U.PORTO

MESTRADO EM ONCOLOGIA

ONCOLOGIA LABORATORIAL

Functional Glycoproteomics Characterization of Gastric Cancer Cells Expressing Immature *O*-Glycans

Martina Dourado Gonçalves

M

2023

U.PORTO

Martina Dourado Gonçalves. Functional Glycoproteomics Characterization of Gastric Cancer Cells Expressing Immature *O*-Glycans

Functional Glycoproteomics Characterization of Gastric Cancer Cells Expressing Immature *O*-Glycans

Martina Dourado Gonçalves

M.ICBAS **2023**

INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR

MARTINA DOURADO GONÇALVES

FUNCTIONAL GLYCOPROTEOMICS CHARACTERIZATION OF GASTRIC CANCER CELLS EXPRESSING IMMATURE *O*-GLYCANS

Dissertação de Candidatura ao grau de Mestre em Oncologia submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto

Orientadora: Andreia Peixoto, PhD
Categoria: Investigador Júnior
Afiliação: Instituto Português de Oncologia do Porto Francisco Gentil EPE (IPO-Porto) e Instituto de Investigação e Inovação em Saúde da Universidade do Porto (i3S)

Co-orientador: José Alexandre Ferreira, PhD
Categoria: Investigador Principal
Afiliação: Instituto Português de Oncologia do Porto Francisco Gentil EPE (IPO-Porto)

Co-orientador: Lúcio Lara Santos, MD, PhD
Categoria: Professor Catedrático Convidado e Coordenador
Afiliação: Instituto de Ciências Biomédicas de Abel Salazar da Universidade do Porto e Instituto Português de Oncologia do Porto Francisco Gentil EPE (IPO-Porto)

“Struggle and determination lead to great things.
*Just ask the **butterfly**.”*

- Unknown author **-**

PRELIMINARY NOTE

This work was financed by national funds through FCT/MCTES within the scope of FCT projects RESOLVE (PTDC/MED-OUT/2512/2021). The authors also thank the financial support of the Portuguese Oncology Institute of Porto—Research Centre (CI-IPOP-29–2017-2020; CI-IPOP-58–2017-2021).

Contributions to the conception and development of the following research work, conducted during the yearlong development of this master's degree Thesis in Oncology:

Oral Communications:

I. **Dourado-Gonçalves, M.**, Ferreira, D., Paulo, P., Santos, LL., Peixoto, A., Ferreira, JA. Functional glycoproteomics characterization of gastric cancer cells expressing immature *O*-Glycans. 7th SYMPOSIUM OF THE MSc IN ONCOLOGY, IPO-PORTO.

II. **Dourado-Gonçalves, M.**, Ferreira, D., Paulo, P., Santos, LL., Peixoto, A., Ferreira, JA. Functional glycoproteomics characterization of gastric cancer cells expressing immature *O*-Glycans. IJUP 2023: Encontro de Investigação Jovem, Universidade do Porto.

AGRADECIMENTOS

Concretização é a palavra que melhor define o sentimento que me envolve neste momento. A ela vêm acompanhadas a gratidão e a felicidade por ter alcançado mais uma grande etapa da minha vida. Muitas foram as vezes que achei que seria impossível aqui chegar, mas afinal quem sempre me disse “o que tem que ser para nós, a nós chegará” tinha e tem toda a razão. Nesta aventura foram muitas as pessoas que me ajudaram a chegar a bom porto e é a elas que quero expressar os meus mais profundos agradecimentos.

À minha orientadora, Doutora Andreia Peixoto, por toda a sua dedicação, excelente orientação e apoio ao longo deste ano. Tenho a agradecer toda a paciência, incentivo e, claro, todos os conhecimentos e expertise que me transmitiu. Obrigada por acreditar nas minhas capacidades e por ser uma inspiração na Ciência!

Ao Doutor José Alexandre Ferreira, meu coorientador, por todo os conhecimentos transmitidos e pela confiança depositada em mim e no meu trabalho. É, sem dúvida, um privilégio aprender consigo e, assim, crescer como futura investigadora que aspiro ser. Obrigada!

Ao Prof. Doutor Lúcio Lara Santos, como meu coorientador e coordenador do Grupo de Patologia e Terapêutica Experimental do Instituto Português de Oncologia do Porto, por me ter recebido no seu laboratório e me ter proporcionado a oportunidade aprender tanto e desenvolver este projeto.

À Prof. Doutora Cármen Jerónimo, na qualidade de diretora do mestrado em Oncologia, e a todos os docentes com quem tanto aprendi.

Ao Grupo de Patologia e Terapêutica Experimental, em especial ao Luís, à Cristiana, à Janine, à Sofia, à Marta, ao Dylan, à Rafaela e ao Eduardo, o meu profundo agradecimento por todo o apoio durante a realização deste trabalho. Obrigada por me terem recebido tão bem e por serem profissionais fantásticos.

Ao Rui, por ser um colega fantástico e por tanto me ter ajudado. Obrigada por estares sempre disponível para responder às minhas dúvidas, independentemente do momento. Foi um privilégio aprender contigo.

À Daniela, a minha parceira da Vila, por todas as viagens que partilhamos, por todos os desabafos e pela amizade que contruímos. Obrigada por todo o carinho e força que me transmitiste sempre que precisei.

Às minhas “irmãs” do laboratório, Beatriz e Mariana, por tornarem esta aventura ainda mais especial. Obrigada por tudo o que vivemos, por estarem presentes quer nos momentos de frustração, quer de alegria e por tudo o que partilhamos, desde as maiores desilusões até às pequenas doces vitórias da vida e da ciência. Bia, a ti que me acompanhaste neste caminho desde o dia 1, agradeço-te por teres sido tantas vezes o meu porto de abrigo e conforto. Mari, a ti que te juntaste às manas uns meses mais tarde, agradeço-te por seres uma pessoa tão especial e preencheres o espaço que faltava na nossa ilha espacial (o nosso cantinho) e emocional. Obrigada por esta amizade tão bonita que as três construímos e que vamos, com certeza, levar para as nossas vidas.

