small amount of purified peptides (1 μl) from each sample was subjected to Qubit assay to determine the concentration, while the remaining samples were dried by vacuum centrifugation. Afterwards, peptides were redissolved in dissolution buffer and a total of 150 μg for each condition was labeled with 4-plex iTRAQ™ (Applied Biosystems, Foster City, CA) as described by the manufacturer. After labeling, the samples were mixed 1:1:1:1 and lyophilized by vacuum centrifugation.

2.13. Sialic acid containing glycopeptide enrichment by TiSH protocol

The method used for sialylated glycopeptides enrichment is a modification of the TiSH protocol [18] described in [19]. Briefly, samples were resuspended in loading buffer (1 M glycolic acid, 80% ACN, 5% TFA) and incubated with TiO_2 beads (GL Sciences, Japan, 10 μm; using a total of 0.6 mg TiO_2 beads per 100 μg of peptides). The supernatant containing the unmodified peptides was carefully separated. The TiO_2 beads were sequentially washed with loading buffer, washing buffer 1 (80% ACN, 1% TFA) and washing buffer 2 (20% ACN, 0.1% TFA), saving the washings with the previous supernatant. The bound peptides were eluted with 1.5% ammonium hydroxide by shaking for

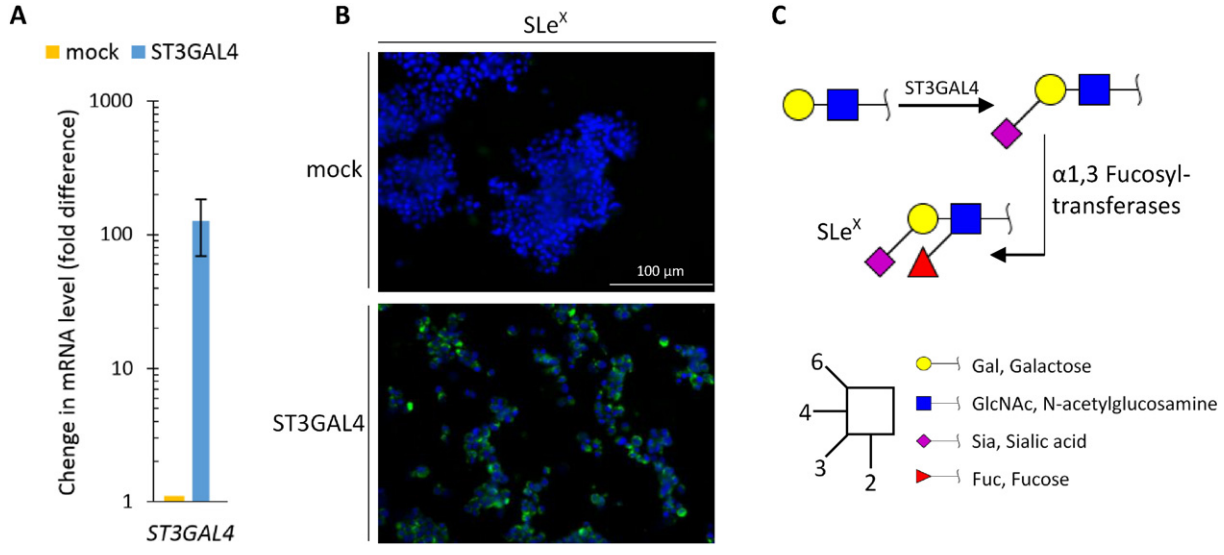


Fig. 1. Gastric carcinoma cell line overexpressing ST3GAL4 shows upregulation of SLe^X expression. A. Relative quantification of *ST3GAL4* mRNA expression using quantitative real-time PCR. MKN45 cells transfected with ST3GAL4 show an approximately 150 times increase when compared with mock transfected MKN45 cells. Values were based on quadruplicates of cell lines. Results are presented as average ± SD. B. Cells that overexpress ST3GAL4 display an increased expression of the SLe^X epitope as shown by immunofluorescence using anti-SLe^X antibody (KM93). C. Depiction of the addition of α2-3 sialic acid to Galβ1-4GlcNAc, which is the major reaction that is catalyzed by the sialyltransferase ST3GAL4 as previously described [18]. The glycan structure Galβ1-4GlcNAc is commonly found in both *N*- and *O*-glycans and is referred to as type 2 chain. This structure can be further fucosylated leading to an increase of the SLe^X glycan epitope.

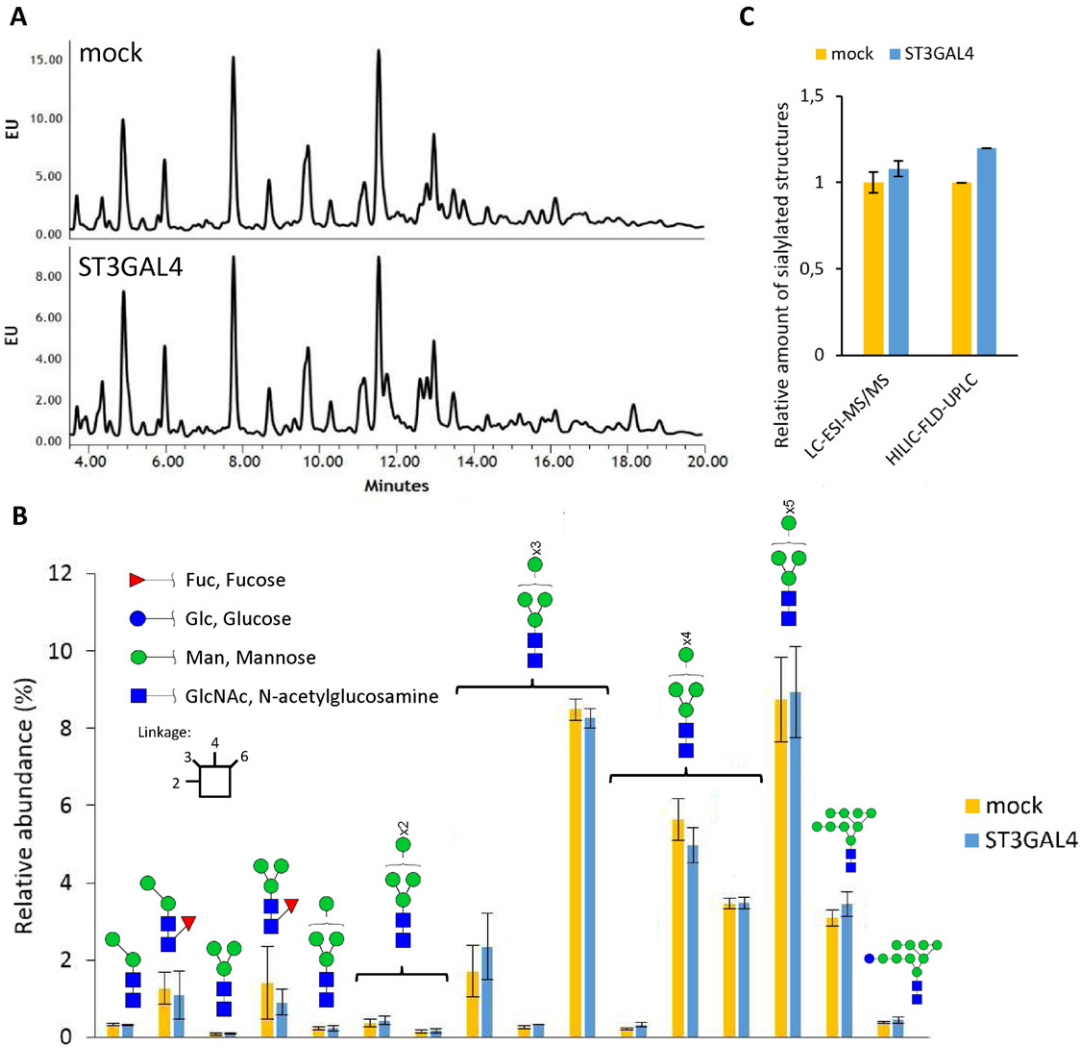


Fig. 2. The overall *N*-glycome of ST3GAL4 overexpressing cells is largely unaltered. A. The quantitative HILIC-FLD-UPLC chromatograms of 2AB labeled N-glycans show highly similar *N*-glycan profiles of *ST3GAL4* transfected and control cell line. EU: Emission unit. B. Relative quantification of high mannose structures reveals that the expression of this group of *N*-glycans is unaffected when compared between ST3GAL4 overexpressing cells and the mock transfected control. *N*-glycans were quantified by evaluating base-peak intensity of extracted ion chromatograms from LC-ESI-MS and were performed in three biological replicates. Values are presented as average ± SD. C. Quantification of the relative amount of sialylated *N*-glycan structures by LC-ESI-MS and HILIC-FLD-UPLC shows no significant increase when comparing sialyltransferase ST3GAL4 overexpressing cells with mock control.

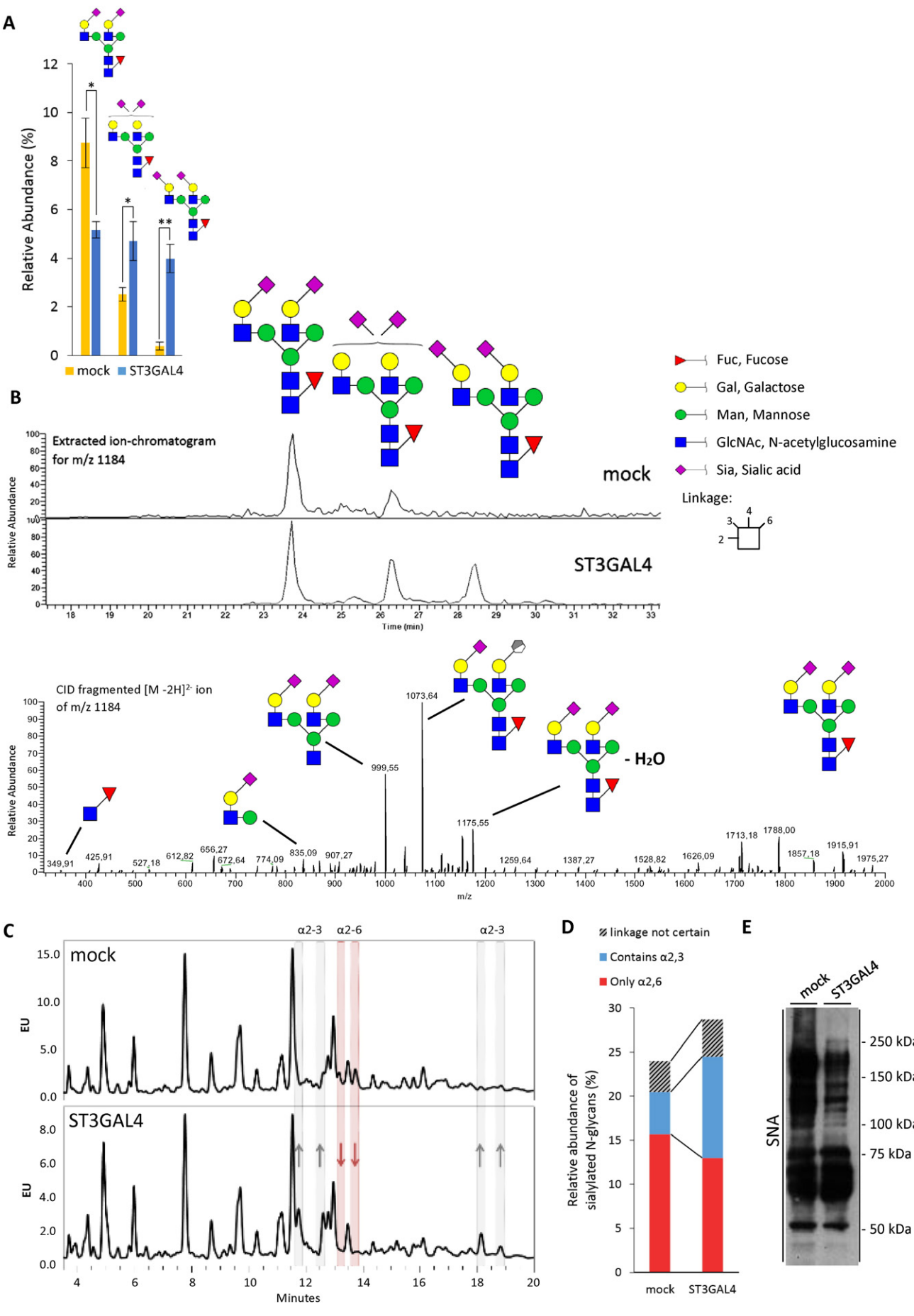
15 min. The eluted fraction containing the phosphopeptides and sialylated glycopeptides was dried by vacuum centrifugation and subjected to an enzymatic deglycosylation in 20 mM TEAB buffer using 500 U of PNGase F (New England Biolabs, Ipswich, MA) and 0.1 U of Sialidase A (Prozyme, Hayward, CA) overnight at 37 °C.

To separate phosphorylated peptides and formerly glycosylated peptides, the samples were subjected to a second TiO_2 enrichment procedure to separate phosphorylated from deglycosylated peptides. The supernatant containing the deglycosylated peptides was saved and the beads were washed with 50% ACN, 0.1% TFA. The washing was added to the supernatant. The deglycosylated fraction was desalted on Oligo R3 staged tip column and dried prior to the HILIC fractionation [19]. All fractions were dried by vacuum centrifugation prior nLC–MS/MS analysis.

2.14. Sialic acid containing glycopeptide analysis by nLC–MS/MS

Samples were resuspended in 6 µL of 0.1% TFA for analysis. Peptides were loaded on an in-house packed Reprosil-Pur C18-AQ (2 cm x 100 µm, 5 µm; Dr. Maisch GmbH, Germany) pre-column and separated on an in-house packed Reprosil-Pur C18-AQ (17 cm x 75 µm, 3 µm; Dr. Maisch GmbH, Germany) column using an Easy-nLC II system (Thermo Scientific, Bremen, Germany) and eluted at a flow of 250 nL/min. Mobile phase was 95% acetonitrile (B) and water (A) both containing 0.1% formic acid. Depending on the samples, gradient was from 1% to 30% solvent B in 80 or 110 min, 30 - 50% B in 10 min, 50 - 100% B in 5 min and 8 min at 100% B. Mass spectrometric analyses were performed in an Orbitrap Fusion Tribrid system (Thermo Scientific, Bremen, Germany). MS scans (400 – 1200 m/z) were acquired in the orbitrap at a resolution of 120000 at 200 m/z for a AGC target of 5 x 10^5 ions and a maximum injection time of 60 ms. Data-dependent HCD MS/MS analysis at top speed of the most intense ions were performed at a resolution of 30,000 at 200 m/z for a AGC target of 5 x 10^4 and a maximum injection time of 150 ms using the quadrupole to isolate the ions and an isolation window of 1.2 m/z, a NCE of 38% and a dynamic exclusion of 20 s.

The raw data were processed and quantified by Proteome Discoverer (version 1.4.1.14, Thermo Scientific) against SwissProt and Uniprot human reference database by using Mascot (v2.3.02, Matrix Science Ltd, London, UK) and Sequest HT, respectively. Database searches were performed using the following parameters: precursor mass tolerance of 10 ppm, product ion mass tolerance of 0.02 Da, 1 missed cleavages for trypsin, carbamidomethylation of Cys and iTRAQ


A
Relative Abundance (%)
mock
ST3GAL4
B
Extracted ion-chromatogram for m/z 1184
mock
ST3GAL4
Relative Abundance
Time (min)
Fuc, Fucose
Gal, Galactose
Man, Mannose
GlcNAc, N-acetylglucosamine
Sia, Sialic acid
Linkage:
CID fragmented [M -2H]2- ion of m/z 1184
- H2O
m/z
C
mock
ST3GAL4
EU
Minutes
α2-3
α2-6
D
linkage not certain
Contains α2,3
Only α2,6
Relative abundance of sialylated N-glycans (%)
E
mock
ST3GAL4
SNA
250 kDa
150 kDa
100 kDa
75 kDa
50 kDa

labelling on protein N-terminal and Lys as fixed modifications, and phosphorylation on S/T/Y and deamidation of Asn as dynamic modifications. The iTRAQ datasets were quantified using the centroid peak intensity with the "reporter ions quantifier" node. Only peptides with up to a q-value of 0.01 (Percolator), Mascot and Sequest HT rank 1, Sequest HT ΔCn of 0.1, cut off value of Mascot score ≥ 18 and a cut-off value of XCorr score for charge states of +1, +2, +3, and +4 higher than 1.5, 2, 2.25 and 2.5, respectively, were considered for further analysis.

2.15. Data normalization and significance analysis

Three biological replicates were analysed and submitted to the statistical analysis. The log2 values of the measured intensities were normalized by the median. Modified peptides were merged with the R Rollup function (http://www.omics.pnl.gov) allowing for one-hit-wonders and using the mean of the normalized intensities for each peptide. Quantification of proteins was obtained by merging the unmodified peptides with the R Rollup function considering at least 2 unique peptides not allowing for one-hit-wonders and using the mean of the intensities. Then the mean over the experimental conditions for each peptide in each replicate was subtracted in order to merge data from different iTRAQ runs. Formerly sialylated glycopeptides containing the consensus motif for N-linked glycosylation (NXS/T/C; where X # P) were normalized based on the protein expression in each of the replicates. Significant up/down-regulations between experimental conditions were calculated allowing a false discovery rate of 0.05. Therefore, we applied combined limma and rank product tests [20], subsequently corrected for multiple testing according to Storey.

Since spontaneous deamidation is frequently observed for asparagine residues, especially when the C-terminal amino acid is glycine (NG), the sites with NGS/T/C are considered as only potential glycosylation. However, in order to reduce the contribution from spontaneous deamidation in the final list, we sort first for the N-linked consensus site (NXS/T/C) and then we filter for proteins that are membrane-associated in order to exclude intracellular proteins that are not N-linked glycosylated.

3. Results

3.1. Overexpression of ST3GAL4 leads to increased expression of SLe^X

The gastric cancer cell line MKN45 shows low expression levels of the sialyltransferase ST3GAL4 and of the terminal glycan epitope sialyl Lewis X (SLe^X) [13]. We have generated clones of MKN45 cells that were stably transfected with the full length *ST3GAL4* resulting in approximately 150 times upregulation of this gene expression (Fig. 1a). ST3GAL4 has been described to be upregulated in the context of gastric cancer and catalyzes the sialylation of Galβ1-3GalNAc on *O*-glycans and of type 2 (Galβ1-4GlcNAcβ-) extensions on both *N*- and *O*-glycans, an intermediate epitope in the biosynthesis of SLe^X [21–23]. As a consequence of the increased expression of ST3GAL4, the terminal glycan epitope SLe^X is overexpressed (Fig. 1). Previous studies on this cell line have demonstrated increased invasive potential both *in vitro* and *in vivo* compared to the mock transfected control cell line [24].

3.2. The N-glycome of ST3GAL4 overexpressing cells is specifically altered in α2-6/3 linked sialic acid content, bisecting and branched structures

To assess the full extent of glycomic alterations in our cell line model we performed a whole *N*-glycomic analysis based on LC-ESI-MS/MS (Table 1 in [25]), to gain semi-quantitative and detailed structural information, and HILIC-FLD-UPLC for quantitative comparisons. The HILIC-FLD-UPLC spectra shows no alteration in the major peaks, which are of high-mannose nature and expectedly unaffected by alterations in sialyltransferase levels. This was confirmed by LC-ESI-MS/MS analysis (Fig. 2a and b). Moreover, the total amount of sialylated structures was not significantly altered (Fig. 2c). It should be noted that the structural analysis revealed also several truncated glycans (Table 1 in [25]) previously described in the MKN45 cell line as free *N*-glycans [26]. Alterations in free *N*-glycans are in conformity with their non-truncated equivalents and are thus not further described in this study.

Despite the overall similarity of the *N*-glycome of cancer cells expressing ST3GAL4 and the control cell line, detailed analysis revealed several specific alterations in the cancer cells expressing ST3GAL4. Our results showed that the sialic acids linked α2-6 were reduced and the α2-3 linked reciprocally increased, indicating a shift of linkage in the *ST3GAL4* transfected cells (Fig. 3). Nevertheless, α2-6 linked sialic acid structures remain a large component of the ST3GAL4 cells' glycome. Since expression levels of sialyltransferases of the ST6 family are unaltered as shown by RNASeq (Supplementary table 1) we hypothesize that the observed isomeric change may be due to competition for the acceptor substrate.

Additionally, bisecting *N*-glycan structures were significantly decreased (Fig. 4). The presence of bisecting *N*-acetylglucosamine excludes the possibility of adding a β1-6 branched arm to the complex *N*-glycan and therefore often inversely correlates with the amount of large branched structures [27,28]. In our cancer cell line model, the amount of large and presumably branched complex *N*-glycans is increased (Fig. 4).

3.3. O-glycomic analysis reveals an increase in truncated O-glycans

Mucin type *O*-glycosylation is commonly altered in cancer with truncation as the most common aberration [29]. MKN45, as a gastric epithelial derived cancer cell line, produces mainly core 2 *O*-glycans [30]. Our analyses showed high amounts of sialylated *O*-glycans in the MKN45 cell line with the majority of structures being decorated with at least one sialic acid (Table 3 in [25]). However, the overexpression of ST3GAL4 leads to an earlier termination by sialylation and thus, to the formation of truncated *O*-glycans (Fig. 5). This is mainly due to the increase in doubly sialylated core 2 structure which accounts for more than 40% of the total *O*-glycan amount in our transfected cells. On the other hand, *O*-glycans composed of more than 2 *N*-acetylhexosamines or hexoses were reduced by approximately 30% (Fig. 5).

3.4. Identification of over 40 glycoproteins with altered glycosylation

The sialoproteomic analysis performed on protein extracts of *ST3GAL4* and mock transfected cells revealed glycoproteins that display aberrant glycosylation. The analysis quantified 1566 unique glycopeptides of which 69 had a significantly altered abundance in sialylated

Fig. 3. Sialylation of *N*-glycans in ST3GAL4 overexpressing cells shows a shift from α2-6 linked towards α2-3 linked. A. Relative quantification of disialylated biantennary structures illustrates the general trend that α2-6 linked sialic acids are reduced whereas α2-3 linked are increased. The quantifications are established by base-peak intensity of extracted ion chromatograms from LC-ESI-MS performed in three biological replicates and presented as average ± SD. Statistical significance was determined by unpaired student's t-test (p-value *<0.05; **<0.01). B. Extracted ion-chromatogram for m/z 1184 shows the relative increase of disialylated biantennary structures with α2-3 linked sialic acids when compared to α2-6 linked. Structures were assigned by a combination of MS^2 spectra (shown below) and α2-3 specific sialidase treatment. C, D. Quantitative HILIC-FLD-UPLC of 2-AB labeled *N*-glycans confirm that α2-3 carrying species are increased and α2-6 are decreased in *ST3GAL4* compared to mock transfected cells. Sialic acid linkage was determined by Sialidase A (releases α2-3,6,8) and Sialidase S (releases α2-3) sialidase treatment experiments. EU: Emission unit. E. *Sambucus nigra* (SNA, binds α2-6 linked sialic acid) lectin blot reveals that especially proteins in the higher molecular mass area (>75kDa) are significantly less decorated with α2,6-linked sialic acid.

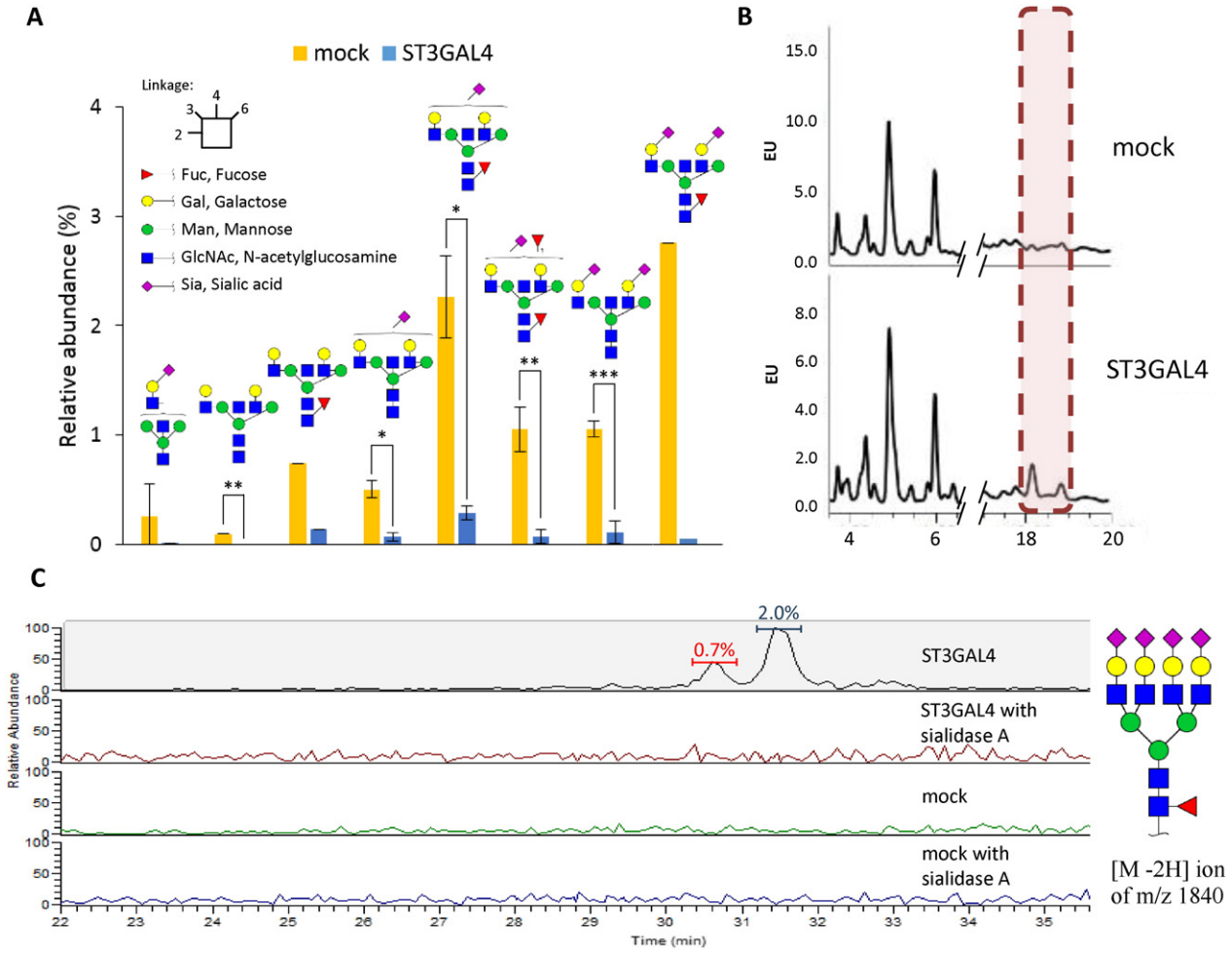


Fig. 4. Bisecting structures are significantly reduced and large branched *N*-glycans are increased in ST3GAL4 overexpressing cells. A. Quantification of relative abundance of bisecting structures shows significant reduction of all examined structures. The glycans were quantified with base-peak intensity of extracted ion chromatograms from LC-ESI-MS performed in three biological replicates. Statistical significance was determined by unpaired student's t-test (p-value *<0.05; **<0.01; ***<0.001). B. The *N*-glycan HILIC-FLD-UPLC data reveals an increase in very large, putative highly branched, α2-3 sialylated *N*-glycans (framed area) in ST3GAL4 overexpressing cells. These structures correspond to approximately 4% of the total *N*-glycome. EU: emission units. C. Core fucosylated, tetrasialylated, tetraantennary structures (depicted on the right side) are very abundant in ST3GAL4 overexpressing cells but are not detectable in the mock control as shown by extracted ion chromatogram of [M–2H] ion of m/z 1840. The two peaks in the upper panel correspond to isomers of the depicted structure. Annotation is based on fragmentation spectra and glycosidase digestion experiments. Relative amount was determined by base-peak intensity and is shown above the corresponding peaks.

N-glycosylation. Of those glycopeptides two showed a decrease in sialylation and the remaining 67 an increase in *ST3GAL4* transfected cells, corresponding to an overall total of 47 glycoproteins (Table 5 and 6 in [25]). Many of the identified glycoproteins have established functions in cellular signaling as either receptors or receptor interacting proteins, in cellular adhesion and migration, or are proteases (Fig. 6). Among the identified glycoproteins, many well characterized targets that are described to be altered in the context of gastric carcinogenesis were found, such as insulin receptor, CEACAM1, CEA, various integrins and RON [12,31–33].

3.5. Activation of altered glycosylated RON receptor tyrosine kinase

Given that the sialoproteomic analysis pointed out a potential alteration of the RON receptor tyrosine kinase glycosylation, we investigated this further, especially since this protein has been shown to be hyperactivated in the context of gastric cancer, contributing in tumorigenesis, malignant progression, angiogenesis, chemoresistance and correlating with bad prognosis [34–38]. Several mechanisms, including overexpression of this receptor and generation of oncogenic variants, have been described that can account for aberrant activation of RON, however, there are cases in which the cause for this abnormal activation remains unknown [38–40]. Comparative analysis between ST3GAL4 overexpressing cells and the mock transfected cells revealed that the expression of RON is unaltered at both RNA (Supplementary table 1) and protein levels (Fig. 7), but shows around 4.5 times increased receptor activation in ST3GAL4 overexpressing cells (Fig. 7a, b), as measured by the increase of phosphorylation of RON. The activated phosphorylated RON protein was observed on the plasma membrane coinciding with the expected localization of fully glycosylated RON (Fig. 7c).

3.6. Altered glycosylation of RON in gastric cancer

In order to confirm the altered glycosylation of RON in the cell line model we used *in situ* proximity ligation assay (PLA). We demonstrated that RON is aberrantly glycosylated carrying the SLe^X epitope in *ST3GAL4* transfected cells (Fig. 8). The localization of RON receptor decorated with SLe^X was observed in the cellular membrane. To assess whether the modification of RON occurs also in gastric carcinoma tissues we screened 15 human gastric tumor samples for the evaluation of the expression of RON and SLe^X. Of these 15 cases, 8 that showed

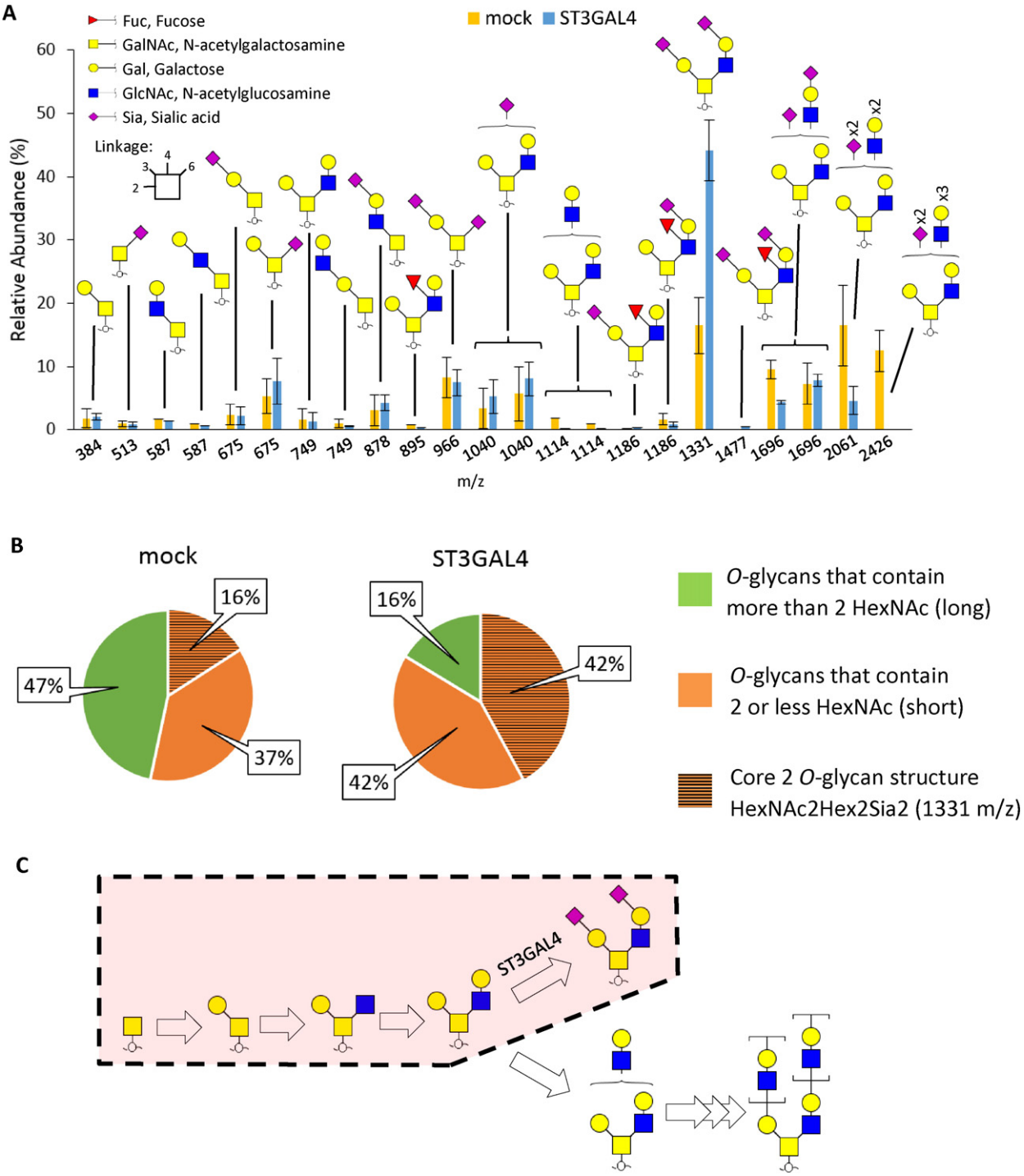


Fig. 5. Increase in disialylated core 2 structures in ST3GAL4 overexpressing cells leads to earlier termination and thus truncation of *O*-glycan structures. A. Relative abundance of all structurally characterized *O*-glycan compounds. Diagram is ordered from left to right by mass. Simple *O*-glycan residues show no significant alteration. Terminated unextended disialylated core 2 structures [M –H] ion of m/z 1331 are significantly increased and as a consequence extended structures are reduced in ST3GAL4 overexpressing cells. The quantifications are established by base-peak intensity of extracted ion chromatograms from LC-ESI-MS performed in three biological replicates and presented as average ± SD. B. Pie-diagram illustrates the approximately equal distribution of short and long *O*-glycans in mock transfected MKN45 cells, and the truncation that is observed in ST3GAL4 overexpressing cells. *O*-glycans with less than or equal 2 HexNAc or Hex were considered short *O*-glycans, structures with 3 or more HexNAcs or Hex were considered long *O*-glycans. The abundance of the simple disialylated core 2 structure (NeuAc2Hex2HexNAc2) is highlighted in dashed orange C. Schematic representation of the impact of ST3GAL4 overexpression on the biosynthesis of *O*-glycan in MKN45. The frame indicates the preferred pathway in the *ST3GAL4* transfected cell line.

expression of both antigens were further analyzed by PLA (Supplementary table 2). All evaluated cases showed PLA positivity signal with 2 cases displaying a high number of signal corresponding to co-expression in proximity of RON and SLe^X. The tumor adjacent mucosa did not show any PLA signal and represents an internal control (Fig. 8).

4. Discussion

The altered glycosylation observed in gastric cancer with overexpression of α2-3 sialylated glycans, including the SLe^X epitope, has long been associated with aggressive features of the disease and poor prognosis for the patients [5,7,41–43]. However, the biological role of

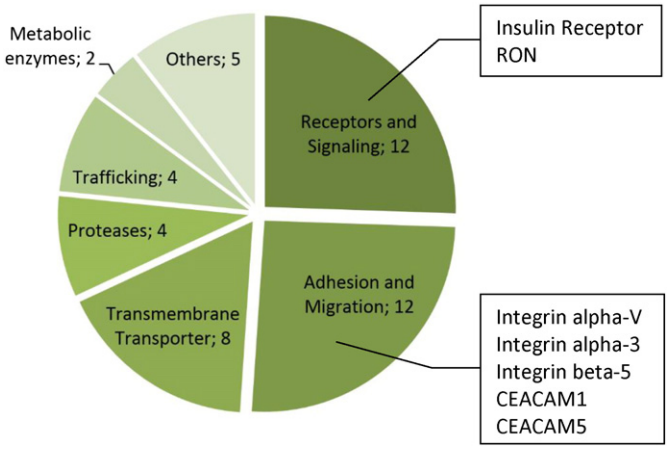


Fig. 6. Representation of the biological functions of the proteins identified with increased sialylation on their N-glycans in *ST3GAL4* transfected cells. Functional clustering reveals that most of the proteins found with increased sialylated N-glycans are involved in signaling, adhesion and migration processes. Remarkably many of these targets have been described to be often altered in gastric cancer. The 47 proteins with significantly increased sialylation were classified according to their main function described in Uniprot (www.uniprot.org).

these sialylated glycans remains to be fully understood. In this study, the detailed molecular characterization of gastric cancer cells displaying altered sialylated terminal glycan structures revealed how these glycosylation modifications can lead to functional changes that confer advantage to cancer cells.

4.1. Glycosylation alterations induced by ST3GAL4 overexpression

The sialyltransferase ST3GAL4, encoded by the *ST3GAL4* gene, catalyzes the addition of α2-3-sialic acid to type 2 extension chains (Galβ1,4GlcNAcβ1-R) generating the precursor of SLe^X [13,21]. In our study we demonstrated that the overexpression of ST3GAL4, in accordance to its substrate specificity, leads to a broad range of glycomic alterations on both *N*- and *O*-glycans. Importantly, ST3GAL4 overexpression induces the expression of SLe^X which has been described to be commonly upregulated in gastric cancer [41,44,45]. Further, we demonstrated for the first time in a gastric cancer context that the upregulation of ST3GAL4, an α2-3 sialyltransferase, not only increases the amount of α2-3 linked sialic acid, but as a consequence, reduces the amount of α2-6-linked sialic acid expressed by the cancer cell, as it has been previously reported in a pancreatic cancer cell line model [46]. The expression levels of other sialyltransferases remained unaltered (Supplementary table 1) supporting the assumption that this linkage-shift is rooted in competition for the same asialylated substrate. Although the general amount of sialylated glycans is described to increase in the context of transformation and malignancy [47,48], it is known that the linkage type is of particular importance. Interestingly, the upregulation of α2-6 sialylation has been shown to increase the adhesion of colon and breast cancer cells to extra cellular matrix components and to reduce the invasive capacity of colon and glioma cancer cell lines [49–51]. Similarly, the increased invasive capacity of ST3GAL4 overexpressing cells previously described [24] could also result from the reduction of α2-6 linked sialic acid modified structures.

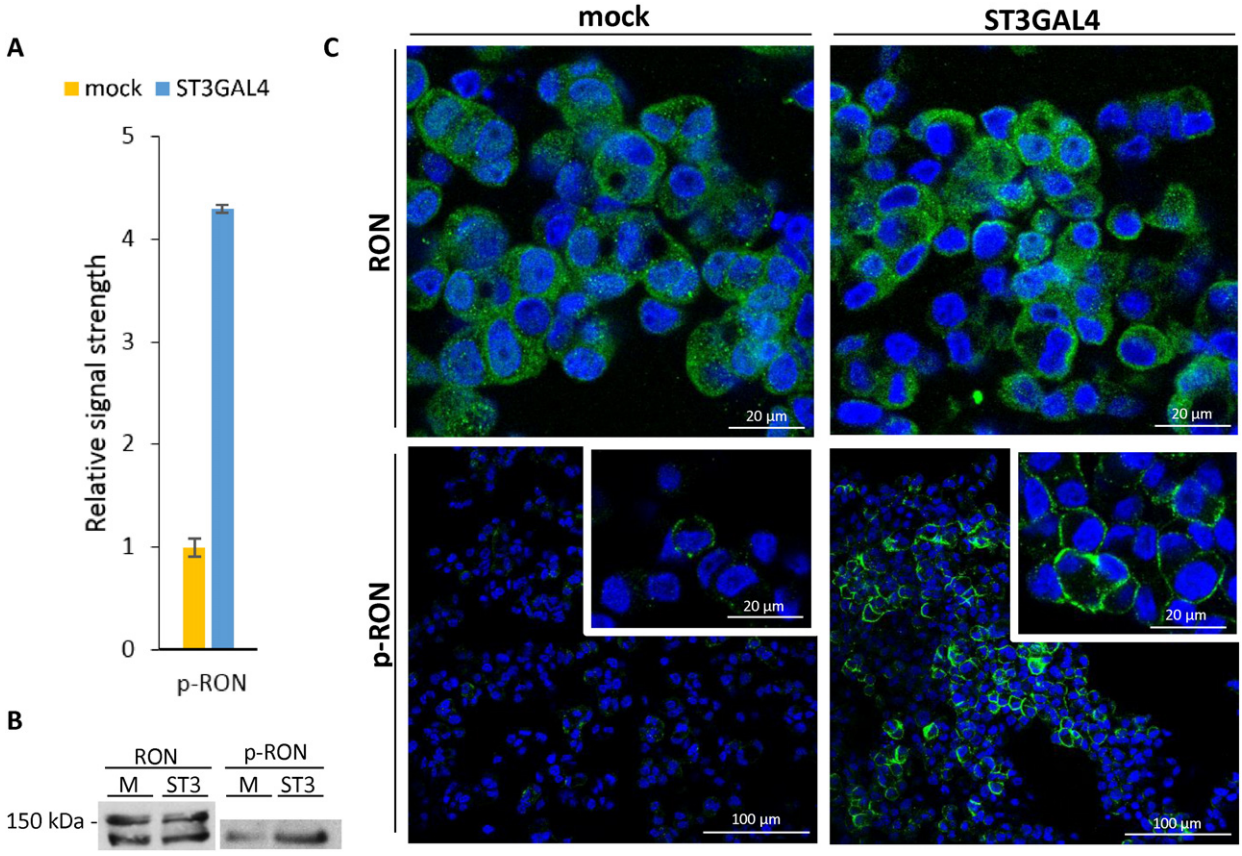


Fig. 7. Receptor tyrosine kinase RON is activated in ST3GAL4 overexpressing cells. A. RON shows approximately 4.5 fold increased activation in ST3GAL4 overexpressing cells as indicated by phospho-RTK array results. The optical density of p-RON signal was measured in duplicates. Average values, with standard deviation are shown. The receptor activation of *ST3GAL4* transfected cells is shown relative to the mock control cells. B. Western blot for RON and phosphorylated RON (p-RON) confirms the increased phosphorylation of RON in ST3GAL4 overexpressing cells and shows no difference in total RON receptor amount. C. Immunofluorescence images of cells stained for RON show cytoplasmic and membrane localization of the receptor with similar amounts for both, *ST3GAL4* transfected and control cell lines. In contrast, activated RON is localized at the plasma membrane and strongly increased in cells overexpressing ST3GAL4.

Our results also showed an almost complete loss of all bisected *N*-glycan structures in cancer cells overexpressing ST3GAL4. A correlation between sialylation and the activity of GnT-III, the GlcNAc-transferase that leads to the synthesis of bisected *N*-glycans, has to our knowledge never been reported. Sialyltransferases are normally located in the trans-Golgi and the sialylation resembles a capping event for mature complex and hybrid *N*-glycans [52], on the other hand the GnT-III is localized in the medial-Golgi and acts early on the biosynthesis of complex *N*-glycans [53]. It is important to consider that the subcellular localization of glycosyltransferases and the organization of the secretory compartments are often altered in the context of cancer [54] and this may lead to deviations of the conventional sequential biosynthesis pathway. Moreover, it is known that bisected structures are often reduced in malignancy and this reduction is correlated with tumor progression and poorer prognosis [27,55,56]. Our results showed that *N*-glycans from *ST3GAL4* transfected cells display a higher degree of branching resulting in enlargement and increased complexity of structures. In addition, no differences were observed on GnT-V, which is the GlcNAc-transferase that catalyzes the formation of β1-6 branched *N*-glycans, and GnT-III expression levels (Supplementary table 1). These enzymes are known to act mutually exclusive on their common glycan acceptor substrate [28,53,56], and therefore the increased branched *N*-glycans could stem from the reduced amount of bisected structures.

Regarding *O*-glycosylation, we demonstrated that ST3GAL*4* upregulation can also lead to the increased truncation of otherwise elongated *O*-glycans. The accumulation of short truncated *O*-glycans like simple carbohydrate antigens are a common feature of gastrointestinal cancer [57]. Our results show that the biosynthesis of early sialylated core-2 structures preclude further elongation of the *O*-glycan chain.

Our data indicates, for the first time, that several cancer relevant glycomic alterations, such as truncation of *O*-glycans, reduction of bisected *N*-glycans, increase of α2-3 sialylation and branching, might not be independent events but may arise coordinated in gastric cancer. The observed changes were further confirmed in an additional glycomic analysis performed in parallel (Tables 2 and 4 in [25]) [58,59]. These glycophenotypic alterations may result from genetic and epigenetic alterations, as well as from tumor microenvironment modifications, occurring in the cancer context [60–63].

4.2. Target proteins displaying aberrant sialylation and its biological implications

In addition, our results revealed that even though the total amount of sialylated *N*-glycans is not significantly increased, 47 glycoproteins showed significantly increased sialylation in a site specific manner in ST3GAL4 overexpressing cancer cells. Considering that only proteins that are translated into the ER are *N*-glycosylated and that only those proteins that are conveyed through the Golgi-network undergo addition of sialic acids [64], it is not surprising that especially transmembrane proteins were found to be affected by the increased sialylation. Furthermore, an interesting observation is that mainly proteins involved

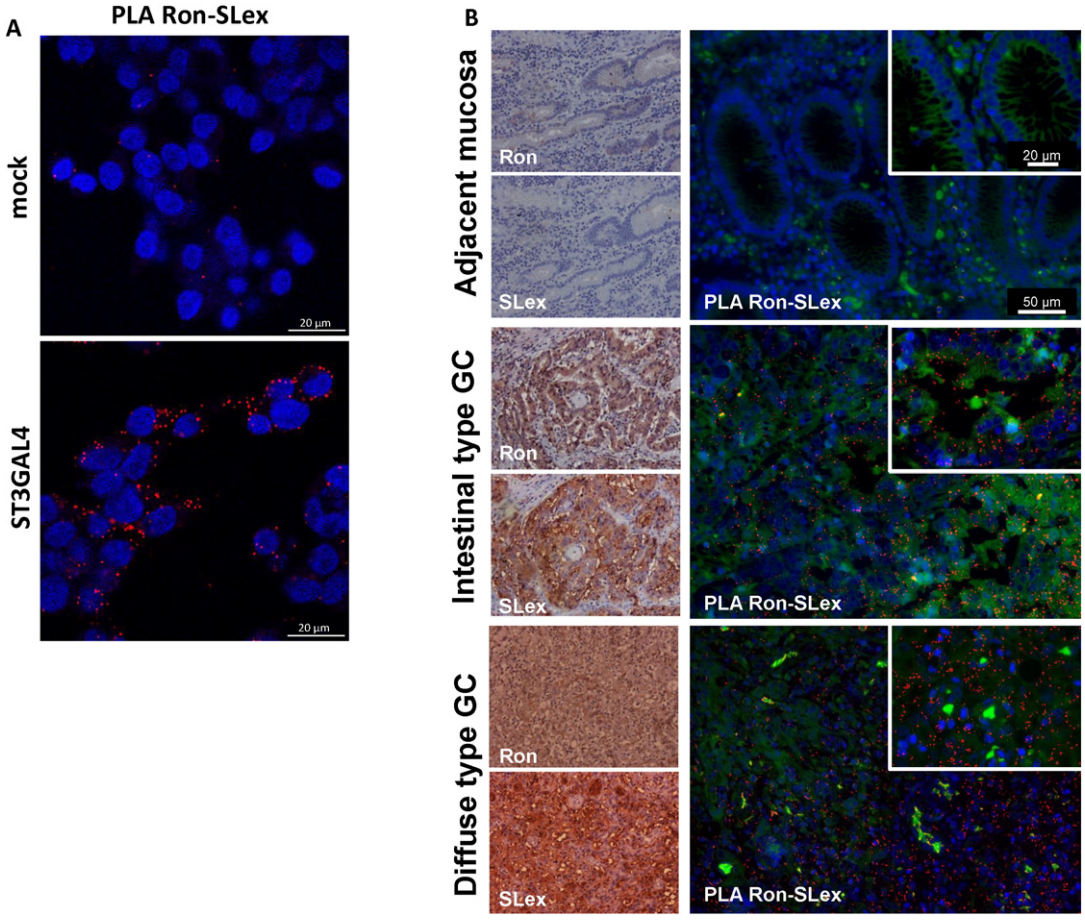


Fig. 8. Receptor tyrosine kinase RON is aberrantly glycosylated with SLe^X in gastric cancer. A. The mock transfected cell line is largely negative in the PLA for RON and SLe^X. On the other hand ST3GAL4 overexpressing cells show strong PLA signal, confirming altered glycosylation of the RON receptor in these cells. The PLA signal is shown in red and nuclei in blue. B. Screening of gastric tumor samples with IHC for RON or SLe^X (left panel) and PLA for RON and SLe^X (right panel). Adjacent mucosa is negative for SLe^X and expresses low level of RON as shown by IHC. The PLA experiment shows no signal in the adjacent mucosa and resembles an internal negative control. The selected cases of intestinal type GC and diffuse type GC are positive for SLe^X and RON. Evaluation of colocalization by PLA reveals that RON is a carrier of SLe^X in gastric carcinomas. The PLA signal is shown in red and nuclei in blue. Tissue autofluoresence has been used to visualize the tissue structure (green).

in cellular adherence and signaling were affected by this altered glycosylation. Important targets included key players in malignancy, such as integrins, which are known to mediate the interaction of cancer cells to the extracellular matrix [33]. Integrin glycosylation has been shown to modulate the cell adhesion to fibronectin and to affect cancer cell invasion and metastasis [65].

In this study we only considered for validation unique peptide sequences that contained *N*-glycosylation sites. The majority of enriched peptides, due to increased sialylation on *O*-glycans or due to common amino acid sequences, were excluded from the analysis. With our approach we could not distinguish between sialic acid linkage and therefore the generation of information of glycopeptides with specific increases of α2-3 sialic acids was precluded. Nonetheless, our approach allowed to identify numerous targets showing altered glycosylation that can be used for follow-up studies. Further, it illustrates that alterations of a single glycosyltransferase can affect a multitude of cancer relevant targets.

4.3. Receptor glycosylation as a modulator of activity in cancer

Receptor tyrosine kinases are glycoproteins and key players of transformation and malignant growth [66]. Glycosylation alterations have been demonstrated to modulate receptor activity [24,67,68]. In the present study we have focused on the alterations leading to the receptor RON activation. In the last years several efforts have been made to target RON in cancer patients. Currently several tyrosine kinase inhibitors and monoclonal antibodies against RON are applied in clinical or pre-clinical trials for therapeutic efficacy [69–76]. Our screening revealed that this receptor, which has been described to be an oncogene in gastric cancer [12], showed an altered glycosylation and concomitant increased activation. RON presents 3 immunoglobulin-like plexin and transcription (IPT) domains that have been described as a cause for constitutive activation in cancer when altered [39,77]. We demonstrated that Asn841 shows increased sialylation in cells overexpressing ST3GAL4 (Table 5 in [25]). The Asn841 is located in one of the IPT domains and therefore may underlie the increased activation of this cellular RTK. Further analysis by proximity ligation assay confirmed the glycosylation of RON with SLe^X in human gastric carcinoma tissues. These observations support that the altered glycosylation observed in this receptor in gastric cancer may lead to its dysfunctional oncogenic activation and therefore modulating the more aggressive cancer phenotype associated with SLe^X expressing tumors. Further studies are warranted to evaluate the biomarker potential of RON glycosylation profile for gastric cancer staging, prognosis and therapeutic response.

4.4. Potential application in cancer therapy and future outlook

Hyper-activation of RTKs due to their upregulation or mutations are common oncogenic events that modern medicine tries to evaluate in the process of tumor screening to design a personalized treatment via inhibitors [9]. However, the present screening methods mostly rely on identifying mutations or gene amplifications [78,79], and therefore, are not considering post-translational modifications, such as glycosylation, as potential activators in a cancer context. We demonstrated in this work that aberrant glycosylation is an alternative route of RTK activation in cancer, emphasizing the necessity for its evaluation in tumor characterization and patient profiling.

A common drawback of whole glycome analysis is its intrinsic complexity, delaying its application in the clinical routine. Much has to be done to simplify and standardize such glycome analysis and to improve our understanding of glycosylation alterations on specific targets. However, the present work demonstrates the importance of such approaches in addressing complex diseases in combination with large scale genomics, transcriptomics and proteomics. We are convinced that the screening of specific target proteins for glycan alterations can be a future corner stone of clinical tumor characterization complementing the conventional methods and therefore leading to improved molecular characterization of cancers for a better diagnosis, prognosis and therapeutics.

Transparency document

The Transparency document is associated with this article can be found, in the online version.

Acknowledgements

We acknowledge the support from the European Union, Seventh Framework Programme, Gastric Glyco Explorer initial training network: grant number 316929. IPATIMUP integrates the i3S Research Unit, which is partially supported by FCT, the Portuguese Foundation for Science and Technology. This work is funded by FEDER funds through the Operational Programme for Competitiveness Factors-COMPETE (FCOMP-01-0124-FEDER028188) and National Funds through the FCT-Foundation for Science and Technology, under the projects: PEst-C/SAU/LA0003/2013, PTDC/BBB-EBI/0786/2012, and PTDC/BBB-EBI/0567/2014 (to CAR). This work was also supported by "Glycoproteomics" project grant number PCIG09-GA-2011-293847 (to DK) and the Danish Natural Science Research Council and a generous grant from the VILLUM Foundation to the VILLUM Center for Bioanalytical Sciences at the University of Southern Denmark (to MRL). Grants were received from FCT, POPH (Programa Operacional Potencial Humano) and FSE (Fundo Social Europeu): SFRH/BPD/75871/2011 to AM; SFRH/BPD/111048/2015 to JAF; SFRH/BPD/96510/2013 to CG. The UPLC instrument was obtained with a grant from the Ingabritt and Arne Lundbergs Research Foundation (to NK). C.J. was supported by the Knut and Alice Wallenberg Foundation. The mass spectrometer (LTQ) was obtained by a grant from the Swedish Research Council (342-2004-4434) (to NK).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbagen.2015.12.016.

References

[1] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2015, CA Cancer J. Clin. 65 (2015) 5–29.
[2] M. May, Statistics: attacking an epidemic, Nature 509 (2014) S50–S51.
[3] J. Ferlay, E. Steliarova-Foucher, J. Lortet-Tieulent, S. Rosso, J.W. Coebergh, H. Comber, D. Forman, F. Bray, Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012, Eur. J. Cancer 49 (2013) 1374–1403.
[4] M. Carpelan-Holmstrom, J. Louhimo, U.H. Stenman, H. Alfthan, C. Haglund, CEA, CA 19-9 and CA 72-4 improve the diagnostic accuracy in gastrointestinal cancers, Anticancer Res. 22 (2002) 2311–2316.
[5] C.A. Reis, H. Osorio, L. Silva, C. Gomes, L. David, Alterations in glycosylation as biomarkers for cancer detection, J. Clin. Pathol. 63 (2010) 322–329.
[6] A. Varki, R. Kannagi, B.P. Toole, Glycosylation Changes in Cancer, in: A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, M.E. Etzler (Eds.), Essentials of Glycobiology, Cold Spring Harbor, NY, 2009.
[7] S.S. Pinho, C.A. Reis, Glycosylation in cancer: mechanisms and clinical implications, Nat. Rev. Cancer 15 (2015) 540–555.
[8] K.M. Wong, T.J. Hudson, J.D. McPherson, Unraveling the genetics of cancer: genome sequencing and beyond, Annu. Rev. Genomics Hum. Genet. 12 (2011) 407–430.
[9] R.L. Schilsky, Personalized medicine in oncology: the future is now, Nat. Rev. Drug Discov. 9 (2010) 363–366.
[10] C. Sawyers, Targeted cancer therapy, Nature 432 (2004) 294–297.
[11] A. Arora, E.M. Scholar, Role of tyrosine kinase inhibitors in cancer therapy, J. Pharmacol. Exp. Ther. 315 (2005) 971–979.
[12] H.P. Yao, Y.Q. Zhou, R. Zhang, M.H. Wang, MSP-RON signalling in cancer: pathogenesis and therapeutic potential, Nat. Rev. Cancer 13 (2013) 466–481.
[13] A.S. Carvalho, A. Harduin-Lepers, A. Magalhaes, E. Machado, N. Mendes, L.T. Costa, R. Matthiesen, R. Almeida, J. Costa, C.A. Reis, Differential expression of alpha-2,3-sialyltransferases and alpha-1,3/4-fucosyltransferases regulates the levels of sialyl Lewis a and sialyl Lewis x in gastrointestinal carcinoma cells, Int. J. Biochem. Cell Biol. 42 (2010) 80–89.
[14] T. Conze, A.S. Carvalho, U. Landegren, R. Almeida, C.A. Reis, L. David, O. Soderberg, MUC2 mucin is a major carrier of the cancer-associated sialyl-Tn antigen in intestinal metaplasia and gastric carcinomas, Glycobiology 20 (2010) 199–206.

[15] P. Lauren, The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. An attempt at a histo-clinical classification, Acta Pathol. Microbiol. Scand. 64 (1965) 31–49.
[16] B.L. Schulz, N.H. Packer, N.G. Karlsson, Small-scale analysis of O-linked oligosaccharides from glycoproteins and mucins separated by gel electrophoresis, Anal. Chem. 74 (2002) 6088–6097.
[17] L. Royle, C.M. Radcliffe, R.A. Dwek, P.M. Rudd, Detailed structural analysis of N-glycans released from glycoproteins in SDS-PAGE gel bands using HPLC combined with exoglycosidase array digestions, Methods Mol. Biol. 347 (2006) 125–143.
[18] K. Engholm-Keller, P. Birck, J. Storling, F. Pociot, T. Mandrup-Poulsen, M.R. Larsen, TiSH—a robust and sensitive global phosphoproteomics strategy employing a combination of TiO_2, SIMAC, and HILIC, J. Proteomics 75 (2012) 5749–5761.
[19] M.N. Melo-Braga, M. Ibanez-Vea, M.R. Larsen, K. Kulej, Comprehensive protocol to simultaneously study protein phosphorylation, acetylation, and N-linked sialylated glycosylation, Methods Mol. Biol. 1295 (2015) 275–292.
[20] V. Schwammle, I.R. Leon, O.N. Jensen, Assessment and improvement of statistical tools for comparative proteomics analysis of sparse data sets with few experimental replicates, J. Proteome Res. 12 (2013) 3874–3883.
[21] A. Harduin-Lepers, M.A. Krzewinski-Recchi, F. Colomb, F. Foulquier, S. Groux-Degroote, P. Delannoy, Sialyltransferases functions in cancers, Front. Biosci. 4 (2012) 499–515.
[22] K. Sasaki, E. Watanabe, K. Kawashima, S. Sekine, T. Dohi, M. Oshima, N. Hanai, T. Nishi, M. Hasegawa, Expression cloning of a novel Gal beta (1-3/1-4) GlcNAc alpha 2,3-sialyltransferase using lectin resistance selection, J. Biol. Chem. 268 (1993) 22782–22787.
[23] H. Kitagawa, J.C. Paulson, Cloning of a novel alpha 2,3-sialyltransferase that sialylates glycoprotein and glycolipid carbohydrate groups, J. Biol. Chem. 269 (1994) 1394–1401.
[24] C. Gomes, H. Osorio, M.T. Pinto, D. Campos, M.J. Oliveira, C.A. Reis, Expression of ST3GAL4 leads to SLe(x) expression and induces c-Met activation and an invasive phenotype in gastric carcinoma cells, PLoS One 8 (2013), e66737.
[25] S. Mereiter, A. Magalhães, B. Adamczyk, C. Jin, A. Almeida, L. Drici, M. Ibanez-Vea, M.R. Larsen, D. Kolarich, N.G. Karlsson, C.A. Reis, Glycomic and sialoproteomic data of gastric carcinoma cells overexpressing ST3GAL4, Data in Brief (2015) (submitted for publication).
[26] A. Ishizuka, Y. Hashimto, R. Naka, M. Kinoshita, K. Kakehi, J. Seino, Y. Funakoshi, T. Suzuki, A. Kameyama, H. Narimatsu, Accumulation of free complex-type N-glycans in MKN7 and MKN45 stomach cancer cells, Biochem. J. 413 (2008) 227–237.
[27] S.S. Pinho, C.A. Reis, J. Paredes, A.M. Magalhaes, A.C. Ferreira, J. Figueiredo, W. Xiaogang, F. Carneiro, F. Gartner, R. Seruca, The role of N-acetylglucosaminyltransferase III and V in the post-transcriptional modifications of E-cadherin, Hum. Mol. Genet. 18 (2009) 2599–2608.
[28] Y. Zhao, T. Nakagawa, S. Itoh, K. Inamori, T. Isaji, Y. Kariya, A. Kondo, E. Miyoshi, K. Miyazaki, N. Kawasaki, N. Taniguchi, J. Gu, N-acetylglucosaminyltransferase III antagonizes the effect of N-acetylglucosaminyltransferase V on alpha3beta1 integrin-mediated cell migration, J. Biol. Chem. 281 (2006) 32122–32130.
[29] M.R. Kudelka, T. Ju, J. Heimburg-Molinaro, R.D. Cummings, Simple sugars to complex disease—mucin-type O-glycans in cancer, Adv. Cancer Res. 126 (2015) 53–135.
[30] Y. Rossez, E. Maes, T. Lefebvre Darroman, P. Gosset, C. Ecobichon, M. Joncquel Chevalier Curt, I.G. Boneca, J.C. Michalski, C. Robbe-Masselot, Almost all human gastric mucin O-glycans harbor blood group A, B or H antigens and are potential binding sites for Helicobacter pylori, Glycobiology 22 (2012) 1193–1206.
[31] C. Boccaccio, P.M. Comoglio, Invasive growth: a MET-driven genetic programme for cancer and stem cells, Nat. Rev. Cancer 6 (2006) 637–645.
[32] N. Beauchemin, A. Arabzadeh, Carcinoembryonic antigen-related cell adhesion molecules (CEACAMs) in cancer progression and metastasis, Cancer Metastasis Rev. 32 (2013) 643–671.
[33] J.S. Desgrosellier, D.A. Cheresh, Integrins in cancer: biological implications and therapeutic opportunities, Nat. Rev. Cancer 10 (2010) 9–22.
[34] M.H. Wang, W. Lee, Y.L. Luo, M.T. Weis, H.P. Yao, Altered expression of the RON receptor tyrosine kinase in various epithelial cancers and its contribution to tumourigenic phenotypes in thyroid cancer cells, J. Pathol. 213 (2007) 402–411.
[35] M.N. Thobe, D. Gurusamy, P. Pathrose, S.E. Waltz, The Ron receptor tyrosine kinase positively regulates angiogenic chemokine production in prostate cancer cells, Oncogene 29 (2010) 214–226.
[36] J. Logan-Collins, R.M. Thomas, P. Yu, D. Jaquish, E. Mose, R. French, W. Stuart, R. McClaine, B. Aronow, R.M. Hoffman, S.E. Waltz, A.M. Lowy, Silencing of RON receptor signaling promotes apoptosis and gemcitabine sensitivity in pancreatic cancers, Cancer Res. 70 (2010) 1130–1140.
[37] R.J. McClaine, A.M. Marshall, P.K. Wagh, S.E. Waltz, Ron receptor tyrosine kinase activation confers resistance to tamoxifen in breast cancer cell lines, Neoplasia 12 (2010) 650–658.
[38] D.V. Catenacci, G. Cervantes, S. Yala, E.A. Nelson, E. El-Hashani, R. Kanteti, M. El Dinali, R. Hasina, J. Bragelmann, T. Seiwert, M. Sanicola, L. Henderson, T.A. Grushko, O. Olopade, T. Karrison, Y.J. Bang, W.H. Kim, M. Tretiakova, E. Vokes, D.A. Frank, H.L. Kindler, H. Huet, R. Salgia, RON (MST1R) is a novel prognostic marker and therapeutic target for gastro-esophageal adenocarcinoma, Cancer Biol. Ther. 12 (2011) 9–46.
[39] C. Collesi, M.M. Santoro, G. Gaudino, P.M. Comoglio, A splicing variant of the RON transcript induces constitutive tyrosine kinase activity and an invasive phenotype, Mol. Cell. Biol. 16 (1996) 5518–5526.
[40] M.M. Santoro, C. Collesi, S. Grisendi, G. Gaudino, P.M. Comoglio, Constitutive activation of the RON gene promotes invasive growth but not transformation, Mol. Cell. Biol. 16 (1996) 7072–7083.
[41] M. Amado, F. Carneiro, M. Seixas, H. Clausen, M. Sobrinho-Simoes, Dimeric sialyl-Le(x) expression in gastric carcinoma correlates with venous invasion and poor outcome, Gastroenterology 114 (1998) 462–470.
[42] R. Kannagi, Carbohydrate-mediated cell adhesion involved in hematogenous metastasis of cancer, Glycoconj. J. 14 (1997) 577–584.
[43] S. Nakamori, M. Kameyama, S. Imaoka, H. Furukawa, O. Ishikawa, Y. Sasaki, Y. Izumi, T. Irimura, Involvement of carbohydrate antigen sialyl Lewis(x) in colorectal cancer metastasis, Dis. Colon Rectum 40 (1997) 420–431.
[44] M. Tatsumi, A. Watanabe, H. Sawada, Y. Yamada, Y. Shino, H. Nakano, Immunohistochemical expression of the sialyl Lewis x antigen on gastric cancer cells correlates with the presence of liver metastasis, Clin. Exp. Metastasis 16 (1998) 743–750.
[45] S.E. Baldus, T.K. Zirbes, S.P. Monig, S. Engel, E. Monaca, K. Rafiqpoor, F.G. Hanisch, C. Hanski, J. Thiele, H. Pichlmaier, H.P. Dienes, Histopathological subtypes and prognosis of gastric cancer are correlated with the expression of mucin-associated sialylated antigens: sialosyl-Lewis(a), sialosyl-Lewis(x) and sialosyl-Tn, Tumour Biol. 19 (1998) 445–453.
[46] M. Perez-Garay, B. Arteta, E. Llop, L. Cobler, L. Pages, R. Ortiz, M.J. Ferri, C. de Bolos, J. Figueras, R. de Llorens, F. Vidal-Vanaclocha, R. Peracaula, alpha2,3-Sialyltransferase ST3Gal IV promotes migration and metastasis in pancreatic adenocarcinoma cells and tends to be highly expressed in pancreatic adenocarcinoma tissues, Int. J. Biochem. Cell Biol. 45 (2013) (1748-1757).
[47] Y.J. Kim, A. Varki, Perspectives on the significance of altered glycosylation of glycoproteins in cancer, Glycoconj. J. 14 (1997) 569–576.
[48] F. Dall'Olio, M. Chiricolo, Sialyltransferases in cancer, Glycoconj. J. 18 (2001) 841–850.
[49] S. Lin, W. Kemmner, S. Grigull, P.M. Schlag, Cell surface alpha 2,6 sialylation affects adhesion of breast carcinoma cells, Exp. Cell Res. 276 (2002) 101–110.
[50] H. Yamamoto, A. Oviedo, C. Sweeley, T. Saito, J.R. Moskal, Alpha2,6-sialylation of cell-surface N-glycans inhibits glioma formation in vivo, Cancer Res. 61 (2001) 6822–6829.
[51] M. Chiricolo, N. Malagolini, S. Bonfiglioli, F. Dall'Olio, Phenotypic changes induced by expression of beta-galactoside alpha2,6 sialyltransferase I in the human colon cancer cell line SW948, Glycobiology 16 (2006) 146–154.
[52] R. Kornfeld, S. Kornfeld, Assembly of asparagine-linked oligosaccharides, Annu. Rev. Biochem. 54 (1985) 631–664.
[53] N. Taniguchi, Y. Ihara, Recent progress in the molecular biology of the cloned N-acetylglucosaminyltransferases, Glycoconj. J. 12 (1995) 733–738.
[54] S. Kellokumpu, R. Sormunen, I. Kellokumpu, Abnormal glycosylation and altered Golgi structure in colorectal cancer: dependence on intra-Golgi pH, FEBS Lett. 516 (2002) 217–224.
[55] Y. Song, J.A. Aglipay, J.D. Bernstein, S. Goswami, P. Stanley, The bisecting GlcNAc on N-glycans inhibits growth factor signaling and retards mammary tumor progression, Cancer Res. 70 (2010) 3361–3371.
[56] S.S. Pinho, P. Oliveira, J. Cabral, S. Carvalho, D. Huntsman, F. Gartner, R. Seruca, C.A. Reis, C. Oliveira, Loss and recovery of Mgat3 and GnT-III Mediated E-cadherin N-glycosylation is a mechanism involved in epithelial-mesenchymal–epithelial transitions, PLoS One 7 (2012), e33191.
[57] S. Hakomori, Aberrant glycosylation in tumors and tumor-associated carbohydrate antigens, Adv. Cancer Res. 52 (1989) 257–331.
[58] P.H. Jensen, N.G. Karlsson, D. Kolarich, N.H. Packer, Structural analysis of N- and O-glycans released from glycoproteins, Nat. Protoc. 7 (2012) 1299–1310.
[59] D. Kolarich, M. Windwarder, K. Alagesan, F. Altmann, Isomer-specific analysis of released N-glycans by LC-ESI MS/MS with porous graphitized carbon, Methods Mol. Biol. 1321 (2015) 427–435.
[60] C. Slawson, G.W. Hart, O-GlcNAc signalling: implications for cancer cell biology, Nat. Rev. Cancer 11 (2011) 678–684.
[61] F. Alisson-Silva, L. Freire-de-Lima, J.L. Donadio, M.C. Lucena, L. Penha, J.N. Sa-Diniz, W.B. Dias, A.R. Todeschini, Increase of O-glycosylated oncofetal fibronectin in high glucose-induced epithelial–mesenchymal transition of cultured human epithelial cells, PLoS One 8 (2013), e60471.
[62] S. Olivier-Van Stichelen, C. Guinez, A.M. Mir, Y. Perez-Cervera, C. Liu, J.C. Michalski, T. Lefebvre, The hexosamine biosynthetic pathway and O-GlcNAcylation drive the expression of beta-catenin and cell proliferation, Am. J. Physiol. Endocrinol. Metab. 302 (2012) E417–E424.
[63] Z. Ma, D.J. Vocadlo, K. Vosseller, Hyper-O-GlcNAcylation is anti-apoptotic and maintains constitutive NF-kappaB activity in pancreatic cancer cells, J. Biol. Chem. 288 (2013) 15121–15130.
[64] P. Stanley, Golgi Glycosylation, Cold Spring Harbor Perspectives In Biology, 3, 2011.
[65] J. Gu, N. Taniguchi, Regulation of integrin functions by N-glycans, Glycoconj. J. 21 (2004) 9–15.
[66] D. Hanahan, R.A. Weinberg, Hallmarks of cancer: the next generation, Cell 144 (2011) 646–674.
[67] K. Kaszuba, M. Grzybek, A. Orlowski, R. Danne, T. Rog, K. Simons, U. Coskun, I. Vattulainen, N-Glycosylation as determinant of epidermal growth factor receptor conformation in membranes, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) 4334–4339.
[68] A. Cazet, S. Julien, M. Bobowski, J. Burchell, P. Delannoy, Tumour-associated carbohydrate antigens in breast cancer, Breast Cancer Res. 12 (2010) 204.
[69] T.K. Choueiri, U. Vaishampayan, J.E. Rosenberg, T.F. Logan, A.L. Harzstark, R.M. Bukowski, B.I. Rini, S. Srinivas, M.N. Stein, L.M. Adams, L.H. Ottesen, K.H. Laubscher, L. Sherman, D.F. McDermott, N.B. Haas, K.T. Flaherty, R. Ross, P. Eisenberg, P.S. Meltzer, M.J. Merino, D.P. Bottaro, W.M. Linehan, R. Srinivasan, Phase II and biomarker study of the dual MET/VEGFR2 inhibitor foretinib in patients with papillary renal cell carcinoma, J. Clin. Oncol. 31 (2013) 181–186.
[70] A. Belalcazar, D. Azana, C.A. Perez, L.E. Raez, E.S. Santos, Targeting the Met pathway in lung cancer, Expert. Rev. Anticancer. Ther. 12 (2012) 519–528.
[71] Y. Zhang, P.J. Kaplan-Lefko, K. Rex, Y. Yang, J. Moriguchi, T. Osgood, B. Mattson, A. Coxon, M. Reese, T.S. Kim, J. Lin, A. Chen, T.L. Burgess, I. Dussault, Identification of a novel recepteur d'origine nantais/c-met small-molecule kinase inhibitor with antitumor activity in vivo, Cancer Res. 68 (2008) 6680–6687.
[72] S. Sharma, J.Y. Zeng, C.M. Zhuang, Y.Q. Zhou, H.P. Yao, X. Hu, R. Zhang, M.H. Wang, Small-molecule inhibitor BMS-777607 induces breast cancer cell polyploidy with increased resistance to cytotoxic chemotherapy agents, Mol. Cancer Ther. 12 (2013) 725–736.

[73] B.S. Pan, G.K. Chan, M. Chenard, A. Chi, L.J. Davis, S.V. Deshmukh, J.B. Gibbs, S. Gil, G. Hang, H. Hatch, J.P. Jewell, I. Kariv, J.D. Katz, K. Kunii, W. Lu, B.A. Lutterbach, C.P. Paweletz, X. Qu, J.F. Reilly, A.A. Szewczak, Q. Zeng, N.E. Kohl, C.J. Dinsmore, MK-2461, a novel multitargeted kinase inhibitor, preferentially inhibits the activated c-Met receptor, Cancer Res. 70 (2010) 1524–1533.
[74] A.B. Northrup, M.H. Katcher, M.D. Altman, M. Chenard, M.H. Daniels, S.V. Deshmukh, D. Falcone, D.J. Guerin, H. Hatch, C. Li, W. Lu, B. Lutterbach, T.J. Allison, S.B. Patel, J.F. Reilly, M. Reutershan, K.W. Rickert, C. Rosenstein, S.M. Soisson, A.A. Szewczak, D. Walker, K. Wilson, J.R. Young, B.S. Pan, C.J. Dinsmore, Discovery of 1-[3-(1-methyl-1H-pyrazol-4-yl)-5-oxo-5H-benzo[4,5]cyclohepta[1,2-b]pyridin-7-yl]-N-(pyridin-2-ylmethyl)methanesulfonamide (MK-8033): a specific c-Met/Ron dual kinase inhibitor with preferential affinity for the activated state of c-Met, J. Med. Chem. 56 (2013) 2294–2310.
[75] J.M. O'Toole, K.E. Rabenau, K. Burns, D. Lu, V. Mangalampalli, P. Balderes, N. Covino, R. Bassi, M. Prewett, K.J. Gottfredsen, M.N. Thobe, Y. Cheng, Y. Li, D.J. Hicklin, Z. Zhu, S.E. Waltz, M.J. Hayman, D.L. Ludwig, D.S. Pereira, Therapeutic implications of a human neutralizing antibody to the macrophage-stimulating protein receptor tyrosine kinase (RON), a c-MET family member, Cancer Res. 66 (2006) 9162–9170.
[76] Z. Gunes, A. Zucconi, M. Cioce, A. Meola, M. Pezzanera, S. Acali, I. Zampaglione, V. De Pratti, L. Bova, F. Talamo, A. Demartis, P. Monaci, N. La Monica, G. Ciliberto, A. Vitelli, Isolation of fully human antagonistic RON antibodies showing efficient block of downstream signaling and cell migration, Transl. Oncol. 4 (2011) 38–46.
[77] Q. Ma, K. Zhang, S. Guin, Y.Q. Zhou, M.H. Wang, Deletion or insertion in the first immunoglobulin-plexin-transcription (IPT) domain differentially regulates expression and tumorigenic activities of RON receptor tyrosine kinase, Mol. Cancer 9 (2010) 307.
[78] M. Ingelman-Sundberg, Personalized medicine into the next generation, J. Intern. Med. 277 (2015) 152–154.
[79] T. Tursz, R. Bernards, Hurdles on the road to personalized medicine, Mol. Oncol. 9 (2015) 935–939.
[80] N. Shibuya, I.J. Goldstein, W.F. Broekaert, M. Nsimba-Lubaki, B. Peeters, W.J. Peumans, The elderberry (*Sambucus nigra* L.) bark lectin recognizes the Neu5Ac(alpha 2-6)Gal/GalNAc sequence, J. Biol. Chem. 262 (1987) 1596–1601.

CHAPTER IV

THE MOLECULAR AND FUNCTIONAL IMPACT OF *O*–GLYCAN TRUNCATION ON CD44 IN CANCER

THE MOLECULAR AND FUNCTIONAL IMPACT OF *O*-GLYCAN TRUNCATION ON CD44 IN CANCER

Stefan Mereiter[1,2,3], Catarina Gomes[1,2], Cintia Barreira[1,2], Joana Macedo[1,2], Lylia Drici[4], Maria Ibáñez-Vea[4], Martin R. Larsen[4], Ana Magalhães[1,2], Celso A. Reis[1,2,3,5]

1. i3S – Instituto de Investigação e Inovação em Saúde, University of Porto, Portugal
2. Institute of Molecular Pathology and Immunology of the University of Porto – IPATIMUP, Porto, Portugal
3. Institute of Biomedical Sciences of Abel Salazar – ICBAS, University of Porto, Portugal
4. Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark
5. Medical Faculty, University of Porto, Portugal

Abstract

CD44 is a highly heterogeneous glycoprotein and among the most frequently overexpressed proteins in gastric cancer. It is a central player of many cellular mechanisms, such as cell receptor activation and cellular adhesion, and hence a potent driver of malignancy. Even though CD44 is carrier of *N*-glycans, Heparan sulfate, Chondroitin sulfate and is particular extensively *O*-glycosylated, the intrinsic diversity of CD44 has been mainly attributed to alternative splicing. In this work we demonstrate that the truncation of *O*-linked glycosylation, as it can be found in virtually all gastric carcinomas, significantly affects CD44 molecular appearance and function. We show that CD44 was the single major carrier of truncated *O*-glycans in our gastric carcinoma cell models and that the truncation of *O*-glycans by several means alters its molecular mass and recognition by site-specific antibodies. CD44 has previously been described to act through its variant form v6 as co-receptor for the RON receptor tyrosine kinase activation. We present that the truncation of *O*-glycans significantly increases the colocalization between CD44v6 and RON. We therefore conclude that *O*-glycan truncation conveys advantageous cancer phenotypes through its impact on CD44 and its increased interaction with oncogenic receptors.

1. Introduction

CD44 is a highly heterogeneous transmembrane protein, ranging in molecular weight from 50 to over 230 kDa kDa [1]. This diversity has been mainly attributed to alternative splicing at the extracellular stem region and its susceptibility to proteolytic cleavage at the transmembrane adjoining sites, but also to its extensive glycosylation [2-4].

CD44 is a key player of cellular adhesion and signal transduction by multiple means: (i) it is the main receptor of hyaluronan, one of the chief components of the extracellular matrix [5], (ii) it may carry heparan sulfate and chondroitin sulfate chains which are docking sites for various growth factors and proteases [6] and (iii) it is co-receptor to cell surface receptor tyrosine kinases (RTKs), such as MET, RON and HER2 [7, 8]. The execution of these functions can vary significantly among different CD44 isoforms. For instance, the efficiency of CD44 to bind hyaluronan is controlled by the alternatively spliced variable (v) chains and its glycosylation status [9, 10]. Further, the single heparan sulfate chain of CD44 is linked to the variant chain v3 and the peptide sequence that acts as co-receptor of MET and RON is encoded in the v6 variant exon [7, 11]. In gastric cancer, CD44 is commonly upregulated and its variant form CD44v6 frequently *de novo* expressed. Cancer cells that overexpress CD44 and CD44v6 display a significant increase in metastatic potential and are associated with poor survival of the patients [12-14]. Its central role in both regulating cellular adhesion and receptor activation has been shown to contribute to several cancer hallmark abilities.

The alterations of CD44 protein levels indicated by antibody binding assays often do not correlate at RNA level, which frequently leads to inconsistent results [15]. In the present work, we demonstrate that alterations in the glycosylation machinery can

have a severe impact on the molecular characteristics of CD44 leading to differential molecular weight and affecting its recognition by antibodies. We additionally show functional implication of the alterations of CD44 such as increased colocalization with the RON receptor tyrosine kinase.

2. Material and methods

2.1 Cell lines

The gastric carcinoma cell line MKN45 was obtained from the Japanese Cancer Research Bank (Tsukuba, Japan) and was stably transfected with the full length human *ST3GAL4* gene (ST3GAL4), the full length human *ST6GALNAC1* (ST3GalNAc-I) gene or the corresponding empty vector pcDNA3.1 (Mock) as previously described [16, 17]. The truncated *O*-glycan SimpleCells were obtained by targeting the COSMC gene of MKN45 cells by zinc-finger nuclease as previously described [18, 19]. All cells were grown in monolayer in uncoated cell culture flasks. Cells were maintained at 37°C in an atmosphere of 5% CO_2, in RPMI 1640 GlutaMAX, HEPES medium supplemented with 10% fetal bovine serum (FBS), 1% penicillin-streptomycin and in the presence of 0.5 mg/ml G418 (all from Invitrogen, Waltham, MA). Cell culture medium was replaced every two days.

2.2 Western blotting and immunoprecipitation

Western blotting (WB) experiments were performed as previously described in [20]. Briefly, cells were washed twice with PBS and directly collected in lysis buffer 17 (R&D Systems, McKinley Place, MN) supplemented with 1 mM sodium orthovanadate, 1 mM phenylmethanesulfonylfluoride (PMSF) and protease inhibitor cocktail (Roche, Basel, Switzerland). The protein concentrations of lysates were determined by the DC protein assay (BioRad, Hercules, CA) and 500 µg of lysate were used for immunoprecipitation (IP) with protein G fast flow sepharose beads (GE healthcare, Little Chalfont, UK). The antibodies used for WB and IP are shown in table 1.

Table 1. Antibodies used for western blotting and immunoprecipitation

Targeted epitope	Clone name	Manufacturer	WB dil.	IP dil.	IF/PLA dil.
CD44	156/3C11	Cell Signaling	1:1000	1:60	1:400
CD44v6	MA54	Invitrogen	1:1000	/	1:100
RON	C-20	Sant Cruz Biotechnology	/	/	1:60
Syndecan 1	B-A38	Abcam	1:125	/	/
Heparan Sulfate	10E4	USBiological	1:5000	/	/
Chondroitin Sulfate	A-7	Sant Cruz Biotechnology	1:100	/	/
Sialyl Tn (STn)	TKH2	[21]	1:3	/	1:3
Sialyl Lewis X (SLe^x)	CSLEX-1	BD Biosciences	/	/	1:80

2.3 iTraq based proteomics

Proteomic analysis was performed as described in [20]. Briefly, cell pellets were dissolved in ice-cold Na_2CO_3 buffer (0.1 M, pH 11) supplemented with protease inhibitor (Roche complete EDTA free), PhosSTOP phosphatase inhibitor cocktail (Roche) and 10 mM sodium pervanadate. The suspensions were sonicated and the soluble fraction was separated from the membrane fraction. The membrane fraction was resuspended in 6 M urea and 2 M thiourea. Soluble and membrane fractions were both reduced with dithiothreitol and alkylated with iodoacetamide. Samples were digested with endoproteinase Lys-C (Wako, Osaka, Japan) and then with trypsin. The digests were purified using in-house packed staged tips with a mixture of Poros R2 and Oligo R3 reversed phase resins (Applied Biosystem, Foster City, CA, USA). Samples were dried and resuspended in dissolution buffer. A total of 150 µg for each condition was labeled with 4-plex iTRAQ™ (Applied Biosystems, Foster City, CA) as described by the manufacturer.

Peptides were loaded on an in-house packed Reprosil-Pur C18-AQ (2 cm x 100 µm, 5 µm; Dr. Maisch GmbH, Germany) pre-column and separated on an in-house packed Reprosil-Pur C18-AQ (17 cm x 75 µm, 3 µm; Dr. Maisch GmbH, Germany) column using an Easy-nLC II system (Thermo Scientific, Bremen, Germany) and eluted at a flow of 250 nL/min. Mass spectrometric analyses were performed in an Orbitrap Fusion Tribrid system (Thermo Scientific, Bremen, Germany). The raw data were processed and quantified by Proteome Discoverer (version 1.4.1.14, Thermo Scientific) against SwissProt and Uniprot human reference database by using Mascot (v2.3.02, Matrix Science Ltd, London, UK) and Sequest HT, respectively. The iTRAQ datasets were quantified using the centroid peak intensity with the "reporter ions quantifier" node. For the quantitative comparison of CD44 we normalized unique CD44 peptides derived from mock lysates to 1 and determined the relative quantity of the same peptides derived from ST3GAL4 transfected cells.

2.4 Transcriptomics

The transcriptomic analysis was performed as previously described [20]. Briefly, total RNA extracts from cell lysates were isolated with TRI Reagent (Sigma-Aldrich, St. Louis, MO). The mRNAs of over 20,000 primed targets were sequenced by using Ion AmpliSeq Transcriptome Human Gene Expression Kit (Life Technologies, Carlsbad, CA). The Ion Chef system was used for templating and the loaded chips sequenced using the Ion Proton System (both from Life Technologies). Sequencing data was automatically transferred to the dedicated Ion Torrent server to generate sequencing reads. Reads quality and trimming was performed using Torrent Server v4.2 before read alignment with TMAP 4.2. The TS plugin CoverageAnalysis v4.2 was used to generate reads count. The sequencing was performed in duplicates and sequence reads were normalized to the total read count.

2.5 PCR

Total RNA extracts from mock and ST3GAL4 transfected cell lysates were isolated with TRI Reagent (Sigma-Aldrich, St. Louis, MO) and converted into cDNA using the SuperScript® II Reverse Transcriptase (Invitrogen) according to the manufacturer's protocol. The primer sequence and conditions were previously described in [22]. The following CD44 primers were applied:

(A) 5'AGTCACAGACCTGCCCAATGCCTTT3';

(B) 3'TTTGCTCCACCTTCTTGACTCCCATG5';

(C) 5'GGGAGCCAAATGAAGAAAATGAAGATGAAAG3';

(D) 3'GGTGCCTGTCTCTTTCATCTTCATTTTCTTCATTT5';

(E) 3'TCTGTTGCCAAACCACTGTTCCTTCTG5'.

2.6 Deglycosylation

The glycosaminoglycans were removed from cell lysates by applying Chondroitinase ABC (120 mU/ml), Heparinase I (2,5 mU/ml) and Heparinase III (2,5 mU/ml) all from Sigma Aldrich per 550 µg protein lysate, the reaction mix was supplemented by 50 µM calcium acetate and incubated overnight at 37°C in a final volume of 250 µl. *N*- and *O*-glycans were removed on-beads after IP of target proteins by sequential application of Neuraminidase (Sigma-Aldrich, St.Louis, MO) for 2h at pH 5 and 37°C, and Protein Deglycosylation Mix containing PNGase F, *O*-glycosidase, Neuraminidase, β1-4 Galactosidase and β-N-Acetylglucosaminidase (New England Biolabs, Ipswich, MA) as described by manufacturer, overnight at pH 7 and 37°C.

2.7 *In silico* analysis

The *in silico* *N*-glycosylation site prediction was performed with NetNGlyc 1.0 software (http://www.cbs.dtu.dk/services/NetNGlyc/; [23]). *N*-glycosylation sites with proline at the N+1 position or negative prediction results were excluded. The *O*-glycosylation sites were predicted by NetOGlyc 4.0

(http://www.cbs.dtu.dk/services/NetOGlyc/; [24]). Only *O*-glycosylation sites with a score higher than 0.5 were considered. Experimentally confirmed *O*-glycosylation sites were extracted from Glycodomainviewer (http://glycodomain.glycomics.ku.dk/; [24]).

2.8 Immunofluorescence (IF) and *in situ* proximity ligation assay (PLA)

Cells were grown in 15 µ-Chamber 12 well glass slides (IBIDI, Martinsried, Germany) and fixed with cold acetone for 5 min. The IF and PLA protocol was performed as previously described [20] using the antibodies and dilutions noted in table 1. Samples were examined under a Zeiss Imager.Z1 Axio fluorescence microscope (Zeiss, Welwyn Garden City, UK). Images were acquired using a Zeiss Axio cam MRm and the AxioVision Release 4.8.1 software.

3. Results

3.1 *In silico* analysis of CD44

CD44 has 11 *N*-glycosylation sequons (N-X-T/S) in the extracellular domain, two of which (N172 and N328) are unlikely glycosylated due to proline at N+1 position. Among the remaining 9 sequons, the sites N25, N57, N110, N120, N599 and N636 showed a positive outcome in the *in silico* *N*-glycosylation prediction (Table 2) and N25, N57, N100, N110 and N120 were previously experimentally confirmed by other authors [25]. In addition, CD44 has numerous *O*-glycosylation sites. The *in silico* prediction resulted in 146 *O*-glycosylation sites, 41 of them have been previously experimentally proven in other studies [24]. The *O*-glycosylation sites are primarily found in the stem region of CD44, flanked by *N*-glycosylation sites at the *N*-terminal domain and adjacent to the transmembrane domain (Fig 1). CD44 may also carry chondroitin sulfate (CS) at S180 and heparan sulfate (HS) at S293 or S295 [26-28].

Site	*In silico* prediction	Experimental demonstrated
N25	+++	Described in [25]
N57	+++	Described in [25]
N100	–	Described in [25]
N110	++	Described in [25]
N120	+	Described in [25]
N350	–	
N548	–	
N599	++	
N636	+	

Table 2. *N*–glycosylation sites of CD44.

Human CD44 asparagines within the *N*–glycosylation consensus sequence are shown with their *in silico* prediction scores. "–" indicates improbable and "+", "++" and "+++" probable *N*–glycosylation sites with increasing confidence. *N*–glycosylation sites that were experimentally demonstrated are shown in the right column.