À minha amiga do coração, Ana, por toda a amizade genuína que já vem de há uns anos. Obrigada por todos os teus conselhos, toda a força e carinho que me transmites quando o mundo parece desabar. Obrigada por, mesmo longe, estares sempre perto em todos os momentos e mais alguns.

À minha tão querida família e ao meu querido Igor. À minha preciosa mãe pelo apoio incondicional, por toda a paciência, coragem e carinho que foram fulcrais para conseguir ultrapassar os momentos mais complicados. Ao meu pai por todo o carinho, apoio e, claro, por todas as boleias para o laboratório ao fim de semana. Aos meus avós e tios por todos os conselhos, todo o apoio, força e carinho. Ao meu Igor, o meu companheiro e pilar em todos os momentos, não há palavras que chegam para te agradecer. Obrigada por estarem na primeira fila a acompanhar e a apoiar todas as etapas da minha vida. É a vocês que, com todo o orgulho, dedico esta dissertação.

Com a maior gratidão,
Martina

ABSTRACT

Gastric cancer (GC) is a global health concern, associated with high mortality rates. Aberrant protein glycosylation, a major hallmark of cancer, is correlated with GC aggressiveness and poor prognosis, especially when tumours express truncated short-chain *O*-glycans such Tn and Sialyl-Tn (STn) antigens. Thus, the presence of these aberrant glycans on the cell surface offers a significant opportunity to investigate new approaches for diagnosis and treatment. This work aimed to develop and characterize a GC cell line overexpressing Tn and STn antigens and determine their functional implications on cell aggressiveness. Moreover, it aimed to establish a robust protocol to study the glycoproteome of cell surface proteins.

Accordingly, the GC AGS cell line was glycoengineered to knock-out the *C1GALT1* gene and overexpress Tn and STn antigens. The genotypic characterization of the model was performed by Indel Detection by Amplicon Analysis and Sanger Sequencing. Moreover, the phenotypic validation of the model was achieved by immunofluorescence and flow cytometric analysis. Functional characterization was assessed through *in vitro* proliferation, migration, invasion, and colony formation assays. Subsequently, a reliable method involving biotin-neutravidin coupling was developed to analyze the glycoproteomic profile of the glycoengineered cells.

Therefore, Tn and STn overexpressing cells exhibited reduced migratory ability while displaying increased invasive potential and resistance to anoikis. Additionally, a robust protocol based on biotin labelling followed by cell scraping and neutravidin enrichment was established.

In summary, the present work provides a comprehensive analysis demonstrating that the overexpression of Tn and STn antigens leads to enhanced invasive and colony formation capabilities, both associated with cancer aggressiveness. The established enrichment protocol targeting plasma membrane proteins will be useful in characterizing the glycoproteomic profile of the glycoengineered cells. Overall, this offers a promising future for the development of targeted interventions, potentially revolutionizing the clinical management of GC.

Keywords: Gastric cancer, Glycosylation, Truncated *O*-glycans, Tn, STn, Glycoproteomics

RESUMO

O cancro gástrico (CG) é uma preocupação global de saúde, associando-se a elevadas taxas de mortalidade. A glicosilação aberrante de proteínas, uma das principais características do cancro, está correlacionada com a agressividade do CG e com um prognóstico desfavorável, especialmente quando os tumores expressam *O*-glicanos truncados como os antigénios Tn e Sialyl-Tn (STn). Desta forma, a presença desses glicanos aberrantes na superfície celular oferece uma oportunidade para investigar novas abordagens de diagnóstico e tratamento. Assim, este trabalho teve como objetivo desenvolver e caracterizar uma linha celular de CG com sobrexpressão dos antigénios Tn e STn e, ainda, determinar as suas implicações funcionais na agressividade celular. Além disso, procurou estabelecer um protocolo robusto para estudar o glicoproteoma de proteínas da superfície celular.

Desta forma, a linha celular de CG AGS foi editada geneticamente para inativar o gene *C1GALT1*, levando à sobrexpressão dos antigénios Tn e STn. O perfil genotípico do modelo foi validado através de Indel Detection by Amplicon Analysis e sequenciação de Sanger, enquanto a validação fenotípica foi realizada recorrendo a técnicas de imunofluorescência e citometria de fluxo. A caracterização funcional foi avaliada através de ensaios *in vitro* de proliferação, migração, invasão e formação de colónias. Posteriormente, desenvolveu-se um método robusto, baseado no acoplamento biotina-neutravidina, para analisar o perfil glicoproteómico do modelo geneticamente editado.

Deste modo, as células com sobrexpressão de Tn e STn exibiram uma capacidade migratória reduzida, enquanto demonstraram um aumento do seu potencial invasivo e da capacidade de resistência à *anoikis*. Adicionalmente, estabeleceu-se um protocolo robusto, baseado na marcação celular com biotina, seguida de destacamento das células por *scraping* e posterior enriquecimento com neutravidina.

Em suma, o presente trabalho demonstra que a sobrexpressão dos antigénios Tn e STn se correlaciona com uma maior capacidade invasiva e de formação de colónias, o que se associa ao aumento da agressividade tumoral. Além disso, o protocolo de enriquecimento para proteínas da membrana plasmática estabelecido, será útil na caracterização do perfil glicoproteómico das células editadas. De uma forma geral, este trabalho oferece uma perspetiva promissora para o desenvolvimento de intervenções direcionadas, potencialmente revolucionárias para o tratamento clínico do CG.