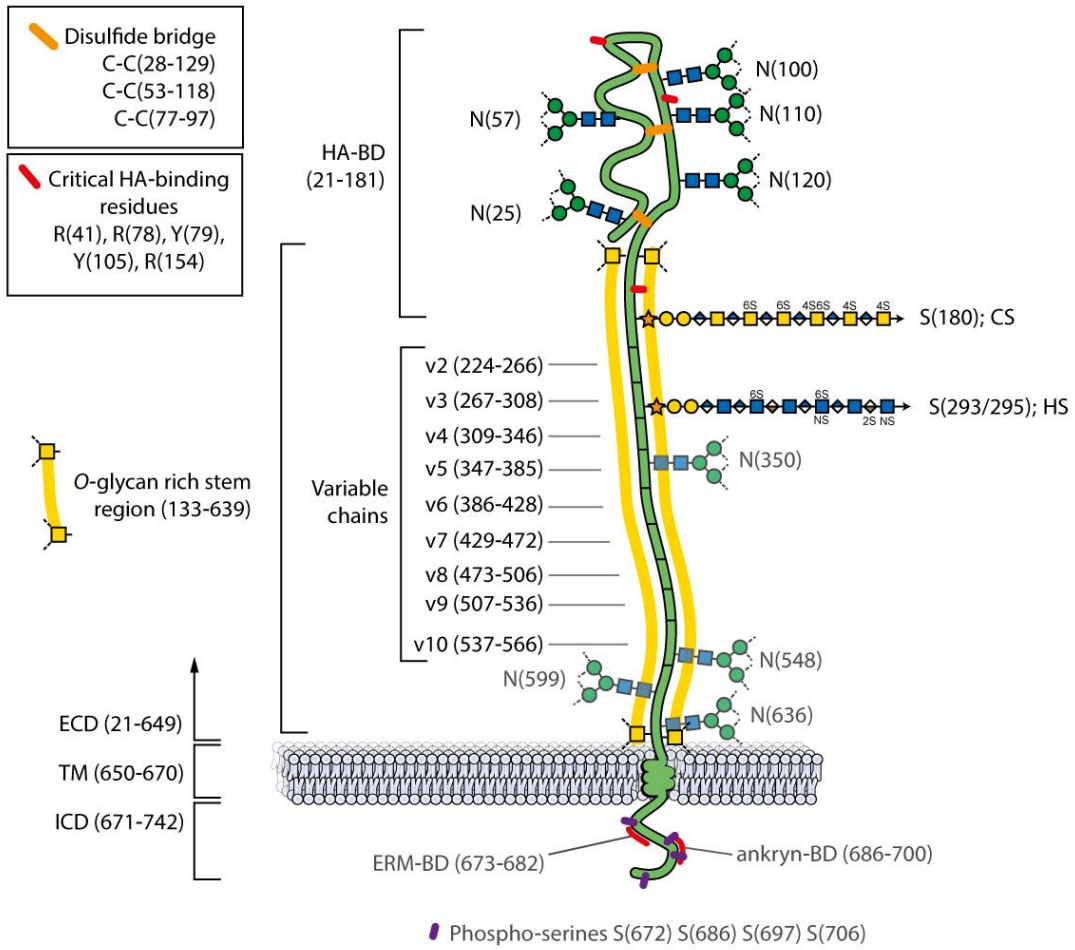


Figure 1. CD44 glycoprotein depiction including its isoforms and post-translational modifications.

Amino acid positions are annotated between brackets. The full length protein stretches 742 amino acids with the first 20 amino acids being however cleaved as part of the signal peptide. Disulfide bridges are illustrated in orange. Amino acids whose mutations lead to loss of hyaluronan binding are annotated as critical HA-binding residues and shown in red. Potential palmitoylation of the cytosolic domain is not shown. HA-BD: Hyaluronan binding domain; ECD: Extracellular domain; TM: Transmembrane domain; ICD: Intracellular domain; ERM-BD: Ezrin-radixin-moesin binding domain; ankryn-BD: Ankryn binding domain; CS: Chondroitin sulfate; HS: Heparan sulfate.

3.2 The modification of the glycosylation machinery has a big impact on CD44 molecular features.

To test whether *O*-glycan truncation has an impact on CD44 features, we performed a comparative analysis on three different glycoengineered cell line models derived from MKN45 gastric carcinoma cells (WT): namely ST3GAL4 overexpressing (ST3), ST6GALNAC1 overexpressing (ST6) and COSMC knock-out MKN45 cells (SC), and their respective control cell lines. Whereas the control MKN45 cells expressed predominantly CD44 with molecular weight above 150 kDa (heavy CD44, Fig 2A), the glycoengineered cell line models expressed a CD44 protein with less than 150 kDa (light CD44). The light CD44 was approximately 50 - 70 kDa smaller than the heavy CD44 and presented a more defined band indicating a more homogenous CD44 isoform. The observed expression of light CD44 was independent of confluency status of the cultured cells (data not shown). Total mRNA levels for CD44 were unaltered in all cell line models (Fig 2B), and proteomic analysis of the ST3 cell line further supported no changes in total CD44 protein levels (Fig 2C). A WB analysis specific for the v6 splice variant did not present the light CD44 band (Fig 2D), indicating an absence of the v6 exon in the light CD44. Despite the different CD44v6 WB pattern displayed by the different cell line models, IF showed similar intensities and subcellular localization (Fig 2E).

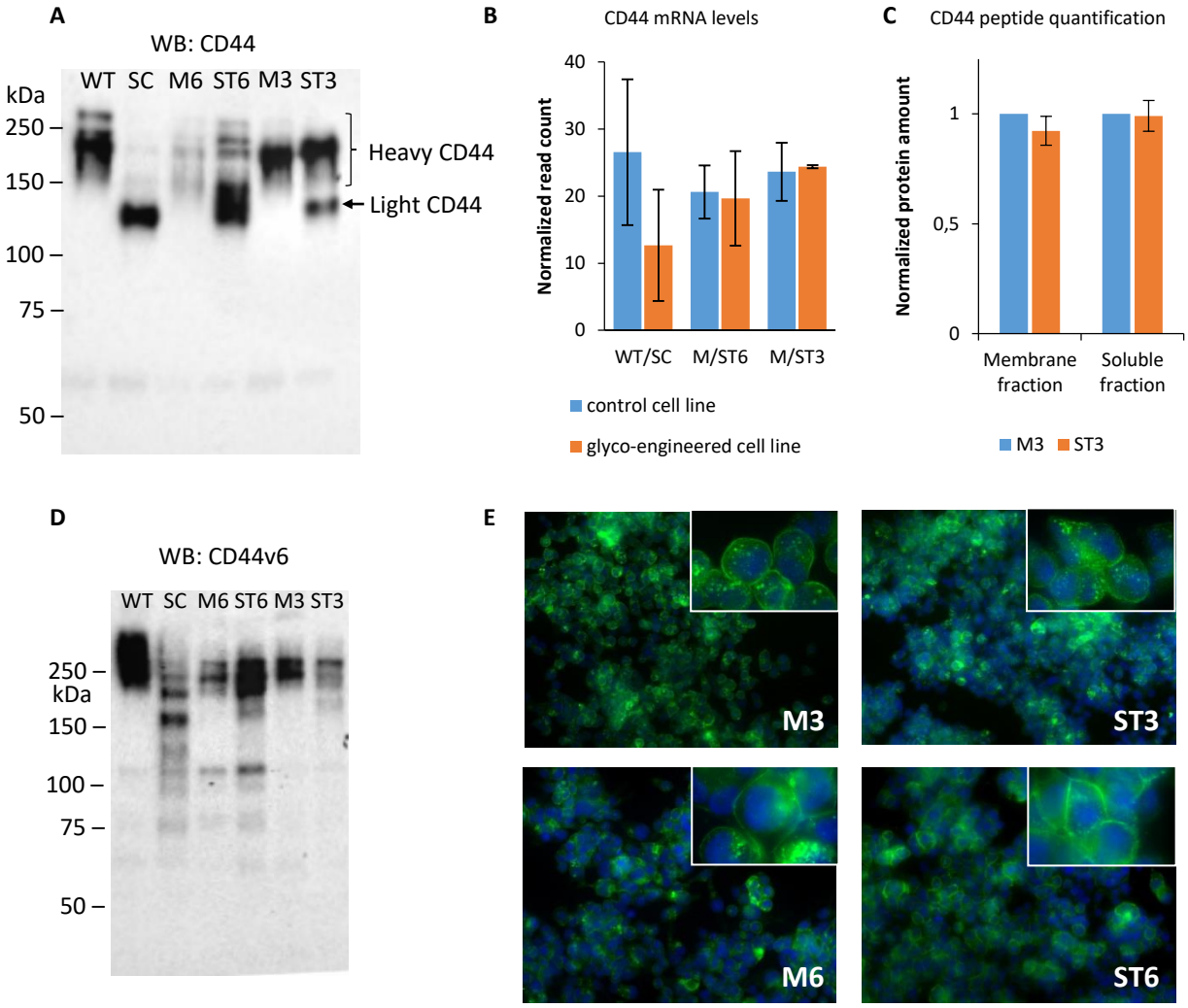


Figure 2. CD44 expression in different glycoengineered cell line models. **A.** Western blot analysis of CD44 expressed by the different cell line models. Same amounts of protein were loaded for all samples. **B.** Transcriptomic data on CD44 mRNA levels of the cell line models. **C.** Quantitative proteomic evaluation of CD44 peptides from membrane and soluble fractions of M3 and ST3 cells. The CD44 protein coverage of membrane and soluble fraction was 18.9% and 6.3%, respectively. **D.** Western blot analysis of CD44v6 expressed by the different cell line models. Same amount of protein were loaded for all samples. **E.** Immunofluorescence analysis of CD44v6 (green) expression of M3, ST3, M6 and ST6 cell line models. Data is shown from MKN45 wild type (WT), SimpleCells (SC), ST3GAL4 transfected (ST3), ST6GALNAC1 transfected (ST6) and empty vector transfected control cell lines (M3 and M6).

3.3 CD44 splicing is unaffected in our glycoengineered cell line models

To assess if changes in the molecular mass of CD44 and CD44v6 stem from alterations in alternative splicing machinery we evaluated the expression levels of splicing factors ESRP1, ESRP2, RBM5, RBM6, SRSF1, SRSF6, hnRNPM, SND1, TRA2B and SRSF2 that have been described to induce differential splicing of CD44 (Fig 3) [29]. Expression levels of all splicing factors were remarkably uniform among the glycoengineered cell lines and their controls. In addition, we evaluated different CD44 transcripts by a series of PCRs that span different number of variable chains (Fig 4A). No differences in CD44 cDNA length and abundance became evident (Fig 4B).

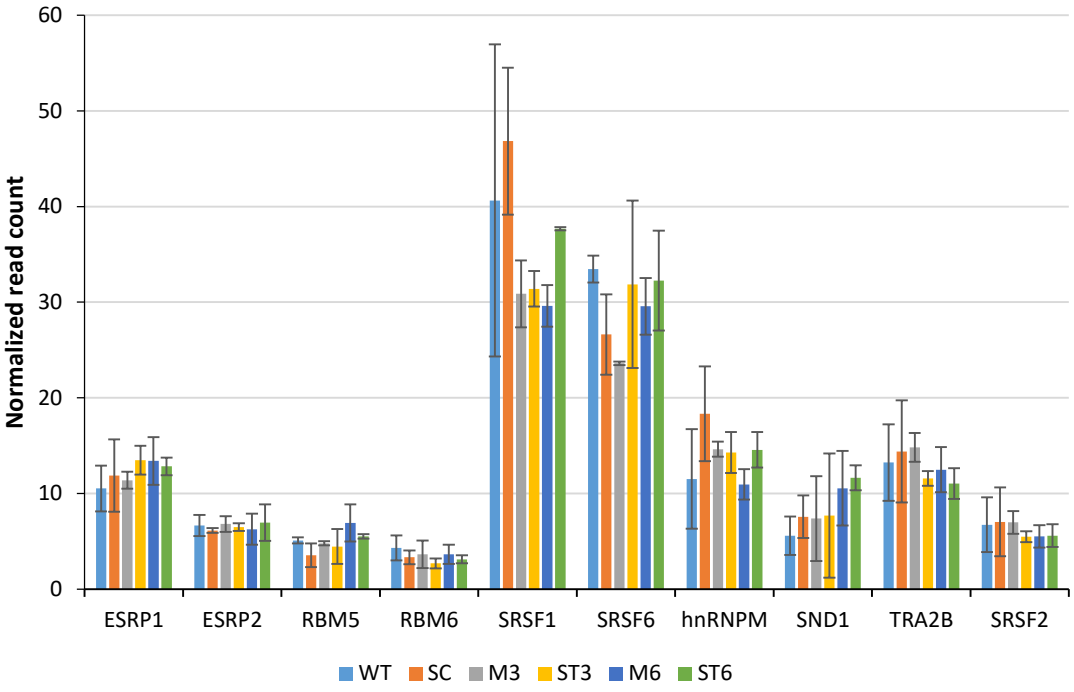


Figure 3. Expression levels of factors involved in CD44 alternative splicing. The mRNA levels of ESRP1, ESRP2, RBM5, RBM6, SRSF1, SRSF6, hnRNPM, SND1, TRA2B and SRSF2 splicing factors were analysed in duplicates. Average and standard deviation are shown for MKN45 wild type (WT), SimpleCells (SC), ST3GAL4 transfected (ST3), ST6GALNAC1 transfected (ST6) and empty vector transfected control cell lines (M3 and M6).

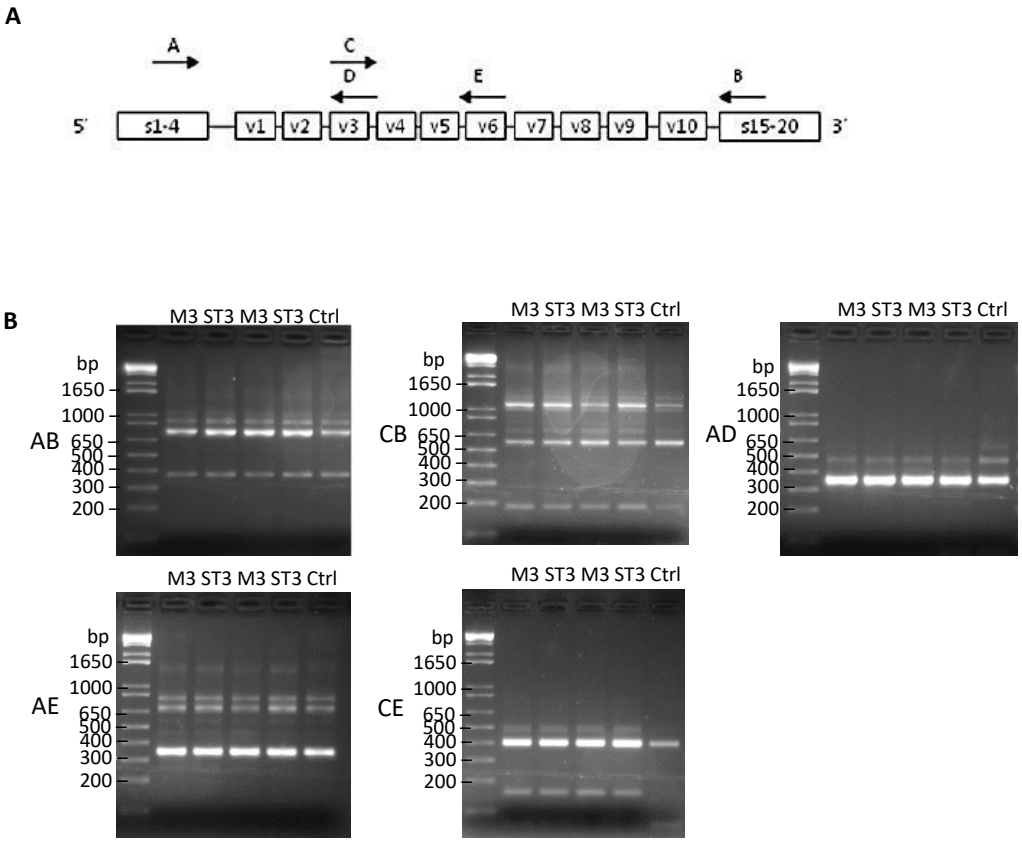


Figure 4. PCR analyses of CD44 transcripts.

A. A series of PCR analyses were performed using primers flanking different variant regions within CD44 transcript. Extracted from [22]. **B.** PCR products of primer pairs (shown at left) resolved in 2% of agarose gels. No differences between the transcripts of ST3GAL4 overexpressing cells (ST3) and mock control (M3) were detected. mRNA from Colo 205 cell line was used as a control (Ctrl) and two independent M3 an ST3 samples were loaded for analysis.

3.4 CD44 as the major carrier of STn

The overexpression of the sialyltransferase ST6GALNAC1 (ST6) or the knock-out of COSMC (SC) resulted in the overexpression of the truncated *O*-glycan epitope STn in these gastric cell line models [16, 19]. The IP for CD44 showed that the light CD44 is the major carrier of STn in these cell lines (Fig 5).

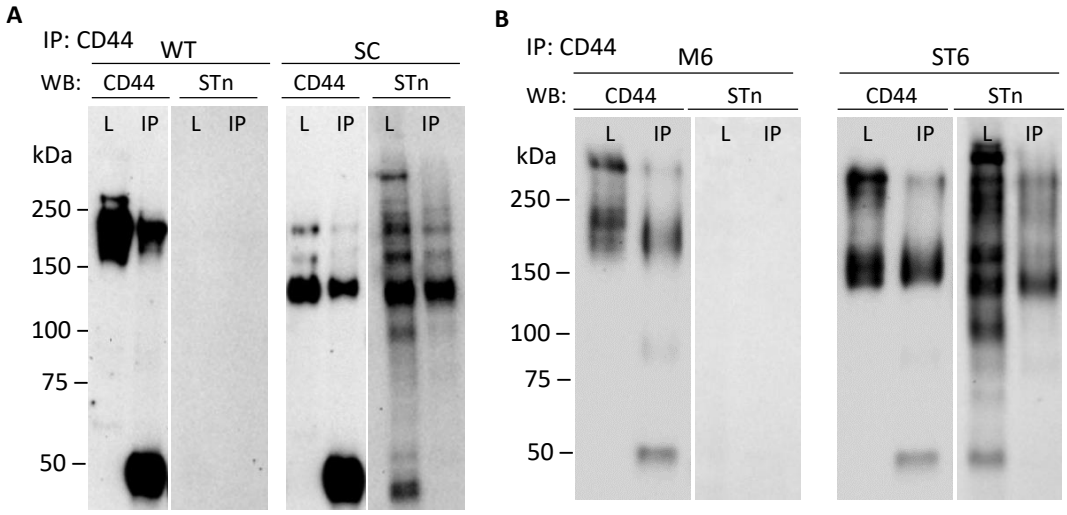


Figure 5. STn expression on CD44.

Western blotting for CD44 and STn in total cell lysate (L) and immunoprecipitated CD44 (IP) from MKN45 SC and WT (**A**) or ST6GALNAC1 overexpressing cells (ST6) and mock control (M6) (**B**).

3.5 Heparan sulfate and chondroitin sulfate are unaltered on CD44 in the glycoengineered cell line models

Despite the different alterations induced in the three models, we observed a similar shift in electrophoretic behaviour of CD44. We therefore hypothesized that the underlying mechanisms that account for the loss of molecular weight are equal among our models. Since we ruled out that this effect was due to alternative splicing, glycosylation remained as the only posttranslational modification that could provide such high molecular weight differences between the heavy and light CD44. Alterations in occupancy and structure of these chains can significantly influence glycoprotein/proteoglycan molecular weight. To test whether this drastic shift was caused by differential addition of the glycosaminoglycans heparan sulfate (HS) and chondroitin sulfate (CS), we performed complete digestion of HS, CS or both on cell lysates (Fig 6). The efficiency of these digestions was shown by the loss of HS-specific and CS-specific antibody binding as shown by dot-blot analysis (Fig 6A). As additional control of HS digestion, Syndecan 1 which is known to carry HS [30] was evaluated. The WB detection of Syndecan 1 at under 100 kDa demonstrates the efficient release of HS (Fig 6B). The digestion of HS and CS had no significant effect on CD44 and therefore excludes alterations in glycosaminoglycan occupancy or length as the cause for the difference between light and heavy CD44 (Fig 6C). Interestingly, Syndecan 1 showed a sharper band with lower molecular weight in the ST6 cell model compared to the control (Fig 6B). This difference in molecular weight is similar to what was observed for CD44. Syndecan 1 is a highly *O*-glycosylated protein, but unlike CD44, Syndecan 1 has no alternative spliced isoforms. This further indicates the involvement of *O*-glycosylation in the change of molecular weight in these cell line models.

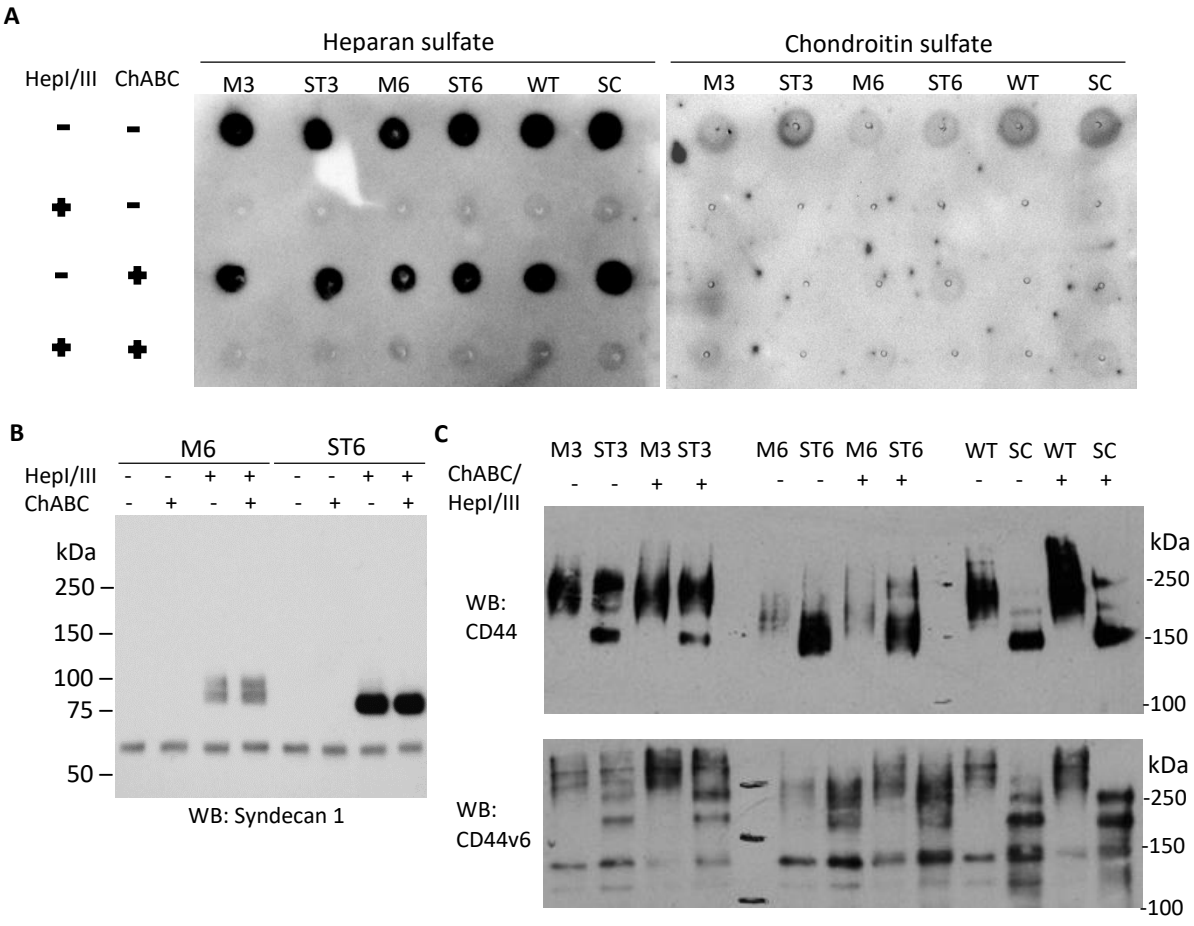


Figure 6. Single and combined digestions of heparan sulfate and chondroitin sulfate.

A. Dot-blot of cell lysates tested with anti-HS or anti-CS antibody. **B.** Syndecan 1 Western blotting with lysates of M6 and ST6. **C.** Western blotting analysis of CD44 and CD44v6 in cell line models before and after CS and HS digestion. ChABC: chondroitinase ABC; CS: chondroitin sulfate; Hep: heparinase; HS: heparan sulfate.

3.6 *N*- and *O*-glycosylation digestion of CD44

Enzymatic deglycosylation of CD44 was performed in order to evaluate whether the different molecular weight of light and heavy CD44 arose from differences in *N*- and *O*-glycosylation. The deglycosylation protocol was optimized using the ST6 cell line (Fig 7A and B). The double treatment with Neuraminidase and Deglycosylation Mix (containing PNGase F, *O*-glycosidase, Neuraminidase, β1-4 Galactosidase and β-N-Acetylglucosaminidase) was found most efficient. Applied on CD44 of M6, the deglycosylation treatment reduced the molecular weight significantly by around 50kDa, aligning it to the light CD44 of ST6 (Fig 7C). The molecular weight of the light CD44 from ST6 was additionally reduced after the deglycosylation by approximately 25kDa.

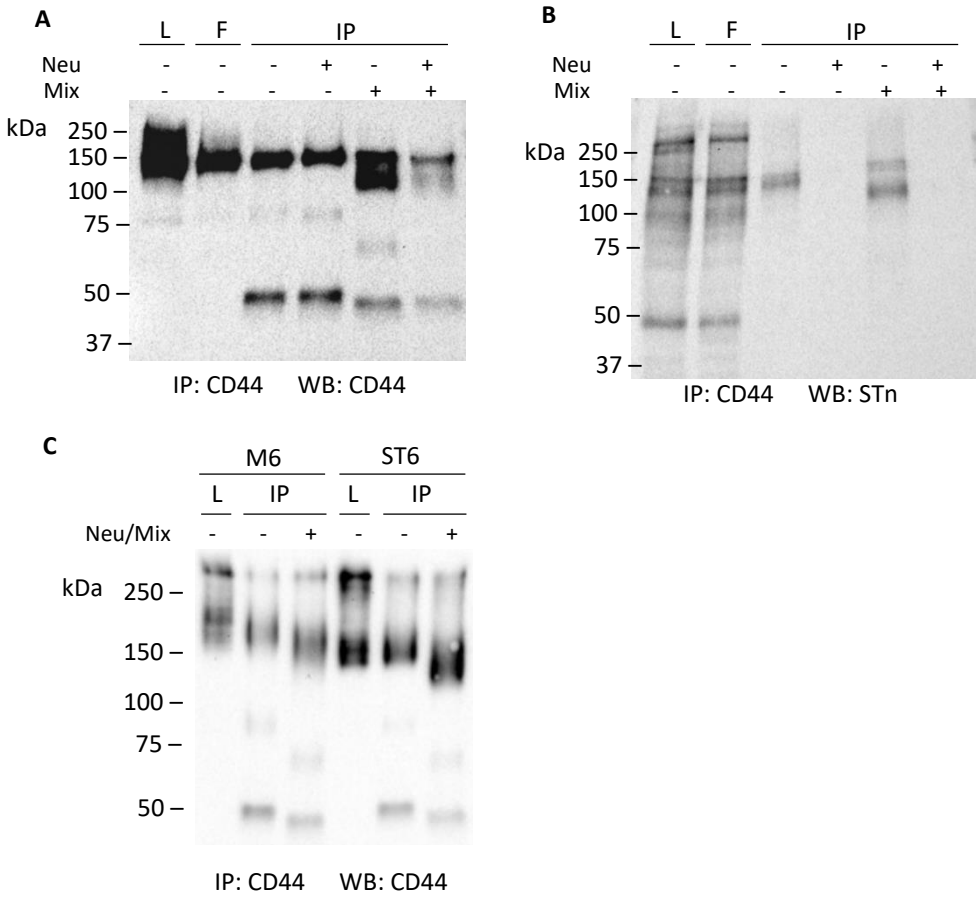


Figure 7. *N*– and *O*–glycosylation digestions of CD44.

A. and B. Optimization of enzymatic deglycosylation process on immunoprecipitated CD44 from ST6 cells analysed by Western blotting for CD44 and STn. **C.** Removal of *N*– and *O*–glycosylation from CD44 of ST6 and M6 cell line models. L: Whole cell lysate; F: Flow through; IP: Immunoprecipitation;

3.7 Altered glycosylation increases CD44v6 and RON colocalization

CD44v6 is one of the co-receptor of the RON receptor tyrosine kinase. The oncogenic hyperactivation of RON has been described as a common event in various cancers, such as pancreatic cancer and gastric cancer [31]. In these cancers, CD44, and in particular the isoform v6, is often overexpressed. In order to test if the altered glycosylation of CD44 influences the v6 association with RON, we performed a series of *in situ* proximity ligation assay (PLA) experiments. ST3 and M3 have previously been described to be SLe^x positive and negative, respectively [17, 20]. In accordance, the results of the PLA of ST3 cells showed colocalization of CD44v6 and SLe^x (Fig 8A). In agreement with the statement of CD44 being the major carrier of STn in ST6 cells, the PLA displayed colocalization between CD44v6 and STn (Fig 8B). Finally, we performed a PLA for CD44v6 and RON. Remarkably, the association between CD44v6 and RON was significantly more frequent in cells expressing light CD44 (ST3 and ST6) than in cells with heavy CD44 (M3 and M6) (Fig 8C).

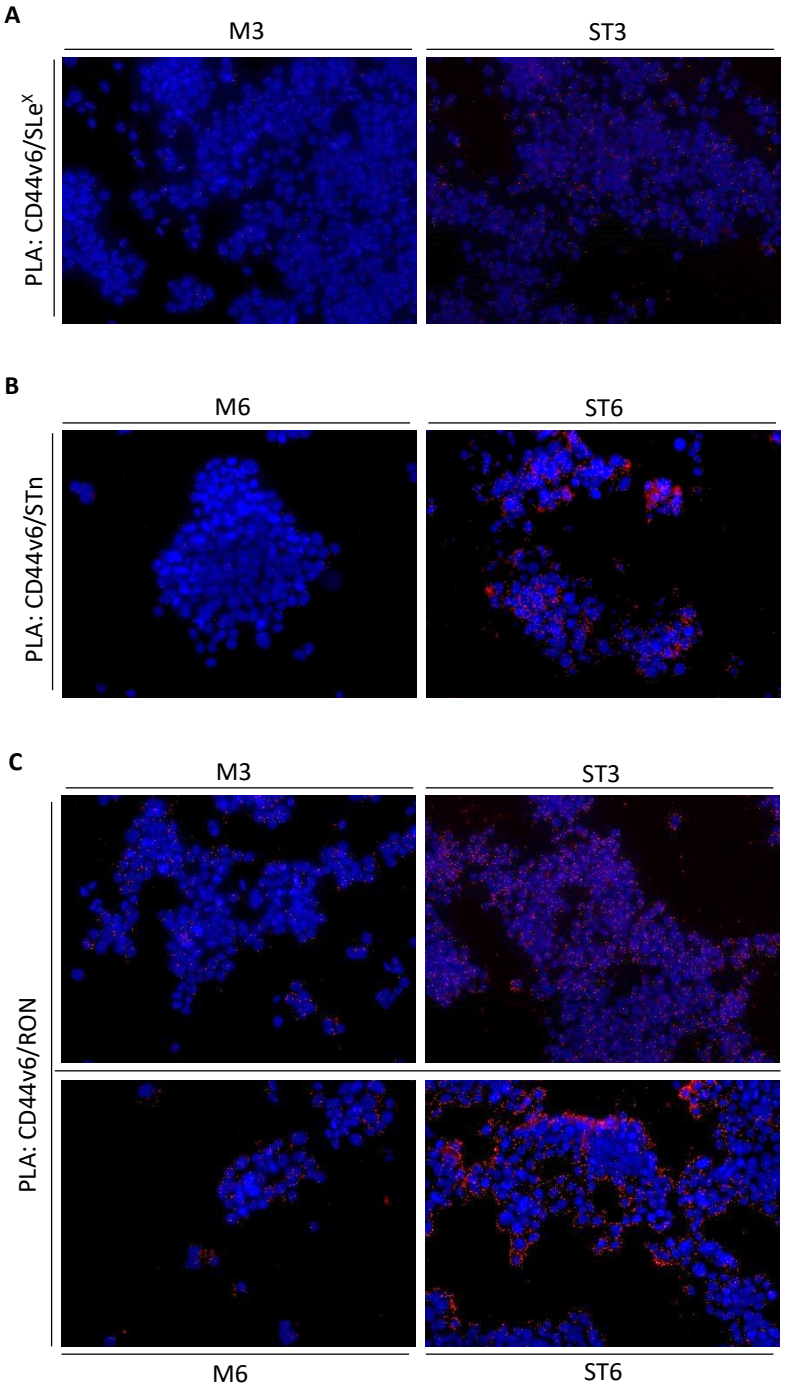


Figure 8. Colocalization analyses of CD44v6 with glycan epitopes and RON.

A. *In situ* proximity ligation assay (PLA) of CD44v6 and SLe^x in M3 and ST3 cells. **B.** PLA of CD44v6 and STn in M6 and ST6 cells. **C.** PLA of CD44v6 and RON in M3 and ST3 (top panel) and M6 and ST6 (bottom panel). Concomitantly, the colocalization between CD44v6 and RON is highly increased. Red dots represent conjugation events due to close molecular proximity of the two tested antigens. The magnification is 200x and nuclei are shown in blue.

4. Discussion

CD44 is a multi-functional transmembrane protein that is abundantly decorated by a variety of glycan moieties. The modulation of glycosylation of CD44 has been shown to alter its function as hyaluronan receptor [10, 25]. In gastric cancer, high expression levels of CD44 are found in about half of all primary tumors and are associated with the presence of distant metastases, tumor recurrence and increased mortality [32]. Moreover, an increasing body of evidence demonstrates that CD44 is a gastric cancer stem cell marker, since CD44 expressing cancer cells are competent to perform self-renewal and to produce differentiated progeny [13]. CD44 has been shown to act as co-receptor for several receptor tyrosine kinases and to elevate cancer cell resistance to therapy. In particular, the variant form CD44v6 was directly linked to RON receptor tyrosine kinase activation [7]. In this manuscript we studied if cancer associated glycan alterations, namely *O*-glycan truncation, can alter molecular features of CD44 in gastric cancer and modulate its function to act as co-receptor to RON.

We performed a comparative analysis on three differently glycoengineered cell line models: ST3GAL4 overexpressing (ST3), ST6GALNAC1 overexpressing (ST6) and COSMC knock-out (SC), and compared them to their unmodified mock and WT cell lines. Although a different gene was targeted, the truncation of *O*-glycans was a common feature of the three cell line models.

ST3GAL4 is a sialyltransferase that affects both *N*- and *O*-glycans. The upregulation of ST3GAL4 in MKN45 leads to an increase of α2,3 and a decrease of α2,6 sialylation on *N*-glycans and significantly decreases the amount of bisected structures [20, 33]. On *O*-glycans ST3GAL4 leads to the earlier termination due to the increase of one specific structure: the di-sialylated core 2 structure [20, 33]. Both the overexpression of ST6GALNAC1 and the COSMC deletion induce alteration on *O*-glycans and lead to

an upregulation of STn by two differential mechanisms. ST6GALNAC1 is a sialyltransferase that adds sialic acid to the Tn epitope to form STn and leads thereby to the early termination of the *O*-glycan chains [16]. COSMC on the other hand is the chaperone of the core 1 galactosyltransferase (C1GALT1). Without its chaperone C1GALT1 is dysfunctional and fails elongating the Tn antigen to core 1. In SC the lack of COSMC causes the accumulation of Tn and STn antigens [19, 34].

The four mechanisms that may contribute to major changes of CD44 are (i) alternative splicing, (ii) proteolytic cleavage, (iii) *N*- and *O*-glycosylation and (iv) glycosaminoglycan chains. MKN45 expresses mostly CD44v8-10 splicing isoforms, known as CD44E because of its common expression in epithelial cells, but also CD44s (without variant chains) and CD44v6 [12]. As the staining intensity and subcellular localization for CD44v6 remained similar among the cell line models (Fig 2 E) we presume that levels of CD44v6 remain unaltered. This was further supported by PCR analyses showing equal results for all tested splice variations including v3 and v6 (Fig 4). Furthermore, although transcriptomic analysis of the cells showed in general very different expression profiles among the glycoengineered cell lines (data not shown), CD44 expression levels remained equal (Fig 2 B). Furthermore, splicing factors known to modulate CD44 isoforms [29] were unaltered at mRNA level such as ESRP1, ESRP2, RBM5, RBM6, SRSF1, SRSF6, hnRNPM, SND1, TRA2B, SRSF2 (Fig 3).

On the other hand, the shift in the molecular weight of CD44 could have been caused by a proteolytic cleavage resulting in the sequential release of the extracellular domain (ECD) and the intracellular domain (ICD). Several proteases can be involved in the shedding of the ECD, such as membrane type 1-matrix metalloproteinase (MT1-MMP) and disintegrin and metalloproteinase domain-containing proteins 10 and 17 (ADAM-10 and ADAM-17) [35, 36]. For some proteases the cleavage sites have been confined to the stem region which releases the ectodomain of CD44. Interestingly, it has been shown that the proteolytic cleavage can be triggered by the binding of CD44

[37]. In that sense, the modulation of the CD44 interaction with other proteins and glycosaminoglycans is expected to alter the cleavage susceptibility. Considering that the methodology employed in this work analysed the membrane associated CD44, the proteolytic cleavage followed by the shedding of the ectodomain can be precluded as the cause of the shift from heavy CD44 (more than 150 kDa) to light CD44 (between 100 and 150 kDa) observed in this work.

We also ruled out that alterations in glycosaminoglycans accounted for the light version of CD44. However, our results point towards the truncation of *O*-glycosylation as the underlying mechanism for the expression of the light CD44 isoform. This is supported by a similar result observed for the highly *O*-glycosylated proteoglycan Syndecan 1 (Fig 6B). The extracellular domain of Syndecan 1 consists of 56 serine or threonine out of a total 232 amino acids, with the majority (50) being predicted to be *O*-glycosylated (data not shown). In accordance with our hypothesis that the lower molecular weight form of CD44 stems from the truncation of *O*-glycans, we observe a lower molecular weight form of Syndecan 1 as well.

We have not completely equalized the electrophoretic behaviour of CD44 from glycoengineered and control cell lines. This is likely due to incomplete release of *O*-glycans from our samples by the enzymatic *O*-glycan removal. Presently, no single glycosidase can release all *O*-glycans. We have therefore used a serial digest with Neuraminidase followed by an enzyme mix containing PNGase F, *O*-glycosidase, Neuraminidase, β1-4 Galactosidase and β-N-Acetylglucosaminidase. We demonstrated that Neuraminidase and PNGase F treatment were successful in removing sialic acids and *N*-glycans, respectively. Given that high density of extended *O*-glycans entails considerable resistance to glycosidases a chemical approach for glycan removal should be carried out in the future to obtain a fully deglycosylated protein.

Alterations of the glycosylation machinery leading to the truncation of *O*-glycans are a common feature in cancer. In this work we demonstrated the big impact that *O*-glycan truncation has on CD44, a key player of cell-matrix interaction and receptor activation. CD44 is among the most amplified genes in gastric cancer, indicating its central role in this malignancy [38, 39]. In addition, CD44 shows frequently *de novo* expression of the v6 splice variant in gastric cancer [12]. The expression of CD44v6 facilitates the receptor tyrosine kinase activation of RON by a yet unknown mechanism [7]. We revealed here that the truncation of *O*-glycan leads to increased colocalization between CD44v6 and RON. These results show that the aberration of *O*-glycans presents a novel oncogenic mechanism of RON activation mediated by CD44. CD44 is also involved in the activation of other receptor tyrosine kinases, such as MET and EGFR [7, 8]. The expression of truncated glycan epitopes bares therefore the potential to be used in the future as biomarker for patient treatment stratification with tyrosine kinase inhibitors.

5. References

1 Marhaba, R. and Zoller, M. (2004) CD44 in cancer progression: adhesion, migration and growth regulation. Journal of molecular histology. 35, 211-231

2 Okamoto, I., Kawano, Y., Murakami, D., Sasayama, T., Araki, N., Miki, T., . . . Saya, H. (2001) Proteolytic release of CD44 intracellular domain and its role in the CD44 signaling pathway. J Cell Biol. 155, 755-762

3 Lesley, J., English, N., Perschl, A., Gregoroff, J. and Hyman, R. (1995) Variant cell lines selected for alterations in the function of the hyaluronan receptor CD44 show differences in glycosylation. J Exp Med. 182, 431-437

4 Screaton, G. R., Bell, M. V., Jackson, D. G., Cornelis, F. B., Gerth, U. and Bell, J. I. (1992) Genomic structure of DNA encoding the lymphocyte homing receptor CD44 reveals at least 12 alternatively spliced exons. Proc Natl Acad Sci USA. 89, 12160-12164

5 Naor, D., Sionov, R. V. and Ish-Shalom, D. (1997) CD44: structure, function, and association with the malignant process. Adv Cancer Res. 71, 241-319

6 Jones, M., Tussey, L., Athanasou, N. and Jackson, D. G. (2000) Heparan sulfate proteoglycan isoforms of the CD44 hyaluronan receptor induced in human inflammatory macrophages can function as paracrine regulators of fibroblast growth factor action. J Biol Chem. 275, 7964-7974

7 Matzke, A., Herrlich, P., Ponta, H. and Orian-Rousseau, V. (2005) A five-amino-acid peptide blocks Met- and Ron-dependent cell migration. Cancer Res. 65, 6105-6110

8 Sherman, L. S., Rizvi, T. A., Karyala, S. and Ratner, N. (2000) CD44 enhances neuregulin signaling by Schwann cells. J Cell Biol. 150, 1071-1084

9 Stamenkovic, I., Aruffo, A., Amiot, M. and Seed, B. (1991) The hematopoietic and epithelial forms of CD44 are distinct polypeptides with different adhesion potentials for hyaluronate-bearing cells. EMBO J. 10, 343-348

10 Bennett, K. L., Modrell, B., Greenfield, B., Bartolazzi, A., Stamenkovic, I., Peach, R., . . . Aruffo, A. (1995) Regulation of CD44 binding to hyaluronan by glycosylation of variably spliced exons. J Cell Biol. 131, 1623-1633

11 Bennett, K. L., Jackson, D. G., Simon, J. C., Tanczos, E., Peach, R., Modrell, B., . . . Aruffo, A. (1995) CD44 isoforms containing exon V3 are responsible for the presentation of heparin-binding growth factor. J Cell Biol. 128, 687-698

12 da Cunha, C. B., Oliveira, C., Wen, X., Gomes, B., Sousa, S., Suriano, G., . . . Seruca, R. (2010) De novo expression of CD44 variants in sporadic and hereditary gastric cancer. Laboratory investigation; a journal of technical methods and pathology. 90, 1604-1614

13 Takaishi, S., Okumura, T., Tu, S., Wang, S. S., Shibata, W., Vigneshwaran, R., . . . Wang, T. C. (2009) Identification of gastric cancer stem cells using the cell surface marker CD44. Stem cells. 27, 1006-1020

14 Yamaguchi, A., Goi, T., Yu, J., Hirono, Y., Ishida, M., Iida, A., . . . Hirose, K. (2002) Expression of CD44v6 in advanced gastric cancer and its relationship to hematogenous metastasis and long-term prognosis. J Surg Oncol. 79, 230-235

15 Uhlen, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., . . . Ponten, F. (2015) Proteomics. Tissue-based map of the human proteome. Science. 347, 1260419

16 Marcos, N. T., Pinho, S., Grandela, C., Cruz, A., Samyn-Petit, B., Harduin-Lepers, A., . . . Reis, C. A. (2004) Role of the human ST6GalNAc-I and ST6GalNAc-II in the synthesis of the cancer-associated sialyl-Tn antigen. Cancer Res. 64, 7050-7057

17 Carvalho, A. S., Harduin-Lepers, A., Magalhaes, A., Machado, E., Mendes, N., Costa, L. T., . . . Reis, C. A. (2010) Differential expression of alpha-2,3-sialyltransferases and alpha-1,3/4-fucosyltransferases regulates the levels of sialyl Lewis a and sialyl Lewis x in gastrointestinal carcinoma cells. The international journal of biochemistry & cell biology. 42, 80-89

18 Steentoft, C., Vakhrushev, S. Y., Vester-Christensen, M. B., Schjoldager, K. T., Kong, Y., Bennett, E. P., . . . Clausen, H. (2011) Mining the O-glycoproteome using zinc-finger nuclease-glycoengineered SimpleCell lines. Nature methods. 8, 977-982

19 Campos, D., Freitas, D., Gomes, J., Magalhaes, A., Steentoft, C., Gomes, C., . . . Reis, C. A. (2015) Probing the O-glycoproteome of gastric cancer cell lines for biomarker discovery. Molecular & cellular proteomics : MCP. 14, 1616-1629

20 Mereiter, S., Magalhaes, A., Adamczyk, B., Jin, C., Almeida, A., Drici, L., . . . Reis, C. A. (2016) Glycomic analysis of gastric carcinoma cells discloses glycans as modulators of RON receptor tyrosine kinase activation in cancer. Biochim Biophys Acta. 1860, 1795-1808

21 Kjeldsen T, Clausen H, Hirohashi S, Ogawa T, Iijima H, Hakomori S. (1988) Preparation and characterization of monoclonal antibodies directed to the tumor-associated O-linked sialosyl-2-6 alpha-N-acetylgalactosaminyl (sialosyl-Tn) epitope. Cancer research. 48:2214-2220.

22 Banky, B., Raso-Barnett, L., Barbai, T., Timar, J., Becsagh, P. and Raso, E. (2012) Characteristics of CD44 alternative splice pattern in the course of human colorectal adenocarcinoma progression. Mol Cancer. 11, 83

23 Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S. and Brunak, S. (2004) Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics. 4, 1633-1649

24 Steentoft, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B., Schjoldager, K. T., . . . Clausen, H. (2013) Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J. 32, 1478-1488

25 Bartolazzi, A., Nocks, A., Aruffo, A., Spring, F. and Stamenkovic, I. (1996) Glycosylation of CD44 is implicated in CD44-mediated cell adhesion to hyaluronan. J Cell Biol. 132, 1199-1208

26 Faassen, A. E., Schrager, J. A., Klein, D. J., Oegema, T. R., Couchman, J. R. and McCarthy, J. B. (1992) A cell surface chondroitin sulfate proteoglycan, immunologically related to CD44, is involved in type I collagen-mediated melanoma cell motility and invasion. J Cell Biol. 116, 521-531

27 Zhang, L., David, G. and Esko, J. D. (1995) Repetitive Ser-Gly sequences enhance heparan sulfate assembly in proteoglycans. J Biol Chem. 270, 27127-27135

28 Greenfield, B., Wang, W. C., Marquardt, H., Piepkorn, M., Wolff, E. A., Aruffo, A. and Bennett, K. L. (1999) Characterization of the heparan sulfate and chondroitin sulfate assembly sites in CD44. J Biol Chem. 274, 2511-2517

29 Prochazka, L., Tesarik, R. and Turanek, J. (2014) Regulation of alternative splicing of CD44 in cancer. Cell Signal. 26, 2234-2239

30 Xian, X., Gopal, S. and Couchman, J. R. (2010) Syndecans as receptors and organizers of the extracellular matrix. Cell Tissue Res. 339, 31-46

31 Yao, H. P., Zhou, Y. Q., Zhang, R. and Wang, M. H. (2013) MSP-RON signalling in cancer: pathogenesis and therapeutic potential. Nat Rev Cancer. 13, 466-481

32 Mayer, B., Jauch, K. W., Gunthert, U., Figdor, C. G., Schildberg, F. W., Funke, I. and Johnson, J. P. (1993) De-novo expression of CD44 and survival in gastric cancer. Lancet. 342, 1019-1022

33 Mereiter, S., Magalhaes, A., Adamczyk, B., Jin, C., Almeida, A., Drici, L., . . . Reis, C. A. (2016) Glycomic and sialoproteomic data of gastric carcinoma cells overexpressing ST3GAL4. Data Brief. 7, 814-833

34 Ju, T., Lanneau, G. S., Gautam, T., Wang, Y., Xia, B., Stowell, S. R., . . . Cummings, R. D. (2008) Human tumor antigens Tn and sialyl Tn arise from mutations in Cosmc. Cancer Res. 68, 1636-1646

35 Okamoto, I., Kawano, Y., Tsuiki, H., Sasaki, J., Nakao, M., Matsumoto, M., . . . Saya, H. (1999) CD44 cleavage induced by a membrane-associated metalloprotease plays a critical role in tumor cell migration. Oncogene. 18, 1435-1446

36 Nagano, O. and Saya, H. (2004) Mechanism and biological significance of CD44 cleavage. Cancer Sci. 95, 930-935

37 Shi, M., Dennis, K., Peschon, J. J., Chandrasekaran, R. and Mikecz, K. (2001) Antibody-induced shedding of CD44 from adherent cells is linked to the assembly of the cytoskeleton. J Immunol. 167, 123-131

38 Fukuda, Y., Kurihara, N., Imoto, I., Yasui, K., Yoshida, M., Yanagihara, K., . . . Inazawa, J. (2000) CD44 is a potential target of amplification within the 11p13 amplicon detected in gastric cancer cell lines. Genes Chromosomes Cancer. 29, 315-324

39 Cancer Genome Atlas Research, N. (2014) Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 513, 202-209

CHAPTER V

GLYCAN SIGNATURE ACTS AS PREDICTIVE MARKER IN GASTRIC CANCER

GLYCAN SIGNATURE ACTS AS PREDICTIVE MARKER IN GASTRIC CANCER

Stefan Mereiter[1,2,3], Coralie Williams[4], Nina Persson[5], António Polonia[1,2], Karol Polom[6], Franco Roviello[6], Ola Blixt[5], Mariana Kuras[4], Ana Magalhães[1,2], Celso A. Reis[1,2,3,7]

1. i3S — Instituto de Investigação e Inovação em Saúde, University of Porto, Portugal
2. Institute of Molecular Pathology and Immunology of the University of Porto — IPATIMUP, Porto, Portugal
3. Institute of Biomedical Sciences of Abel Salazar — ICBAS, University of Porto, Portugal
4. Ariana Pharmaceuticals, Paris, France
5. Department of Chemistry, University of Copenhagen, Denmark
6. Department of Medicine, Surgery and Neurosciences, University of Siena, Italy
7. Medical Faculty, University of Porto, Portugal

Abstract

Malignancies of the stomach remain to be the third leading cause of cancer deaths worldwide. Patients are typically diagnosed in advanced stages in which curative attempts are often failing. Predictive biomarkers are needed for the stratification of patients and the improvement of patient outcome. Cancer associated glycan epitopes already play a pivotal role as independent diagnostic and predictive biomarkers and for monitoring treatment progress and cancer recurrence. However, the association of these glycan epitopes with other molecular alterations and clinical parameters is not understood yet. In an attempt to link the expression of glycan epitopes with other indicative factors of the tumor we first characterized the expression of sialyl Lewis a (SLe^a), sialyl Lewis x (SLe^x), sialyl Tn (STn) and Tn in a series of gastric carcinomas. We found that 61%, 43%, 46% and 83% of the cases expressed SLe^a, SLe^x, STn and Tn in more than a quarter of the tumor area, respectively. The epitopes of SLe^a and SLe^x as well as STn and Tn were statistically coexpressed. We then studied the association between glycan markers and 57 other variables of the patients, including clinical data, markers of genetic screening, family history, surgical data, histopathological

classifications and blood biomarkers. We thereby identified that SLe^a expression is a prerequisite of nerve infiltration of tumor cells and that the absence of SLe^a goes together with *MLK3* mutation and microsatellite instability. Furthermore, an inverse correlation between SLe^x and CA125 became evident. Finally, we describe in this work a completely new negative association between STn and E-cadherin expression. Overall, this work provides a list of novel associations that puts glycan alteration into context with other malignant processes and is a first step to improve their utility as gastric cancer marker for patient stratification in the clinical setting.

1. Introduction

Gastric cancer remains the fifth most common cancer and the third leading cause of cancer death worldwide [1]. The absence of symptoms in early stages and the lack of good diagnostic biomarkers lead to usually late diagnosis of this disease. Surgical resection of the stomach (gastrectomy) remains the best curative attempt to treat gastric cancer [2]. However, the surgical intervention frequently fails to cure patients that present advanced stages. Gastric cancer rapidly spreads to local lymph nodes, invades into adjacent organs and ultimately metastasizes to other peritoneal surfaces and to the liver [2]. Therefore, aggressive therapies including chemotherapy, radiotherapy and targeted therapy are usually being applied [2].

Thus, there is a demand in the clinical setting for prognostic biomarkers that predict disease course, cancer recurrence, molecular tumor subtype, preferred sites of metastasis and way of spread. In addition, biomarkers for treatment stratification are needed to improve personalized medical approaches. These demands could be met by unraveling associations between immunohistological markers and pathological features of the cancer. In general, identification of oncogenic changes that are concomitant with the expression of cancer biomarkers would improve our comprehension of the disease and facilitate future clinical applications.

Glycans represent an ideal type of molecule for biomarker application [3, 4]. Glycosylation is abundantly presented on the cell surface of cells and is altered in the course of carcinogenesis, leading to the expression of abnormal glycan epitopes [5]. These glycoconjugates are commonly shed and secreted into the blood stream where they can be detected by minimal invasive serological assays. Therefore, glycosylation plays a pivotal role as gastric cancer marker, with several assays for glycan epitopes being in clinical use, such as CA 19.9 and CA 72.4 [6]. The CA 19.9 clinical assay,

which evaluates sialyl Lewis A (SLe^a) expression levels, is being used to monitor treatment response and to detect cancer recurrence [7]. The CA 72.4 clinical assay, which evaluates sialyl Tn (STn) expression levels, has prognostic potential for the outcome of gastric cancer patients [8]. In addition, other glycan epitopes such as sialyl lewis X (SLe^x) and Tn are abnormally expressed in gastric cancer [9]. The *de novo* expression of SLe^x and the overexpression of Tn have been associated with poor prognosis of patients [10, 11]. Great progress has been made in recent years in the field of immune therapy by targeting the ectopic expression of the Tn antigen [12].

In this work we aimed to assess the expression profile of the gastric cancer associated glycan markers and their association with patients' clinical data and molecular features of their tumors. For this purpose, we evaluated the epitopes SLe^a, SLe^x, STn and Tn in a series of 80 formalin fixed paraffin embedded gastric carcinomas. The generated antigen expression profiles were used for cutting edge association analysis with KEM® (Knowledge Extraction and Management) using a total of 57 clinicopathological and molecular variables of the patients.

2. Material and methods

2.1 Immunohistochemistry

Formalin fixed paraffin embedded (FFPE) tissue samples from 80 gastric carcinomas were provided by department of surgical oncology of the University of Siena (Italy). All procedures were performed after patients' written informed consent and approved by the local ethical committee. FFPE blocks were cut into 3 µm sections and mounted onto glass slides. The sections were dewaxed, rehydrated and endogenous peroxidases were blocked with 3% hydrogen peroxide (H_2O_2) in methanol. Tissue sections were blocked for 30 min with normal rabbit serum in PBS with 10% BSA. Primary antibodies were incubated overnight at 4°C (Table 1). Biotin-labeled secondary antibodies (Dako, Glostrup, Denmark) were applied for 30 min and the ABC kit (Vector Labs,´Burlingame, CA) for 30 min. In case of the single chain variable fragment (scFv) G2D11 the tissue samples were incubated with a mouse anti-His-tag antibody for 30 min before applying the biotin-labeled antibody. Finally, sections were stained by 3,3′-diaminobenzidine tetrahydrochloride (DAB) and counterstained with Gill's hematoxylin solution. Slides were examined using a Zeiss Optical Microscope. The cases included all gastric cancer subtypes of the Lauren classification and all stages (Table 2).

Table 1. Antibodies used in this study.

Targeted epitope	Antibody	Manufacturer	Antibody type	Dilution	Ref.
SLe^a	CA19.9	Santa Cruz	Mouse IgG 1	1:500	[13]
SLe^x	CSLEX-1	BD Biosciences	Mouse IgM	1:80	[14]
STn	B72.3	/	Mouse IgG 1	1:5	[15]
Tn	G2D11	/	scFv	2.5 µg/ml	

Table 2. Clinical characteristics of gastric cancer cohort of this study.

Gender			
	Female	34 (42%)	
	Male	46 (58%)	
Age			
	Range	30-89	
	Median	72	
Subtype (Lauren)			
	Intestinal	53 (66%)	
	Diffuse	18 (23%)	
	Mixed	5 (6%)	
	Not classified	4 (5%)	
Stage			
	IA	1 (1%)	11 (14%)
	IB	10 (13%)	
	II	3 (4%)	16 (20%)
	IIA	9 (11%)	
	IIB	4 (5%)	
	IIIA	8 (10%)	41 (51%)
	IIIB	12 (15%)	
	IIIC	21 (26%)	
	IV	12 (15%)	12 (15%)

2.2 Co-expression analysis and association rules:

The co-expression of antigens was calculated by χ^2 statistical test. Association rules were established by KEM® (Knowledge Extraction and Management) decision support technology [16]. The proprietary data mining tool KEM® is based on the Formal Concept Analysis (FCA) and founded on the Galois lattices theory [17]. The sets of formal concepts and their connections were visualized in a condensed representation (hierarchical lattice) by extracting non-redundant groups of variables and accounting for missing values in the data [18, 19]. IHC variables were discretized for the KEM® data mining analysis.

The association rule is an implication X→Y, with X being the antecedent and Y the consequent of the rule, that allows inferring X as an explanation of Y. Four main quality measures were used: support, probability, lift and P-value. These quality measures are described below, by considering a dataset composed by N patients with n_X, n_Y and n_{XY} as the numbers of patients satisfying the antecedent X of a rule, the consequent Y or both parts of the rule, respectively. P-values were calculated by using a Fisher exact test. Support, which represents the generality of the rule, was defined as the number of records satisfying both the antecedent and the consequent of the rule: $Support(X \to Y) = n_{XY}$. Probability was defined as the percentage of the characterized patients verifying the rule. It represents the predictive ability of the association rule and is given by: $Probability(X \to Y) = \frac{n_{XY}}{n_X}$. Lift was defined as the ratio of observed support to that expected if X and Y were independent. It measures the performance of a rule to identify a subgroup from a larger population: $Lift\ (X \to Y) = \frac{N \times n_{XY}}{n_X \times n_y}$.

3. Results

3.1 Profiling of Glycosignature

The sialylated antigens SLe^a, SLe^x and STn were evaluated using antigen specific monoclonal antibodies CA19.9, CSLEX-1 and B72.3 respectively (Table 1, 3). The normal mucosas adjacent to tumors were negative for all three sialylated antigens, underlining their absence from the healthy gastric epithelium. Highly inflamed gastric epithelia with high amount of infiltrating immune cells were frequently SLe^x positive. Intestinal metaplasia (IM), in particular the secretory vesicles of goblet cells, showed strong positivity for SLe^a and STn and in some cases sparse staining for SLe^x.

Table 3. Number of cases per category of glycan epitope distribution.

		SLe^a	*SLe^x*	*STn*	*Tn*
Distribution	**Negative**	11	3	1	0
	Rare (Sparse cells)	13	22	16	1
	<25%	7	20	14	12
	25-50%	19	10	14	19
	75-50%	10	8	7	22
	>75%	19	16	5	24
	Total	79	79	57	78

Evaluation of gastric carcinomas revealed that 86% of all tumors were positive for SLe^a (Fig 1) with typically cytoplasmic and membranous staining and often strong positivity for what appears to be secreted proteins. SLe^a positivity involving more than 25% of the tumor area was observed for 61% of all tumors (Fig 1).

Most tumors were positive for SLe^x. Of the 96% SLe^x positive carcinomas, 53% showed staining in minor tumor areas (<25%) or sparse cells and 43% showed area-wide SLe^x expression involving more than 25% of the tumor area (Fig 2). The SLe^x epitope was

typically located at the cytoplasm and membrane and often included what appeared to be secreted proteins present in the extracellular mucus.

The truncated *O*-glycan epitope STn was detected in 98% of all evaluated tumors (Fig 3). Similar to SLe^x, most STn positive cases, namely 53%, showed staining in confined regions of the whole tumor (<25%) or in sparse tumor cells. The remaining 45% cases showed area-wide STn expression involving more than 25% of the tumor area. The subcellular localization of the STn epitope was equal as for SLe^a and SLe^x, including the cytoplasm, the membrane and secreted proteins. Interestingly, among the evaluated sialylated epitopes, STn showed clearly the fewest cases characterized for >75% cancer cell positivity.

The fourth evaluated antigen was the Tn antigen. This truncated *O*-glycan was evaluated by the highly Tn specific single-chain variable fragment (scFv) G2-D11 (unpublished data, manuscript in preparation). The Tn epitope is an intermediate product of the normal *O*-glycan biosynthesis. Epithelial cells of adjacent normal mucosa and IM showed therefore punctuated Tn staining in the perinuclear region, presumably cis-Golgi. Vesicles of goblet cells were negative, proving no cross reactivity of G2-D11 with STn. All assessed gastric carcinomas were to some amount Tn positive and 71% showed overexpression of Tn (Fig 4). Additionally, cancer cells showed deviation from non-malignant epithelial cells in subcellular localization of the Tn antigen. The Tn staining of carcinoma cells was not restricted to perinuclear compartments but was found in the cytoplasm and stretched commonly out over the cellular frontiers including membrane and on what appeared to be secreted proteins.

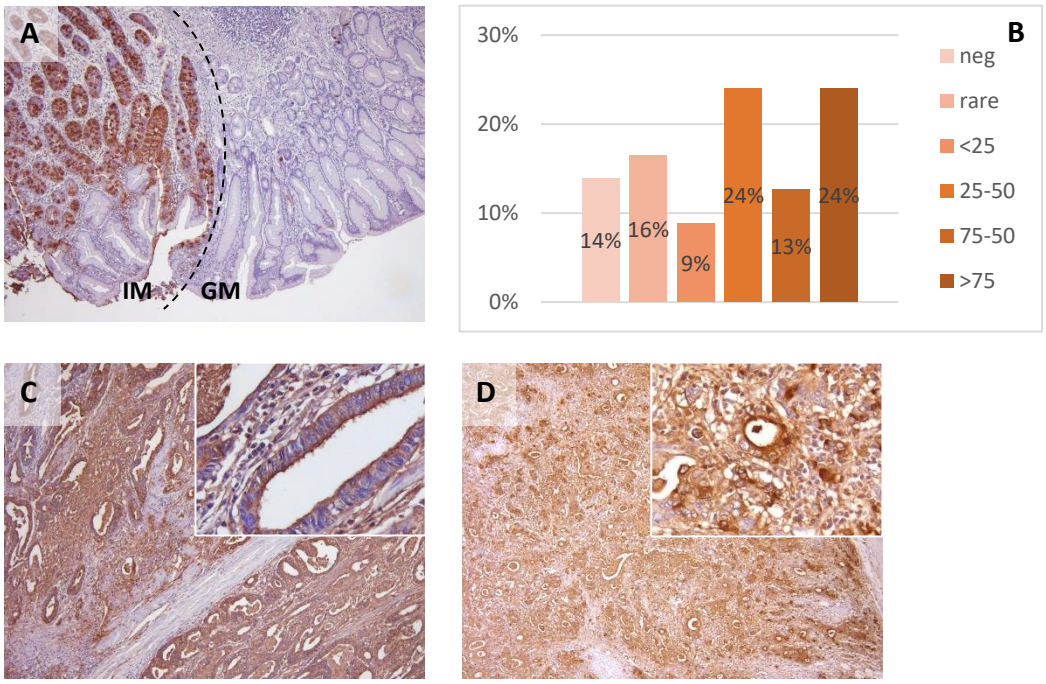


Figure 1. SLe^a profiling of gastric carcinomas.

A. SLe^a staining of intestinal metaplasia (IM) and adjacent gastric mucosa (GM) at 50x magnification. **B.** Diagram of SLe^a staining in gastric carcinomas. The extent of SLe^a positivity was categorized for each tumor into negative (neg) and positive for sparse cells (rare), less than 25% (<25), 25 to 50%, (25-50), 50 to 75% (75-50) and more than 75% (75) tumor area. The bars and their respective percentage show the distribution of the evaluated tumors among these categories. **C. D.** Two representative SLe^a positive carcinomas at 50x and 400x magnification.

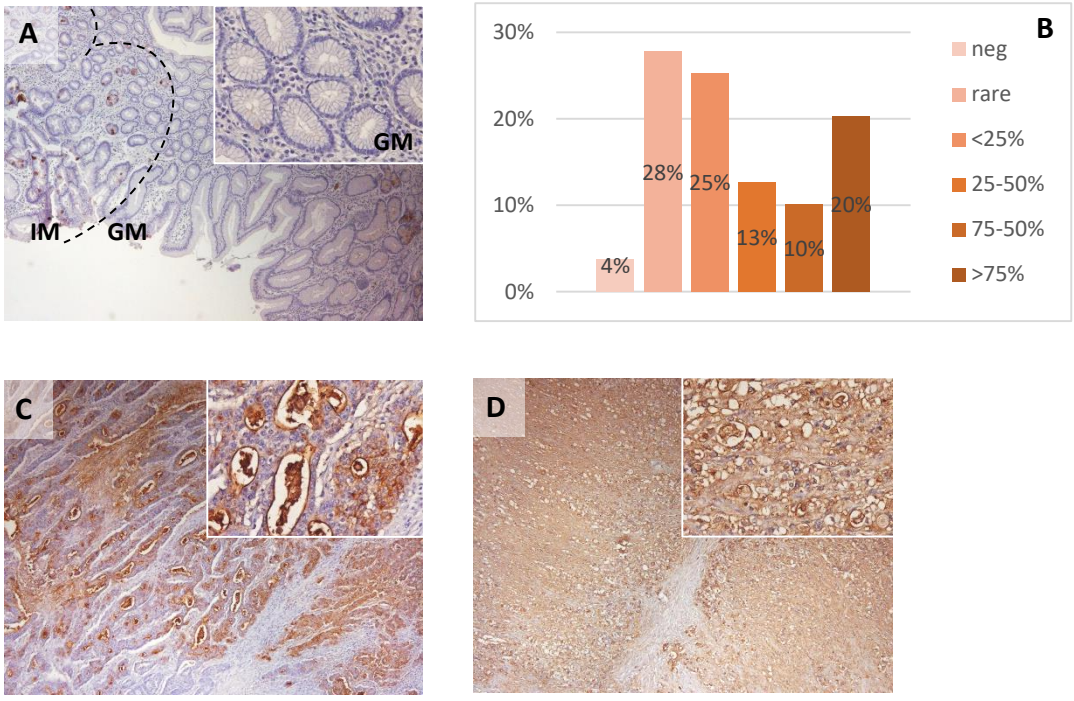


Figure 2. SLe^x profiling of gastric carcinomas.

A. SLe^x staining of intestinal metaplasia (IM) and adjacent gastric mucosa (GM) at 50x magnification and GM also at 200x magnification. **B.** Diagram of SLe^x staining in gastric carcinomas. The extent of SLe^x positivity was categorized for each tumor into negative (neg) and positive for sparse cells (rare), less than 25% (<25), 25 to 50%, (25–50), 50 to 75% (75–50) and more than 75% (75) tumor area. The bars and their respective percentage show the distribution of the evaluated tumors among these categories. **C. D.** Two representative SLe^x positive carcinomas at 50x and 400x magnification.

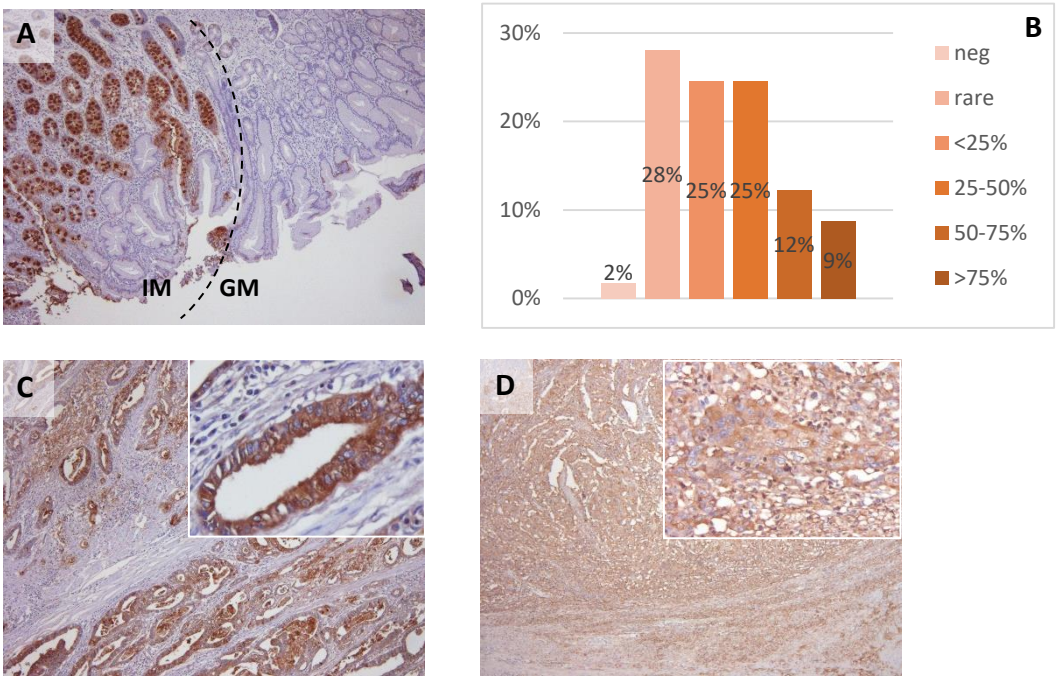


Figure 3. STn profiling of gastric carcinomas.

A. STn staining of intestinal metaplasia (IM) and adjacent gastric mucosa (GM) at 50x magnification. **B.** Diagram of STn staining in gastric carcinomas. The extent of STn positivity was categorized for each tumor into negative (neg) and positive for sparse cells (rare), less than 25% (<25), 25 to 50%, (25-50), 50 to 75% (75-50) and more than 75% (75) tumor area. The bars and their respective percentage show the distribution of the evaluated tumors among these categories. **C. D.** Two representative STn positive carcinomas at 50x and 400x magnification.

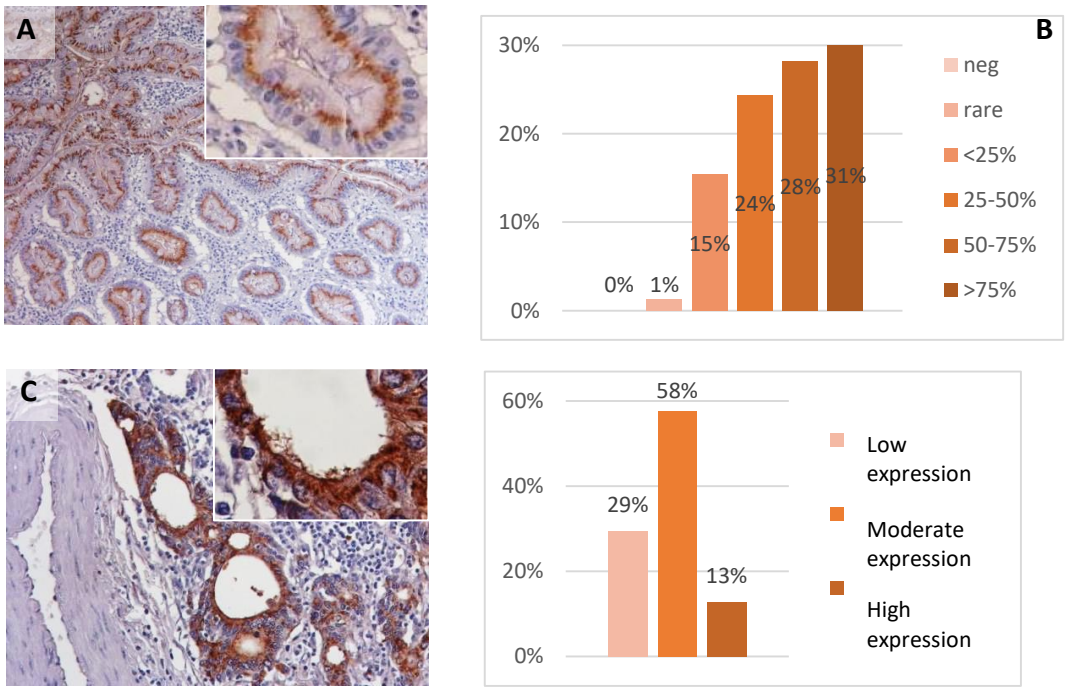


Figure 4. Tn profiling of gastric carcinomas.

A. Tn staining of intestinal metaplasia (IM) at 50x and 400x magnification. **B.** Diagram of Tn staining in gastric carcinomas. The extent of Tn positivity was categorized for each tumor into negative (neg) and positive for sparse cells (rare), less than 25% (<25), 25 to 50%, (25-50), 50 to 75% (75-50) and more than 75% (75) tumor area. The bars and their respective percentage show the distribution of the evaluated tumors among these categories. **C.** A representative Tn positive carcinomas at 100x and 400x magnification. **D.** Diagram of Tn staining intensity in gastric carcinomas, divided into the categories low, moderate and high expression. Cases categorized as low expression resemble similar or weaker staining intensities than that of the adjacent mucosa. Moderate and high expression contain cases that showed stronger Tn staining than normal mucosa does.

3.2 Co-expression of glycomarkers

All four analyzed glycomarkers were overexpressed or *de novo* expressed in gastric carcinomas. Among them, Tn presented the highest number of tumor positivity, but was also the only of the four markers that was present by default in the healthy gastric mucosa. Comparing the *de novo* expressed epitopes, SLe^a was most frequently presenting more than 50% tumor positivity in gastric cancer, followed by SLe^x (Fig 5A). The formation of SLe^a and SLe^x seems to occur partly coordinated in cancer as these two epitopes showed significant association in their expression (Fig 5B). Among the four analyzed epitopes STn had the least cases with more than 50% tumor positivity and Tn the most. Remarkably, a highly significant association was still evident for STn and Tn (Fig 5C). Especially, STn positivity seems to strongly imply Tn positivity. Out of 14 cases with over 50% STn positivity only a single case did not also show over 50% Tn positivity, but was 25-50% positive for Tn. These associations are probably based on overlapping molecular mechanisms that can lead to the upregulation of SLe^a and SLe^x or Tn and STn. No other marker correlations reached significance.

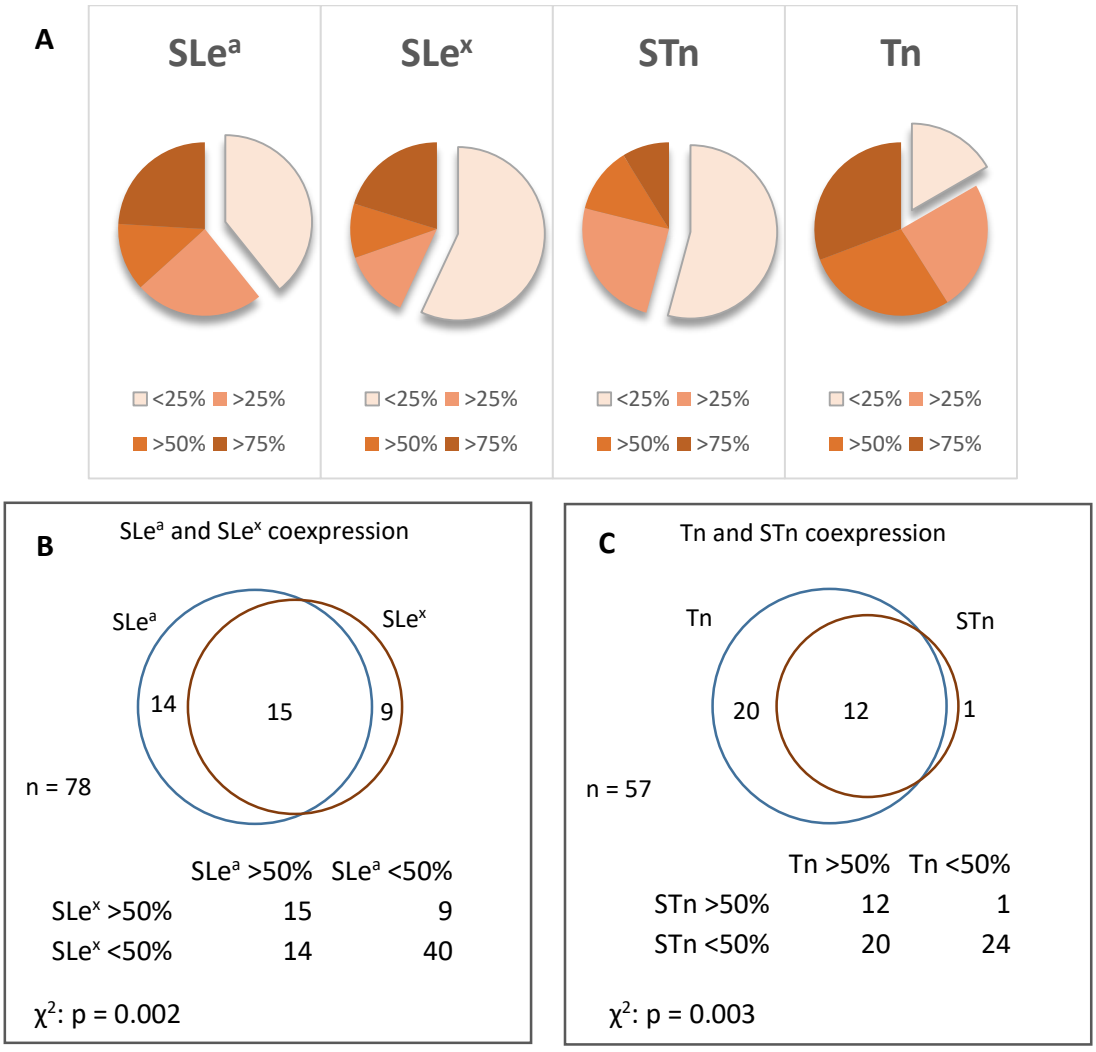
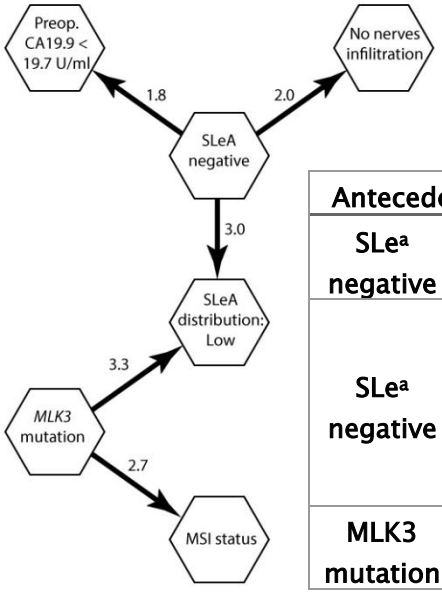


Figure 5. Distribution of glycan marker positivity and associations among them. **A.** Pie-diagram of SLe^a, SLe^x, STn and Tn positivity in quartiles. **B and C.** Association analysis between SLe^a and SLe^x or Tn and STn distribution with cut-off of 50%. On the top, Venn diagram of cases with >50% positivity. On the bottom, χ^2 statistical test with p-value.

3.3 Glycomarker associate with various clinical variables

We evaluated the association of SLe^a, SLe^x, STn and Tn expression with 57 other variables of the patients, including clinicopathological data, markers of genetic screening, family history, surgical data and serological biomarkers.

The absence of SLe^a expression in the tumor was significantly associated with low level of SLe^a in patients' sera (Fig 6). This association was according to expectations but demonstrates the capability of our approach to highlight causalities in this cohort. Further, SLe^a negative tumors associated with the absence of nerve infiltration of tumor cells. This association was highly significant and supported by 8 out of 8 SLe^a negative cases. Lastly, we identified a statistical significant association of *MLK3* mutation with tumors that presented absence or low SLe^a expression in sparse tumor cells.



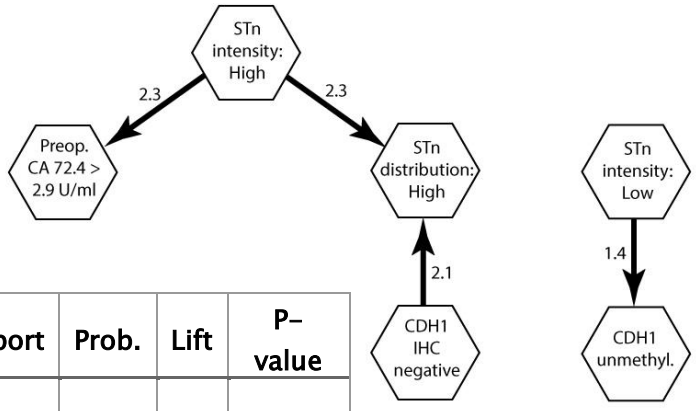
Antecedent		Consequent		Support	Prob.	Lift	P-value
SLe^a negative	8	No nerve infiltration	25	8	100%	2.0	0.002
SLe^a negative	9	Preoperative CA19.9 serum levels lower than 19.7 U/ml	35	9	100%	1.8	0.003
MLK3 mutation	3	SLe^a negative/rare	16	3	100%	3.3	0.025

Figure 6. Association rules of SLe^a.

Association map connecting SLe^a negativity, low distribution of SLe^a (contains cases that were classified as negative or rare), absence of nerve infiltration, preoperative CA19.9 serum levels lower than 19.7 U/ml, presence of *MLK3* mutation and microsatellite instability (MSI status). The numerical evaluations are shown in the table, containing number of cases with antecedent and consequent variable, support, probability (prob.), lift and Fisher's exact test p-value. Arrow and number display the direction and the Lift value of the association, respectively.

Regarding STn expression, we were able to identify 3 significant association rules (Fig 7). High intensity of STn staining at the tumor site was associated with greater than 2.9 U/ml of the CA 72.4 serological assay. This is according to expectation as CA 72.4 assay evaluates serum levels of STn. More interesting, the low intensity of STn entailed in our cohort that the *CDH1* gene was unmethylated. *CDH1* methylation is one of the mechanisms leading to downregulation of E-cadherin (the protein encoded by CDH1) expression in the context of gastric cancer [20]. In support of this finding we identified that all E-cadherin negative tumors expressed high levels of STn. This strongly indicates a negative correlation between E-cadherin and STn expression in gastric cancer.

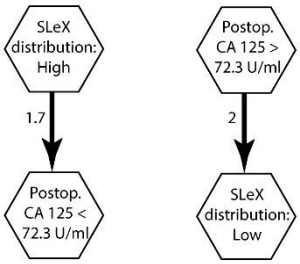
Tumors with few SLe^x positive cancer cells (<25%) associated with postoperative high CA 125 serum levels (greater than 72.3 U/ml) (Fig 8). CA 125 assay quantifies the mucin MUC16. Reciprocally, tumors with more than >25% SLe^x positive cancer cells associated with postoperative CA 125 serum levels lower than 72.3 U/ml.



Antecedent		Consequent		Support	Prob.	Lift	P-value
STn intensity low	18	*CDH1* gene unmethylated	32	18	100%	1.4	0.0005
CDH1 negative	5	STn distribution high: >25% positive	20	5	100%	2.1	0.018
STn intensity high	4	Preoperative CA 72.4 serum level greater than 2.9 U/ml	18	4	100%	2.3	0.027

Figure 7. Association rules of STn.

A. Association maps with STn distribution high (positivitiy in more than 25% of the tumor), high intensity of STn tumor staining, low intensity of STn tumor staining, CDH1 negativity, *CDH1* unmethylated and preoperative CA 72.4 serum level greater than 2.9 U/ml. The numerical evaluations are shown in the table, containing number of cases with antecedent and consequent variable, support, probability (prob.), lift and Fisher's exact test p-value. Arrow and number display the direction and the Lift value of the association, respectively.



Antecedent		Consequent		Support	Prob.	Lift	P-value
Postoperative CA 125 serum level greater than 72.3 U/ml	4	SLe^x distribution low: <25% positive	5	4	100%	2	0.024
SLe^x distribution high: >25% positive	5	Postoperative CA 125 serum level lower than 72.3 U/ml	6	5	100%	1.7	0.024

Figure 8. Association rules of SLe^x.

A. Association maps with SLe^x distribution high (positivitiy in more than 25% of the tumor), SLe^x distribution low (positivitiy in less than 25% of the tumor), postoperative CA125 serum level greater than 72.3 U/ml and postoperative CA125 serum level lower than 72.3 U/ml. The numerical evaluations are shown in the table, containing number of cases with antecedent and consequent variable, support, probability (prob.), lift and Fisher's exact test p-value. Arrow and number display the direction and the Lift value of the association, respectively.

4. Discussion

Gastric cancer remains the third leading cause of cancer death worldwide, owing this placement predominantly to the absence of symptoms in early stages and the lack of good biomarkers for diagnosis, treatment stratification and prognosis. Many potential cancer biomarkers were primarily tested on their diagnostic aptitude and their ability to predict patients' outcome. Most of those markers ultimately do not meet the clinical criteria, because of limitations in specificity or sensitivity, despite being clearly associated with gastric cancer disease severity. These markers might be associated with clinicopathological features of the tumor that have not been elucidated yet. Moreover, the identification of associations between markers can identify causally connected alteration networks and thereby contribute to our understanding of tumor cell biology.

In this work we evaluated the expression of four cancer associated glycan markers, namely SLe^a, SLe^x, STn and Tn, in a large series of gastric carcinomas. We characterized the distribution and the intensity of these epitopes within the tumor and evaluated their pattern of expression. As such Tn showed the highest number of positive cases, but is the only of the evaluated markers that is also expressed in healthy epithelia. A clear increase in staining and deviation of the healthy subcellular localization including membrane and extracellular space was evident in cancer. Therefore, Tn bares great potential for immunotherapy [21]. SLe^a was among the *de novo* expressed epitopes with 61% the most commonly expressed, considering cases with positivity in more than 25% of the tumor area. Under the same criteria, SLe^x and STn were expressed in 43% and 46% of the evaluated gastric carcinomas, respectively. These evaluations are in accordance to previous studies [22, 23].

We were able to identify a significant association between STn and Tn expression. In particular, cases that expressed STn in more than half of the tumor area were in 12 out of 13 cases also Tn positive in more than half of the tumor area. This supports that Tn accumulation might be a common mechanism for STn overexpression in gastric cancer. The two epitopes have previously been described to frequently occur concomitantly because of some common mechanisms of overexpression [24]. However, the upregulation of ST6GALNAC1 is a Tn independent mechanism of STn overexpression in gastric cancer [25, 26]. We also found that SLe^a and SLe^x expression was significantly associated. These two sialofucosylated epitopes are structurally similar and have biosynthetically overlapping pathways [27].

The high percentage of gastric tumors that acquires SLe^a, SLe^x and STn expression foreshadows that these epitopes might convey a positive selection pressure when expressed. The association of SLe^a, SLe^x and STn with clinical features supports this hypothesis [7, 8]. Similarly, the ectopic expression of Tn throughout the cytoplasm of virtually all gastric carcinoma cells might be accompanied by an advantageous phenotype.

The second aim of this work was to test SLe^a, SLe^x, STn and Tn for predictive potential. We performed multivariate analysis comparing the expression these glycan markers with a data set of 57 other variables for each patient. This data set contained 14 categories with patient information (e.g. gender, age and blood groups), 12 variables on markers of genetic screening, 19 variables with surgical and pathological evaluations and information of pre- and postoperative blood biomarker levels. With our approach we anticipated to discover novel associations for two reasons: (i) use of unprecedented large number of patient specific information to associate our analysis with; (ii) employment of patented algorithm that can fully explore complex datasets in order to reveal hidden associations.

The expression of the STn epitope associated through two independent association rules with the loss of E-cadherin expression (Fig 7). E-cadherin is an important cell-cell adhesion molecule and tumor suppressor gene [28]. Loss of function of the E-cadherin gene, *CDH1*, is a common event in carcinogenesis and predominantly linked to diffuse type gastric cancer [20]. Hypermethylation of the *CDH1* promoter region is a frequent cause of E-cadherin repression [29, 30]. Whether the association between STn expression and E-cadherin absence is rooted in a direct causality or these two events have a common underlying mechanism is not known. To our knowledge, this is the first time that these two epitopes have been inversely correlated.

The inverse correlation between SLe^x and CA 125 remains elusive as well. The serological CA 125 assay determines the MUC16 concentration in the blood of patients [31]. Elevated MUC16 serum levels are typical in advanced gastric cancer, where it is a marker of poor prognosis [32]. MUC16, as the largest member of the mucin family, is carrier of numerous *O*-glycosylation sites [33]. It is therefore surprising that the expression of potential major carrier of SLe^x inversely correlates with this modification. Further studies are needed to unravel this phenomenon.

We revealed a novel association between SLe^a and nerve infiltration (Fig 6). Perineural invasion is a key pathological feature of gastric cancer [34] and leads to significantly worse prognosis for the patients [35]. Perineural spread impedes complete tumor removal because cancer cells stretch across safe tumor margins and hence, optimal resection in such cases is rare [36]. In the so called neural tracking, tumor cells actively migrate along the perineural space and conduct extensive crosstalk with cells of the neural niche [36, 37]. Our finding that SLe^a expression might be a prerequisite for neural tracking might explain SLe^a function in cancer invasion and progression and its commonly described association with poor prognosis [38].

Moreover, we identified an indirect association between absence of SLe^a and microsatellite instability (MSI) via *MLK3* mutation (Fig 6). The applied screening method only allowed associations with 100% probability. This explains the absence of the direct association between the two common events SLe^a negativity and MSI. MSI is one of the four molecular subgroups of gastric cancer suggested by the The Cancer Genome Atlas (TCGA) [39] and is linked to favorable prognosis of the patients [40]. The hallmark of MSI tumors is a dysfunctional DNA mismatch repair system leading to high mutation rates and increase in tandem oligonucleotide repeats [41]. These cancers are particularly immunogenic and are therefore accompanied by high amount of inflammation and infiltrative immune cells. Additionally, MSI cancers accumulate activating hot spot mutations at receptor tyrosine kinases such EGFR, HER2 and HER3 [39]. SLe^a negative tumors might therefore be particularly susceptible for therapies targeting these receptors.

Overall, this work has established a set of novel associations between glycan markers and tumor features. These linkages provide the basis for new lines of research and ultimately bare the potential to improve the utility of the gastric cancer marker SLe^a, SLe^x, STn and Tn for patient stratification in the clinical setting.