Palavras-chave: Cancro gástrico, Glicosilação, *O*-glicanos truncados, Tn, STn, Glicoproteómica

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS & ACRONYMS

Asn Asparagine
C1GalT1 Core 1 β1–3 galactosyltransferase
C1GalT1C1 C1GalT1-specific chaperone 1 (COSMC)
C2GnT Core 2 β1-6 N-acetylglucosaminyltransferase
CIN Chromosomal Instability
CT Computed Tomography
Dol-PP Dolichol Pyrophosphate
EBV Epstein-Barr virus
EMT Epithelial-Mesenchymal Transition
ER Endoplasmic Reticulum
Fuc Fucose
FUT fucosyltransferase
GA Golgi Apparatus
GAC Gastric Adenocarcinoma
GAGs Glycosaminoglycans
Gal Galactose
GalNAc *N*-acetylgalactosamine
GC Gastric Cancer
Glc Glucose
GlcA Glucuronic Acid
GlcNAc *N*-acetylglucosamine
GPI Glycosylphosphatidylinositol
GS Genomically Stable
H. pylori *Helicobacter pylori*
HDGC Hereditary Diffuse Gastric Cancer
IDAA Indel Detection by Amplicon Analysis
IdoA Iduronic acid
Le^{A} Lewis A
Le^{X} Lewis X
m/z Mass to charge ratio
MALDI Matrix-Assisted Laser Desorption/Ionization
Man Mannose
MS Mass Spectrometry
MS/MS Tandem Mass Spectrometry
MSI Microsatellite Instable Tumours

nanoLC Nano Liquid Chromatography

NeuAse Neuraminidase

PM Plasma Membrane

PNA *Peanut agglutinin*

PNGaseF Peptide *N*-glycosidase F

RT Room Temperature

RT-PCR Reverse Transcription Polymerase Chain Reaction

Ser Serine

Sia Sialic Acid

SLe^A Sialyl Lewis A

SLe^X Sialyl Lewis X

ST Sialyl-T

ST6GALNAC1 α-*N*-acetylgalactosaminide α-2,6-sialyltransferase

STn Sialyl-Tn

TACAs Tumour-Associated Carbohydrate Antigens

TCGA The Cancer Genome Atlas

Thr Threonine

Tn GalNAcα-O-Ser/Thr

TNM Tumor Node Metastasis

WHO World Heath Organization

Xyl Xylose

CHAPTER I

INTRODUCTION

1. GASTRIC CANCER

1.1. Epidemiology & Risk Factors

Gastric cancer (GC) is the fifth most common cancer worldwide, with around 1.1 million new cases, representing 5.6% of all cancer cases diagnosed in 2020, according to Globocan data. GC is also the fourth deadliest cancer, with 768,793 deaths, representing 7.7% of all cancer-related deaths in 2020 (2). The high mortality associated with this disease is related to late diagnoses since most patients only become symptomatic in more advanced stages of the disease, limiting the therapeutic options and their effectiveness (3, 4).

Several well-known risk factors are associated with the development of GC, such as *Helicobacter pylori* (*H. pylori*) infection, which is considered the principal cause of non-cardia GC (2, 5, 6). Chronic or recurrent *H. pylori* colonization in the gastric mucosa is present in more than 50% of the world population (7). This chronic infection is associated with gastric mucosal inflammation, which leads to chronic atrophic gastritis, characterized by glandular architecture loss (5, 8). Atrophy may evolve into intestinal metaplasia, promoted by epithelial-mesenchymal transition (EMT), resulting in the replacement of the gastric epithelium for an intestinal-type epithelium, eventually leading to GC progression (9, 10). Furthermore, Epstein-Barr virus (EBV) infection is the cause of about 10% of all GC cases, being more frequently found in men than women (11, 12).

One to 3% of all GC cases are associated with the existence of inherited syndromes. Some genetic disorders related to the development of GC are hereditary diffuse gastric cancer (HDGC), Lynch syndrome, Peutx-Jeghers syndrome, familial adenomatous polyposis, and Li-Fraumeni syndrome (5, 9). Moreover, the presence of mutations or the deletion of genes as p53, BRCA2, MSH2, and MSH1 is associated with an increased risk of developing GC (13).

Alcohol consumption and tobacco smoking play significant roles in GC development (14). There is evidence that alcohol can irritate and erode the stomach lining, resulting in gastritis, which can be a precursor for GC (15).

Dietary factors are another crucial topic associated with GC. It is well known that diets with a low intake of fruit and vegetables and/or with a high intake of salty and smoked foods are considered risk factors for GC development. As such, high ingestion of salt may damage the mucin layer that protects the stomach epithelium, leading to chronic atrophic gastritis and intestinal metaplasia, which are precursors for GC (5, 9, 16). There is also a correlation between GC development, low socioeconomic status, and low levels of physical activity (16-18).

The risk of developing GC increases with age since more than 80% of the cases occur between 60 and 80 years old. Being male is also considered a risk factor for developing GC, with some theories suggesting that these differences may be explained by environmental exposures, lifestyle factors, or even the potentially protective role of female sex-specific hormones (5).

Globocan 2020 registries indicate that, in Portugal, GC is the sixth most common cancer, affecting more men than women. The total number of new cases of GC in Portugal was 2950, representing 4.9% of all new cancer cases, for both sexes and ages in 2020 (19) (**Figure 1**).

The highest GC incidence rates were observed in Eastern Asia and Eastern Europe, being higher in men than women (2) (**Figure 2**). In developed countries, the incidence and mortality rates of GC have been steadily decreasing over the last decades due to better hygiene habits and the better management of *H. pylori* infection (11, 16, 17).

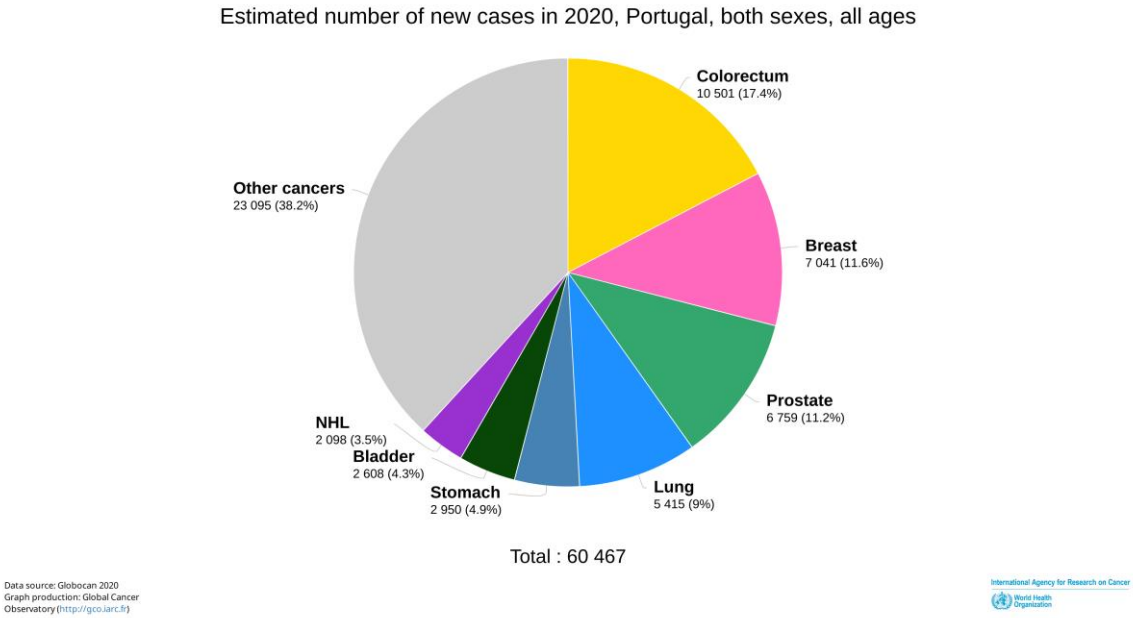

Figure 1. Estimated number of new cancer cases in 2020 in Portugal. Data Source: Globocan 2020; Graph production: Global Cancer Observatory.

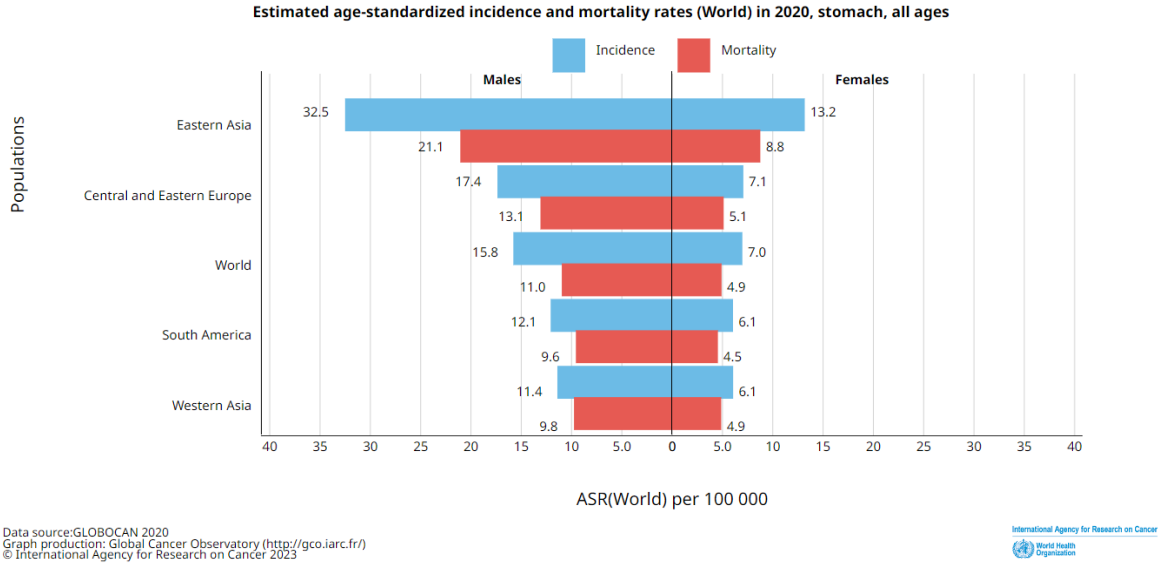

Figure 2. World distribution of gastric cancer incidence and mortality. Data Source: Globocan 2020; Graph production: Global Cancer Observatory.

1.2. Pathophysiology & Classification Systems

GC includes a range of tumours with diverse causes, resulting in significant molecular heterogeneity. The most common histotype of GC is gastric adenocarcinoma (GAC), which represents more than 90% of gastric tumours. GACs arise from the stomach mucosa, which represents the glands of the most superficial layer of the stomach (6).

The oldest and most used classification for GACs is the Lauren classification, which divides gastric adenocarcinomas into three types: intestinal (54%), diffuse (32%) and indeterminate (15%) (**Table 1**) (20-22). This classification is based on the histological characteristics of GACs and has an important role in surgical decisions. Intestinal adenocarcinoma is often associated with *H. pylori* infection and intestinal metaplasia (22). This tumour type, which is more common in older male patients, is marked by visible glands and cohesive tumour cells (5, 9, 23). On the other hand, the diffuse type is characterized by discohesive, small and round cells, associated with poor or no gland formation. The diffuse type adenocarcinomas, which are more frequent in young and female patients, are originated in the normal gastric mucosa and can penetrate the stroma as single or small clusters of cells, facilitating progression to peritoneal dissemination (9, 21). Overall, the intestinal subtype is better differentiated than the diffuse subtype, being associated with a better prognosis (14, 24).