5. References

1 Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., . . . Bray, F. (2015) Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 136, E359-386

2 Van Cutsem, E., Sagaert, X., Topal, B., Haustermans, K. and Prenen, H. (2016) Gastric cancer. Lancet

3 Adamczyk, B., Tharmalingam, T. and Rudd, P. M. (2012) Glycans as cancer biomarkers. Biochimica et biophysica acta. 1820, 1347-1353

4 Reis, C. A., Osorio, H., Silva, L., Gomes, C. and David, L. (2010) Alterations in glycosylation as biomarkers for cancer detection. J Clin Pathol. 63, 322-329

5 Pinho, S. S. and Reis, C. A. (2015) Glycosylation in cancer: mechanisms and clinical implications. Nat Rev Cancer. 15, 540-555

6 Marrelli, D., Roviello, F., De Stefano, A., Farnetani, M., Garosi, L., et al. (1999) Prognostic significance of CEA, CA 19-9 and CA 72-4 preoperative serum levels in gastric carcinoma. Oncology. 57, 55-62

7 Marrelli, D., Pinto, E., De Stefano, A., Farnetani, M., Garosi, L., et al. (2001) Clinical utility of CEA, CA 19-9, and CA 72-4 in the follow-up of patients with resectable gastric cancer. American journal of surgery. 181, 16-19

8 Ucar, E., Semerci, E., Ustun, H., Yetim, T., Huzmeli, C., et al. (2008) Prognostic value of preoperative CEA, CA 19-9, CA 72-4, and AFP levels in gastric cancer. Advances in therapy. 25, 1075-1084

9 Mereiter, S., Balmana, M., Gomes, J., Magalhaes, A. and Reis, C. A. (2016) Glycomic Approaches for the Discovery of Targets in Gastrointestinal Cancer. Front Oncol. 6, 55

10 Amado, M., Carneiro, F., Seixas, M., Clausen, H. and Sobrinho-Simoes, M. (1998) Dimeric sialyl-Le(x) expression in gastric carcinoma correlates with venous invasion and poor outcome. Gastroenterology. 114, 462-470

11 Springer, G. F. (1997) Immunoreactive T and Tn epitopes in cancer diagnosis, prognosis, and immunotherapy. J Mol Med (Berl). 75, 594-602

12 Kulasingam, V. and Diamandis, E. P. (2008) Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. Nat Clin Pract Oncol. 5, 588-599

13 Magnani, J. L., Steplewski, Z., Koprowski, H. and Ginsburg, V. (1983) Identification of the gastrointestinal and pancreatic cancer-associated antigen detected by monoclonal antibody 19-9 in the sera of patients as a mucin. Cancer Res. 43, 5489-5492

14 Fukushima, K., Hirota, M., Terasaki, P. I., Wakisaka, A., Togashi, H., Chia, D., . . . Hakomori, S. (1984) Characterization of sialosylated Lewisx as a new tumor-associated antigen. Cancer Res. 44, 5279-5285

15 Thor, A., Ohuchi, N., Szpak, C. A., Johnston, W. W. and Schlom, J. (1986) Distribution of oncofetal antigen tumor-associated glycoprotein-72 defined by monoclonal antibody B72.3. Cancer Res. 46, 3118-3124

16 Moremen, K. W., Tiemeyer, M. and Nairn, A. V. (2012) Vertebrate protein glycosylation: diversity, synthesis and function. Nat Rev Mol Cell Biol. 13, 448-462

17 Ganter, B. and Wille, R. (2012) Formal concept analysis: mathematical foundations. Springer Science & Business Media

18 Valtchev, P., Missaoui, R. and Godin, R. (2004) Formal concept analysis for knowledge discovery and data mining: The new challenges. In International Conference on Formal Concept Analysis ed.)^eds.). pp. 352-371, Springer

19 Lakhal, L. and Stumme, G. (2005) Efficient mining of association rules based on formal concept analysis. In Formal concept analysis. pp. 180-195, Springer

20 Graziano, F., Humar, B. and Guilford, P. (2003) The role of the E-cadherin gene (CDH1) in diffuse gastric cancer susceptibility: from the laboratory to clinical practice. Ann Oncol. 14, 1705-1713

21 Posey, A. D., Jr., Schwab, R. D., Boesteanu, A. C., Steentoft, C., Mandel, U., Engels, B., . . . June, C. H. (2016) Engineered CAR T Cells Targeting the Cancer-Associated Tn-Glycoform of the Membrane Mucin MUC1 Control Adenocarcinoma. Immunity. 44, 1444-1454

22 Filella, X., Fuster, J., Molina, R., Grau, J. J., Garcia-Valdecasas, J. C., et al. (1994) TAG-72, CA 19.9 and CEA as tumor markers in gastric cancer. Acta oncologica. 33, 747-751

23 Ju, T., Lanneau, G. S., Gautam, T., Wang, Y., Xia, B., et al. (2008) Human tumor antigens Tn and sialyl Tn arise from mutations in Cosmc. Cancer research. 68, 1636-1646

24 Ju, T., Lanneau, G. S., Gautam, T., Wang, Y., Xia, B., Stowell, S. R., . . . Cummings, R. D. (2008) Human tumor antigens Tn and sialyl Tn arise from mutations in Cosmc. Cancer Res. 68, 1636-1646

25 Marcos, N. T., Bennett, E. P., Gomes, J., Magalhaes, A., Gomes, C., David, L., . . . Reis, C. A. (2011) ST6GalNAc-I controls expression of sialyl-Tn antigen in gastrointestinal tissues. Front Biosci (Elite Ed). 3, 1443-1455

26 Marcos, N. T., Pinho, S., Grandela, C., Cruz, A., Samyn-Petit, B., Harduin-Lepers, A., . . . Reis, C. A. (2004) Role of the human ST6GalNAc-I and ST6GalNAc-II in the synthesis of the cancer-associated sialyl-Tn antigen. Cancer Res. 64, 7050-7057

27 Harduin-Lepers, A., Vallejo-Ruiz, V., Krzewinski-Recchi, M. A., Samyn-Petit, B., Julien, S. and Delannoy, P. (2001) The human sialyltransferase family. Biochimie. 83, 727-737

28 Carneiro, P., Fernandes, M. S., Figueiredo, J., Caldeira, J., Carvalho, J., Pinheiro, H., . . . Seruca, R. (2012) E-cadherin dysfunction in gastric cancer--cellular consequences, clinical applications and open questions. FEBS Lett. 586, 2981-2989

29 Tamura, G., Yin, J., Wang, S., Fleisher, A. S., Zou, T., Abraham, J. M., . . . Meltzer, S. J. (2000) E-Cadherin gene promoter hypermethylation in primary human gastric carcinomas. J Natl Cancer Inst. 92, 569-573

30 Corso, G., Carvalho, J., Marrelli, D., Vindigni, C., Carvalho, B., Seruca, R., . . . Oliveira, C. (2013) Somatic mutations and deletions of the E-cadherin gene predict poor survival of patients with gastric cancer. J Clin Oncol. 31, 868-875

31 Yin, B. W. and Lloyd, K. O. (2001) Molecular cloning of the CA125 ovarian cancer antigen: identification as a new mucin, MUC16. J Biol Chem. 276, 27371-27375

32 Webb, A., Scott-Mackie, P., Cunningham, D., Norman, A., Andreyev, J., O'Brien, M. and Bensted, J. (1996) The prognostic value of serum and immunohistochemical tumour markers in advanced gastric cancer. Eur J Cancer. 32A, 63-68

33 Hanson, R. L. and Hollingsworth, M. A. (2016) Functional Consequences of Differential O-glycosylation of MUC1, MUC4, and MUC16 (Downstream Effects on Signaling). Biomolecules. 6

34 Liebig, C., Ayala, G., Wilks, J. A., Berger, D. H. and Albo, D. (2009) Perineural invasion in cancer: a review of the literature. Cancer. 115, 3379-3391

35 Bilici, A., Seker, M., Ustaalioglu, B. B., Kefeli, U., Yildirim, E., Yavuzer, D., . . . Gumus, M. (2010) Prognostic significance of perineural invasion in patients with gastric cancer who underwent curative resection. Ann Surg Oncol. 17, 2037-2044

36 Amit, M., Na'ara, S. and Gil, Z. (2016) Mechanisms of cancer dissemination along nerves. Nat Rev Cancer. 16, 399-408

37 Jobling, P., Pundavela, J., Oliveira, S. M., Roselli, S., Walker, M. M. and Hondermarck, H. (2015) Nerve-Cancer Cell Cross-talk: A Novel Promoter of Tumor Progression. Cancer Res. 75, 1777-1781

38 Kochi, M., Fujii, M., Kanamori, N., Kaiga, T., Kawakami, T., Aizaki, K., . . . Yamagata, M. (2000) Evaluation of serum CEA and CA19-9 levels as prognostic factors in patients with gastric cancer. Gastric Cancer. 3, 177-186

39 Cancer Genome Atlas Research, N. (2014) Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 513, 202-209

40 Falchetti, M., Saieva, C., Lupi, R., Masala, G., Rizzolo, P., Zanna, I., . . . Ottini, L. (2008) Gastric cancer with high-level microsatellite instability: target gene mutations, clinicopathologic features, and long-term survival. Hum Pathol. 39, 925-932

41 Thibodeau, S. N., French, A. J., Roche, P. C., Cunningham, J. M., Tester, D. J., Lindor, N. M., . . . Halling, K. C. (1996) Altered expression of hMSH2 and hMLH1 in tumors with microsatellite instability and genetic alterations in mismatch repair genes. Cancer Res. 56, 4836-4840

CHAPTER VI

GENERAL DISCUSSION

CONCLUDING REMARKS

1. GENERAL DISCUSSION

Cancer is among the leading causes of death worldwide, affecting every geographic region [1]. While many other disease burden are constantly decreasing, each year the number of cancer deaths is ever-expanding [2]. Cancer is for several reasons an extraordinarily difficult disease to diagnose and treat: (i) every patients cancer is unique in its accumulated genotypic and phenotypic alterations [3], (ii) a tumor is intrinsically diverse and heterogeneous, composed of cancer cell subpopulations that exhibit different treatment susceptibilities and resistances [4], (iii) as cancer cells are derived from body cells, they display mostly self-antigens which makes both their antigen-based detection and targeted treatment challenging [5, 6], (iv) the pursuit of cancer cells to spread beyond the location of origin turns them eventually into systemic diseases, impeding curative surgical interventions [7].

In order to tackle a malady as complex as cancer that exploits all cellular mechanisms to gain growth advantage, it is necessary to understand all molecular aspects of cancer. One of the most intricate processes that is altered in the context of cancer is protein glycosylation [8, 9].

Protein glycosylation is a highly regulated process that involves the interplay of glycosyltransferases and glycosidases in the assembly of glycan structures [10]. This process requires additionally a well-functioning cell metabolism and enzymatic cascades in the formation of glycan donor substrates and transporters to localize these substrates in the right compartments [11]. A process that incorporates so many pathways is easily altered by molecular and cellular changes that occur in diseases. In cancer, several specific glycosylation alterations such as the *de novo* expression of cancer associated glycan epitopes can frequently be observed [9]. Deciphering the cause of these alterations is challenging as they may stem from various changes in

the underlying biosynthetic pathways. Additionally, glycosyltransferases and glycosidases are often involved in the buildup and trimming of numerous different glycan structures. This makes investigations on specific glycan epitopes in cell models complicated given that the modulation of the biosynthetically involved enzymes likewise induces collateral changes on many other structures.

In this thesis we aimed to broadly tackle the role of glycan alterations in gastric cancer. We focused on the oncogenic expression of sialofucosylated epitopes and truncation of *O*-glycan structures, and sought to understand the impact of these alterations at glycoprotein, cancer cell and tumor level. Three main approaches were chosen to achieve this general aim: (i) mimicking different malignant glycophentoypes using glycoengineered cell models to unravel concomitantly appearing glycan alterations and to identify altered glycoprotein targets; (ii) dissecting the impact of altered glycosylation on selected glycoproteins that are key players of carcinogenesis and cancer progression (iii) analyzing associations between glycan alterations and tumor features in a cohort of patients' carcinomas.

1.1 Cell models as tools to explore cancer associated glycan changes

We studied in chapter 2 and 3 the effect of increased SLe^x and STn formation by modulating the expression of key proteins of their biosynthesis. For this approach we used the MKN45 gastric carcinoma cell line, which has previously been demonstrated to be SLe^x and STn negative [12]. The MKN45 cells have been stably transfected with an expression vector containing the full length *ST3GAL4* or *ST6GALNAC1* gene controlled by constitutively active enhancer-promoter. The α2,3-sialyltransferase ST3GAL4 has been selected as it is one of the major enzymes involved in the SLe^x formation [13] and has been described to be commonly overexpressed in gastric cancer [14, 15]. ST6GALNAC1 on the other hand is an α2,6-sialyltransferase involved

in the formation of STn [16]. ST6GALNAC1 overexpression has previously been associated with a more aggressive cancer cell phenotype [17, 18]. ST3GAL4 and ST6GALNAC1 overexpression led in our model to the increased formation of SLe^x and STn, respectively. To understand the full scope of the induced alterations by ST3GAL4 or ST6GALNAC1, we performed *N*– and *O*–glycomic analyses of these cell lines. Indeed, we were able to reveal systemic changes on the glycome. As such, ST3GAL4 directly modulated numerous *N*– and *O*–glycan acceptor structures. On the other hand, ST6GALNAC1 has a more restricted role, as it predominantly sialylates the Tn epitope. Both enzymes showed, additionally to the increased synthesis of their products, further glycomic changes presumably by altering the availability of acceptors and donors for alternative glycosylation pathways. Interestingly, the overexpression of ST3GAL4 thereby induced simultaneously several cancer associated glycan alterations that have been described independently in previous publications [19-21], suggesting that these changes might arise, at least in part, coordinately in cancer.

Many of the glycan changes in cancer that have been associated with clinical implications for the patients were based on antibody recognition [9, 22]. This leaves the possibility that these immunogenic epitopes, that gave rise to antibodies, merely act as flags for other glycan alterations that could be even more functionally relevant and that arise concomitantly.

For the glycomic analysis we chose two complementary strategies of structural analysis based on ESI-MS/MS and a more quantitative approach by UPLC-FLD, both leading to the same results demonstrating the suitability and reliability of these techniques. Regarding the structural analysis, we used the extracted ion chromatograms of the ESI-MS in negative mode for relative quantification. This approach does not allow to draw conclusion on abundance of individual glycan structures for several reasons: (i) the porous graphitized carbon (PGC) separation column tends to poorly recover large oligosaccharides, which might introduce a

structural bias [23]; (ii) the negative mode ionization used bares advantages for the detection of negatively charged structures and the formation of diagnostic cross ring fragmentations, but fails producing realistic proportions [24, 25]; (iii) the ionization of glycans is in general inefficient and co-eluting molecules might impinge strong ion suppression effects [23]. Therefore, the results taken from the PGC-ESI-MS/MS analysis were exclusively comparative between relative abundances of structures from samples that were prepared and analyzed in parallel.

In the *N*- and *O*-glycomic analysis of ST3GAL4 overexpressing cells we were not able to detect major differences in the relative abundance of sialofucosylated structures. This was inconsistent with our findings that these cells show a significant increase in KM93 (anti-SLe^x antibody) binding. We confirmed the increased expression of SLe^x by using another SLe^x binding antibody, CSLEX1. All experiments provided the same results with CSLEX1 as they did with KM93. False positivity due to antibody cross-reactivity is therefore unlikely. We hypothesize that the increase in SLe^x could not be reproduced by PGC-ESI-MS/MS because the epitope is found on large *N*-glycans in our cell line model. We were not able to obtain structural information on large *N*-glycans, such as tri- and tetra-antennary structures with our analysis. Especially β1,6 GlcNAcs have been described to a be common carrier of SLe^x [26]. We have demonstrated a remarkable decrease of bisected *N*-glycans in ST3GAL4 overexpressing cells that may consequently lead to an increase in β1,6 branching, as bisecting GlcNAc has been shown to preclude the β1,6 GlcNAc [27]. In fact, the ST3GAL4 overexpressing cells evidently show a strong increase in very large α2,3 sialylated *N*-glycans as detected by HILIC-FLD. Taking all these findings into account we conclude that β1,6 branched tri- and tetra-antennary *N*-glycans are probable carriers of SLe^x in the ST3GAL4 overexpressing cell line model.

Considering that the analysis of a gastric carcinoma cell line overexpressing ST3GAL4 revealed several specific glycan alterations independent of SLe^x, it is disputable

whether SLe^x itself is in this cell line model of biological relevance. In the tumor, SLe^x expression has previously been associated with patients' poor outcome and was shown to enhance the metastatic potential mediated through direct interaction of SLe^x and selectins at the vascular endothelium [28, 29]. The SLe^x epitope bares additional the advantage in most experimental set-ups of being specifically detectable by antibodies, such as KM93 and CSLEX1. The α2,3 sialylation of type 2 chains catalyzed by ST3GAL4 is an essential part of the SLe^x epitope [30]. We therefore used the anti-SLe^x antibodies as tools to detect α2,3 sialylation of type 2 chains and to detect their co-expression with RON and CD44 in chapter 3 and 4. We further suspect that the expression of SLe^x in tumor samples may imply other alterations, described in chapter 2 and 3, such as the increased expression of sialylated short core 2 *O*-glycans, the decrease of bisected structures, and the decrease of α2,6 sialylated *N*-glycans. The association of these glycan alterations remains to be shown in other glyco-engineered cell line models and in gastric tumor specimens.

1.2 Identification of altered glycosylated protein targets

Virtually all proteins that pass through the secretory pathway are glycosylated. The alteration of glycan structures affects therefore a multitude of proteins simultaneously, complicating the association of these alterations with concrete functions. Hence, our understanding on how aberrant glycosylation conveys advantageous phenotypes to cancer cells is limited. In chapter 3 we unraveled critical protein targets whose altered glycosylation may favor malignancy. We achieved this by a quantitative glycoproteomic analysis of gastric carcinoma cells that upon ST3GAL4 overexpression exhibited a more invasive phenotype [31]. This method quantitatively assessed peptides that were modified with sialylated *N*-glycans, which means we were not able to evaluate proteins that were carriers of altered *O*-

glycosylation or sialylation independent altered *N*-glycosylation. In addition, regarding sialylated *N*-glycans, we demonstrated in chapter 2 by quantitative HILIC-FLD-UPLC that the major alteration induced by ST3GAL4 overexpression was rather the change of sialic acid linkage than an overall increase of sialylation. Therefore, the applied method probably identified only a subset of all aberrantly glycosylated proteins. Moreover, only unique peptides are unambiguously assigned to a protein, which has been previously described for this method to be around 15% of all identified peptides in a human cell line [32].

Despite the aforementioned limitations, we were able to identify 67 *N*-glycosylation sites in 47 glycoproteins with significantly increased sialylation in cells overexpressing ST3GAL4. Remarkably, among them, many key players of malignancy such as Insulin receptor, RON, Integrin αV, Integrin α3, Integrin β6, Carcinoembryonic antigen-related cell adhesion molecule 1 (CEACAM1), CEACAM5, Plexin B2 and Receptor-type tyrosine-protein phosphatase eta (PTPRJ). Altering the glycosylation of these proteins bares the potential of modulating their interactions with other proteins, their stability or their cellular localization, and thereby changing the malignant phenotype of the cell.

Our experiments on RON confirmed the biological relevance of the glycosylation by showing increased activation of this receptor tyrosine kinase in the model with elevated ST3GAL4 levels. Although the other identified protein targets remain to be analysed for functional implications, it is expected that they are also to a certain degree affected. For instance, integrins function has been shown to be critically regulated through their glycosylation [33-35].

Independent of the functional impact, the list of site-specific altered glycosylation is a valuable register for potential biomarkers. These glycosylation sites might be particularly sensitive in being differentially sialylated when expression levels of

sialyltransferases change. The increase in sialylation is prevalent in cancer [9] and might be faithfully presented on these protein *N*-glycan sites. The protein CEACAM5 (known as CEA) has to be highlighted in this regard. We identified in chapter 2 and 3 the *N*-glycosylation sites N204 and N208 of CEA being extensively sialylated. Elevated CEA levels in serum are used in the clinical setting as a marker for various cancers, including gastric cancer [36]. The CEA assay lacks sensitivity with elevated levels found in only around 25% of all gastric cancer patients [37]. It should be tested whether specific evaluation of N204 and N208 *N*-glycan sialylation increases the diagnostic potential of CEA. These glycan sites can be assessed by glycoproteomics as tryptic digest generates unique peptide sequences around these glycosylation sites.

1.3 RON and CD44: a sweet relationship

The increased activation of the RON receptor tyrosine kinase was shown in chapter 3 with three complementary methods in the cell line model. Ligand-independent aberrant RON activation in cancer has been mainly attributed to receptor overexpression [38, 39], expression of constitutively activated splice variants and protein truncation [40, 41]. RON activation results in the autophosphorylation of Tyr1238 and Tyr1239 in the kinase domain [42]. Based on these phosphorylation sites, we evaluated in chapter 3 the activation of RON. Phosphorylation of the regulatory residues activates the kinase activity, leading to further phosphorylation of Tyr1353 and Tyr1360 and initiation of downstream signaling [42, 43]. This event is classically mediated by RAS-ERK and PI3K-AKT pathways [44], which are responsible for RON-mediated cell proliferation and survival, as well as migration and invasion, angiogenesis and chemoresistance [45-48]. RON actively crosstalks with other RTKs including MET, EGFR, insulin-like growth factor receptors and platelet-derived growth

factor receptor (PDGFR) but also viral oncoproteins derived from Epstein-Barr virus (EBV) [49].

In tissue the observation of the receptor activation is more difficult since phospho-tyrosines are labile moieties, readily removed by endogenous phosphatases and lost over time during samples storage [50]. We found the phospho-RON in formalin fixated paraffin embedded (FFPE) gastric carcinoma specimens only in the nuclei of the carcinoma cells, where the phoshpo-tyrosine seems to be better conserved. The cytosolic domain of RON is translocated to the nucleus under hypoxic conditions and acts as a transcription factor, similar to MET [51, 52]. As the nuclear localization is not associated with the canonical signal transduction of the receptor, conclusions from the nuclear phopsho-RON staining about the receptor activation cannot be drawn. Further studies using fresh frozen biopsies are warranted to assess the activation of RON in tumors of patients and the role of altered glycosylation in this process.

The mechanism by which RON is increasingly activated in cells that overexpress ST3GAL4 remains to be fully characterized. Generally, there are two alternative mechanisms for the RON activation: (i) Autonomous activation, such as conformational changes of RON induced through altered glycosylation that enhance the ligand independent activation; (ii) other protein dependent activation, such as alteration of interaction with other glycoproteins.

The increased sialylation of RON *N*-glycosylation site N841 appears likely to be involved in its activation, as it is located in the IPT domain, known to be important for receptor activation [53, 54]. This may be tested in the future by site directed mutagenesis of N841 in ST3GAL4 and mock transfected cells. If the alteration of this *N*-glycan site is the sole cause of hyperactivation, abolishing it in both cell lines would normalize RON activation levels. However, we suspect that the interaction with co-

receptors or other proteins is also involved in RON activation upon glycosylation changes. Several glycoproteins that are involved in RON signaling show likewise an increase in sialylation. For instance, Receptor-type tyrosine-protein phosphatase eta (PTPRJ, DEP-1) and Receptor-type tyrosine-protein phosphatase H (PTPRH, SAP-1) were shown to be increased sialylated in ST3GAL4 overexpressing cells. Both protein tyrosine phosphatases have been shown to be involved in the deactivation of receptor tyrosine kinases and to be relevant in gastric cancer [55-57]. These two transmembrane glycoproteins showed the most altered *N*-glycan sites of our sialoproteomic analysis, DEP-1 and SAP-1 with 6 and 3 sites, respectively. The phosphatase function of protein tyrosine phosphatases requires close interaction with the receptor tyrosine kinase [55]. Sialic acids as negatively charged moieties at the forefront of glycans can increase electrostatic repulsion between these glycoproteins. The increased sialylation of RON and DEP-1 or SAP-1 may therefore decrease the dephosphorylating turnover of phospho-RON.

Plexin B2 is another protein that can contribute to RON hyperactivation in our ST3GAL4 overexpressing cell line model. Plexins are semaphorin receptors that initiate signaling by the recruitment and activation of cytosolic GTPases. Plexin B2 and RON have been shown to directly interact and to have synergistic effects on their activation [58]. The sialoproteomic analysis revealed increased sialylation of Plexin B2 at N844 and N127 upon ST3GAL4 overexpression. The N844 *N*-glycan site drew our attention as it is located in one of the IPT domains. Plexin B2 and RON are among the few transmembrane proteins that contain IPT domains [30], and both proteins are increased sialylated in this particular domain in ST3GAL4 overexpressing cells. It is not known how Plexin B2 and RON bind to one another but the involvement of the IPT domains is apparent, as most IPT containing transmembrane proteins have been shown to interact [58].

As previously discussed, our glycoproteomic approach did not assess proteins with altered *O*-glycosylation. Likely, the truncation of *O*-glycans induced by ST3GAL4 or ST6GALNAC1 overexpression functionally affects *O*-glycosylated key proteins. This is supported by the Western blot analysis of syndecan 1 and CD44, two highly *O*-glycosylated proteins with critical cellular functions. In cell lines with truncated *O*-glycans both, CD44 and syndecan 1, showed an isoform with significantly lower molecular mass as described in chapter 4. Interestingly, this indicates that the modification of different glycosylation mechanisms can induce similar alterations on specific glycoproteins and proteoglycans.

We are starting to understand the functional consequences of this truncation. As such we demonstrated in chapter 4 that the number of colocalization events between CD44v6 and RON significantly increased in cells overexpressing ST3GAL4 or ST6GALNAC1. CD44v6 is in almost 50% of all gastric cancers *de novo* expressed and is especially in more differentiated carcinoma associated with poor survival of the patients [59]. CD44v6 is a known co-receptor of RON activation through a 5 amino-acid long peptide sequence [60]. CD44 is highly *O*-glycosylated in the CD44v6 adjoining region [61]. The truncation of *O*-glycans induced by high expression levels of ST3GAL4 or ST6GALNAC1 could therefore facilitate the interaction of CD44v6 and RON. We consider to examine in depth in the future if the truncation of *O*-glycans on CD44 can modulate RON receptor activation. We envision an approach applying antibody inhibition experiments that can be conducted to probe the direct involvement of truncated *O*-glycan of CD44 on RON activation. This can be achieved by blockage of the CD44v6-RON interaction by a v6 specific inhibitory antibody. This will reveal if abolishing the CD44v6-RON interaction is sufficient to normalize RON activation. Subsequently, it could be tested if the incubation with antibodies specific for truncated glycan epitopes, such as an anti-STn antibody in the ST6GALNAC1 overexpressing cells, also results in the normalization of the RON activation.

A recent study has shown that hypoxia can induce the increased formation of truncated *O*-glycans such as STn [62]. Hypoxic conditions are widespread in tumors and might therefore be a frequent trigger of aberrant *O*-glycan expression. As a major carrier of STn, we hypothesize that CD44 is a sentinel of this event and that the aberrantly glycosylated CD44 isoforms increase the malignant phenotype of gastric cancer cells through the oncogenic activation of RTKs such as RON or MET, as illustrated in figure 1.

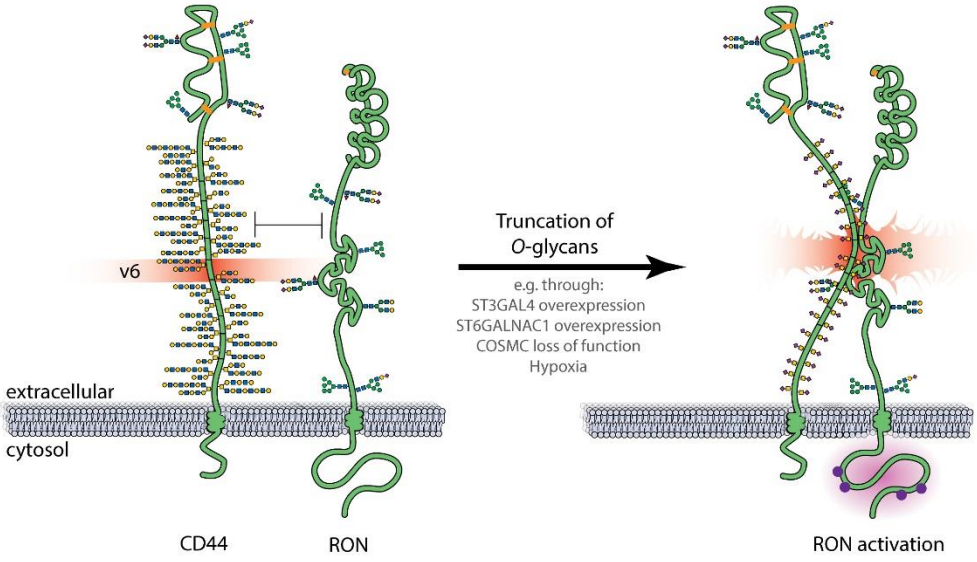


Figure 1. Proposed mechanism for RON activation by CD44 in a glycosylation dependent manner.

On the left, impediment of CD44v6 and RON interaction through elongated *O*-glycans. On the right, CD44v6 with truncated *O*-glycans promoting RON activation.

1.4 Novel glycan associations for biomarker discovery

There is a demand in the clinical setting for prognostic biomarkers that predict the gastric cancer disease course such as cancer recurrence, preferred sites of metastasis or favored way of spread. In addition, biomarkers that can be linked to molecular tumor subtypes are required. These biomarkers could guide medical doctors to select ideal therapies for each patient and determine monitoring strategies. In chapter 5 we characterized a cohort of 80 gastric carcinomas for the expression of the sialofucosylated epitopes SLe^x and SLe^a, as well as the truncated *O*-glycan epitopes STn and Tn. We used the expression profiles for KEM® based association studies with clinical features of the tumors. We obtained 3 novel and clinical relevant associations: (i) The absence of SLe^a expression precludes perineural invasion; (ii) Low SLe^a expression levels in tumors can be an indicator of *MLK3* mutation, which are linked to microsatellite instable (MSI) tumors; (iii) there seems to be an inverse correlation between STn and CDH1 expression levels.

This indicates that the oncogenic expression of SLe^a is essential for the cancer cell interaction with neurons or glial cells. The invasive capacity of cancer cells is accelerated through active migration along the perineural tract [63, 64]. SLe^a expression levels of tumors have been used as marker for aggressiveness [65] and might be partly explained by this result. Also the association with MSI is of clinical relevance. MSI tumors are chemotherapy resistant [66] and tend to accumulate hotspot mutations on growth promoting cell receptors such as ERBB2, ERBB3 and EGFR [67]. The identification of this molecular subtype might be assisted by SLe^a expression analysis and are of highest importance for the selection of optimal therapeutic strategies.

For the generation of these association rules we screened the data set for events with 100% probability. This means, only cases in which all antecedents were supporting

the hypothesis were considered. Biological samples in general and tumors in particular present tremendous heterogeneity. Even the strongest associations will inevitably have outliers if the sample number is large enough. We therefore have not assessed numerous very relevant associations yet, comprised of larger antecedent numbers and therefore presenting probabilities below 100%. These associations remain to be explored in the future. Nevertheless, even with these stringent settings we were able to establish several association rules with antecedent numbers above 5. Many of these associations were supported by more than one rule, such as the link between STn and E-cadherin, increasing the confidence of this result. We believe that these identified association rules are based on especially strong correlation of two parameters. To our knowledge this was the first absolutely blind association analysis with such a well characterized data set. As a result, many of the established rules have never been described before and might pave the way for new lines of research, in which these associations are experimentally confirmed and further exploited.

2. CONCLUDING REMARKS

Glycosylation alterations are a common feature in gastric cancer leading to the modulation of cancer development and progression. This altered glycosylation can be found on specific proteins in cancer cells, modulating the functional role of these proteins and constituting an excellent source of biomarkers for screening and prognosis purposes.

The general aim of the present work was the characterization of the gastric cancer glycosylation profile and the identification of aberrantly glycosylated protein targets to unravel the role of glycan alterations in carcinogenesis and cancer progression. The work presented in this thesis allowed us to draw the following conclusions:

1. In gastric carcinoma cells, ST3GAL4 overexpression leads to systemic glycosylation alterations, including the *de novo* expression of SLe^x, reduced *O*-glycan extension and decreased bisected and increased branched *N*-glycans. In addition, sialic acid linkage on *N*-glycans was shifted from α2-6 towards α2-3. In contrast, ST6GALNAC1 overexpression predominantly increased the formation of *O*-glycan epitope STn, leaving the *N*-glycome largely unaffected.

2. Oncogenic alterations in expression levels of glycosyltransferases affect numerous cancer relevant glycoproteins simultaneously. In ST3GAL4 overexpressing gastric carcinoma cells 47 glycoproteins were identified with significantly increased sialylated *N*-glycans. Almost half of which were involved in migration and adhesion or integral parts of the cell signaling machinery. Among them key players of malignancy such as Insulin receptor, RON, Integrin αV, Integrin α3, Integrin β6, CEACAM1, CEACAM5 and Receptor-type tyrosine-protein phosphatases DEP-1 and SAP-1.

3. The RON receptor tyrosine kinase is hyperactivated and colocalizes with SLe^x in gastric carcinoma cells overexpressing ST3GAL4. The colocalization of SLe^x and RON can be found in gastric carcinomas of patients but is absent from the healthy gastric mucosa.

4. CD44 is a major carrier of truncated *O*-glycans at the surface of gastric carcinoma cells. Alterations in the glycosylation machinery can have a severe impact on the molecular weight of CD44 and its recognition by antibodies. The truncation of *O*-glycans has functional implication on CD44 and increases the colocalization of CD44v6 and RON receptor tyrosine kinase.

5. The absence of SLe^a expression in tumors is associated with no perineural infiltration of the cancer cells and the presence of *MLK3* mutations. In addition, STn inversely associates with *CDH1* expression. This proofs that powerful association algorithms can reveal novel associations in data sets filled with well-studied parameters.

3. REFERENCES

1 Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., . . . Bray, F. (2015) Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 136, E359-386

2 Jemal, A., Center, M. M., DeSantis, C. and Ward, E. M. (2010) Global patterns of cancer incidence and mortality rates and trends. Cancer Epidemiol Biomarkers Prev. 19, 1893-1907

3 Greaves, M. and Maley, C. C. (2012) Clonal evolution in cancer. Nature. 481, 306-313

4 Meacham, C. E. and Morrison, S. J. (2013) Tumour heterogeneity and cancer cell plasticity. Nature. 501, 328-337

5 Houghton, A. N. and Guevara-Patino, J. A. (2004) Immune recognition of self in immunity against cancer. J Clin Invest. 114, 468-471

6 Kulasingam, V. and Diamandis, E. P. (2008) Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. Nat Clin Pract Oncol. 5, 588-599

7 Van Cutsem, E., Sagaert, X., Topal, B., Haustermans, K. and Prenen, H. (2016) Gastric cancer. Lancet

8 Moremen, K. W., Tiemeyer, M. and Nairn, A. V. (2012) Vertebrate protein glycosylation: diversity, synthesis and function. Nat Rev Mol Cell Biol. 13, 448-462

9 Pinho, S. S. and Reis, C. A. (2015) Glycosylation in cancer: mechanisms and clinical implications. Nat Rev Cancer. 15, 540-555

10 Rini, J., Esko, J. and Varki, A. (2009) Glycosyltransferases and Glycan-processing Enzymes. In Essentials of Glycobiology (Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W. and Etzler, M. E., eds.), Cold Spring Harbor (NY)

11 Freeze, H. H. and Elbein, A. D. (2009) Glycosylation Precursors. In Essentials of Glycobiology (Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W. and Etzler, M. E., eds.), Cold Spring Harbor (NY)

12 Marcos, N. T., Cruz, A., Silva, F., Almeida, R., David, L., Mandel, U., . . . Reis, C. A. (2003) Polypeptide GalNAc-transferases, ST6GalNAc-transferase I, and ST3Gal-

transferase I expression in gastric carcinoma cell lines. J Histochem Cytochem. 51, 761-771

13 Mondal, N., Buffone, A., Jr., Stolfa, G., Antonopoulos, A., Lau, J. T., Haslam, S. M., . . . Neelamegham, S. (2015) ST3Gal-4 is the primary sialyltransferase regulating the synthesis of E-, P-, and L-selectin ligands on human myeloid leukocytes. Blood. 125, 687-696

14 Petretti, T., Schulze, B., Schlag, P. M. and Kemmner, W. (1999) Altered mRNA expression of glycosyltransferases in human gastric carcinomas. Biochim Biophys Acta. 1428, 209-218

15 Jun, L., Yuanshu, W., Yanying, X., Zhongfa, X., Jian, Y., Fengling, W., . . . Shuxiang, C. (2012) Altered mRNA expressions of sialyltransferases in human gastric cancer tissues. Med Oncol. 29, 84-90

16 Marcos, N. T., Bennett, E. P., Gomes, J., Magalhaes, A., Gomes, C., David, L., . . . Reis, C. A. (2011) ST6GalNAc-I controls expression of sialyl-Tn antigen in gastrointestinal tissues. Front Biosci (Elite Ed). 3, 1443-1455

17 Ozaki, H., Matsuzaki, H., Ando, H., Kaji, H., Nakanishi, H., Ikehara, Y. and Narimatsu, H. (2012) Enhancement of metastatic ability by ectopic expression of ST6GalNAcI on a gastric cancer cell line in a mouse model. Clin Exp Metastasis. 29, 229-238

18 Sewell, R., Backstrom, M., Dalziel, M., Gschmeissner, S., Karlsson, H., Noll, T., . . . Taylor-Papadimitriou, J. (2006) The ST6GalNAc-I sialyltransferase localizes throughout the Golgi and is responsible for the synthesis of the tumor-associated sialyl-Tn O-glycan in human breast cancer. J Biol Chem. 281, 3586-3594

19 Julien, S., Ivetic, A., Grigoriadis, A., QiZe, D., Burford, B., Sproviero, D., . . . Burchell, J. M. (2011) Selectin ligand sialyl-Lewis x antigen drives metastasis of hormone-dependent breast cancers. Cancer Res. 71, 7683-7693

20 Chia, J., Goh, G. and Bard, F. (2016) Short O-GalNAc glycans: regulation and role in tumor development and clinical perspectives. Biochim Biophys Acta. 1860, 1623-1639

21 Kizuka, Y. and Taniguchi, N. (2016) Enzymes for N-Glycan Branching and Their Genetic and Nongenetic Regulation in Cancer. Biomolecules. 6

22 Amado, M., Carneiro, F., Seixas, M., Clausen, H. and Sobrinho-Simoes, M. (1998) Dimeric sialyl-Le(x) expression in gastric carcinoma correlates with venous invasion and poor outcome. Gastroenterology. 114, 462-470

23 Zaia, J. (2010) Mass spectrometry and glycomics. OMICS. 14, 401-418

24 Pabst, M. and Altmann, F. (2008) Influence of electrosorption, solvent, temperature, and ion polarity on the performance of LC-ESI-MS using graphitic carbon for acidic oligosaccharides. Anal Chem. 80, 7534-7542

25 Stavenhagen, K., Kolarich, D. and Wuhrer, M. (2015) Clinical Glycomics Employing Graphitized Carbon Liquid Chromatography-Mass Spectrometry. Chromatographia. 78, 307-320

26 Beum, P. V., Singh, J., Burdick, M., Hollingsworth, M. A. and Cheng, P. W. (1999) Expression of core 2 beta-1,6-N-acetylglucosaminyltransferase in a human pancreatic cancer cell line results in altered expression of MUC1 tumor-associated epitopes. J Biol Chem. 274, 24641-24648

27 Kizuka Y, Taniguchi N: Enzymes for N-Glycan Branching and Their Genetic and Nongenetic Regulation in Cancer. Biomolecules 2016, 6(2).

28 Takada, A., Ohmori, K., Yoneda, T., Tsuyuoka, K., Hasegawa, A., Kiso, M. and Kannagi, R. (1993) Contribution of carbohydrate antigens sialyl Lewis A and sialyl Lewis X to adhesion of human cancer cells to vascular endothelium. Cancer Res. 53, 354-361

29 Perez-Garay, M., Arteta, B., Pages, L., de Llorens, R., de Bolos, C., Vidal-Vanaclocha, F. and Peracaula, R. (2010) alpha2,3-sialyltransferase ST3Gal III modulates pancreatic cancer cell motility and adhesion in vitro and enhances its metastatic potential in vivo. PLoS One. 5

30 Harduin-Lepers, A., Vallejo-Ruiz, V., Krzewinski-Recchi, M. A., Samyn-Petit, B., Julien, S. and Delannoy, P. (2001) The human sialyltransferase family. Biochimie. 83, 727-737

31 Gomes, C., Osorio, H., Pinto, M. T., Campos, D., Oliveira, M. J. and Reis, C. A. (2013) Expression of ST3GAL4 leads to SLe(x) expression and induces c-Met activation and an invasive phenotype in gastric carcinoma cells. PLoS One. 8, e66737

32 Palmisano G, Lendal SE, Engholm-Keller K, Leth-Larsen R, Parker BL, Larsen MR: Selective enrichment of sialic acid-containing glycopeptides using titanium dioxide chromatography with analysis by HILIC and mass spectrometry. Nature protocols 2010, 5(12):1974-1982.

33 Bassaganas, S., Carvalho, S., Dias, A. M., Perez-Garay, M., Ortiz, M. R., Figueras, J., . . . Peracaula, R. (2014) Pancreatic cancer cell glycosylation regulates cell adhesion

and invasion through the modulation of alpha2beta1 integrin and E-cadherin function. PLoS One. 9, e98595

34 Guo, H. B., Lee, I., Kamar, M., Akiyama, S. K. and Pierce, M. (2002) Aberrant N-glycosylation of beta1 integrin causes reduced alpha5beta1 integrin clustering and stimulates cell migration. Cancer Res. 62, 6837-6845

35 Janik, M. E., Litynska, A. and Vereecken, P. (2010) Cell migration-the role of integrin glycosylation. Biochim Biophys Acta. 1800, 545-555

36 Adamczyk B, Tharmalingam T, Rudd PM: Glycans as cancer biomarkers. Biochimica et biophysica acta 2012, 1820(9):1347-1353.

37 Carpelan-Holmstrom M, Louhimo J, Stenman UH, Alfthan H, Haglund C: CEA, CA 19-9 and CA 72-4 improve the diagnostic accuracy in gastrointestinal cancers. Anticancer research 2002, 22(4):2311-2316.

38 Thomas, R. M., Toney, K., Fenoglio-Preiser, C., Revelo-Penafiel, M. P., Hingorani, S. R., Tuveson, D. A., . . . Lowy, A. M. (2007) The RON receptor tyrosine kinase mediates oncogenic phenotypes in pancreatic cancer cells and is increasingly expressed during pancreatic cancer progression. Cancer Res. 67, 6075-6082

39 Maggiora, P., Marchio, S., Stella, M. C., Giai, M., Belfiore, A., De Bortoli, M., . . . Comoglio, P. M. (1998) Overexpression of the RON gene in human breast carcinoma. Oncogene. 16, 2927-2933

40 Collesi, C., Santoro, M. M., Gaudino, G. and Comoglio, P. M. (1996) A splicing variant of the RON transcript induces constitutive tyrosine kinase activity and an invasive phenotype. Mol Cell Biol. 16, 5518-5526

41 Zhou, Y. Q., He, C., Chen, Y. Q., Wang, D. and Wang, M. H. (2003) Altered expression of the RON receptor tyrosine kinase in primary human colorectal adenocarcinomas: generation of different splicing RON variants and their oncogenic potential. Oncogene. 22, 186-197

42 Ponzetto, C., Bardelli, A., Zhen, Z., Maina, F., dalla Zonca, P., Giordano, S., . . . Comoglio, P. M. (1994) A multifunctional docking site mediates signaling and transformation by the hepatocyte growth factor/scatter factor receptor family. Cell. 77, 261-271

43 Iwama, A., Yamaguchi, N. and Suda, T. (1996) STK/RON receptor tyrosine kinase mediates both apoptotic and growth signals via the multifunctional docking site conserved among the HGF receptor family. EMBO J. 15, 5866-5875

44 Wang, D., Shen, Q., Xu, X. M., Chen, Y. Q. and Wang, M. H. (2005) Activation of the RON receptor tyrosine kinase attenuates transforming growth factor-beta1-induced apoptotic death and promotes phenotypic changes in mouse intestinal epithelial cells. Carcinogenesis. 26, 27-36

45 Wang, M. H., Dlugosz, A. A., Sun, Y., Suda, T., Skeel, A. and Leonard, E. J. (1996) Macrophage-stimulating protein induces proliferation and migration of murine keratinocytes. Exp Cell Res. 226, 39-46

46 Cote, M., Miller, A. D. and Liu, S. L. (2007) Human RON receptor tyrosine kinase induces complete epithelial-to-mesenchymal transition but causes cellular senescence. Biochem Biophys Res Commun. 360, 219-225

47 Thobe, M. N., Gurusamy, D., Pathrose, P. and Waltz, S. E. (2010) The Ron receptor tyrosine kinase positively regulates angiogenic chemokine production in prostate cancer cells. Oncogene. 29, 214-226

48 Logan-Collins, J., Thomas, R. M., Yu, P., Jaquish, D., Mose, E., French, R., . . . Lowy, A. M. (2010) Silencing of RON receptor signaling promotes apoptosis and gemcitabine sensitivity in pancreatic cancers. Cancer Res. 70, 1130-1140

49 Yao, H. P., Zhou, Y. Q., Zhang, R. and Wang, M. H. (2013) MSP-RON signalling in cancer: pathogenesis and therapeutic potential. Nat Rev Cancer. 13, 466-481

50 O'Hurley G, Sjostedt E, Rahman A, Li B, Kampf C, Ponten F, Gallagher WM, Lindskog C: Garbage in, garbage out: a critical evaluation of strategies used for validation of immunohistochemical biomarkers. Molecular oncology 2014, 8:783-798.

51 Mahdavi, J., Sonden, B., Hurtig, M., Olfat, F. O., Forsberg, L., Roche, N., . . . Boren, T. (2002) Helicobacter pylori SabA adhesin in persistent infection and chronic inflammation. Science. 297, 573-578

52 Gomes, D. A., Rodrigues, M. A., Leite, M. F., Gomez, M. V., Varnai, P., Balla, T., . . . Nathanson, M. H. (2008) c-Met must translocate to the nucleus to initiate calcium signals. J Biol Chem. 283, 4344-4351

53 Ma, Q., Zhang, K., Guin, S., Zhou, Y. Q. and Wang, M. H. (2010) Deletion or insertion in the first immunoglobulin-plexin-transcription (IPT) domain differentially regulates expression and tumorigenic activities of RON receptor Tyrosine Kinase. Mol Cancer. 9, 307

54 Zhou DH, Li C, Yang LN: Variant RONDelta160 of the RON receptor tyrosine kinase promotes the growth and invasion in vitro and in vivo in gastric cancer cell lines. Cancer cell international 2015, 15(1):9.