To better stratify GC patients and aid detailed information about histopathological features, the World Health Organization (WHO) has proposed a more refined classification. According to WHO classification, GACs can be subdivided into tubular, papillary, mucinous, and signet-ring cell carcinoma (Table 1) (9, 25). The tubular adenocarcinoma is the most prevalent histologic subtype in this classification (45-65% of GACs), which is associated with early gastric carcinoma (21, 26). The papillary adenocarcinoma variant, also found in early gastric carcinomas, is correlated with poor survival (25, 26). Mucinous adenocarcinomas consist of tumour cells and mucin pools, which occupy more than 50% of the tumour area. The poorly cohesive carcinoma subtype includes signet ring cells, which are similar to lymphocytes and plasma cells, and also incorporate non-signet ring cells (21). The mixed subtype, which is equivalent to the indeterminate type of Lauren classification, includes the most uncommon histologic variants, such as squamous cell carcinoma (24, 26). This heterogeneous subtype, considered more invasive and metastatic, consists of two different components, glandular and poorly cohesive (14, 21).

Table 1. Gastric cancer histological classification systems. Adapted from (2019) Johnston,F. & Beckman,M. (24).

WHO classification	Lauren classification
Papillary adenocarcinoma Tubular adenocarcinoma Mucinous adenocarcinoma	Intestinal type
Signet-ring cell carcinoma Another poorly cohesive carcinoma	Diffuse type
Mixed carcinoma	Indeterminate type

Adding to this classifications systems, the AJCC 8th tumour-node-metastasis (TNM) staging (**Figure 3**) presents an improved approach marked by staging uniformity, discrimination capacity, prognostic accuracy, and consistency in predicting GC prognosis (27). Accordingly, prognosis is determined by considering the depth of the tumour (T), the invasion of local lymph nodes (N), and the presence of distant metastases (M) (14, 28, 29). The T and N stages of the classification system are the most significant indicators of successful GC surgery (30). This classic system, the most used classification system for

GC, has been useful in guiding clinical decision and improving patient outcomes, however, it has certain limitations because clinical decisions may be modified by operator subjectivity and data heterogeneity (29, 31).

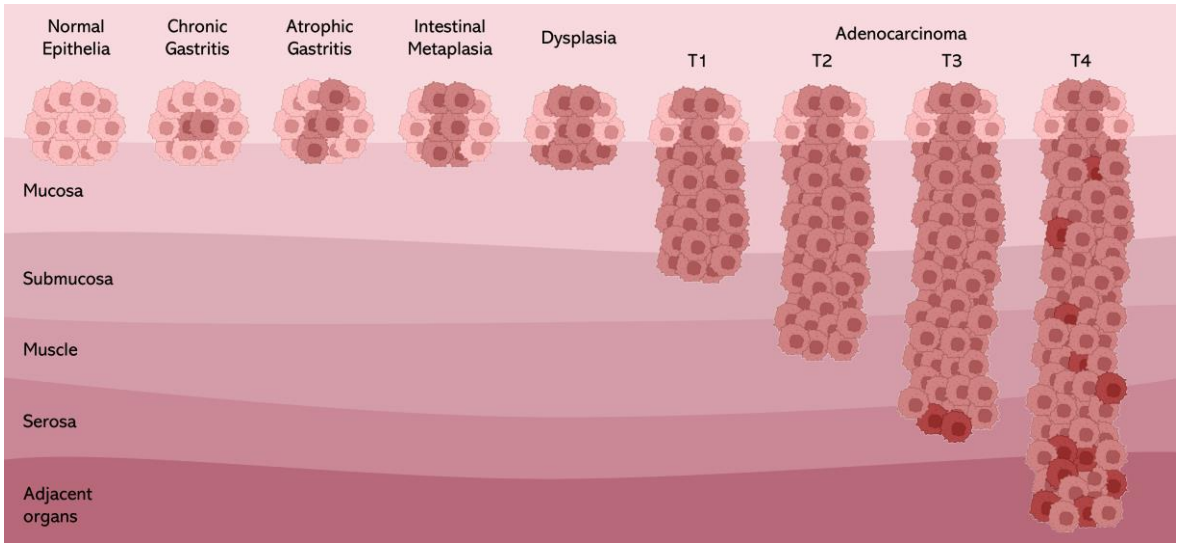

Figure 3. Summary of gastric carcinogenesis and TNM staging system. The transformation of cells, influenced by environmental and host characteristics, progresses from normal cells to adenocarcinoma through the process of carcinogenesis. Concerning the T stage of the tumour, T1 is determined by the invasion into the gastric submucosa; T2 signifies penetration into the muscle layer; T3 indicates invasion into the serosa; and T4 encompasses tumours that invade nearby organs. Image created using Biorender.com (accessed on 23 September 2023).

The GC classification systems previously described translate the histological heterogeneity of gastric tumours, however, they lack robust molecular characterization (32). To overcome this need, The Cancer Genome Atlas (TCGA) project offers a molecular classification, which divides gastric tumours into four subcategories. Namely, tumours positive for Epstein-Barr virus (EBV), tumours with microsatellite instability (MSI), genomically stable tumours (GS) and tumours with chromosomal instability (CIN) (33). According to this classification system, patients with the EBV subtype showed the best prognosis and improved overall survival, whereas the GS subtype exhibited a worse prognosis and decreased overall survival (34). TCGA classification enables the identification of significant genetic phenotypes, nevertheless, clinical routine has not yet benefited from this due to the absence of a robust genetic-clinical correlation and the concurrent occurrence of various molecular subtypes (28, 34, 35).

Globally, the classifications mentioned above face challenges in accurately stratifying GC based on histological and, more recently, molecular characteristics. Nevertheless, the increasing demand for more therapies and tools for patient care has

encouraged research towards integrating molecular data into the currently employed clinical categories. Integrating glycoproteomic data into these well-established categories in the future may present a compelling approach to enhance clinical classifications and provide guidance for therapeutic decisions.

1.3. Gastric Cancer Diagnosis and Therapy

Most GCs are diagnosed in the late stages since patients are often asymptomatic in early stages. Some symptoms associated with the disease are weight loss, anaemia, indigestion, dysphagia and early satiety (20, 28). These are often non-specific symptoms, hampering early diagnoses. In advanced stages, more severe symptoms are often presented namely Virchow's node (left supraclavicular lymphadenopathy), Sister Mary Joseph's node (umbilical node) or even Blumer's shelf (24).