55 Ostman A, Hellberg C, Bohmer FD: Protein-tyrosine phosphatases and cancer. Nature reviews Cancer 2006, 6(4):307-320.

56 Ruivenkamp, C. A., van Wezel, T., Zanon, C., Stassen, A. P., Vlcek, C., Csikos, T., . . . Demant, P. (2002) Ptprj is a candidate for the mouse colon-cancer susceptibility locus Scc1 and is frequently deleted in human cancers. Nat Genet. 31, 295-300

57 Noguchi, T., Tsuda, M., Takeda, H., Takada, T., Inagaki, K., Yamao, T., . . . Kasuga, M. (2001) Inhibition of cell growth and spreading by stomach cancer-associated protein-tyrosine phosphatase-1 (SAP-1) through dephosphorylation of p130cas. J Biol Chem. 276, 15216-15224

58 Conrotto, P., Corso, S., Gamberini, S., Comoglio, P. M. and Giordano, S. (2004) Interplay between scatter factor receptors and B plexins controls invasive growth. Oncogene. 23, 5131-5137

59 Yamaguchi, A., Goi, T., Yu, J., Hirono, Y., Ishida, M., Iida, A., . . . Hirose, K. (2002) Expression of CD44v6 in advanced gastric cancer and its relationship to hematogenous metastasis and long-term prognosis. J Surg Oncol. 79, 230-235

60 Matzke, A., Herrlich, P., Ponta, H. and Orian-Rousseau, V. (2005) A five-amino-acid peptide blocks Met- and Ron-dependent cell migration. Cancer Res. 65, 6105-6110

61 Steentoft, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B., Schjoldager, K. T., . . . Clausen, H. (2013) Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J. 32, 1478-1488

62 Peixoto, A., Fernandes, E., Gaiteiro, C., Lima, L., Azevedo, R., Soares, J., . . . Ferreira, J. A. (2016) Hypoxia enhances the malignant nature of bladder cancer cells and concomitantly antagonizes protein O-glycosylation extension. Oncotarget

63 Bilici, A., Seker, M., Ustaalioglu, B. B., Kefeli, U., Yildirim, E., Yavuzer, D., . . . Gumus, M. (2010) Prognostic significance of perineural invasion in patients with gastric cancer who underwent curative resection. Ann Surg Oncol. 17, 2037-2044

64 Amit, M., Na'ara, S. and Gil, Z. (2016) Mechanisms of cancer dissemination along nerves. Nat Rev Cancer. 16, 399-408

65 Jo, J. C., Ryu, M. H., Koo, D. H., Ryoo, B. Y., Kim, H. J., Kim, T. W., . . . Kang, Y. K. (2013) Serum CA 19-9 as a prognostic factor in patients with metastatic gastric cancer. Asia Pac J Clin Oncol. 9, 324-330

66 An, J. Y., Kim, H., Cheong, J. H., Hyung, W. J., Kim, H. and Noh, S. H. (2012) Microsatellite instability in sporadic gastric cancer: its prognostic role and guidance for 5-FU based chemotherapy after R0 resection. Int J Cancer. 131, 505-511

67 Cancer Genome Atlas Research, N. (2014) Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 513, 202-209

ANNEX I

GLYCOMIC APPROACHES FOR THE DISCOVERY OF TARGETS IN GASTROINTESTINAL CANCER

frontiers in Oncology

REVIEW
published: 09 March 2016
doi: 10.3389/fonc.2016.00055

Glycomic Approaches for the Discovery of Targets in Gastrointestinal Cancer

Stefan Mereiter[1,2,3], Meritxell Balmaña[4], Joana Gomes[1,2], Ana Magalhães[1,2] and Celso A. Reis[1,2,3,5]**

[1] Instituto de Investigação e Inovação em Saúde (I3S), University of Porto, Porto, Portugal, [2] Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP), Porto, Portugal, [3] Institute of Biomedical Sciences of Abel Salazar (ICBAS), University of Porto, Porto, Portugal, [4] Biochemistry and Molecular Biology Unit, Department of Biology, University of Girona, Girona, Spain, [5] Medical Faculty, University of Porto, Porto, Portugal

Gastrointestinal (GI) cancer is the most common group of malignancies and many of its types are among the most deadly. Various glycoconjugates have been used in clinical practice as serum biomarker for several GI tumors, however, with limited diagnose application. Despite the good accessibility by endoscopy of many GI organs, the lack of reliable serum biomarkers often leads to late diagnosis of malignancy and consequently low 5-year survival rates. Recent advances in analytical techniques have provided novel glycoproteomic and glycomic data and generated functional information and putative biomarker targets in oncology. Glycosylation alterations have been demonstrated in a series of glycoconjugates (glycoproteins, proteoglycans, and glycosphingolipids) that are involved in cancer cell adhesion, signaling, invasion, and metastasis formation. In this review, we present an overview on the major glycosylation alterations in GI cancer and the current serological biomarkers used in the clinical oncology setting. We further describe recent glycomic studies in GI cancer, namely gastric, colorectal, and pancreatic cancer. Moreover, we discuss the role of glycosylation as a modulator of the function of several key players in cancer cell biology. Finally, we address several state-of-the-art techniques currently applied in this field, such as glycomic and glycoproteomic analyses, the application of glycoengineered cell line models, microarray and proximity ligation assay, and imaging mass spectrometry, and provide an outlook to future perspectives and clinical applications.

Keywords: gastric cancer, colorectal cancer, pancreatic cancer, glycomics, glycan biomarkers, microarray, proximity ligation assay, imaging mass spectrometry

OPEN ACCESS

Edited by:
Leonardo Freire-de-Lima, Federal University of Rio de Janeiro, Brazil

Reviewed by:
Stephan Von Gunten, University of Bern, Switzerland
Ali Mobasheri, University of Surrey, UK

****Correspondence:***
Ana Magalhães
amagalhaes@ipatimup.pt;
Celso A. Reis
celsor@ipatimup.pt

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *03 February 2016*
Accepted: *24 February 2016*
Published: *09 March 2016*

Citation:
Mereiter S, Balmaña M, Gomes J, Magalhães A and Reis CA (2016) Glycomic Approaches for the Discovery of Targets in Gastrointestinal Cancer. Front. Oncol. 6:55. doi: 10.3389/fonc.2016.00055

GLYCOBIOLOGY IN CANCER

The cells' glycocalix constitutes an important interface with the extracellular milieu and plays critical roles in physiological and pathological conditions. This glycan-rich coating of the cells' plasma membrane is composed by different classes of glycoconjugates, including glycoproteins, glycolipids, and proteoglycans, which participate in key regulatory events for cellular and organ homeostasis. Alterations in glycosylation can interfere with normal molecular functions such as cell–cell recognition, communication, and adhesion, leading to acquisition of malignant features.

Moreover, the shedding of aberrant glycoconjugates, uniquely expressed by tumor cells, into circulation provides one valuable source of biomarkers for cancer diagnosis and prognosis (1).

Substantial advances in the frontiers of cancer glycobiology have been possible in the recent past due to the combination of novel tumor cell biology concepts with cutting-edge glycomic technologies. Specific glycosylation alterations have been identified in tumors and some of the molecular pathways underlying these modifications have been disclosed (2). In addition, aberrant glycoforms have been demonstrated to be molecularly associated with more aggressive cancer cell and tumor features, including increased migration, invasion, and metastization potential, providing novel targets for therapeutic intervention (3–5).

This review describes the recent progress in gastric cancer, colorectal cancer (CRC), and pancreatic ductal adenocarcinoma (PDAC) glycobiology and discusses the clinical value of aberrant glycosylation as a source of screening biomarkers and therapeutic targets. A comprehensive overview of the advances in glycomic and glycoproteomic tools is also provided and their possible applications for tumor glycan-profiling and discovery of novel targets for improving gastrointestinal (GI) tumors' clinical management are discussed.

GLYCOSYLATION ALTERATION IN GASTROINTESTINAL CANCERS

Despite the large amount of glycan epitopes that can be found in the human GI tract and the complex and manifold alterations of the glycosylation machinery during the process of carcinogenesis and cancer progression, the current knowledge makes it possible to group the most common glycan alterations. Expression of truncated simple *O*-glycans, changes in *N*-glycan branching, and increase in sialylation and fucosylation are three major *N*- and *O*-glycosylation events involved in GI cancer that will be described in detail (**Figure 1**). Furthermore, we give an overview on other common glycosylation alterations in cancer, such as changes in *O*-GlcNAcylation, modified glycosphingolipids, and glycosaminoglycans (GAGs) and proteoglycans.

Truncated Simple *O*-Glycans

One common feature observed in GI tumors is the overexpression and exposure of short, truncated *O*-glycans. Mucin-type *O*-glycans are found on most transmembrane and secreted proteins. A single *O*-glycan oligosaccharide chain can present more than 20 monosaccharide constituents (6). In malignancy, *O*-glycans are often shortened resulting in an increase of the monosaccharide Tn antigen (GalNAcα1-Ser/Thr), the disaccharide T antigen (also known as Thomsen–Friedenreich antigen or core 1 structure, Galβ1-3GalNAcα1-Ser/Thr) and their sialylated forms, STn (Neu5Acα2-6GalNAcα1-Ser/Thr), and ST (Neu5Acα2-3Galβ1-3GalNAcα1-Ser/Thr), respectively (**Figure 1**) (7, 8).

Polypeptide GalNAc transferases (ppGalNAcTs), which are the initiating enzymes of the mucin-type *O*-glycosylation (9, 10), show often altered expression in cancer (11–13). A total of 20 different ppGalNAcTs are known in human and their expression profile and subcellular localization determine *O*-glycosylation sites and densities (9, 14). In CRC, for example, the ppGalNAc-T3 is associated with tumor differentiation, disease aggressiveness, and prognosis (12). In gastric cancer, the expression of ppGalNAc-T6 is associated with venous invasion and the downregulation of ppGalNAc-T2 increases the cancer cell proliferation, adhesion, and invasion (11, 15).

In addition, enzymes competing for the same substrate can also induce expression of truncated glycans and exposure of protein epitopes that would be hidden otherwise. For instance, the relative enzymatic activities of C2GnT (*N*-acetylglucosaminyltransferase) and ST3Gal-I (sialyltransferase), two glycosyltransferases that compete for the same substrate, have been shown to determine the *O*-glycan structure in cancer cells (16).

STn is expressed in most GI carcinomas correlating with decreased cancer cell adhesion, increased cancer cell invasion, and poor prognosis of the patients (17–23). The terminal STn epitope is synthesized by the sialylation of Tn by the ST6GalNAc-I sialyltransferase (17, 18). In cancer, the formation of STn may occur due to ST6GalNAc-I upregulation, early sialylation caused by glycosyltransferase misslocalization in the secretory pathway, or the impairment of the elongation of the Tn antigen (14, 17, 18, 24).

In gastric cancer, expression of STn is a common feature associated with more malignant phenotypes. The overexpression of ST6GalNAc-I has been shown to induce migration and invasion in gastric carcinoma cells *in vitro* (20).

In this regard, another gene that can underlie the synthesis of truncated *O*-glycans is *COSMC*, which encodes for a C1GalT1 dedicated chaperone (25). The galactosyltransferase C1GalT1 is responsible for the elongation of the Tn antigen to form the core 1 structure also known as the T antigen. The absence of a functional *COSMC* entails the dysfunction of C1GalT1. In PDAC, it has been shown that hypermethylation of *COSMC*, and not somatic mutations, is the prevalent cause of truncated *O*-glycans (23). In addition, the downregulation of C1GalT1 in combination with the upregulation of ST6GalNAc-I has been associated with increased STn expression in CRC cell lines and epithelial cells derived from resected CRC tumor tissue (26). In contrary, the overexpression of C1GalT1 is associated with invasion, metastization, and poor survival in CRC. In C1GalT1 overexpressing CRC cells, the knockdown of C1GalT1 suppresses the malignant phenotype *in vitro* and *in vivo* (27). Increased levels of C2GnT, a glycosyltransferase responsible for the biosynthesis of core 2 structures, are also frequent in CRC (28). This enzyme has also a critical role in the biosynthesis of terminal sialylated Lewis antigens on *O*-glycans that will be further discussed in Section "Increased Sialylation and Fucosylation."

Normal pancreas does not express the Tn antigen and its corresponding sialylated epitope STn (21). The Tn antigen is detected in 75–90% of PDACs and up to 67% in precursor lesions (24). The appearance of the STn in mucins, on the other hand, is a late event in PDAC disease progression (29). These truncated *O*-glycans are associated with cancer cell growth and tumor invasion in PDAC (23, 24). The situation is slightly different in CRC, where the overexpression of T antigen is associated with early events in cancer progression and both Tn and STn antigens are frequently overexpressed in advanced and poorly differentiated adenocarcinomas

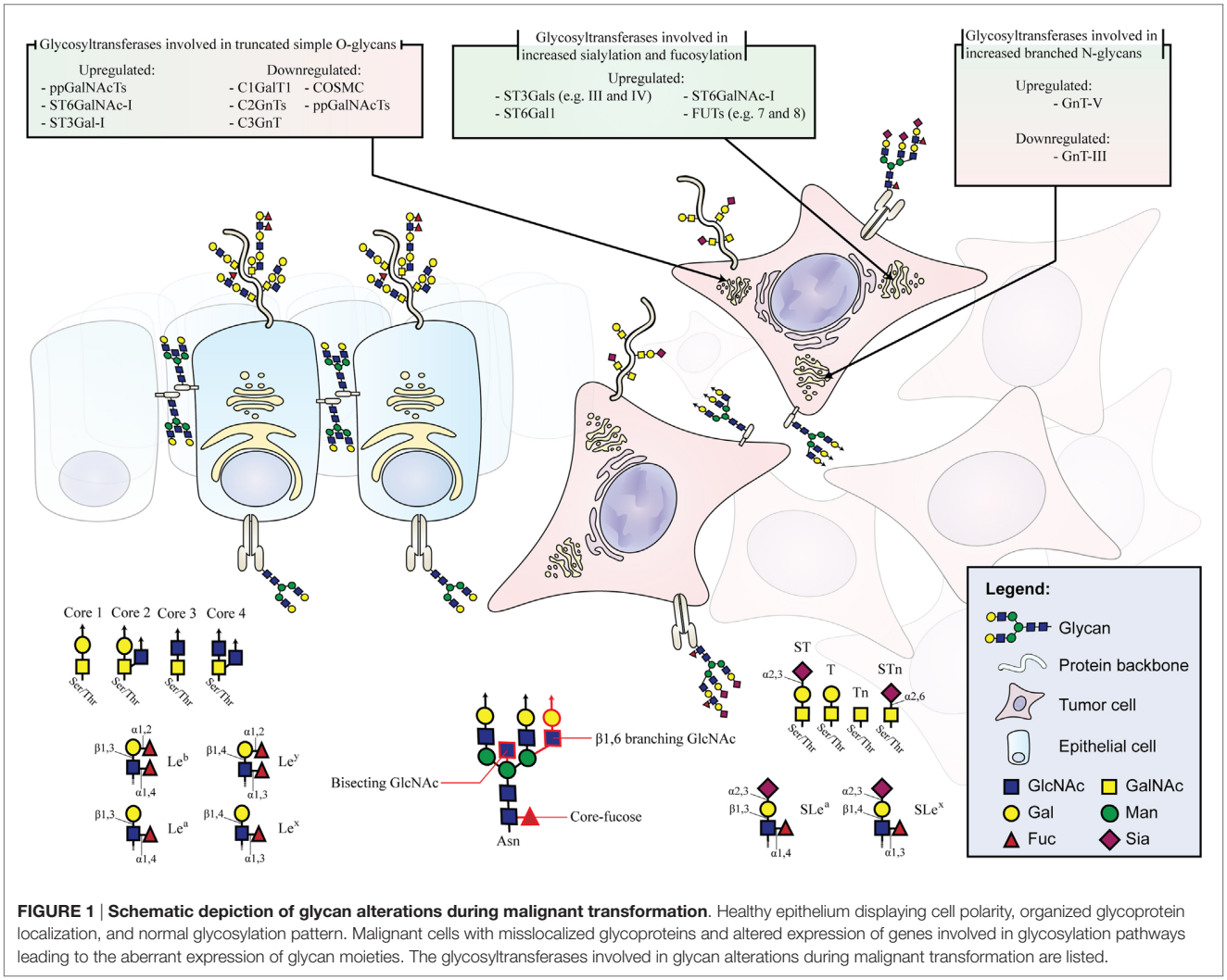


FIGURE 1 | Schematic depiction of glycan alterations during malignant transformation. Healthy epithelium displaying cell polarity, organized glycoprotein localization, and normal glycosylation pattern. Malignant cells with misslocalized glycoproteins and altered expression of genes involved in glycosylation pathways leading to the aberrant expression of glycan moieties. The glycosyltransferases involved in glycan alterations during malignant transformation are listed.

and also in mucinous carcinomas. Therefore, these antigens are considered useful markers for poor outcome (22).

Besides cores 1 and 2, *O*-glycans with cores 3 and 4 are also often expressed in normal GI epithelia, especially in colon, but are significantly decreased in malignancy due to downregulation of the core 3 and 4-synthetase (8, 30–32).

The *de novo* expression of truncated *O*-glycans in GI cancers is of avail in the search of specific cancer biomarkers. It leads to the expression of unique glycopeptide structures and, because of the small steric size of these truncated *O*-glycan moieties, to the exposure of protein regions that would otherwise be masked, and, therefore, not detected by specific antibodies (29, 33).

Branched *N*-Glycans

The biosynthesis and maturation of *N*-glycan structures is defined by a complex interplay of numerous glycosidases and glycosyltransferases in the endoplasmic reticulum and Golgi. Among *N*-glycan types, the complex *N*-glycans display the largest structural diversity. Two structural features of complex *N*-glycans are the β1,6-branching, catalyzed by the glycosyltransferase GnT-V, and the bisecting-GlcNAc, added by the glycosyltransferase GnT-III. GnT-V is known to be upregulated in gastric carcinoma (**Figure 1**) (34), leading to the increased branching of *N*-glycans and contributing to cancer cell invasion and metastases (35, 36). Analogically, normal colon epithelium presents high levels of bisecting-GlcNAc, due to high expression levels of GnT-III, which is associated with suppression of the tumor progression. However, during cancer progression, these bisecting structures are decreased (37) and it has been described a general increase of β1,6-branched in complex *N*-linked glycans that are also associated with tumor invasion and metastasis (38). Histochemical studies using specific lectins for the detection of β1,6-branched structures showed increased staining concomitant with tumor CRC staging (39), and an association with lymph node metastasis and decreased survival rates in CRC patients (40). GnT-V, the enzyme responsible for the synthesis of β1,6-branched *N*-glycans, is commonly upregulated in CRC correlating to the metastatic potential and consequently

considered an important prognosis factor to detect poor CRC patients' outcome (41). A recent study demonstrated that GnT-V levels modulate CRC stem cells and tumor formation by Wnt signaling (42). Increased extension of β1,6-branched complex *N*-glycans by long polymers of *N*-acetyllactosamine (LacNAc) due to upregulation of β3GnT8 has also been described in CRC cells (43).

Regarding PDAC progression, little has been described about bisecting structures, although the increase in highly branched *N*-glycans is well established. The number of tri- and tetra-antennary glycans is significantly increased in both pancreatic cancer cells and PDAC patients' serum and correlate with cancer progression (44, 45).

Another mechanism leading to increased branching is the downregulation of GnT-III and the addition of bisecting-GlcNAc. Bisected structures cannot be further modified by GnT-V and, therefore, preclude the formation of branched *N*-glycans under healthy conditions.

The interplay of GnT-III and GnT-V modulates cell adhesion and migration in a gastric cancer context (46, 47). This has been shown to be particularly important for cell–cell and cell–matrix interactions in gastric cancer by altering the functionality of E-cadherin and integrins in malignant transformation. E-cadherin promotes adherence junction formation and, thus, maintains intercellular adhesion. E-cadherin is stabilized by bisected *N*-glycans delaying its endocytosis and turnover (47–49). Furthermore, bisecting-GlcNAc on E-cadherin is gastric tumor suppressive by downregulating signaling pathways involved in cell motility and the EMT process (50–54). Conversely, E-cadherin is dysregulated when glycosylated with branched *N*-glycans by GnT-V in the context of gastric cancer (34, 52, 53). GnT-V is commonly upregulated in gastric carcinomas contributing to cell invasion and metastases (35, 36). The overexpression of GnT-V leads to destabilization of adherence junctions, delocalization of E-cadherin into the cytoplasm, and mesenchymal appearance of the cells with increased metastatic capability (34, 52, 55).

Integrins convey adhesion to extracellular matrix components and are often altered in GI carcinomas. In gastric cancer, the modification of α3β1 integrin with branched *N*-glycans increases cell migration (56). The modification of α3β1 integrin with bisecting-GlcNAc has the opposite effect by inhibiting cell migration (56). Consistently, the overexpression of GnT-III resulted in the inhibition of α5β1 integrin-mediated cell migration and reduced binding to fibronectin due to a specific *N*-glycosylation site on the α5 integrin (57, 58).

Increased Sialylation and Fucosylation

Sialic acids are the largest and the only intrinsically negatively charged monosaccharides present in human glycosylation. As a terminal event, sialylation caps glycosylation chains usually resulting in exposed locations of the negative charge at the forefront of the oligosaccharides and first encounter point for adjacent glycans, proteins, and cells. Sialylation has, therefore, been shown to play important roles in modulating cellular recognition, cell adhesion, and cell signaling (59). Moreover, cancer cell sialylation patterns define sialic acid-binding lectins (Siglecs) interactions and modulate immune response (60, 61).

An increase in global sialylation, especially in α2,6- and α2,3-linked sialylation, owing to altered glycosyltransferases expression, has been closely associated with cancer and commonly described as one of the main modifications in GI cancers (62, 63). For example, ST6Gal-I, the enzyme that adds α2,6-linked sialic acid to lactosamine chains (Neu5Acα2,6Galβ1,4GlcNAc), is commonly overexpressed in GI cancers correlating with poor prognosis (59, 64, 65). Additionally, α2,3-sialyltransferases, such as ST3Gal-III and ST3Gal-IV, are often upregulated in the course of gastric cancer and PDAC progression leading to a more invasive and metastatic phenotypes of the cancer cells (65–69). Furthermore, sialylation, in particular α2,3 and α2,6-linked, can modulate the ECM adhesion and migration. Specifically, it has been described that while the overexpression of terminal α2,6-linked sialic acid leads to increased ECM adhesion, the overexpression of α2,3-linked terminal sialic acid epitopes in PDAC cancer cell lines results in a more migratory phenotype (70). Similarly, in gastric cancer cells, the overexpression of α2,3-linked terminal sialic acid epitopes causes a more invasive phenotype *in vitro* and *in vivo* (67).

The major α2,3-sialylated antigens associated with cancer are SLe^a and SLe^x (**Figure 1**). Although these structures can also be present in non-neoplastic cells, SLe^a and SLe^x have been demonstrated to be highly expressed in many malignant tissues, including GI tumors, both in glycoproteins and glycosphigolipids (71–74). SLe^x-increased expression levels are associated with advanced stages and have been correlated with poor survival in GI cancer patients (75–77). SLe^x is the well-known ligand for selectins (78). During inflammation, selectins mediate the initial attachment of leukocytes to the endothelium during the process of leukocyte extravasation. In cancer, SLe^x interactions with selectins favor metastasis by forming emboli of cancer cells and platelets and promoting their arrest on endothelia (77).

The overexpression of SLe^x in a gastric carcinoma cell line transfected with *ST3GAL4* has shown to increase the cells invasive potential both *in vitro* and *in vivo* due to the activation of the oncogenic c-Met receptor tyrosine kinase (67). Moreover, overexpression of *ST3GAL4* has been shown to result in RON receptor tyrosine kinase activation and co-expression of RON and SLe^x is observed in gastric tumors (79). This is of particular biological relevance since it has been described that RON activation contributes to tumor progression, angiogenesis, and therapy resistance and correlates with bad prognosis (80–84).

Sialylated Lewis epitopes are potential good markers for prognosis due to their high incidence of recurrence or presence in metastasis and correlation with the tumor stage. For example, a recent work described the increase of the SLe^x epitope on ceruloplasmin in PDAC. The increased ceruloplasmin with the SLe^x epitope in chronic pancreatitis was lower, suggesting good specificity for pancreatic malignancy (85). Moreover, studies using high-density antibody microarray also detected increased levels of SLe^x and SLe^a antigens on glycoproteins in serum or plasma of CRC patients (86).

Overexpression of the enzyme β-galactoside α2,6-sialyltransferase I (ST6Gal-I), especially in *N*-glycans and not in *O*-glycans, has been associated with CRC progression, increased invasion, and metastization and consequently poor prognosis in

CRC patients (64, 87). Further studies taking into consideration the low levels of ST6Gal-I in healthy individuals and upregulation in CRC patients could lead to the development of new diagnostic and therapeutic targets.

Fucosylation is also an important modification involved in cancer and inflammation (88). The attachment of fucose to *N*-glycans, *O*-glycans, and glycolipids has been reported in cancer tissues, regulating different biological processes, and being also responsible for the increased expression of Lewis antigens (89) (and sialylated-Lewis antigens, as previously described). Different studies link the presence of fucosylated epitopes on specific glycoproteins with cancer. In particular, research performed on some acute phase proteins suggest the suitability of fucosylated epitopes for cancer management. It has been demonstrated that fucosylated alpha-fetoprotein (AFP) is more specific as a hepatocellular carcinoma biomarker than AFP itself. Nowadays fucosylated AFP (AFP-L3) is used for hepatocellular carcinoma risk assessment (90, 91). Acute phase proteins, such as AFP, are proteins synthesized by hepatocytes and have shown clinical value as markers for liver and pancreatic-related diseases. For example, haptoglobin and AGP have revealed an increase in fucosylated epitopes that could help to improve PDAC diagnosis (89, 92).

Increased activity in α1,3 and α1,4 fucosyltransferases (FUTs) was described in CRC patients, resulting in the synthesis of SLe^x and SLe^a epitopes, respectively (89). Particularly, FUT6 was more recently reported as the major regulator of SLe^x biosynthesis in CRC (93). Increased levels of fucosylation in plasma samples of CRC patients compared to normal controls were also described using methods for *N*-glycoproteomics analysis to identify plasma markers (94). In addition, increased levels of α1,2-FUT1 and FUT2, which add fucose to terminal galactose and are essential for the synthesis of Lewis Y and B antigens, were shown in CRC tumors (95). Alterations of FUT expression have also been described in the process of gastric carcinogenesis (96). In particular, the downregulation of FUT3 and FUT5 changes the Lewis antigens expression and reduces the adhesion capacities of gastric cancer cells (97). This is contrary to what is observed in gastric inflamed mucosa, where FUT3 is upregulated (98).

Increased core-fucosylation of *N*-glycans catalyzed by α1,6-FUT8 has been described in CRC patients and is associated with tumor aggressiveness (37, 99). The core-fucosylation of E-cadherin enhances the cellular adhesion of CRC cells (100). However, in gastric cancer, the decrease of core-fucosylation has been demonstrated to be a common event contributing to cancer cell proliferation (101).

Other Relevant Glycosylation Alterations

In addition to the mucin-type *O*-glycosylation, there are further forms of protein *O*-glycosylation, including the modification of nuclear and cytoplasmic proteins with *O*-linked β-*N*-acetylglucosamine (*O*-GlcNAc). Noteworthy, increased *O*-GlcNAcylation is a general feature of cancer and the modification of proteins with *O*-GlcNAc has been shown to play key regulatory roles in tumor cell signaling (102). The addition of *O*-GlcNAc to nuclear and cytosolic proteins is mediated by the *O*-GlcNAc transferase (OGT), whereas the enzyme *O*-GlcNAc-selective *N*-acetyl-beta-D-glucosaminidase (*O*-GlcNAcase) removes the *O*-GlcNAc, returning the protein to its basal state (103). *O*-GlcNAcylation has been shown to have extensive crosstalk with phosphorylation and to antagonize phosphorylation-mediated cell signaling (104).

In the pancreas, beta-cells are characterized by expressing high levels of the OGT enzyme. This allows these cells to dynamically respond to physiological increases in the extracellular glucose levels by converting glucose to UDP-GlcNAc, which is the OGT substrate, and therefore modulating intracellular *O*-linked protein glycosylation (105). In PDAC, hyper-*O*-GlcNAcylation has been associated with increased expression of the OGT enzyme and reduction of the *O*-Glc-NAcase glycosidase and has been demonstrated to block cancer cell apoptosis and to lead to the oncogenic activation of the NF-kB signaling pathway (106). Similarly, increased *O*-GlcNAcylation in colon has been demonstrated to contribute for the development of colitis and colitis-associated cancer by enhancing NF-kB-mediated signaling (107).

Along with aberrant protein glycosylation, cancer cells also display major glycosylation alterations on other classes of glycoconjugates, including the proteoglycans and the glycosphingolipids. Proteoglycans consist of a core protein with one or more covalently attached large GAG chains, and can be either located at the cell membrane or secreted. The syndecans are a family of transmembrane proteoglycans that carry heparan sulfate GAG chains and that can also be additionally modified with chondroitin sulfate chains (108). The heparan sulfate-rich proteoglycan syndecan-4, a critical partner of integrins for the establishment of focal adhesion complexes, has been shown to be upregulated in gastric mucosa in response to the oncogenic bacteria *Helicobacter pylori*. However, its functional role in the gastric carcinogenesis process remains to be disclosed (109, 110). Another syndecan family member, the syndecan-1, has been reported to be differently regulated and expressed in GI tumors. Loss of syndecan-1 expression has been described in gastric adenocarcinomas of higher stages (111), while in CRC and PDAC the expression of this proteoglycan has been shown to be upregulated, suggesting its possible application as a biomarker (112, 113).

Heparan sulfate GAG chains can also be carried by glypicans, a family of glycosylphosphatidylinositol (GPI)-anchored proteoglycans. Glypicans have been shown to bind a wide range of signaling molecules and to regulate the signaling of the Wnt, Hedgehog, fibroblast growth factor, and bone morphogenetic protein (BMP) pathways (114). Glypican-1 has been shown to be overexpressed in PDAC cell models and patient tumors (115). Moreover, the key role of glypican-1 in PDAC progression has been well documented using mouse models (116, 117). Recently, glypican-1 has been shown to be specifically expressed by cancer circulating exosomes and, therefore, to have potential to be used as a minimal-invasive diagnostic and screening tool to detect early PDAC stages (118).

The CD44 proteoglycan has also been on the focus of tumor biology research because the expression of specific splice variants is strongly associated with malignancy. Specifically, the exon v6-containing CD44 isoform (CD44v6) is highly expressed in premalignant and malignant gastric lesions (119). Modification of CD44v6 with STn was demonstrated in gastric mucosa and serum of cancer patients, indicating its potential as a biomarker

for early diagnosis of gastric tumors (120). Different strategies aiming the impairment of CD44-dependent cancer cell migration have been proposed. The ceramide nanoliposome (CNL) was shown to induce anoikis and to limit metastasis by inducing lysosomal degradation of CD44 in PDAC cells (121).

Another glycosylation modification frequently observed in cancer is the altered sialylation of glycosphingolipids that can lead to the appearance of tumor-associated antigens. The human plasma membrane-associated sialidase NEU3, which catalyzes the removal of sialic acids from glycoproteins and glycolipids, is a key enzyme for ganglioside degradation. NEU3 has been shown to be overexpressed in many tumors, including CRC (122). Modulation of ganglioside expression by increased NEU3 activity has been proposed as a mechanism of protection against programed cell death and has, therefore, a critical implication in therapeutic strategies (123). Recently, NEU3 was demonstrated to regulate the Wnt signaling pathway, therefore contributing for the malignant transformation of CRC cells (124). These findings suggest NEU as a relevant target for diagnosis and therapy of CRC. During CRC progression, besides reduced expression of sialylated gangliosides, overall alteration in glycosphingolipids glycosylation includes increased fucosylation, decreased acetylation and sulfation, and reduced expression of globo-series glycans (125).

GLYCAN CANCER BIOMARKERS

Glycosylation changes on glycoconjugates either expressed on the cell surface or secreted by cancer cells are potential sources of cancer biomarkers. The overexpression of these altered glycosylated structures and the loss of polarity of carcinoma cells lead to the shedding of glycoconjugates with altered glycosylation into the circulation. Currently, several serological assays used in the clinics are based on the quantification of glycoconjugate levels in the serum of cancer patients. Most of these biomarkers have been useful for prognostic and monitoring purposes. These include well-established serological biomarkers, such as the CA15-3 assay, detecting mucin MUC1 glycoprotein used for breast cancer (126–130), the CA125 assay, which detects the circulating mucin MUC16 in ovarian cancer (131, 132), and the prostate-specific antigen (PSA), which is used to detect prostate diseases (133, 134).

Regarding GI cancer, one of the most used serological assay detects the SLe^a carbohydrate antigen. SLe^a is present on circulating glycolipids and glycoproteins and is detected by the CA19-9 assay. This serological assay is applied in patients with a previously established diagnosis of PDAC, CRC, gastric, or biliary cancers and used to monitor their clinical response to therapy (135–138).

Another important serological test used in the clinics for GI tumors is the carcinoembryonic antigen (CEA) assay, which detects the CEA glycoprotein produced by carcinoma cells. In GI cancer, CEA is expressed at high levels and shed into the bloodstream being useful for prognosis evaluation and follow-up of these patients (129, 137, 139, 140).

In general, most of these serological assays have primarily been useful for prognosis and patients' monitoring applications. Unfortunately, some of these biomarkers can also be detected due to benign lesions or other factors, such as smoking, which has limited their use in cancer screening strategies for diagnostic purposes. Given the usually late diagnosis of GI cancer, highly specific serum markers for cancer detection and screening are highly needed. Recent developed strategies and advanced technologies are contributing to the definition of novel and more specific glycoconjugate targets. Several of these new targets are currently evaluated and hold potential for improving the cancer detection and early diagnosis.

INNOVATIVE GLYCOBIOLOGICAL STRATEGIES

The difficulty of glycobiological research lies in the intrinsic complexity of glycosylation and its versatile conjugates. Whereas genomic and proteomic analysis made a leap forward by DNA sequencing and mass-spectrometric protein sequencing, respectively, that enabled the reading of a linear code with limited number of variabilities; for the more complex glycans, no comparable tool exists.

Nevertheless, the recent years have brought up many innovative approaches and methods that enable the unraveling of glycobiological challenges. With the development of glycoengineered cell strategies, glycan complexities have been reduced and the effects of specific glycan epitopes have been pinpointed. On the other hand, analytical methods and protocols for glycomic and glycoproteomic analyses have improved and new approaches, such as the adaptation of the array technology on glycans and lectins or novel antibody-based assays, have accelerated the acquisition of glycobiological knowledge. The following sections discuss several promising strategies in the glycobiology field.

Glycoengineered Cell Line Models

The characterization of the function of glycans in cancer has been a major challenge in the field due to technical difficulties related to the complexity and heterogeneity of glycans synthesized in eukaryotic cells.

Genetic engineered cell models have been developed to study the functions of specific glycan epitopes in cancer. Some of these models include the overexpression of glycosyltransferases, which has allowed the characterization of the biosynthesis and function of simple cancer-associated carbohydrate epitopes, such as Tn, STn, T, and ST (17, 18, 20, 141, 142). Similar strategies have used stably transfected cell lines with glycosyltransferases to characterize the function of branched glycan structures (52, 56) as well as terminal sialylated/fucosylated structures frequently overexpressed by cancer cells, as previously explained in Section "Increased Sialylation and Fucosylation" (67, 143, 144).

Another major challenge in the field was related to the identification of structures at individual glycosylation sites. Major efforts have been done in this discipline with the generation of site-specific mutants of important proteins in cancer. One example is the use of site-specific mutants of *N*-glycosylation sites of the human epidermal growth factor receptor. This strategy has allowed the demonstration that Asn-420-linked oligosaccharide chain in this receptor interferes with its activation in cancer cell lines (145). Another cell line model has addressed the role of

E-cadherin *N*-glycosylation sites in gastric cancer (146–148). The use of E-cadherin constructs engineered to lack specific *N*-glycosylation sites has demonstrated the effect of specific *N*-glycosylation structures on cell adhesion (149, 150).

The recent use of genomic editing tools has allowed the development of isogenic cell systems that along with extensive application of mass spectrometry (MS) methods is utilized for high-throughput site-specific *O*-Glycosylation (*O*-GalNAc and *O*-Mannose) proteomics. These technologies have enabled the precise determination of protein *O*-glycosylation sites in cells (151, 152). These strategies have greatly evolved in the past years and are showing vast potential in the glycobiology field. One approach has used the zinc-finger nucleases targeting the knockout of *COSMC* gene and has been applied in several human cancer cell lines originated from different organs (152). These so-called SimpleCell models produce stable cells expressing homogeneous truncated *O*-glycosylation with Tn and/or STn *O*-glycans (24, 120, 151, 153).

These cell models have provided a source of unlimited material for isolation and identification of GalNAc *O*-glycopeptides from cell lysates or secretomes using lectin chromatography followed by advanced MS, enabling the identification of hundreds of unique *O*-glycoproteins and *O*-glycosylation sites in several cell line models from different tissues (120, 153, 154). In addition, this approach has provided a versatile method for the functional analysis of different ppGalNAc-Ts (153, 155). Furthermore, similar strategies have been applied targeting the *O*-mannose glycoproteome. To reduce the structural heterogeneity of *O*-mannosylation (*O*-Man), the nuclease-mediated gene editing of a human cell line was performed by zinc-finger nuclease targeting of the *POMGNT1* gene. This gene encodes for the enzyme POMGnT1 that controls the first step in the elongation of *O*-Man glycans. The *O*-Man glycoproteome has been characterized using both chromatography and advanced MS (156).

The knowledge of *O*-glycosites in specific cancer cell types allows for the analysis of novel biological functions of glycosylation and for potential cancer cell-specific *O*-glycosites. This is particularly important given the complexity of *O*-glycosylation and that the various ppGalNAc-Ts that control the protein *O*-glycosylation sites may determine large variation at protein, cell, and tissue levels (9).

Glycomic Strategy

Glycomics is the study of all glycan structures of a given cell, tissue, or organism. The intrinsic complexity of glycan structures and their versatile conjugates render this field particularly challenging. Due to the constant advancement of analytical instruments and methods, the *N*- and *O*-glycomic characterization of cancer cell lines, tumors, and cancer patients' body fluids has rendered possible. Still, there is no single ideal method for this analysis and, thus, today a large variety of analytical methods is available for the glycomic characterization, resulting from different combinations of initial sample preparation, derivatization, glycan separation, and detection. Each method bares advantages and disadvantages.

For the glycomic analysis of cells or tumors, the sample is usually homogenized and proteins are denatured, followed by the release of glycans (**Figure 2A**). The study of the glycans of serum or plasma is more challenging and requires often purification steps for glycoproteins prior to the release of their glycan structures. There are several methods to release glycans from the protein backbone to facilitate their characterization. The release of glycans is not a prerequisite as the analysis of whole glycopeptides is also possible (covered in Section "Glycoproteomic Strategy"). The most prominent technique to release *N*-glycans is by Peptide-*N*-Glycosidase F (PNGase F). The release of *N*-glycans via PNGase F is robust, fast, and efficient and is capable of liberating all types of human *N*-glycan structures. PNGase F-released glycans can be chemically labeled. On the other hand, no enzyme has so far been characterized that enables the efficient release of all types of *O*-glycans. For instance, the enzyme *O*-glycanase releases only core 1 *O*-glycans from their peptide backbone. Therefore, chemical techniques have to be utilized for whole *O*-glycomic analyses, such as reductive β-elimination.

Released glycans can be analyzed after derivatization or in their native form (underivatized). The derivatization of glycans bares several advantages, such as adding fluorescent tags for the photometric detection or chemical modification of side groups to stabilize glycan constituents. Despite these advantages, it is preferred in some cases to work with the native glycan, avoiding several time-consuming preparation steps and sample losses.

Complex mixtures of glycans, as they arise from clinical samples or even cell lines are usually separated by chromatography or capillary electrophoresis [reviewed in Ref. (157)] and detected by fluorescence detector (FLD) or MS. The sensitive and quantitative fluorescence detection requires fluorescently tagged glycans, and gives on its own only limited structural information derived from chromatographic or electrophoretic retention times. MS, on the other hand, can be applied on both native and derivatized oligosaccharides and may yield detailed structural information of the glycans. Since the analytes are not consumed by FLD, a sequential setup with MS is possible and often advantageous.

Three successful glycomic workflows that have revealed in the past years several findings in GI cancer are porous graphitized carbon separation with electrospray ionization and tandem MS (PGC-ESI-MS/MS), hydrophilic interaction ultra/high performance liquid chromatography with FLD (HILIC-FLD-UPLC/HPLC), and matrix-assisted laser desorption/ionization MS (MALDI-MS).

Porous graphitized carbon separation with electrospray ionization and tandem MS is a workflow used for both *N*- and *O*-glycomic analyses. First, the *N*-glycans are liberated from the glycoproteins with PNGase F, followed by reductive β-elimination of the glycoproteins to release the remaining O-glycans. The *N*- and *O*-glycans are separated by liquid chromatography with a PGC column, which resolves most isomeric structures and complements, therefore, ideally the subsequent MS and MS/MS structural analysis. A recent glycomic study by PGC-ESI-MS/MS has described structural glycan alterations in CRC, including several unique glycans found solely in the tumor region and indicated a correlation between EGFR expression and sialylation in CRC (158). This method has lately been further utilized for the *N*- and *O*-glycomic characterization of CRC cell lines and tumors, revealing great *O*-glycomic differences between tumors and all tested cell line models (26).

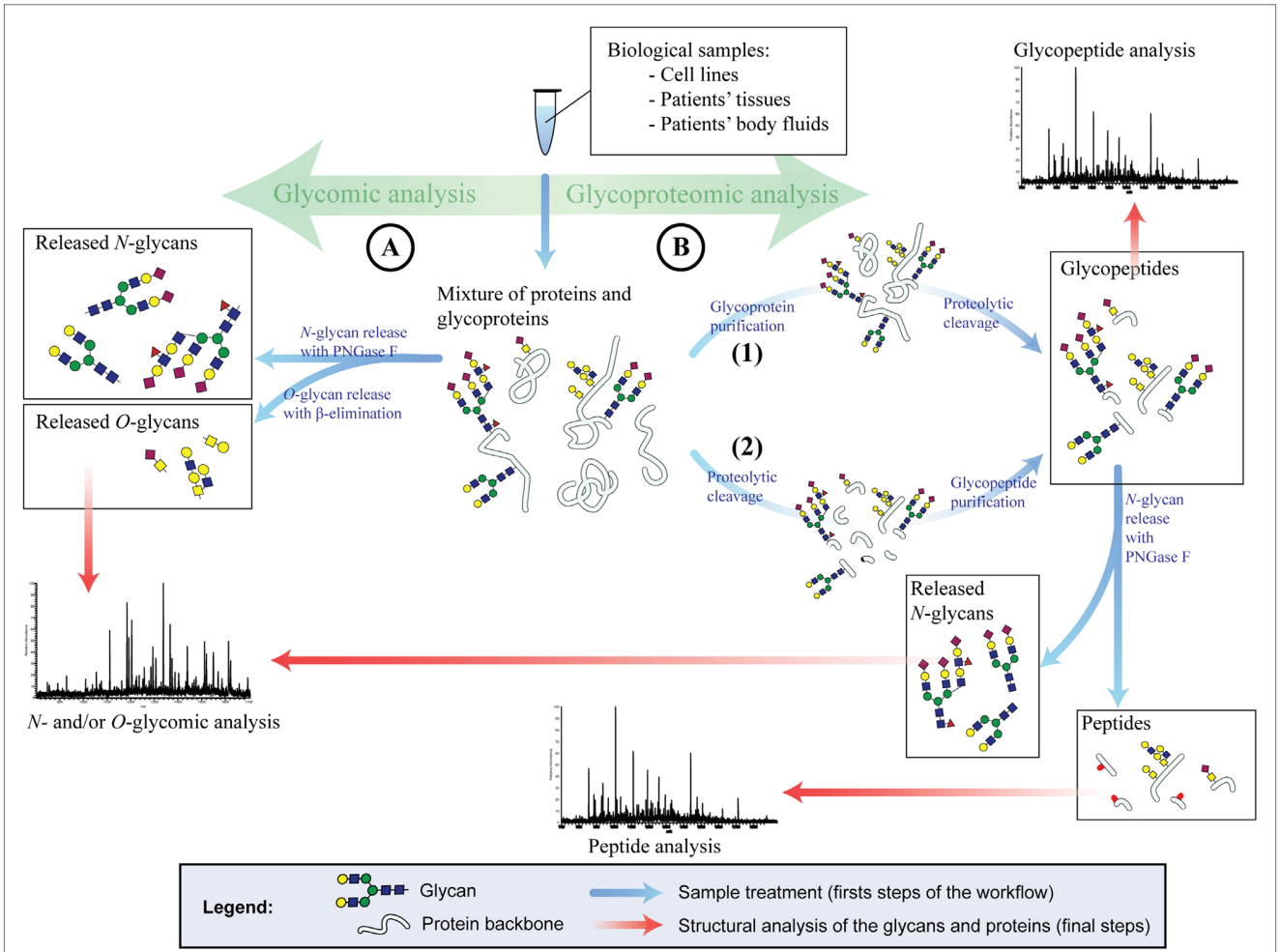


FIGURE 2 | Glycomic and glycoproteomic strategies for glycan analysis. **(A)** Glycomic analysis. *N*- and *O*-glycans are PNGase F or chemically released, respectively. Once the glycans are released, they are structurally analyzed (usually by mass spectrometry). **(B)** Glycoproteomic analysis. The glycoproteins are either purified from a complex mixture and then enzymatically digested (1) or first digested and then purified (2). Method 1 will result in glycopeptides and non-glycosylated peptides originated solely from glycoproteins. Method 2 will result in glycopeptides only. These glycopeptides can be directly analyzed to generate site-specific glycan structures. Alternatively, the glycans of the glycopeptides can be released as described in **(A)** and the peptides are analyzed.

Hydrophilic interaction ultra/high performance liquid chromatography with FLD is used for the quantitative profiling of *N*-glycans. The *N*-glycans are released by PNGase F and reductively aminated with a fluorophore. The labeled *N*-glycans are applied on a HILIC-ultra performance chromatography (UPLC), which separates the glycans according to their size and monosaccharide composition. The retention time can be converted to glucose units (GU) by comparing it with a dextran ladder, yielding reproducible results. Hence, due to the few sample preparation steps, the high recovery of the HILIC column, the quantitative detection via the fluorescent tag, and the possibility of multiplexing, this analysis can be applied for large-scale *N*-glycomic studies. The HILIC-FLD-UPLC *N*-glycomic analysis has recently been applied in a large-scale discovery study on serum of gastric cancer patients revealing an increase in certain SLe^x carrying *N*-glycan structures that correlated with disease progression. Furthermore, in this study other structures, such as bisected *N*-glycans, have been shown to decrease with disease progression (159, 160).