At first diagnosis, GC is identified through imaging techniques (positron emission tomography, computed tomography, laparoscopy) as well as biopsy and cytology to disease assessment and histological classification (18, 24, 28, 32).

Currently, surgical resection is the only possible curative opportunity for patients in the early stages of GC (9). However, endoscopic resection can be used in patients who present early GCs, confined to the mucosa, non-ulcerated and well-differentiated (28). In many cases, adjuvant chemotherapy or chemoradiotherapy are used in resectable localised disease to prevent relapse (32). On the other hand, for locally advanced disease it is recommended neoadjuvant chemoradiotherapy or perioperative chemotherapy (14). Chemotherapy is also used to improve the quality of life for patients with metastatic GC, and first-line treatment is based on a platinum-fluoropyrimidine doublet, associated with an overall response rate ranging between 30% and 60%, but with a median survival time of only 6-9 months (18, 28, 36). Overall, chemotherapy strategies are associated with severe side effects, leading to the development and application of novel approaches to enhance treatment efficacy and minimize adverse effects (18). Advances in genetic profiling of GACs and the discovery of tumour-specific markers have allowed the development of targeted therapies in the form of monoclonal antibodies and small molecule inhibitors (24). As such, patients with advanced gastric carcinomas overexpressing HER2, benefit from chemotherapy together with trastuzumab (anti-HER2 monoclonal antibody), which has improved disease-free survival and overall survival (28). Approaches using immune checkpoint inhibitors such as nivolumab and pembrolizumab, two monoclonal anti-PD-1

antibodies, have shown an improvement in the overall survival of patients, however, their overall response rate is approximately 15% only (37).

The available treatment modalities for individuals with locally advanced and metastatic gastric cancers are limited, emphasizing the urgent need for innovative therapies to improve quality of life and survival.

2. GLYCOSYLATION AND GASTRIC CANCER

2.1. General Features of Protein Glycosylation

Glycosylation is a non-templated enzymatic process present in cells and takes place in the endoplasmic reticulum (ER) and the Golgi apparatus (GA) of eukaryotic cells (38-40). This frequent post-translational modification is achieved by the addition of oligosaccharides to proteins or lipids, giving rise to glycoconjugates, which can be glycoproteins, proteoglycans or glycosphingolipids (1, 41, 42).

Those biomolecules, also known as glycans, share some structural features and terminal modifications between them. The human glycome is very diverse and heterogeneous, even though it is constructed from only ten different monosaccharides: fucose (Fuc), galactose (Gal), glucose (Glc), N-acetylglucosamine (GlcNAc), N-acetyl-galactosamine (GalNAc), glucuronic acid (GlcA), iduronic acid (IdoA), sialic acid (Sia), mannose (Man) and xylose (Xyl) (43-45). Despite all the diversity, glycosylation does not occur randomly, since only specific sites of a protein are glycosylated (46).

Glycans can be expressed on the cell surface or in secreted proteins, playing different physiological roles (46, 47). In the secretory pathway, they allow quality control of protein synthesis, governing folding and targeting to different organelles (45). At the cell surface, glycans influence different cellular functions such as cell growth (48), differentiation and adhesion (39, 42), as well as signal transduction (38, 49), host-pathogen recognition (50, 51) and immune surveillance (52-54). Cell surface glycans are also able to modulate their carrier protein conformation and provide ligands for glycan-binding proteins such as selectins, galectins and siglecs (45, 53, 55).

The most common and better-studied forms of glycosylation are *N*-glycosylation and *O*-glycosylation. *N*-glycans biosynthesis begins in the cytoplasmic side of the ER, where a branched carbohydrate structure, composed of *N*-acetylglucosamine and mannose, is assembled in the lipid carrier dolichol-phosphate (Dol-P). In turn, this product is flipped to the ER lumen, where glucose and mannose saccharides are added. The synthesis continues with the addition of the Dol-P – carbohydrate structure to the amino group of

asparagine (Asn) residue, which belongs to the consensus peptide Asn-X-Ser/Thr (Serine/Threonine) sequence, in which X represents any amino acid except proline. The core structure of *N*-glycans consists of two GlcNAc linked to three mannoses. This structure might undergo several further modifications in the ER and the GA, by the action of different glycosyltransferases, which can add residues of GlcNAc, galactose, sialic acid and fucose. The final products of *N*-glycosylation are very diverse, and they can be grouped as oligomannoses, complex *N*-glycans or hybrid *N*-glycans (38, 39, 56-59).

O-glycosylation also plays an important role in cell biology, and it occurs typically in the GA, with the addition of sugars in the *N*-terminal of the residues, establishing glycosidic linkages in an α or β configuration. *O*-glycans can be classified as mucin-type (also called *O*-GalNAc glycans) or non-mucin-type. The so-called mucin type *O*-glycans caught their name due to their high prevalence in mucin proteins, since more than 80% of their mass consists of *O*-glycans (60). *O*-GalNAc glycans biosynthesis is initiated in the GA by an *N*-acetylgalactosaminyltransferase, which links an *N*-acetylgalactosamine residue from a sugar donor (UDP-GalNAc; Uridine Diphosphate *N*-acetylgalactosamine) to the hydroxyl group of serine, threonine or tyrosine residues, originating the Tn antigen (*α*-GalNAc) (39, 61-63) (**Figure 4**). Tn antigen might be extended to core one structure by the action of core 1 β1-3galactosyltransferase (C1GalT1; T synthase). However, the formation of core 1 structure can be prevented when an α2-6 sialyltransferase ST6GalNAc1 adds a sialic acid to the α-GalNAc residue and generates the sialyl-Tn antigen (STn). On the other hand, β1-6 *N*-acetylglucosaminyltransferases 1, 2, and 3 (C2GnT) can convert the core 1 *O*-glycan into the core 2 structure, with the addition of a GlcNAc residue (48, 62, 64). Elongation of the core 1 structure can be prevented when a terminal sialic acid is added to the T antigen on the galactose residue, forming the Sialyl-T (ST). Similarly to core 1, core 3 *O*-glycans synthesis uses the α-GalNAc structure as substrate, which is added to a β1-3GlcNAc by the β1-3 *N*-acetylglucosaminyltransferase 6 (B3GNT6) enzyme (45, 61, 65). Core 3 structure can also be further branched by adding a β1-6GlcNAc residue through the action of core 4 β-1,6 *N*-acetylglucosaminyltransferase (C4GnT), originating the core 4 structure (60, 66, 67). However, core 3 and core 4 *O*-glycans are less common and more restricted to mucus epithelia from the gastrointestinal and respiratory tracts as well as the salivary glands (39, 62). Additionally, core 1 and core 3 *O*-glycans can be further modified by elongation or by the formation of terminally capped structures with sialic acid and fucose residues, integrating the group of complex *O*-GalNAc glycans (60, 62). One of the most common complex structures of *O*-glycans and *N*-glycans are the Lewis blood group antigens, which clinical relevance will also be explored in subsequent sections (68).

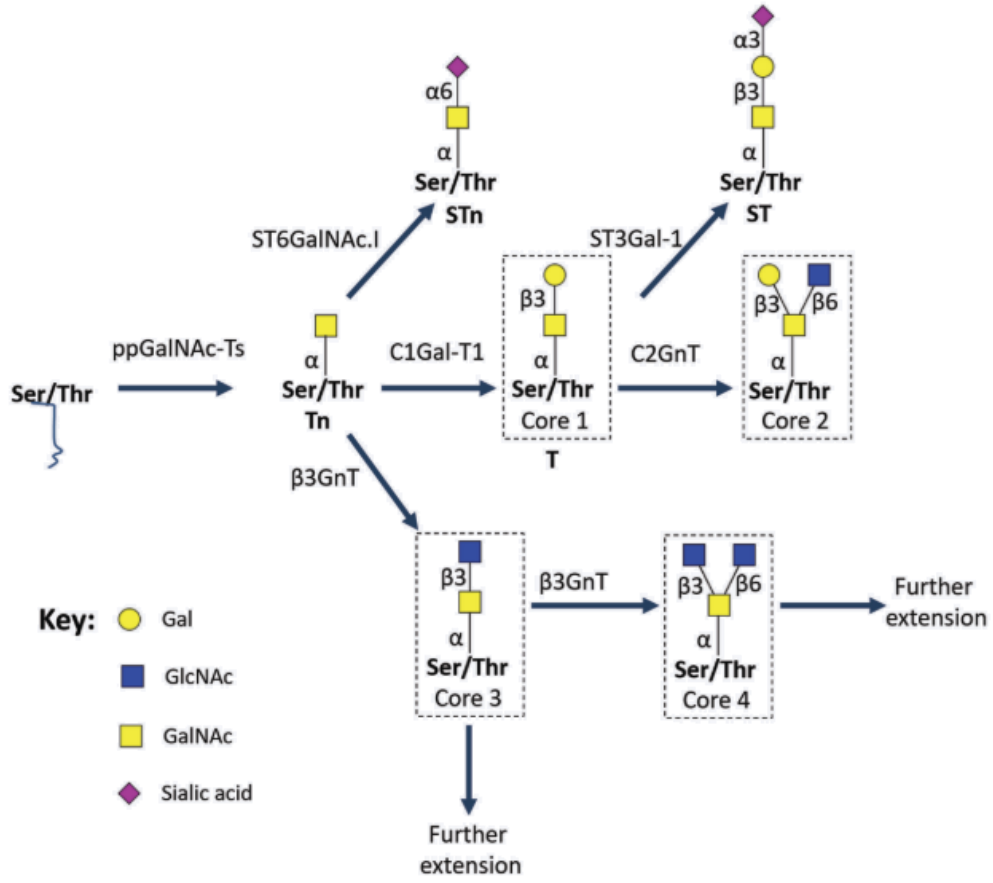

Figure 4. O-glycosylation pathways. The synthesis of Tn antigen bound to Ser/Thr residues of a protein backbone triggers the start of mucin-type *O*-glycosylation on GA. Complex *O*-glycans are created by further modification of Tn. Alternatives include sialylating Tn and T antigens by the sialyltransferases ST6GalNAc1 and ST3Gal1, respectively, resulting in shortened structures, STn and ST antigens. Adapted from Azevedo et al (1).

The dysregulation of glycosylation pathways originates abnormal changes in the expression of glycans. Those alterations include incomplete synthesis of glycans, enhanced expression of complex branched *N*-glycans, expression of truncated *O*-glycans, and altered expression of fucosylated and sialylated structures (39, 69, 70). Glycosylation dysregulation is associated with numerous diseases, such as infection and inflammation (71, 72), congenital and neurological disorders (73, 74), as well as cancer (42, 52, 75). In fact, aberrant glycosylation phenotypes are prevalent in different solid tumours, including gastric cancer (42, 76-79). Changes in glycosylation patterns impact various aspects of cancer cell behaviour, including cell growth, mobility, adhesion, and survival (10, 80). These alterations also influence critical cancer-related processes such as tumour-associated immune responses, epithelial-mesenchymal transition, angiogenesis, cell-cell contact, and metastasis (10, 52, 80-82).

Tumours are characterized by a profound deregulation of *O*-glycosylation pathways, which can be determined by several factors, such as the deregulated expression of glycosyltransferases and glycosidases, the availability of activated sugar donors and the presence of glycoprotein substrates (42, 53, 83). The biosynthesis of tumour-associated antigens can be influenced by the loss of the tightly coordinated expression of specific fucosyltransferases and sialylases, with their substrates. Moreover, the premature stop of *O*-glycosylation extension often causes the overexpression of immature and truncated *O*-glycans, in solid tumours, such as Tn, STn, T and ST antigens (84-86). The overexpression of Tn and STn antigens, may arise due to a mutation or suppression of the C1GalT1-specific chaperone 1 protein (also known as COSMC). This protein serves as an essential mediator for the proper folding and functioning of the C1GalT1 glycosyltransferase, thereby obstructing the C1GalT1 pathway and ultimately leading to an increase in Tn and STn antigens expression (87, 88). The biosynthesis of Sialyl Tn can also be augmented through the impact of the overexpression of ST6GalNAc1 glycosyltransferase, which catalyse STn formation (3, 86).