Recently, the combination of HILIC-FLD-UPLC and PGC-ESI-MS/MS has been used for *N*- and *O*-glycomic analysis of a gastric cancer cell line overexpressing the sialyltransferase ST3Gal-IV (79). This cell line has previously been shown to present a more invasive phenotype (67). The glycomic analysis revealed a broad range of cancer-associated alterations, such as decreased bisected and increased branched structures, truncation of *O*-glycans, and a shift from α2,6- to α2,3-sialylated *N*-glycans (79).

The use of MALDI-MS is another very successful approach of analyzing the *N*-glycome of clinical samples, such as body fluids. MALDI is based on a laser impulse that excites a solid matrix in which the analytes are embedded which in turn desorbs and ionizes the analytes for MS analysis. MALDI is relatively tolerant

to salt and other contaminants, which allows uncomplicated sample preparation after the release of *N*-glycans. This method has recently been applied on the serum of gastric cancer patients and control groups and has been able to identify *N*-glycomic differences between the serum of gastric cancer patients and that of non-atrophic gastritis patients (161). In a large-scale study on sera of PDAC patients, a tendency toward higher branched and fucosylated *N*-glycans has been observed when compared to sera from healthy individuals. The major part of the significantly altered *N*-glycan structures were specifically increased in patients with distant metastases and the ratio of the quantity of two glycans has been proposed as a robust diagnostic marker for PDAC (162). Another recent study utilizing MALDI-MS has revealed in pancreatic cyst fluids, of which certain subtypes bare a high risk of undergoing malignant transformation, the hyperfucosylation of *N*-glycans (163).

A broad range of alterations in CRC tissues versus controls have been identified by sequential analyses of fluorescently tagged *N*-glycans by HILIC-FLD-HPLC and MALDI-MS. Additionally, multivariate statistical evaluation and further MS-based structure elucidation have been applied and revealed among others the decrease of bisected structures and the increase of glycans with sialylated lewis epitopes. Furthermore, abnormal core-fucosylated high mannose *N*-glycans have been uniquely found in cancer tissue (37).

Glycoproteomic Strategy

Glycoproteomics is the study of proteins that carry glycan modifications. It usually focuses on the identification and quantification of glycoproteins and the characterization of protein glycosylation sites. Given that most clinical cancer biomarkers are glycoproteins, this field is particularly promising for the identification of new biomarker targets in cancer. Biological samples, such as cell lines, tissues, and body fluids, can be analyzed. However, the glycoproteomic analysis of complex biological samples, such as tissues or sera, is analytically challenging due to the large complexity and vast dynamic range of concentrations of glycoproteins.

The glycoproteomic pipeline typically consists of numerous steps, such as glycoprotein or glycopeptide enrichment, isotopic labeling (optional), multidimensional protein or peptide separation, tandem mass-spectrometric analysis, and bioinformatic data interpretation (**Figure 2B**). In cancer, the vast majority of glycoproteomic findings are based on bottom-up analysis of peptides ("shotgun proteomics"). For this purpose, glycoproteins are proteolytically cleaved into glycopeptides before or after the enrichment step. The enrichment of glycoproteins or glycopeptides is a critical step of the glycoproteomic analysis. Even though this field is rapidly evolving, so far no method has been established that captures unbiased every glycoprotein or glycopeptide and enables full glycoproteomic coverage. Currently, most popular enrichment methods are based on lectins (164–166) or on hydrazide solid-phase extraction (167, 168) and sometimes applied in combination to increase the glycoproteomic coverage (168). Alternative strategies are boronic acid functionalized beads (169), size exclusion chromatography (170), hydrophilic interaction (171), and graphite powder micro column (172). Due to the difficulties of covering the whole glycoproteome many cancer studies pursue a different strategy of enriching specifically glycoproteins and glycopeptides carrying cancer-relevant glycan epitopes, such as sialic acids or sialylated Lewis epitopes. These methods are usually based on lectins [such as SNA, WGA, and MAL (173)], antibodies (159), enrichment by titanium dioxide (79, 174), or affinity purification of metabolic labeled glycoproteins (175–177). After the enrichment and proteolytic digestion (not necessarily in this order), glycopeptides may be deglycosylated and are multidimensional separated via chromatography and/or electrophoresis and analyzed by tandem MS. The deglycosylation is a requirement of some enrichment methods, such as hydrazide solid-phase extraction, but may be also applied for all *N*-glycoproteomic analysis. The PNGase F release of *N*-glycans leads to the conversion of the *N*-glycan carrying asparagine to aspartic acid and can, thus, be spotted on the peptide backbone by MS. For the generation of site-specific structural information of *N*- and *O*-glycans, whole glycopeptides are analyzed utilizing a combination of different MS fragmentation methods or collision energies that either fragment peptides or glycans (178–180). This strategy is being optimized in recent years and bares great potential for the discovery of new cancer biomarkers because it unravels site-specific glycan alterations in cancer.

Glycoproteomic analyses have been applied in GI cancer mainly for the identification of biomarkers, such diagnostic biomarkers or biomarkers for multidrug resistance in gastric cancer (181, 182). Glycoproteomics in combination with glycoengineered cell line models was in recent years able to increase the coverage of *O*-glycosylated proteins and to identify numerous novel *O*-glycosylation sites in gastric cancer and PDAC, generating several new potential biomarkers (24, 120).

Other Glycoanalytical Techniques

MS-based glycomic and glycoproteomic analyses require expensive equipments and a fair amount of expertise. MS-independent methods, such as glycoprotein, antibody-lectin-sandwich, and lectin arrays, are capable of rapid data acquisition of glycomic alterations in cancer samples. Glycan arrays, on the other hand, enable a screening for specificities of glycan-binding proteins, improving the data interpretation of antibody and lectin-based research. Regarding tumor biology, it is very relevant to determine not only the glycosylation modifications harbored by tumor cells but also to disclose the topographic distribution of these alterations within the tumor and adjacent tissue. Novel approaches for the identification of *in situ* glycan modification of specific proteins include proximity ligation assay (PLA) and imaging mass spectometry (IMS).

Arrays

The binding of biological molecules to solid matrixes was an idea first described by Chang in 1983 (183). This technology initially consisted of coating glass cover slips with different antibodies in close proximity forming a matrix-like array. Arrays recognize partners from large amounts of biological material using high-throughput screening miniaturized, multiplexed and parallel processing, and detection methods based on multiple probes covalently attached to a solid substrate. Depending on the molecule that is deposited on the surface,

different microarrays exist. To analyze glycan-containing structures, the most common classification is glycan, glycoprotein, or lectin microarray, and also a variant of the latter called antibody-lectin sandwich array (**Figure 3**) (184). The advantages that the microarray technology offers are the small volume of sample required for the analyses, the high reproducibility, and the reduced cost and time to process many samples. Therefore, microarray platforms have been highlighted by its extensive application in the field of biomarker validation, where a large number of samples must be analyzed multiple times (185). Moreover, depending on the type of microarray assay performed, information about the glycan-linkage configuration can be obtained.

Glycan microarrays are used mainly to characterize the binding specificities and affinities of proteins (mostly antibodies and lectins) toward glycans (**Figure 3A**) (186–188). However, they can also be applied for the screening of inhibitors of carbohydrate-mediated interactions and of sugar interactions of an entire organism, such as a whole cell or virus (185, 189, 190). Current available platforms consist of approximately 20,000 microspots of antigens reaching the capacity to include most known human microbial pathogens, autoantigens, and tumor-associated antigens (191–193). The diversity and scope of glycan arrays are continuously increasing allowing a better characterization of glycan-binding proteins but leading to more complex data. Different software tools are currently available for data interpretation (194–196). Glycan arrays present oligosaccharides that were either purified from a biological source or *de novo* synthesized. Regarding the latest ones, it is important to highlight recent works describing new methodologies that allow sialic acid (197, 198) and GAG synthesis (199).

As an alternative to the direct binding of glycans to the array surface, glycans can be presented on proteins or peptides that are attached to the array. A recent advancement in this approach is the coiled coil-based technology, which allows the presentation of the antigens at high densities while mimicking the *in vivo* orientation attached to a fiber-forming peptide. This platform showed increased sensitivity for the identification of antibodies against parasitic glycan antigens and might be adapted in the future for cancer diagnostic (200).

Glycoprotein microarrays are based on printing purified or enriched glycoproteins onto the slides and screening these proteins for glycan epitopes using different lectins or glycan-recognizing antibodies (**Figure 3B**). This approach is usually followed by analytical techniques to identify the spotted proteins and to verify the glycan epitopes found by the array analysis. A recently performed glycoprotein array analysis of lectin-enriched sera from PDAC patients, chronic pancreatitis patients (benign pancreatic disease), and healthy individuals has correctly clustered these three groups, being the PDAC group significantly different from the other two (201). In addition, the glycoprotein microarray may use synthesized peptides and recombinant protein fragments that have been *in vitro* glycosylated for the detection of human autoantibodies (202, 203).

Lectin microarrays, where different lectins are spotted onto the slide, enable a rapid and high-sensitivity profiling of glycan

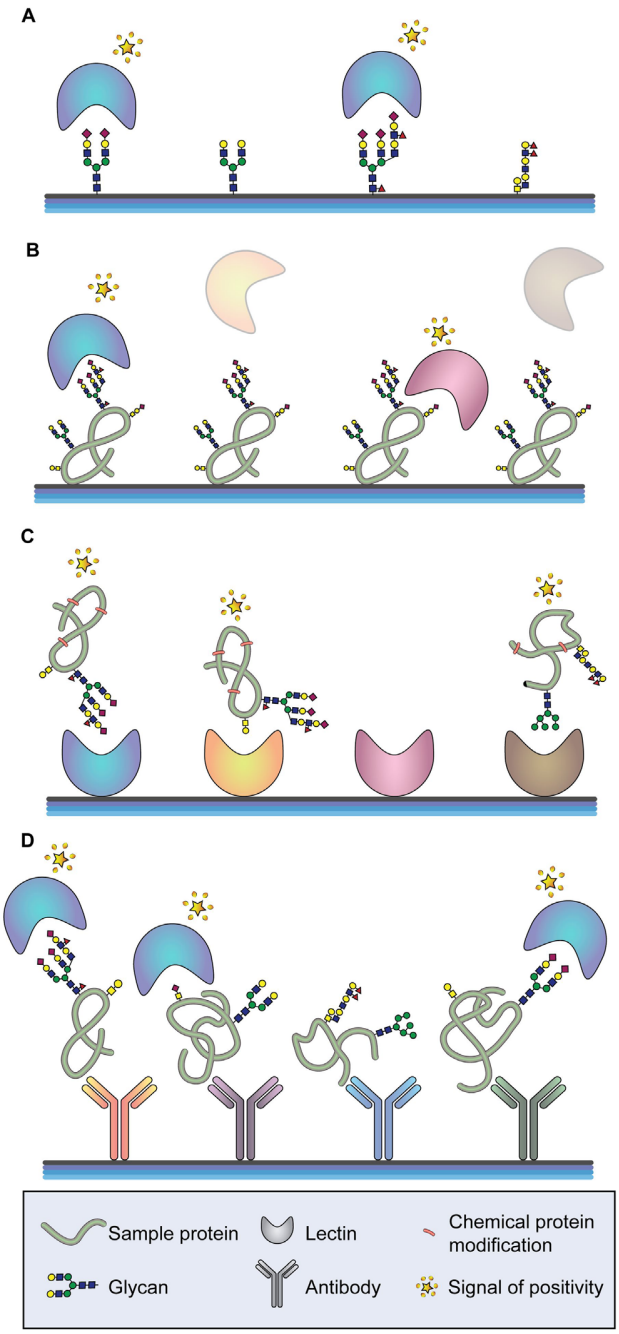


FIGURE 3 | Arrays for glycan analysis. **(A)** Glycan arrays. Different glycan moieties are spotted onto the array to determine specificities of labeled glycan-binding molecules (typically lectins or antibodies). **(B)** Glycoprotein arrays. Glycoproteins are enriched from a sample and spotted onto the array. The glycosylation moieties of the spotted glycoproteins are determined by screening with different glycan-binding molecules. **(C)** Lectin arrays. A series of lectins with well-defined glycan-binding properties are spotted onto the array and different labeled proteins are tested. **(D)** Antibody-lectin sandwich arrays. Multiple antibodies for a series of proteins of interest are spotted onto the array, and the glycan epitopes of the captured proteins are probed using labeled lectins and glycan-binding antibodies.

features found in complex samples, such as cells, tissues, body fluids, and synthetic glycans and their mimics (**Figure 3C**) (204, 205). Lectin arrays offer a general view of the glycan structures on a complex sample and integrate the information from all proteins with the disadvantage that no information about specific glycan changes of the respective protein constituents will be obtained (206). A recent work displayed different glycopatterns in gastric cancer compared to gastric ulcer applying Cy3-labeled proteins extracted from tissues to lectin microarrays (207). Another recent lectin array approach has identified differences in α2-macroglobulin glycosylation between healthy individuals and patients with CRC. The spotted serum purified α2-macroglobulin has displayed, among other changes, significant differences in the content of branched *N*-glycans and α2,6 sialylation (208).

A variant of the lectin array is the antibody-lectin sandwich array. Antibodies to known glycoproteins are spotted on a solid support, and complex glycoprotein samples, which can be crude or prefractionated, are bound to the microarray (**Figure 3D**) (184, 185). The glycosylation of the captured target proteins are then screened by labeled lectins and glycan-specific antibodies. Antibody-lectin sandwich arrays are highly effective for profiling variation in specific glycans on multiple target proteins. Performing this technology, specific glycoforms of MUC5AC and endorepellin glycoproteins in the cyst fluid of patients with precancerous pancreatic cysts have been found (209). In these assays, the glycoprotein nature of the antibodies must be considered and different approaches to prevent glycan recognition of the antibodies by the secondary antibody or lectin applied exist. The chemical derivatization of the glycans of the spotted antibodies also prevents their ability to be recognized by glycan-binding molecules, both antibodies and lectins. One efficient method to study glycans on individual proteins from complex mixtures uses chemically derivatized capture antibodies and tests the glycosylation of captured target proteins by lectins and glycan-binding antibodies. Applying this approach, cancer-associated glycan alterations on the proteins MUC1 and CEA in the serum of PDAC patients have been identified (210). Another strategy consists on producing recombinant antibodies in organisms that do not carry out post-translational modification, such as glycosylation. This approach has been performed for the detection of glycans linked to CEA by ELISA coating the microplate with recombinantly scFv expressed in *Escherichia coli* and using lectins as detection probes (211).

Regarding GI cancer research, arrays have been widely used to discover new biomarkers consisting of proteins bearing aberrant glycosylation that could lead to a more accurate diagnostic.

In situ Proximity Ligation Assay

The association of the glycan expression and location with clinical and molecular characteristics of cancer tissues has rendered possible by histochemistry techniques using glycan-binding antibodies or lectins (205). However, one major limitation of this technique is the lack of capacity to identify the proteins *in situ* on which these glycan motifs are localized. This limitation has been surpassed by the development of the *in situ* proximity ligation assay (PLA) (**Figure 4**) (212, 213).

This sensitive antibody-based method reveals the colocalization between specific proteins and specific glycan structures in tissues and cell samples. The identification of protein glycoforms is of utmost importance for the understanding of glycobiological cellular processes in cancer. PLA could also detect other post-translational modifications of proteins in tissue samples with subcellular resolution. The PLA technology is based on the binding of two specific PLA probes, each containing a unique oligonucleotide, to two targets of interest (**Figure 4A**). Antibodies, lectins, and other binding proteins can act as probes. A ligation solution, containing bridging oligonucleotides and a ligase, will hybridize the oligonucleotides of the PLA probes if they are in close molecular proximity to form a closed circle (**Figure 4B**). This closed nucleotide circle will be amplified by a DNA polymerase generating repeated copies of the circular DNA strands. Finally, fluorescent or chromogenic oligonucleotides hybridize to the amplification product and can be detected as individual spot by microscopy (**Figure 4C**) (213). The first study using this innovative PLA strategy applied to glycobiology has showed that the mucin MUC2 is a major carrier of the cancer-associated STn glycan antigen both in intestinal metaplasia and gastric carcinoma (214). The use of *in situ* PLA for the identification of a mucin glycosylation profile in cancer lesions is being extended, opening new opportunities for the development of novel diagnostic and prognostic markers. One recent study screened for tissue-specific aberrant mucin glycoforms in mucinous adenocarcinomas from different organs (stomach, ampulla of Vater, colon, lung, breast, and ovary). In GI tissues mucins carrying a set of truncated, simple O-glycans and sialylated Lewis antigens have been detected by this approach (215).

More recently, PLA has been used in combination with different glycoproteomics strategies to identify specific glycoforms as potential biomarkers in gastric cancer, leading to the identification of CD44v6/STn (120) and RON/SLe^x (79).

The PLA technique will further improve our understanding of specific protein glycosylation changes that occurs in cancer tissues and that could be applied in clinic as new markers for GI cancer progression (216).

Imaging Mass Spectrometry

Imaging mass spectrometry is a very novel and promising technology that was first developed in 1997 by Caprioli and colleagues for the analysis of proteins (217). This method is based on MALDI-MS and utilizes the laser ionization of a localized area for the two-dimensional screening of a tissue sample. IMS generates for each ionization point of the tissue a spectra that yields structural information and, thus, reveals the spatial distribution of analytes (**Figure 5**). Recently, this technique has been adapted for glycomic analysis and has allowed to create *N*-glycosylation maps of several different frozen tissue (218). Following, *N*-glycan IMS has been also applied on formalin-fixed paraffin-embedded tissue (219). The *N*-glycan IMS workflow consists of four steps. First, *N*-glycans are liberated by PNGase F incubation of the deparaffinized or thawed tissue slide. Second, a thin layer of MALDI matrix is sprayed on top of the tissue slide. Third, the slide is two-dimensionally screened by multiple MALDI-MS analysis. Lastly, each identified *N*-glycan structure

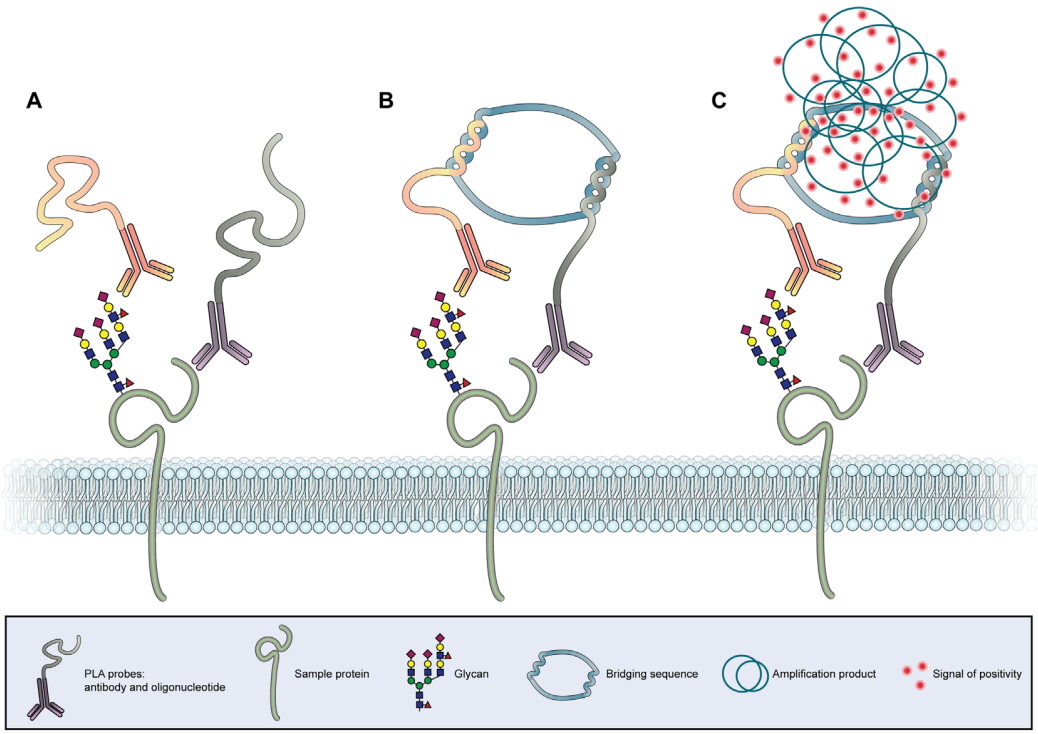


FIGURE 4 | ***In situ* proximity ligation assay**. **(A)** Two antibodies, one specific for the protein backbone and another specific for a glycan epitope, are conjugated with two different oligonucleotide chains (PLA probes). **(B)** In case the antibodies bind to molecules in close proximity a bridging sequence links the two oligonucleotide sequences. **(C)** A subsequent polymerase induced amplification in combination with labeled nucleotides leads to the formation of a fluorescent or chromogenic signal at the co-expression site of the protein and glycan, allowing the *in situ* detection of the PLA signal.

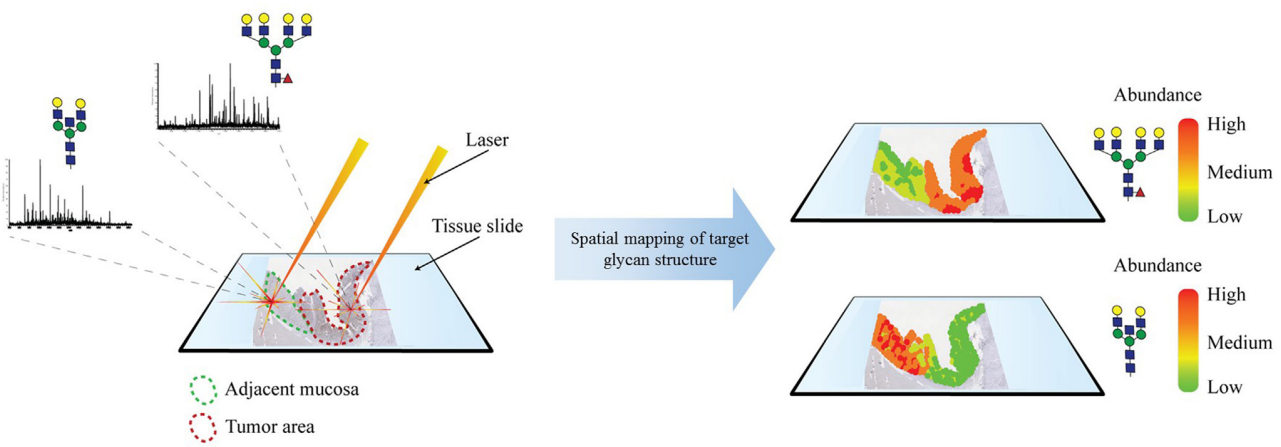


FIGURE 5 | **Imaging mass spectrometry**. MALDI-based ionization generates an MS glycan spectrum of a selected spot of a tissue section. This process is repeated throughout the tissue. Identified glycan structures are mapped on the tissue and can be correlated with histopathological features.

can be computationally visualized on the tissue, generating an epitope map.

IMS applied on hepatocellular carcinomas has shown to be capable of spatially defining glycan compositions and distinguishing malignant tissue from healthy tissue (220). Preliminary IMS results on other tumors, such as PDAC, have been able to differentiate between histopathological areas, such as fibroconnective tissue (220).

The IMS application on *N*-glycan analysis is still in its early stages of development, but bares enormous potential as a next-generation *N*-glycomic tumor characterization tool.

FUTURE PERSPECTIVES AND CLINICAL APPLICATIONS

The recent advances in the glycomic and glycoproteomic fields are currently providing crucial information on the understanding of the role that glycans play in the biology of cells, tissues, and organisms, both in physiological and pathological conditions. However, many issues still remain to be understood, particularly in complex diseases, such as cancer. Advances in the glycobiology field could contribute to disclose key information regarding cancer biological properties, including the identification of prognostic and therapeutic response biomarkers.

In addition, the recent developments in this field could contribute to overcome the limitations of the current serological assays. The set of novel strategies presented in this review provide a clear view for future validation of potential biomarkers and points toward the translation of these strategies in the clinical setting.

AUTHOR CONTRIBUTIONS

All authors have contributed to the conception, drafting, and revision of the manuscript. SM and MB designed the figures. All authors approved the review final form.

FUNDING

We acknowledge the support from the European Union, Seventh Framework Programme, Gastric Glyco Explorer initial training network: grant number 316929. IPATIMUP integrates the i3S Research Unit, which is partially supported by FCT, the Portuguese Foundation for Science and Technology. This work is funded by FEDER funds through the Operational Programme for Competitiveness Factors-COMPETE (FCOMP-01-0124-FEDER028188) and National Funds through the FCT-Foundation for Science and Technology, under the projects: PEst-C/SAU/LA0003/2013, PTDC/BBB-EBI/0786/2012, and PTDC/BBB-EBI/0567/2014. AM acknowledges the grant received from FCT, POPH (Programa Operacional Potencial Humano), and FSE (Fundo Social Europeu) (SFRH/BPD/75871/2011). MB acknowledges the University of Girona for pre-doctoral fellowship.

REFERENCES

1. Reis CA, Osorio H, Silva L, Gomes C, David L. Alterations in glycosylation as biomarkers for cancer detection. *J Clin Pathol* (2010) **63**(4):322–9. doi:10.1136/jcp.2009.071035
2. Pinho SS, Reis CA. Glycosylation in cancer: mechanisms and clinical implications. *Nat Rev Cancer* (2015) **15**(9):540–55. doi:10.1038/nrc3982
3. Stowell SR, Ju T, Cummings RD. Protein glycosylation in cancer. *Annu Rev Pathol* (2015) **10**:473–510. doi:10.1146/annurev-pathol-012414-040438
4. Rabinovich GA, Croci DO. Regulatory circuits mediated by lectin-glycan interactions in autoimmunity and cancer. *Immunity* (2012) **36**(3):322–35. doi:10.1016/j.immuni.2012.03.004
5. Boligan KF, Mesa C, Fernandez LE, von Gunten S. Cancer intelligence acquired (CIA): tumor glycosylation and sialylation codes dismantling antitumor defense. *Cell Mol Life Sci* (2015) **72**(7):1231–48. doi:10.1007/s00018-014-1799-5
6. Sangadala S, Bhat UR, Mendicino J. Quantitation and structures of oligosaccharide chains in human trachea mucin glycoproteins. *Mol Cell Biochem* (1992) **118**(1):75–90. doi:10.1007/BF00249697
7. Kudelka MR, Ju T, Heimburg-Molinaro J, Cummings RD. Simple sugars to complex disease – mucin-type O-glycans in cancer. *Adv Cancer Res* (2015) **126**:53–135. doi:10.1016/bs.acr.2014.11.002
8. Brockhausen I. Mucin-type O-glycans in human colon and breast cancer: glycodynamics and functions. *EMBO Rep* (2006) **7**(6):599–604. doi:10.1038/sj.embor.7400705
9. 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**(6):736–56. doi:10.1093/glycob/cwr182
10. Clausen H, Bennett EP. A family of UDP-GalNAc: polypeptide N-acetylgalactosaminyl-transferases control the initiation of mucin-type O-linked glycosylation. *Glycobiology* (1996) **6**(6):635–46. doi:10.1093/glycob/6.6.635
11. 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* (2009) **57**(1):79–86. doi:10.1369/jhc.2008.952283
12. Shibao K, Izumi H, Nakayama Y, Ohta R, Nagata N, Nomoto M, 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* (2002) **94**(7):1939–46. doi:10.1002/cncr.10423
13. Taniuchi K, Cerny RL, Tanouchi A, Kohno K, Kotani N, Honke K, et al. Overexpression of GalNAc-transferase GalNAc-T3 promotes pancreatic cancer cell growth. *Oncogene* (2011) **30**(49):4843–54. doi:10.1038/onc.2011.194
14. Gill DJ, Clausen H, Bard F. Location, location, location: new insights into O-GalNAc protein glycosylation. *Trends Cell Biol* (2011) **21**(3):149–58. doi:10.1016/j.tcb.2010.11.004
15. Hua D, Shen L, Xu L, Jiang Z, Zhou Y, Yue A, et al. Polypeptide N-acetylgalactosaminyltransferase 2 regulates cellular metastasis-associated behavior in gastric cancer. *Int J Mol Med* (2012) **30**(6):1267–74. doi:10.3892/ijmm.2012.1130
16. Dalziel M, Whitehouse C, McFarlane I, Brockhausen I, Gschmeissner S, Schwientek T, et al. The relative activities of the C2GnT1 and ST3Gal-I glycosyltransferases determine O-glycan structure and expression of a tumor-associated epitope on MUC1. *J Biol Chem* (2001) **276**(14):11007–15. doi:10.1074/jbc.M006523200
17. Marcos NT, Pinho S, Grandela C, Cruz A, Samyn-Petit B, Harduin-Lepers A, et al. Role of the human ST6GalNAc-I and ST6GalNAc-II in the synthesis of the cancer-associated sialyl-Tn antigen. *Cancer Res* (2004) **64**(19):7050–7. doi:10.1158/0008-5472.CAN-04-1921
18. Marcos NT, Bennett EP, Gomes J, Magalhaes A, Gomes C, David L, et al. ST6GalNAc-I controls expression of sialyl-Tn antigen in gastrointestinal tissues. *Front Biosci (Elite Ed)* (2011) **3**:1443–55.
19. David L, Nesland JM, Clausen H, Carneiro F, Sobrinho-Simoes M. Simple mucin-type carbohydrate antigens (Tn, sialosyl-Tn and T) in gastric mucosa, carcinomas and metastases. *APMIS Suppl* (1992) **27**:162–72.
20. Pinho S, Marcos NT, Ferreira B, Carvalho AS, Oliveira MJ, Santos-Silva F, et al. Biological significance of cancer-associated sialyl-Tn antigen: modulation of malignant phenotype in gastric carcinoma cells. *Cancer Lett* (2007) **249**(2):157–70. doi:10.1016/j.canlet.2006.08.010
21. Itzkowitz S, Kjeldsen T, Friera A, Hakomori S, Yang US, Kim YS. Expression of Tn, sialosyl Tn, and T antigens in human pancreas. *Gastroenterology* (1991) **100**(6):1691–700.
22. Itzkowitz SH, Yuan M, Montgomery CK, Kjeldsen T, Takahashi HK, Bigbee WL, et al. Expression of Tn, sialosyl-Tn, and T antigens in human colon cancer. *Cancer Res* (1989) **49**(1):197–204.
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 U S A* (2014) **111**(39):E4066–75. doi:10.1073/pnas.1406619111
24. Hofmann BT, Schluter L, Lange P, Mercanoglu B, Ewald F, Folster A, et al. COSMC knockdown mediated aberrant O-glycosylation promotes oncogenic properties in pancreatic cancer. *Mol Cancer* (2015) **14**:109. doi:10.1186/s12943-015-0386-1
25. Ju T, Cummings RD. A unique molecular chaperone Cosmc required for activity of the mammalian core 1 beta 3-galactosyltransferase. *Proc Natl Acad Sci U S A* (2002) **99**(26):16613–8. doi:10.1073/pnas.262438199
26. Chik JH, Zhou J, Moh ES, Christopherson R, Clarke SJ, Molloy MP, et al. Comprehensive glycomics comparison between colon cancer cell cultures and tumours: implications for biomarker studies. *J Proteomics* (2014) **108**:146–62. doi:10.1016/j.jprot.2014.05.002
27. Hung JS, Huang J, Lin YC, Huang MJ, Lee PH, Lai HS, et al. C1GALT1 overexpression promotes the invasive behavior of colon cancer cells through modifying O-glycosylation of FGFR2. *Oncotarget* (2014) **5**(8):2096–106. doi:10.18632/oncotarget.1815
28. Shimodaira K, Nakayama J, Nakamura N, Hasebe O, Katsuyama T, Fukuda M. Carcinoma-associated expression of core 2 beta-1,6-N-acetylglucosaminyltransferase gene in human colorectal cancer: role of O-glycans in tumor progression. *Cancer Res* (1997) **57**(23):5201–6.
29. Remmers N, Anderson JM, Linde EM, DiMaio DJ, Lazenby AJ, Wandall HH, et al. Aberrant expression of mucin core proteins and o-linked glycans associated with progression of pancreatic cancer. *Clin Cancer Res* (2013) **19**(8):1981–93. doi:10.1158/1078-0432.CCR-12-2662
30. Iwai T, Kudo T, Kawamoto R, Kubota T, Togayachi A, Hiruma T, et al. Core 3 synthase is down-regulated in colon carcinoma and profoundly suppresses the metastatic potential of carcinoma cells. *Proc Natl Acad Sci U S A* (2005) **102**(12):4572–7. doi:10.1073/pnas.0407983102
31. Rossez Y, Maes E, Lefebvre Darroman T, Gosset P, Ecobichon C, Joncquel Chevalier Curt M, et al. Almost all human gastric mucin O-glycans harbor blood group A, B or H antigens and are potential binding sites for *Helicobacter pylori*. *Glycobiology* (2012) **22**(9):1193–206. doi:10.1093/glycob/cws072
32. Kenny DT, Skoog EC, Linden SK, Struwe WB, Rudd PM, Karlsson NG. Presence of terminal N-acetylgalactosaminebeta1-4N-acetylglucosamine residues on O-linked oligosaccharides from gastric MUC5AC: involvement in *Helicobacter pylori* colonization? *Glycobiology* (2012) **22**(8):1077–85. doi:10.1093/glycob/cws076
33. Reis CA, David L, Seixas M, Burchell J, Sobrinho-Simoes M. Expression of fully and under-glycosylated forms of MUC1 mucin in gastric carcinoma. *Int J Cancer* (1998) **79**(4):402–10. doi:10.1002/(SICI)1097-0215(19980821)79:4<402::AID-IJC16>3.0.CO;2-6
34. Tian H, Miyoshi E, Kawaguchi N, Shaker M, Ito Y, Taniguchi N, et al. The implication of N-acetylglucosaminyltransferase V expression in gastric cancer. *Pathobiology* (2008) **75**(5):288–94. doi:10.1159/000151709
35. Granovsky M, Fata J, Pawling J, Muller WJ, Khokha R, Dennis JW. Suppression of tumor growth and metastasis in Mgat5-deficient mice. *Nat Med* (2000) **6**(3):306–12. doi:10.1038/73163
36. Taniguchi N, Korekane H. Branched N-glycans and their implications for cell adhesion, signaling and clinical applications for cancer biomarkers and in therapeutics. *BMB Rep* (2011) **44**(12):772–81. doi:10.5483/BMBRep.2011.44.12.772
37. Balog CI, Stavenhagen K, Fung WL, Koeleman CA, McDonnell LA, Verhoeven A, et al. N-glycosylation of colorectal cancer tissues: a liquid chromatography and mass spectrometry-based investigation. *Mol Cell Proteomics* (2012) **11**(9):571–85. doi:10.1074/mcp.M111.011601
38. Dennis JW, Granovsky M, Warren CE. Glycoprotein glycosylation and cancer progression. *Biochim Biophys Acta* (1999) **1473**(1):21–34. doi:10.1016/S0304-4165(99)00167-1
39. Fernandes B, Sagman U, Auger M, Demetrio M, Dennis JW. Beta 1-6 branched oligosaccharides as a marker of tumor progression in human breast and colon neoplasia. *Cancer Res* (1991) **51**(2):718–23.
40. Seelentag WK, Li WP, Schmitz SF, Metzger U, Aeberhard P, Heitz PU, et al. Prognostic value of beta1,6-branched oligosaccharides in human colorectal carcinoma. *Cancer Res* (1998) **58**(23):5559–64.
41. Murata K, Miyoshi E, Kameyama M, Ishikawa O, Kabuto T, Sasaki Y, et al. Expression of N-acetylglucosaminyltransferase V in colorectal cancer correlates with metastasis and poor prognosis. *Clin Cancer Res* (2000) **6**(5):1772–7.
42. Guo H, Nagy T, Pierce M. Post-translational glycoprotein modifications regulate colon cancer stem cells and colon adenoma progression in Apc(min/+) mice through altered Wnt receptor signaling. *J Biol Chem* (2014) **289**(45):31534–49. doi:10.1074/jbc.M114.602680
43. Ishida H, Togayachi A, Sakai T, Iwai T, Hiruma T, Sato T, et al. A novel beta1,3-N-acetylglucosaminyltransferase (beta3Gn-T8), which synthesizes poly-N-acetyllactosamine, is dramatically upregulated in colon cancer. *FEBS Lett* (2005) **579**(1):71–8. doi:10.1016/j.febslet.2004.11.037
44. Park HM, Hwang MP, Kim YW, Kim KJ, Jin JM, Kim YH, et al. Mass spectrometry-based N-linked glycomic profiling as a means for tracking pancreatic cancer metastasis. *Carbohydr Res* (2015) **413**:5–11. doi:10.1016/j.carres.2015.04.019
45. Zhao J, Qiu W, Simeone DM, Lubman DM. N-linked glycosylation profiling of pancreatic cancer serum using capillary liquid phase separation coupled with mass spectrometric analysis. *J Proteome Res* (2007) **6**(3):1126–38. doi:10.1021/pr0604458
46. Gu J, Isaji T, Xu Q, Kariya Y, Gu W, Fukuda T, et al. Potential roles of N-glycosylation in cell adhesion. *Glycoconj J* (2012) **29**(8–9):599–607. doi:10.1007/s10719-012-9386-1
47. Pinho SS, Reis CA, Paredes J, Magalhaes AM, Ferreira AC, Figueiredo J, et al. The role of N-acetylglucosaminyltransferase III and V in the post-transcriptional modifications of E-cadherin. *Hum Mol Genet* (2009) **18**(14):2599–608. doi:10.1093/hmg/ddp194
48. Pinho SS, Osorio H, Nita-Lazar M, Gomes J, Lopes C, Gartner F, et al. Role of E-cadherin N-glycosylation profile in a mammary tumor model. *Biochem Biophys Res Commun* (2009) **379**(4):1091–6. doi:10.1016/j.bbrc.2009.01.024
49. Gu J, Sato Y, Kariya Y, Isaji T, Taniguchi N, Fukuda T. A mutual regulation between cell-cell adhesion and N-glycosylation: implication of the bisecting GlcNAc for biological functions. *J Proteome Res* (2009) **8**(2):431–5. doi:10.1021/pr800674g
50. Yoshimura M, Ihara Y, Matsuzawa Y, Taniguchi N. Aberrant glycosylation of E-cadherin enhances cell-cell binding to suppress metastasis. *J Biol Chem* (1996) **271**(23):13811–5. doi:10.1074/jbc.271.23.13811
51. Kitada T, Miyoshi E, Noda K, Higashiyama S, Ihara H, Matsuura N, et al. The addition of bisecting N-acetylglucosamine residues to E-cadherin down-regulates the tyrosine phosphorylation of beta-catenin. *J Biol Chem* (2001) **276**(1):475–80. doi:10.1074/jbc.M006689200
52. Pinho SS, Figueiredo J, Cabral J, Carvalho S, Dourado J, Magalhaes A, et al. E-cadherin and adherens-junctions stability in gastric carcinoma: functional implications of glycosyltransferases involving N-glycan branching biosynthesis, N-acetylglucosaminyltransferases III and V. *Biochim Biophys Acta* (2013) **1830**(3):2690–700. doi:10.1016/j.bbagen.2012.10.021
53. Pinho SS, Oliveira P, Cabral J, Carvalho S, Huntsman D, Gartner F, et al. Loss and recovery of Mgat3 and GnT-III mediated E-cadherin N-glycosylation is a mechanism involved in epithelial-mesenchymal-epithelial transitions. *PLoS One* (2012) **7**(3):e33191. doi:10.1371/journal.pone.0033191
54. Xu Q, Isaji T, Lu Y, Gu W, Kondo M, Fukuda T, et al. Roles of N-acetylglucosaminyltransferase III in epithelial-to-mesenchymal transition induced by transforming growth factor beta1 (TGF-beta1) in epithelial cell lines. *J Biol Chem* (2012) **287**(20):16563–74. doi:10.1074/jbc.M111.262154
55. Ihara S, Miyoshi E, Ko JH, Murata K, Nakahara S, Honke K, et al. Prometastatic effect of N-acetylglucosaminyltransferase V is due to modification and stabilization of active matriptase by adding beta 1-6 GlcNAc branching. *J Biol Chem* (2002) **277**(19):16960–7. doi:10.1074/jbc.M200673200
56. Zhao Y, Nakagawa T, Itoh S, Inamori K, Isaji T, Kariya Y, et al. N-acetylglucosaminyltransferase III antagonizes the effect of N-acetylglucosaminyltransferase V on alpha3beta1 integrin-mediated cell migration. *J Biol Chem* (2006) **281**(43):32122–30. doi:10.1074/jbc.M607274200
57. Isaji T, Gu J, Nishiuchi R, Zhao Y, Takahashi M, Miyoshi E, et al. Introduction of bisecting GlcNAc into integrin alpha5beta1 reduces ligand binding and down-regulates cell adhesion and cell migration. *J Biol Chem* (2004) **279**(19):19747–54. doi:10.1074/jbc.M311627200
58. Sato Y, Isaji T, Tajiri M, Yoshida-Yamamoto S, Yoshinaka T, Somehara T, et al. An N-glycosylation site on the beta-propeller domain of the integrin alpha5 subunit plays key roles in both its function and site-specific modification by beta1,4-N-acetylglucosaminyltransferase III. *J Biol Chem* (2009) **284**(18):11873–81. doi:10.1074/jbc.M807660200

59. Dall'Olio F, Chiricolo M. Sialyltransferases in cancer. *Glycoconj J* (2001) **18**(11–12):841–50. doi:10.1023/A:1022288022969
60. Jandus C, Boligan KF, Chijioke O, Liu H, Dahlhaus M, Demoulins T, et al. Interactions between Siglec-7/9 receptors and ligands influence NK cell-dependent tumor immunosurveillance. *J Clin Invest* (2014) **124**(4):1810–20. doi:10.1172/JCI65899
61. Laubli H, Pearce OM, Schwarz F, Siddiqui SS, Deng L, Stanczak MA, et al. Engagement of myelomonocytic Siglecs by tumor-associated ligands modulates the innate immune response to cancer. *Proc Natl Acad Sci U S A* (2014) **111**(39):14211–6. doi:10.1073/pnas.1409580111
62. Kim YJ, Varki A. Perspectives on the significance of altered glycosylation of glycoproteins in cancer. *Glycoconj J* (1997) **14**(5):569–76. doi:10.1023/A:1018580324971
63. Dall'Olio F, Malagolini N, Trinchera M, Chiricolo M. Sialosignaling: sialyltransferases as engines of self-fueling loops in cancer progression. *Biochim Biophys Acta* (2014) **1840**(9):2752–64. doi:10.1016/j.bbagen.2014.06.006
64. Lise M, Belluco C, Perera SP, Patel R, Thomas P, Ganguly A. Clinical correlations of alpha2,6-sialyltransferase expression in colorectal cancer patients. *Hybridoma* (2000) **19**(4):281–6. doi:10.1089/027245700429828
65. Gretschel S, Haensch W, Schlag PM, Kemmner W. Clinical relevance of sialyltransferases ST6GAL-I and ST3GAL-III in gastric cancer. *Oncology* (2003) **65**(2):139–45. doi:10.1159/000072339
66. Petretti T, Schulze B, Schlag PM, Kemmner W. Altered mRNA expression of glycosyltransferases in human gastric carcinomas. *Biochim Biophys Acta* (1999) **1428**(2–3):209–18. doi:10.1016/S0304-4165(99)00080-X
67. Gomes C, Osorio H, Pinto MT, Campos D, Oliveira MJ, Reis CA. Expression of ST3GAL4 leads to SLe(x) expression and induces c-Met activation and an invasive phenotype in gastric carcinoma cells. *PLoS One* (2013) **8**(6):e66737. doi:10.1371/journal.pone.0066737
68. Perez-Garay M, Arteta B, Llop E, Cobler L, Pages L, Ortiz R, et al. alpha2,3-sialyltransferase ST3Gal IV promotes migration and metastasis in pancreatic adenocarcinoma cells and tends to be highly expressed in pancreatic adenocarcinoma tissues. *Int J Biochem Cell Biol* (2013) **45**(8):1748–57. doi:10.1016/j.biocel.2013.05.015
69. Perez-Garay M, Arteta B, Pages L, de Llorens R, de Bolos C, Vidal-Vanaclocha F, et al. alpha2,3-sialyltransferase ST3Gal III modulates pancreatic cancer cell motility and adhesion in vitro and enhances its metastatic potential in vivo. *PLoS One* (2010) **5**(9):e12524. doi:10.1371/journal.pone.0012524
70. Bassaganas S, Perez-Garay M, Peracaula R. Cell surface sialic acid modulates extracellular matrix adhesion and migration in pancreatic adenocarcinoma cells. *Pancreas* (2014) **43**(1):109–17. doi:10.1097/MPA.0b013e31829d9090
71. Nakagoe T, Fukushima K, Nanashima A, Sawai T, Tsuji T, Jibiki M, et al. Expression of Lewis(a), sialyl Lewis(a), Lewis(x) and sialyl Lewis(x) antigens as prognostic factors in patients with colorectal cancer. *Can J Gastroenterol* (2000) **14**(9):753–60.
72. Satomura Y, Sawabu N, Takemori Y, Ohta H, Watanabe H, Okai T, et al. Expression of various sialylated carbohydrate antigens in malignant and nonmalignant pancreatic tissues. *Pancreas* (1991) **6**(4):448–58. doi:10.1097/00006676-199107000-00012
73. Peracaula R, Tabares G, Lopez-Ferrer A, Brossmer R, de Bolos C, de Llorens R. Role of sialyltransferases involved in the biosynthesis of Lewis antigens in human pancreatic tumour cells. *Glycoconj J* (2005) **22**(3):135–44. doi:10.1007/s10719-005-0734-2
74. Nakagoe T, Sawai T, Tsuji T, Jibiki MA, Nanashima A, Yamaguchi H, et al. Predictive factors for preoperative serum levels of sialy Lewis(x), sialyl Lewis(a) and sialyl Tn antigens in gastric cancer patients. *Anticancer Res* (2002) **22**(1A):451–8.
75. Amado M, Carneiro F, Seixas M, Clausen H, Sobrinho-Simoes M. Dimeric sialyl-Le(x) expression in gastric carcinoma correlates with venous invasion and poor outcome. *Gastroenterology* (1998) **114**(3):462–70. doi:10.1016/S0016-5085(98)70529-3
76. Baldus SE, Zirbes TK, Monig SP, Engel S, Monaca E, Rafiqpoor K, et al. Histopathological subtypes and prognosis of gastric cancer are correlated with the expression of mucin-associated sialylated antigens: sialosyl-Lewis(a), sialosyl-Lewis(x) and sialosyl-Tn. *Tumour Biol* (1998) **19**(6):445–53. doi:10.1159/000030036
77. Nakamori S, Kameyama M, Imaoka S, Furukawa H, Ishikawa O, Sasaki Y, et al. Increased expression of sialyl Lewisx antigen correlates with poor survival in patients with colorectal carcinoma: clinicopathological and immunohistochemical study. *Cancer Res* (1993) **53**(15):3632–7.
78. Rosen SD, Bertozzi CR. The selectins and their ligands. *Curr Opin Cell Biol* (1994) **6**(5):663–73. doi:10.1016/0955-0674(94)90092-2
79. Mereiter S, Magalhaes A, Adamczyk B, Jin C, Almeida A, Drici L, et al. Glycomic analysis of gastric carcinoma cells discloses glycans as modulators of RON receptor tyrosine kinase activation in cancer. *Biochim Biophys Acta* (2015). doi:10.1016/j.bbagen.2015.12.016
80. Wang MH, Lee W, Luo YL, Weis MT, Yao HP. Altered expression of the RON receptor tyrosine kinase in various epithelial cancers and its contribution to tumourigenic phenotypes in thyroid cancer cells. *J Pathol* (2007) **213**(4):402–11. doi:10.1002/path.2245
81. Thobe MN, Gurusamy D, Pathrose P, Waltz SE. The Ron receptor tyrosine kinase positively regulates angiogenic chemokine production in prostate cancer cells. *Oncogene* (2010) **29**(2):214–26. doi:10.1038/onc.2009.331
82. Logan-Collins J, Thomas RM, Yu P, Jaquish D, Mose E, French R, et al. Silencing of RON receptor signaling promotes apoptosis and gemcitabine sensitivity in pancreatic cancers. *Cancer Res* (2010) **70**(3):1130–40. doi:10.1158/0008-5472.CAN-09-0761
83. McClaine RJ, Marshall AM, Wagh PK, Waltz SE. Ron receptor tyrosine kinase activation confers resistance to tamoxifen in breast cancer cell lines. *Neoplasia* (2010) **12**(8):650–8. doi:10.1593/neo.10476
84. Catenacci DV, Cervantes G, Yala S, Nelson EA, El-Hashani E, Kanteti R, et al. RON (MST1R) is a novel prognostic marker and therapeutic target for gastroesophageal adenocarcinoma. *Cancer Biol Ther* (2011) **12**(1):9–46. doi:10.4161/cbt.12.1.15747
85. Balmana M, Sarrats A, Llop E, Barrabes S, Saldova R, Ferri MJ, et al. Identification of potential pancreatic cancer serum markers: increased sialyl-Lewis X on ceruloplasmin. *Clin Chim Acta* (2015) **442**:56–62. doi:10.1016/j.cca.2015.01.007
86. Rho JH, Mead JR, Wright WS, Brenner DE, Stave JW, Gildersleeve JC, et al. Discovery of sialyl Lewis A and Lewis X modified protein cancer biomarkers using high density antibody arrays. *J Proteomics* (2014) **96**:291–9. doi:10.1016/j.jprot.2013.10.030
87. Park JJ, Lee M. Increasing the alpha 2, 6 sialylation of glycoproteins may contribute to metastatic spread and therapeutic resistance in colorectal cancer. *Gut Liver* (2013) **7**(6):629–41. doi:10.5009/gnl.2013.7.6.629
88. Miyoshi E, Moriwaki K, Terao N, Tan CC, Terao M, Nakagawa T, et al. Fucosylation is a promising target for cancer diagnosis and therapy. *Biomolecules* (2012) **2**(1):34–45. doi:10.3390/biom2010034
89. Miyoshi E, Moriwaki K, Nakagawa T. Biological function of fucosylation in cancer biology. *J Biochem* (2008) **143**(6):725–9. doi:10.1093/jb/mvn011
90. Breborowicz J, Mackiewicz A, Breborowicz D. Microheterogeneity of alpha-fetoprotein in patient serum as demonstrated by lectin affino-electrophoresis. *Scand J Immunol* (1981) **14**(1):15–20. doi:10.1111/j.1365-3083.1981.tb00179.x
91. Wong RJ, Ahmed A, Gish RG. Elevated alpha-fetoprotein: differential diagnosis – hepatocellular carcinoma and other disorders. *Clin Liver Dis* (2015) **19**(2):309–23. doi:10.1016/j.cld.2015.01.005
92. Balmana M, Gimenez E, Puerta A, Llop E, Figueras J, Fort E, et al. Increased alpha1-3 fucosylation of alpha-1-acid glycoprotein (AGP) in pancreatic cancer. *J Proteomics* (2015) **132**:144–54. doi:10.1016/j.jprot.2015.11.006
93. Trinchera M, Malagolini N, Chiricolo M, Santini D, Minni F, Caretti A, et al. The biosynthesis of the selectin-ligand sialyl Lewis x in colorectal cancer tissues is regulated by fucosyltransferase VI and can be inhibited by an RNA interference-based approach. *Int J Biochem Cell Biol* (2011) **43**(1):130–9. doi:10.1016/j.biocel.2010.10.004
94. Qiu Y, Patwa TH, Xu L, Shedden K, Misek DE, Tuck M, et al. Plasma glycoprotein profiling for colorectal cancer biomarker identification by lectin glycoarray and lectin blot. *J Proteome Res* (2008) **7**(4):1693–703. doi:10.1021/pr700706s
95. Muinelo-Romay L, Gil-Martin E, Fernandez-Briera A. alpha(1,2) Fucosylation in human colorectal carcinoma. *Oncol Lett* (2010) **1**(2):361–6. doi:10.3892/ol_00000064
96. Lopez-Ferrer A, de Bolos C, Barranco C, Garrido M, Isern J, Carlstedt I, et al. Role of fucosyltransferases in the association between apomucin and Lewis antigen expression in normal and malignant gastric epithelium. *Gut* (2000) **47**(3):349–56. doi:10.1136/gut.47.3.349

97. Duell EJ, Bonet C, Munoz X, Lujan-Barroso L, Weiderpass E, Boutron-Ruault MC, et al. Variation at ABO histo-blood group and FUT loci and diffuse and intestinal gastric cancer risk in a European population. *Int J Cancer* (2015) **136**(4):880–93. doi:10.1002/ijc.29034
98. Magalhaes A, Marcos-Pinto R, Nairn AV, Dela Rosa M, Ferreira RM, Junqueira-Neto S, et al. *Helicobacter pylori* chronic infection and mucosal inflammation switches the human gastric glycosylation pathways. *Biochim Biophys Acta* (2015) **1852**(9):1928–39. doi:10.1016/j.bbadis.2015.07.001
99. Muinelo-Romay L, Vazquez-Martin C, Villar-Portela S, Cuevas E, Gil-Martin E, Fernandez-Briera A. Expression and enzyme activity of alpha(1,6)fucosyltransferase in human colorectal cancer. *Int J Cancer* (2008) **123**(3):641–6. doi:10.1002/ijc.23521
100. Osumi D, Takahashi M, Miyoshi E, Yokoe S, Lee SH, Noda K, et al. Core fucosylation of E-cadherin enhances cell-cell adhesion in human colon carcinoma WiDr cells. *Cancer Sci* (2009) **100**(5):888–95. doi:10.1111/j.1349-7006.2009.01125.x
101. Zhao YP, Xu XY, Fang M, Wang H, You Q, Yi CH, et al. Decreased core-fucosylation contributes to malignancy in gastric cancer. *PLoS One* (2014) **9**(4):e94536. doi:10.1371/journal.pone.0094536
102. Ma Z, Vosseller K. Cancer metabolism and elevated O-GlcNAc in oncogenic signaling. *J Biol Chem* (2014) **289**(50):34457–65. doi:10.1074/jbc.R114.577718
103. Fardini Y, Dehennaut V, Lefebvre T, Issad T. O-GlcNAcylation: a new cancer hallmark? *Front Endocrinol* (2013) **4**:99. doi:10.3389/fendo.2013.00099
104. Hart GW, Slawson C, Ramirez-Correa G, Lagerlof O. Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. *Annu Rev Biochem* (2011) **80**:825–58. doi:10.1146/annurev-biochem-060608-102511
105. Konrad RJ, Kudlow JE. The role of O-linked protein glycosylation in beta-cell dysfunction. *Int J Mol Med* (2002) **10**(5):535–9. doi:10.3892/ijmm.10.5.535
106. Ma Z, Vocadlo DJ, Vosseller K. Hyper-O-GlcNAcylation is anti-apoptotic and maintains constitutive NF-kappaB activity in pancreatic cancer cells. *J Biol Chem* (2013) **288**(21):15121–30. doi:10.1074/jbc.M113.470047
107. Yang YR, Kim DH, Seo YK, Park D, Jang HJ, Choi SY, et al. Elevated O-GlcNAcylation promotes colonic inflammation and tumorigenesis by modulating NF-kappaB signaling. *Oncotarget* (2015) **6**(14):12529–42. doi:10.18632/oncotarget.3725
108. Couchman JR, Gopal S, Lim HC, Norgaard S, Multhaupt HA. Syndecans: from peripheral coreceptors to mainstream regulators of cell behaviour. *Int J Exp Pathol* (2015) **96**(1):1–10. doi:10.1111/iep.12112
109. Magalhaes A, Marcos NT, Carvalho AS, David L, Figueiredo C, Bastos J, et al. *Helicobacter pylori* cag pathogenicity island-positive strains induce syndecan-4 expression in gastric epithelial cells. *FEMS Immunol Med Microbiol* (2009) **56**(3):223–32. doi:10.1111/j.1574-695X.2009.00569.x
110. Marcos NT, Magalhaes A, Ferreira B, Oliveira MJ, Carvalho AS, Mendes N, et al. *Helicobacter pylori* induces beta3GnT5 in human gastric cell lines, modulating expression of the SabA ligand sialyl-Lewis x. *J Clin Invest* (2008) **118**(6):2325–36. doi:10.1172/JCI34324
111. Wiksten JP, Lundin J, Nordling S, Lundin M, Kokkola A, von Boguslawski K, et al. Epithelial and stromal syndecan-1 expression as predictor of outcome in patients with gastric cancer. *Int J Cancer* (2001) **95**(1):1–6. doi:10.1002/1097-0215(20010120)95:1<1::AID-IJC1000>3.0.CO;2-5
112. Kim SY, Choi EJ, Yun JA, Jung ES, Oh ST, Kim JG, et al. Syndecan-1 expression is associated with tumor size and EGFR expression in colorectal carcinoma: a clinicopathological study of 230 cases. *Int J Med Sci* (2015) **12**(2):92–9. doi:10.7150/ijms.10497
113. Conejo JR, Kleeff J, Koliopanos A, Matsuda K, Zhu ZW, Goecke H, et al. Syndecan-1 expression is up-regulated in pancreatic but not in other gastrointestinal cancers. *Int J Cancer* (2000) **88**(1):12–20. doi:10.1002/1097-0215(20001001)88:1<12::AID-IJC3>3.0.CO;2-T
114. Filmus J, Capurro M, Rast J. Glypicans. *Genome Biol* (2008) **9**(5):224. doi:10.1186/gb-2008-9-5-224
115. Kleeff J, Ishiwata T, Kumbasar A, Friess H, Buchler MW, Lander AD, et al. The cell-surface heparan sulfate proteoglycan glypican-1 regulates growth factor action in pancreatic carcinoma cells and is overexpressed in human pancreatic cancer. *J Clin Invest* (1998) **102**(9):1662–73. doi:10.1172/JCI4105
116. Whipple CA, Young AL, Korc M. A KrasG12D-driven genetic mouse model of pancreatic cancer requires glypican-1 for efficient proliferation and angiogenesis. *Oncogene* (2012) **31**(20):2535–44. doi:10.1038/onc.2011.430
117. Aikawa T, Whipple CA, Lopez ME, Gunn J, Young A, Lander AD, et al. Glypican-1 modulates the angiogenic and metastatic potential of human and mouse cancer cells. *J Clin Invest* (2008) **118**(1):89–99. doi:10.1172/JCI32412
118. Melo SA, Luecke LB, Kahlert C, Fernandez AF, Gammon ST, Kaye J, et al. Glypican-1 identifies cancer exosomes and detects early pancreatic cancer. *Nature* (2015) **523**(7559):177–82. doi:10.1038/nature14581
119. da Cunha CB, Oliveira C, Wen X, Gomes B, Sousa S, Suriano G, et al. De novo expression of CD44 variants in sporadic and hereditary gastric cancer. *Lab Invest* (2010) **90**(11):1604–14. doi:10.1038/labinvest.2010.155
120. Campos D, Freitas D, Gomes J, Magalhaes A, Steentoft C, Gomes C, et al. Probing the O-glycoproteome of gastric cancer cell lines for biomarker discovery. *Mol Cell Proteomics* (2015) **14**(6):1616–29. doi:10.1074/mcp.M114.046862
121. Haakenson JK, Khokhlatchev AV, Choi YJ, Linton SS, Zhang P, Zaki PM, et al. Lysosomal degradation of CD44 mediates ceramide nanoliposome-induced anoikis and diminished extravasation in metastatic carcinoma cells. *J Biol Chem* (2015) **290**(13):8632–43. doi:10.1074/jbc.M114.609677
122. Miyagi T, Wada T, Yamaguchi K. Roles of plasma membrane-associated sialidase NEU3 in human cancers. *Biochim Biophys Acta* (2008) **1780**(3):532–7. doi:10.1016/j.bbagen.2007.09.016
123. Kakugawa Y, Wada T, Yamaguchi K, Yamanami H, Ouchi K, Sato I, et al. Up-regulation of plasma membrane-associated ganglioside sialidase (Neu3) in human colon cancer and its involvement in apoptosis suppression. *Proc Natl Acad Sci U S A* (2002) **99**(16):10718–23. doi:10.1073/pnas.152597199
124. Takahashi K, Hosono M, Sato I, Hata K, Wada T, Yamaguchi K, et al. Sialidase NEU3 contributes neoplastic potential on colon cancer cells as a key modulator of gangliosides by regulating Wnt signaling. *Int J Cancer* (2015) **137**(7):1560–73. doi:10.1002/ijc.29527
125. Holst S, Stavenhagen K, Balog CI, Koeleman CA, McDonnell LM, Mayboroda OA, et al. Investigations on aberrant glycosylation of glycosphingolipids in colorectal cancer tissues using liquid chromatography and matrix-assisted laser desorption time-of-flight mass spectrometry (MALDI-TOF-MS). *Mol Cell Proteomics* (2013) **12**(11):3081–93. doi:10.1074/mcp.M113.030387
126. Ebeling FG, Stieber P, Untch M, Nagel D, Konecny GE, Schmitt UM, et al. Serum CEA and CA 15-3 as prognostic factors in primary breast cancer. *Br J Cancer* (2002) **86**(8):1217–22. doi:10.1038/sj.bjc.6600248
127. Harris L, Fritsche H, Mennel R, Norton L, Ravdin P, Taube S, et al. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *J Clin Oncol* (2007) **25**(33):5287–312. doi:10.1200/JCO.2007.14.2364
128. Kumpulainen EJ, Keskikuru RJ, Johansson RT. Serum tumor marker CA 15.3 and stage are the two most powerful predictors of survival in primary breast cancer. *Breast Cancer Res Treat* (2002) **76**(2):95–102. doi:10.1023/A:1020514925143
129. Uehara M, Kinoshita T, Hojo T, Akashi-Tanaka S, Iwamoto E, Fukutomi T. Long-term prognostic study of carcinoembryonic antigen (CEA) and carbohydrate antigen 15-3 (CA 15-3) in breast cancer. *Int J Clin Oncol* (2008) **13**(5):447–51. doi:10.1007/s10147-008-0773-3
130. Lauro S, Trasatti L, Bordin F, Lanzetta G, Bria E, Gelibter A, et al. Comparison of CEA, MCA, CA 15-3 and CA 27-29 in follow-up and monitoring therapeutic response in breast cancer patients. *Anticancer Res* (1999) **19**(4C):3511–5.
131. Pauler DK, Menon U, McIntosh M, Symecko HL, Skates SJ, Jacobs IJ. Factors influencing serum CA125II levels in healthy postmenopausal women. *Cancer Epidemiol Biomarkers Prev* (2001) **10**(5):489–93.
132. Gostout BS, Brewer MA. Guidelines for referral of the patient with an adnexal mass. *Clin Obstet Gynecol* (2006) **49**(3):448–58. doi:10.1097/00003081-200609000-00005
133. Gilgunn S, Conroy PJ, Saldova R, Rudd PM, O'Kennedy RJ. Aberrant PSA glycosylation – a sweet predictor of prostate cancer. *Nat Rev Urol* (2013) **10**(2):99–107. doi:10.1038/nrurol.2012.258
134. Saldova R, Fan Y, Fitzpatrick JM, Watson RW, Rudd PM. Core fucosylation and alpha2-3 sialylation in serum N-glycome is significantly increased in prostate cancer comparing to benign prostate hyperplasia. *Glycobiology* (2011) **21**(2):195–205. doi:10.1093/glycob/cwq147
135. Safi F, Schlosser W, Kolb G, Beger HG. Diagnostic value of CA 19-9 in patients with pancreatic cancer and nonspecific gastrointestinal symptoms. *J Gastrointest Surg* (1997) **1**(2):106–12. doi:10.1016/S1091-255X(97)80097-2
136. Locker GY, Hamilton S, Harris J, Jessup JM, Kemeny N, Macdonald JS, et al. ASCO 2006 update of recommendations for the use of tumor markers in

gastrointestinal cancer. *J Clin Oncol* (2006) **24**(33):5313–27. doi:10.1200/JCO.2006.08.2644
137. Duraker N, Celik AN. The prognostic significance of preoperative serum CA 19-9 in patients with resectable gastric carcinoma: comparison with CEA. *J Surg Oncol* (2001) **76**(4):266–71. doi:10.1002/jso.1044
138. Shah UA, Saif MW. Tumor markers in pancreatic cancer: 2013. *JOP* (2013) **14**(4):318–21. doi:10.6092/1590-8577/1653
139. Goldstein MJ, Mitchell EP. Carcinoembryonic antigen in the staging and follow-up of patients with colorectal cancer. *Cancer Invest* (2005) **23**(4):338–51. doi:10.1081/CNV-58878
140. Steele SR, Chang GJ, Hendren S, Weiser M, Irani J, Buie WD, et al. Practice guideline for the surveillance of patients after curative treatment of colon and rectal cancer. *Dis Colon Rectum* (2015) **58**(8):713–25. doi:10.1097/DCR.0000000000000410
141. Julien S, Adriaenssens E, Ottenberg K, Furlan A, Courtand G, Vercoutter-Edouart AS, et al. ST6GalNAc I expression in MDA-MB-231 breast cancer cells greatly modifies their O-glycosylation pattern and enhances their tumourigenicity. *Glycobiology* (2006) **16**(1):54–64. doi:10.1093/glycob/cwj033
142. Julien S, Lagadec C, Krzewinski-Recchi MA, Courtand G, Le Bourhis X, Delannoy P. Stable expression of sialyl-Tn antigen in T47-D cells induces a decrease of cell adhesion and an increase of cell migration. *Breast Cancer Res Treat* (2005) **90**(1):77–84. doi:10.1007/s10549-004-3137-3
143. Ma H, Zhou H, Li P, Song X, Miao X, Li Y, et al. Effect of ST3GAL 4 and FUT 7 on sialyl Lewis X synthesis and multidrug resistance in human acute myeloid leukemia. *Biochim Biophys Acta* (2014) **1842**(9):1681–92. doi:10.1016/j.bbadis.2014.06.014
144. Yang XS, Liu S, Liu YJ, Liu JW, Liu TJ, Wang XQ, et al. Overexpression of fucosyltransferase IV promotes A431 cell proliferation through activating MAPK and PI3K/Akt signaling pathways. *J Cell Physiol* (2010) **225**(2):612–9. doi:10.1002/jcp.22250
145. Tsuda T, Ikeda Y, Taniguchi N. The Asn-420-linked sugar chain in human epidermal growth factor receptor suppresses ligand-independent spontaneous oligomerization. Possible role of a specific sugar chain in controllable receptor activation. *J Biol Chem* (2000) **275**(29):21988–94. doi:10.1074/jbc.M003400200
146. van Roy F, Berx G. The cell-cell adhesion molecule E-cadherin. *Cell Mol Life Sci* (2008) **65**(23):3756–88. doi:10.1007/s00018-008-8281-1
147. Paredes J, Figueiredo J, Albergaria A, Oliveira P, Carvalho J, Ribeiro AS, et al. Epithelial E- and P-cadherins: role and clinical significance in cancer. *Biochim Biophys Acta* (2012) **1826**(2):297–311. doi:10.1016/j.bbcan.2012.05.002
148. Oliveira C, Seruca R, Carneiro F. Genetics, pathology, and clinics of familial gastric cancer. *Int J Surg Pathol* (2006) **14**(1):21–33. doi:10.1177/106689690601400105
149. Carvalho S, Catarino TA, Dias AM, Kato M, Almeida A, Hessling B, et al. Preventing E-cadherin aberrant N-glycosylation at Asn-554 improves its critical function in gastric cancer. *Oncogene* (2015). doi:10.1038/onc.2015.225
150. Pinho SS, Seruca R, Gartner F, Yamaguchi Y, Gu J, Taniguchi N, et al. Modulation of E-cadherin function and dysfunction by N-glycosylation. *Cell Mol Life Sci* (2011) **68**(6):1011–20. doi:10.1007/s00018-010-0595-0
151. Steentoft C, Bennett EP, Schjoldager KT, Vakhrushev SY, Wandall HH, Clausen H. Precision genome editing: a small revolution for glycobiology. *Glycobiology* (2014) **24**(8):663–80. doi:10.1093/glycob/cwu046
152. 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. *EMBO J* (2013) **32**(10):1478–88. doi:10.1038/emboj.2013.79
153. 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. *Proc Natl Acad Sci U S A* (2012) **109**(25):9893–8. doi:10.1073/pnas.1203563109
154. Vakhrushev SY, Steentoft C, Vester-Christensen MB, Bennett EP, Clausen H, Levery SB. Enhanced mass spectrometric mapping of the human GalNAc-type O-glycoproteome with SimpleCells. *Mol Cell Proteomics* (2013) **12**(4):932–44. doi:10.1074/mcp.O112.021972
155. Schjoldager KT, Clausen H. Site-specific protein O-glycosylation modulates proprotein processing – deciphering specific functions of the large polypeptide GalNAc-transferase gene family. *Biochim Biophys Acta* (2012) **1820**(12):2079–94. doi:10.1016/j.bbagen.2012.09.014
156. Vester-Christensen MB, Halim A, Joshi HJ, Steentoft C, Bennett EP, Levery SB, et al. Mining the O-mannose glycoproteome reveals cadherins as major O-mannosylated glycoproteins. *Proc Natl Acad Sci U S A* (2013) **110**(52):21018–23. doi:10.1073/pnas.1313446110
157. Mechref Y. Analysis of glycans derived from glycoconjugates by capillary electrophoresis-mass spectrometry. *Electrophoresis* (2011) **32**(24):3467–81. doi:10.1002/elps.201100342
158. Sethi MK, Kim H, Park CK, Baker MS, Paik YK, Packer NH, et al. In-depth N-glycome profiling of paired colorectal cancer and non-tumorigenic tissues reveals cancer-, stage- and EGFR-specific protein N-glycosylation. *Glycobiology* (2015) **25**(10):1064–78. doi:10.1093/glycob/cwv042
159. Bones J, Byrne JC, O'Donoghue N, McManus C, Scaife C, Boissin H, et al. Glycomic and glycoproteomic analysis of serum from patients with stomach cancer reveals potential markers arising from host defense response mechanisms. *J Proteome Res* (2011) **10**(3):1246–65. doi:10.1021/pr101036b
160. Bones J, Mittermayr S, O'Donoghue N, Guttman A, Rudd PM. Ultra performance liquid chromatographic profiling of serum N-glycans for fast and efficient identification of cancer associated alterations in glycosylation. *Anal Chem* (2010) **82**(24):10208–15. doi:10.1021/ac102860w
161. Ozcan S, Barkauskas DA, Renee Ruhaak L, Torres J, Cooke CL, An HJ, et al. Serum glycan signatures of gastric cancer. *Cancer Prev Res (Phila)* (2014) **7**(2):226–35. doi:10.1158/1940-6207.CAPR-13-0235
162. Nouso K, Amano M, Ito YM, Miyahara K, Morimoto Y, Kato H, et al. Clinical utility of high-throughput glycome analysis in patients with pancreatic cancer. *J Gastroenterol* (2013) **48**(10):1171–9. doi:10.1007/s00535-012-0732-7
163. Mann BF, Goetz JA, House MG, Schmidt CM, Novotny MV. Glycomic and proteomic profiling of pancreatic cyst fluids identifies hyperfucosylated lactosamines on the N-linked glycans of overexpressed glycoproteins. *Mol Cell Proteomics* (2012) **11**(7):M111015792. doi:10.1074/mcp.M111.015792
164. Kaji H, Saito H, Yamauchi Y, Shinkawa T, Taoka M, Hirabayashi J, et al. Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins. *Nat Biotechnol* (2003) **21**(6):667–72. doi:10.1038/nbt829
165. Yang Z, Hancock WS. Approach to the comprehensive analysis of glycoproteins isolated from human serum using a multi-lectin affinity column. *J Chromatogr A* (2004) **1053**(1–2):79–88. doi:10.1016/S0021-9673(04)01433-5
166. Durham M, Regnier FE. Targeted glycoproteomics: serial lectin affinity chromatography in the selection of O-glycosylation sites on proteins from the human blood proteome. *J Chromatogr A* (2006) **1132**(1–2):165–73. doi:10.1016/j.chroma.2006.07.070
167. Zhang H, Li XJ, Martin DB, Aebersold R. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat Biotechnol* (2003) **21**(6):660–6. doi:10.1038/nbt827
168. Liu T, Qian WJ, Gritsenko MA, Camp DG II, Monroe ME, Moore RJ, et al. Human plasma N-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry. *J Proteome Res* (2005) **4**(6):2070–80. doi:10.1021/pr0502065
169. Rawn JD, Lienhard GE. The binding of boronic acids to chymotrypsin. *Biochemistry* (1974) **13**(15):3124–30. doi:10.1021/bi00712a019
170. Alvarez-Manilla G, Atwood J III, Guo Y, Warren NL, Orlando R, Pierce M. Tools for glycoproteomic analysis: size exclusion chromatography facilitates identification of tryptic glycopeptides with N-linked glycosylation sites. *J Proteome Res* (2006) **5**(3):701–8. doi:10.1021/pr050275j
171. Hagglund P, Bunkenborg J, Elortza F, Jensen ON, Roepstorff P. A new strategy for identification of N-glycosylated proteins and unambiguous assignment of their glycosylation sites using HILIC enrichment and partial deglycosylation. *J Proteome Res* (2004) **3**(3):556–66. doi:10.1021/pr034112b
172. Larsen MR, Hojrup P, Roepstorff P. Characterization of gel-separated glycoproteins using two-step proteolytic digestion combined with sequential microcolumns and mass spectrometry. *Mol Cell Proteomics* (2005) **4**(2):107–19. doi:10.1074/mcp.M400068-MCP200
173. Zhao J, Simeone DM, Heidt D, Anderson MA, Lubman DM. Comparative serum glycoproteomics using lectin selected sialic acid glycoproteins with mass spectrometric analysis: application to pancreatic cancer serum. *J Proteome Res* (2006) **5**(7):1792–802. doi:10.1021/pr060034r
174. Melo-Braga MN, Ibanez-Vea M, Larsen MR, Kulej K. Comprehensive protocol to simultaneously study protein phosphorylation, acetylation, and

N-linked sialylated glycosylation. *Methods Mol Biol* (2015) **1295**:275–92. doi:10.1007/978-1-4939-2550-6_21
175. Hang HC, Yu C, Kato DL, Bertozzi CR. A metabolic labeling approach toward proteomic analysis of mucin-type O-linked glycosylation. *Proc Natl Acad Sci U S A* (2003) **100**(25):14846–51. doi:10.1073/pnas.2335201100
176. Laughlin ST, Bertozzi CR. Metabolic labeling of glycans with azido sugars and subsequent glycan-profiling and visualization via Staudinger ligation. *Nat Protoc* (2007) **2**(11):2930–44. doi:10.1038/nprot.2007.422
177. Liang Y, Hua Q, Pan P, Yang J, Zhang Q. Development of a novel method to evaluate sialylation of glycoproteins and analysis of gp96 sialylation in Hela, SW1990 and A549 cell lines. *Biol Res* (2015) **48**:52. doi:10.1186/s40659-015-0041-8
178. Wuhrer M, Catalina MI, Deelder AM, Hokke CH. Glycoproteomics based on tandem mass spectrometry of glycopeptides. *J Chromatogr B Analyt Technol Biomed Life Sci* (2007) **849**(1–2):115–28. doi:10.1016/j.jchromb.2006.09.041
179. An HJ, Froehlich JW, Lebrilla CB. Determination of glycosylation sites and site-specific heterogeneity in glycoproteins. *Curr Opin Chem Biol* (2009) **13**(4):421–6. doi:10.1016/j.cbpa.2009.07.022
180. Hinneburg H, Stavenhagen K, Schweiger-Hufnagel U, Pengelley S, Jabs W, Seeberger PH, et al. The art of destruction: optimizing collision energies in quadrupole-time of flight (Q-TOF) instruments for glycopeptide-based glycoproteomics. *J Am Soc Mass Spectrom* (2016) **27**(3):507–19. doi:10.1007/s13361-015-1308-6
181. Gomes C, Almeida A, Ferreira JA, Silva L, Santos-Sousa H, Pinto-de-Sousa J, et al. Glycoproteomic analysis of serum from patients with gastric precancerous lesions. *J Proteome Res* (2013) **12**(3):1454–66. doi:10.1021/pr301112x
182. Li K, Sun Z, Zheng J, Lu Y, Bian Y, Ye M, et al. In-depth research of multidrug resistance related cell surface glycoproteome in gastric cancer. *J Proteomics* (2013) **82**:130–40. doi:10.1016/j.jprot.2013.02.021
183. Chang TW. Binding of cells to matrixes of distinct antibodies coated on solid surface. *J Immunol Methods* (1983) **65**(1–2):217–23. doi:10.1016/0022-1759(83)90318-6
184. Rakus JF, Mahal LK. New technologies for glycomic analysis: toward a systematic understanding of the glycome. *Annu Rev Anal Chem (Palo Alto Calif)* (2011) **4**:367–92. doi:10.1146/annurev-anchem-061010-113951
185. Patwa T, Li C, Simeone DM, Lubman DM. Glycoprotein analysis using protein microarrays and mass spectrometry. *Mass Spectrom Rev* (2010) **29**(5):830–44. doi:10.1002/mas.20269
186. Gao C, Liu Y, Zhang H, Zhang Y, Fukuda MN, Palma AS, et al. Carbohydrate sequence of the prostate cancer-associated antigen F77 assigned by a mucin O-glycome designer array. *J Biol Chem* (2014) **289**(23):16462–77. doi:10.1074/jbc.M114.558932
187. Blixt O, Head S, Mondala T, Scanlan C, Huflejt ME, Alvarez R, et al. Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc Natl Acad Sci U S A* (2004) **101**(49):17033–8. doi:10.1073/pnas.0407902101
188. Schneider C, Smith DF, Cummings RD, Boligan KF, Hamilton RG, Bochner BS, et al. The human IgG anti-carbohydrate repertoire exhibits a universal architecture and contains specificity for microbial attachment sites. *Sci Transl Med* (2015) **7**(269):269ra1. doi:10.1126/scitranslmed.3010524
189. Horlacher T, Seeberger PH. Carbohydrate arrays as tools for research and diagnostics. *Chem Soc Rev* (2008) **37**(7):1414–22. doi:10.1039/b708016f
190. Seeberger PH. Automated carbohydrate synthesis as platform to address fundamental aspects of glycobiology – current status and future challenges. *Carbohydr Res* (2008) **343**(12):1889–96. doi:10.1016/j.carres.2008.05.023
191. Wang D, Liu S, Trummer BJ, Deng C, Wang A. Carbohydrate microarrays for the recognition of cross-reactive molecular markers of microbes and host cells. *Nat Biotechnol* (2002) **20**(3):275–81. doi:10.1038/nbt0302-275
192. Liu Y, Palma AS, Feizi T. Carbohydrate microarrays: key developments in glycobiology. *Biol Chem* (2009) **390**(7):647–56. doi:10.1515/BC.2009.071
193. Stevens J, Blixt O, Paulson JC, Wilson IA. Glycan microarray technologies: tools to survey host specificity of influenza viruses. *Nat Rev Microbiol* (2006) **4**(11):857–64. doi:10.1038/nrmicro1530
194. Xuan P, Zhang Y, Tzeng TR, Wan XF, Luo F. A quantitative structure-activity relationship (QSAR) study on glycan array data to determine the specificities of glycan-binding proteins. *Glycobiology* (2012) **22**(4):552–60. doi:10.1093/glycob/cwr163
195. Kletter D, Singh S, Bern M, Haab BB. Global comparisons of lectin-glycan interactions using a database of analyzed glycan array data. *Mol Cell Proteomics* (2013) **12**(4):1026–35. doi:10.1074/mcp.M112.026641
196. Cholleti SR, Agravat S, Morris T, Saltz JH, Song X, Cummings RD, et al. Automated motif discovery from glycan array data. *OMICS* (2012) **16**(10):497–512. doi:10.1089/omi.2012.0013
197. Fair RJ, Hahm HS, Seeberger PH. Combination of automated solid-phase and enzymatic oligosaccharide synthesis provides access to alpha(2,3)-sialylated glycans. *Chem Commun (Camb)* (2015) **51**(28):6183–5. doi:10.1039/c5cc01368b
198. Lai CH, Hahm HS, Liang CF, Seeberger PH. Automated solid-phase synthesis of oligosaccharides containing sialic acids. *Beilstein J Org Chem* (2015) **11**:617–21. doi:10.3762/bjoc.11.69
199. Liang CF, Hahm HS, Seeberger PH. Automated synthesis of chondroitin sulfate oligosaccharides. *Methods Mol Biol* (2015) **1229**:3–10. doi:10.1007/978-1-4939-1714-3_1
200. Zacco E, Anish C, Martin CE, V Berlepsch H, Brandenburg E, Seeberger PH, et al. A self-assembling peptide scaffold for the multivalent presentation of antigens. *Biomacromolecules* (2015) **16**(7):2188–97. doi:10.1021/acs.biomac.5b00572
201. Zhao J, Patwa TH, Qiu W, Shedden K, Hinderer R, Misek DE, et al. Glycoprotein microarrays with multi-lectin detection: unique lectin binding patterns as a tool for classifying normal, chronic pancreatitis and pancreatic cancer sera. *J Proteome Res* (2007) **6**(5):1864–74. doi:10.1021/pr070062p
202. Pedersen JW, Blixt O, Bennett EP, Tarp MA, Dar I, Mandel U, et al. Seromic profiling of colorectal cancer patients with novel glycopeptide microarray. *Int J Cancer* (2011) **128**(8):1860–71. doi:10.1002/ijc.25778
203. Burford B, Gentry-Maharaj A, Graham R, Allen D, Pedersen JW, Nudelman AS, et al. Autoantibodies to MUC1 glycopeptides cannot be used as a screening assay for early detection of breast, ovarian, lung or pancreatic cancer. *Br J Cancer* (2013) **108**(10):2045–55. doi:10.1038/bjc.2013.214
204. Hirabayashi J, Yamada M, Kuno A, Tateno H. Lectin microarrays: concept, principle and applications. *Chem Soc Rev* (2013) **42**(10):4443–58. doi:10.1039/c3cs35419a
205. Tang H, Hsueh P, Kletter D, Bern M, Haab B. The detection and discovery of glycan motifs in biological samples using lectins and antibodies: new methods and opportunities. *Adv Cancer Res* (2015) **126**:167–202. doi:10.1016/bs.acr.2014.11.003
206. Yue T, Haab BB. Microarrays in glycoproteomics research. *Clin Lab Med* (2009) **29**(1):15–29. doi:10.1016/j.cll.2009.01.001
207. Huang WL, Li YG, Lv YC, Guan XH, Ji HF, Chi BR. Use of lectin microarray to differentiate gastric cancer from gastric ulcer. *World J Gastroenterol* (2014) **20**(18):5474–82. doi:10.3748/wjg.v20.i18.5474
208. Sunderic M, Sediva A, Robajac D, Miljus G, Gemeiner P, Nedic O, et al. Lectin-based protein microarray analysis of differences in serum alpha-2-macroglobulin glycosylation between patients with colorectal cancer and persons without cancer. *Biotechnol Appl Biochem* (2015). doi:10.1002/bab.1407
209. Cao Z, Maupin K, Curnutte B, Fallon B, Feasley CL, Brouhard E, et al. Specific glycoforms of MUC5AC and endorepellin accurately distinguish mucinous from nonmucinous pancreatic cysts. *Mol Cell Proteomics* (2013) **12**(10):2724–34. doi:10.1074/mcp.M113.030700
210. Chen S, LaRoche T, Hamelinck D, Bergsma D, Brenner D, Simeone D, et al. Multiplexed analysis of glycan variation on native proteins captured by antibody microarrays. *Nat Methods* (2007) **4**(5):437–44. doi:10.1038/nmeth1035
211. Kumada Y, Ohigashi Y, Emori Y, Imamura K, Omura Y, Kishimoto M. Improved lectin ELISA for glycosylation analysis of biomarkers using PS-tag-fused single-chain Fv. *J Immunol Methods* (2012) **385**(1–2):15–22. doi:10.1016/j.jim.2012.07.021
212. Gustafsdottir SM, Schallmeiner E, Fredriksson S, Gullberg M, Soderberg O, Jarvius M, et al. Proximity ligation assays for sensitive and specific protein analyses. *Anal Biochem* (2005) **345**(1):2–9. doi:10.1016/j.ab.2005.01.018
213. Soderberg O, Gullberg M, Jarvius M, Ridderstrale K, Leuchowius KJ, Jarvius J, et al. Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat Methods* (2006) **3**(12):995–1000. doi:10.1038/nmeth947
214. Conze T, Carvalho AS, Landegren U, Almeida R, Reis CA, David L, et al. MUC2 mucin is a major carrier of the cancer-associated sialyl-Tn antigen in intestinal metaplasia and gastric carcinomas. *Glycobiology* (2010) **20**(2):199–206. doi:10.1093/glycob/cwp161

215. Pinto R, Carvalho AS, Conze T, Magalhaes A, Picco G, Burchell JM, et al. Identification of new cancer biomarkers based on aberrant mucin glycoforms by in situ proximity ligation. *J Cell Mol Med* (2012) **16**(7):1474–84. doi:10.1111/j.1582-4934.2011.01436.x
216. Koos B, Kamali-Moghaddam M, David L, Sobrinho-Simoes M, Dimberg A, Nilsson M, et al. Next-generation pathology – surveillance of tumor microecology. *J Mol Biol* (2015) **427**(11):2013–22. doi:10.1016/j.jmb.2015.02.017
217. Caprioli RM, Farmer TB, Gile J. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal Chem* (1997) **69**(23):4751–60. doi:10.1021/ac970888i
218. Powers TW, Jones EE, Betesh LR, Romano PR, Gao P, Copland JA, et al. Matrix assisted laser desorption ionization imaging mass spectrometry workflow for spatial profiling analysis of N-linked glycan expression in tissues. *Anal Chem* (2013) **85**(20):9799–806. doi:10.1021/ac402108x
219. Powers TW, Neely BA, Shao Y, Tang H, Troyer DA, Mehta AS, et al. MALDI imaging mass spectrometry profiling of N-glycans in formalin-fixed paraffin embedded clinical tissue blocks and tissue microarrays. *PLoS One* (2014) **9**(9):e106255. doi:10.1371/journal.pone.0106255
220. Powers TW, Holst S, Wuhrer M, Mehta AS, Drake RR. Two-dimensional N-glycan distribution mapping of hepatocellular carcinoma tissues by MALDI-imaging mass spectrometry. *Biomolecules* (2015) **5**(4):2554–72. doi:10.3390/biom5042554

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ANNEX II

RECIPROCAL MODULATION OF TERMINAL SIALYLATION AND BISECTING *N*-GLYCANS: A NEW AXIS OF CANCER-CELL GLYCOME REGULATION?

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 291, NO. 15, p. 8308, April 8, 2016
Published in the U.S.A.

LETTER

Reciprocal Modulation of Terminal Sialylation and Bisecting *N*-Glycans: A New Axis of Cancer-Cell Glycome Regulation?

Lu *et al.* (1) have investigated the influence of cellular sialylation on the GnT-III-mediated regulation of cancer cell metastatic potential. The authors demonstrated that GnT-III (GlcNAc-bisecting glycosyltransferase) overexpression results in a significant reduction of α2,3-sialylation, with no major alteration of α2,6-sialylation (1). Interestingly, a reciprocal correlation between terminal α2,3-sialylation and bisected *N*-glycans has also been recently reported (2). Glycomic analysis of cancer cells overexpressing the α2,3-sialyltransferase ST3GAL4 showed that increased terminal α2,3-sialylation is accompanied by a substantial loss of bisected *N*-glycans (2).

This coordinated regulation of α2,3-sialylation by bisecting *N*-glycans and vice versa adds a new level of complexity to the regulation of the cancer cell glycome and raises new questions about the molecular mechanisms underlying these glycosylation shifts. Noteworthy, the down-regulation of bisected *N*-glycans and α2,3-sialylation by ST3GAL4 and GnT-III, respectively, do not stem from alterations at the glycosyltransferase transcript levels (1, 2). The supramolecular organization of the Golgi glycosylation pathways is plastic and can be altered in cancer cells (3); thus, we can hypothesize that altered localization of the overexpressed glycosyltransferases could functionally impact the sequential glycan biosynthetic pathways and therefore interfere with bisecting *N*-glycans and terminal sialylation.

Increased α2,3-sialylation is associated with malignancy and patients' poorer prognosis, whereas bisected *N*-glycans suppress metastization (4). Moreover, *N*-glycan branching and sialylation patterns determine galectin lattice dynamics with critical impact on tumor cell signaling and definition of aggressiveness features (5). Therefore, deciphering the mechanisms underlying reciprocal regulation of terminal sialylation and bisecting *N*-glycans is fundamental for our understanding of tumor cell biology.

***Ana Magalhães*[‡§1], *Stefan Mereiter*[‡§¶], *and Celso Reis*[‡§¶‖1]**
[‡]*i3S-Instituto de Investigação e Inovação em Saúde,* [¶]*Institute of Biomedical Sciences of Abel Salazar-ICBAS, and* [‖]*Medical Faculty, University of Porto, Porto, Portugal and* [§]*Institute of Molecular Pathology and Immunology, University of Porto-IPATIMUP, Porto, Portugal*

1 Lu, J., Isaji, T., Im, S., Fukuda, T., Kameyama, A., and Gu, J. (2016) Expression of *N*-acetylglucosaminyltransferase III suppresses α2,3-sialylation and its distinctive functions in cell migration are attributed to α2,6-sialylation levels. *J. Biol. Chem.* 291, 5708–5720
2 Mereiter, S., Magalhães, A., Adamczyk, B., Jin, C., Almeida, A., Drici, L., Ibáñez-Vea, M., Gomes, C., Ferreira, J. A., Afonso, L. P., Santos, L. L., Larsen, M. R., Kolarich, D., Karlsson, N. G., and Reis, C. A. (2015) Glycomic analysis of gastric carcinoma cells discloses glycans as modulators of RON receptor tyrosine kinase activation in cancer. *Biochim. Biophys. Acta* 10.1016/j.bbagen.2015.12.016
3 Hassinen, A., Pujol, F. M., Kokkonen, N., Pieters, C., Kihlström, M., Korhonen, K., and Kellokumpu, S. (2011) Functional organization of Golgi *N*- and *O*-glycosylation pathways involves pH-dependent complex formation that is impaired in cancer cells. *J. Biol. Chem.* **286,** 38329–38340
4 Pinho, S. S., and Reis, C. A. (2015) Glycosylation in cancer: mechanisms and clinical implications. *Nat. Rev. Cancer* **15,** 540–555
5 Croci, D. O., Cerliani, J. P., Dalotto-Moreno, T., Méndez-Huergo, S. P., Mascanfroni, I. D., Dergan-Dylon, S., Toscano, M. A., Caramelo, J. J., García-Vallejo, J. J., Ouyang, J., Mesri, E. A., Junttila, M. R., Bais, C., Shipp, M. A., Salatino, M., and Rabinovich, G. A. (2014) Glycosylation-dependent lectin-receptor interactions preserve angiogenesis in anti-VEGF refractory tumors. *Cell* **156,** 744–758

DOI 10.1074/jbc.L116.722462

[1] E-mail: amagalhaes@ipatimup.pt or celsor@ipatimup.pt