This intricate interplay of glycosylation dysregulation in cancer underscores its pivotal role in disease progression and highlights the potential for targeted interventions aimed at modulating these glycan alterations for improved clinical outcomes.

2.2. Gastric Cancer-Associated Glycosylation

The glycosylation patterns found in advanced gastric cancer play pivotal roles in its progression and aggressiveness. As such, upregulation of the fucosyltransferase FUT3 has been associated with aberrant expression of Lewis blood group antigens, specifically Le^A, in GC (10). Additionally, the overexpression of sialylated Le^A and Le^X structures is a widely recognized occurrence in the development of GC, being associated with tumour aggressiveness and overall poor survival (47, 89). Moreover, the increased expression of sialylated Lewis antigens has been correlated with elevated levels of ST6Gal3, ST3Gal4 and FUT5 enzymes in both *in vitro* and *in vivo* studies (10).

Tn antigen is expressed in over 70% of most human solid tumours, and has been associated with cancer progression, metastasis and poor prognosis in several cancers. However the molecular mechanisms by which the Tn antigen modulates tumour progression remains poorly understood (45). It is noteworthy that the implications associated with Tn antigen expression in gastric tumours have received limited attention.

Only one study, conducted on canine tissues, addressed this issue and found no significant results regarding the role of Tn antigen in neoplastic transformation (90).

STn antigen is rarely observed in normal tissues, while it is aberrantly expressed in several tumours, including GC (86, 91, 92). The presence of the STn antigen in gastric tumours is associated with a more aggressive phenotype (10) and correlated with increased cellular migration (93) and invasion (94), which is linked to enhanced metastatic potential (95). Moreover, STn overexpressing cells have been associated with decreased cell-cell aggregation (96), resistance to chemotherapy (65) and an overall poorer prognosis linked to cancer aggressiveness (76, 84, 92, 97). Accordingly, *in vivo* studies have shown GC cells overexpressing STn antigen displaying heightened invasive and metastatic potential (96, 98). Another study provided evidence that the expression of short-chain *O*-glycans was accompanied by the acquisition of mesenchymal-like characteristics in cells, which led to the remodelling of gene expression and the activation of tyrosine kinase receptors. These findings suggest that the presence of STn is involved in the continuous activation of oncogenic signalling in gastric cancer (94). Moreover, the decreased survival observed in COSMC knock-out mouse xenografts suggests a more aggressive phenotype in GC cells with overexpression of short-chain *O*-glycans (94).

Overall, aberrant glycosylation patterns observed in advanced gastric cancer, particularly the overexpression of sialyl-Tn antigens, underscores the critical role of glycoproteins in the progression and aggressiveness of the disease. However, further research in glycoproteomics holds great potential for refining diagnostic and prognostic strategies in the management of gastric cancer, ultimately offering improved outcomes for patients facing this challenging disease.

3. GLYCOPROTEOMICS RELEVANCE IN GASTRIC CANCER MOLECULAR CHARACTERIZATION

Glycoproteomics consists of the study of both glycans and carrier proteins in a comprehensive and omics fashion (99). Glycoproteomics relies on a sophisticated array of techniques to dissect the intricate interplay between glycans and carrier proteins, demanding a meticulous examination of the glycome, followed by the proteome, enabling the investigation of specific proteomic aspects (99). These techniques encompass high-resolution mass spectrometry, which allows for precise characterization of glycoproteins and their glycan structures. Additionally, glycoprotein enrichment methods, such as lectin

affinity chromatography and antibody-based strategies, allow isolating glycosylated proteins for in-depth analysis (100, 101). Despite the remarkable progress in glycoproteomic methodologies, certain limitations persist, which are primarily characterized by the requirement for specialized training and access to advanced instrumentation. Moreover, the sheer intricacy and heterogeneity of glycan structures, coupled with their dynamic nature, are also challenges in this field. Furthermore, the acquisition of sufficient quantities of high-quality glycoproteins, combined with the absence of standardized protocols and methods, can influence glycoproteomic analysis and contribute to increased variability in the results (102, 103).

Approximately half of eukaryotic proteins undergo glycosylation, especially those on the cell surface, displaying significant roles on cell signalling, adhesion, and immune response modulation (38, 104). In the context of gastric cancer, aberrant glycosylation patterns have emerged as hallmark features, being closely associated with the development, progression, and aggressiveness of gastric tumours (10, 42). Understanding the glycoproteomic profile of GC cells provides critical insights into the underlying molecular mechanisms driving this malignancy, highlighting potential biomarkers for diagnosis and prognosis, offering a window into the progression of the disease and response to treatment. Previous studies have identified two *O*-glycoproteins found in glycoengineered GC cell lines, as well as in the serum of GC patients and GC tissues (105). While additional validation is required to confirm their potential as biomarkers for GC, these findings underscore the significant potential of glycoproteomics in enhancing clinical strategies and refining disease management.

Overall, the composition of the glycoproteome, which corresponds to a subset of proteins highly dependent on the microenvironment, is intricately influenced by factors spanning the glycome, transcriptome, proteome, and metabolome (99, 106). Therefore, glycoproteomics is a powerful tool in unravelling the intricate biology of glycans and their associated proteins, offering valuable insights into the molecular mechanisms underlying diseases like gastric cancer. As such, we hypothesise that modifications in the glycoproteomic profile on the cell surface of GC cells will influence cell aggressiveness. In this context, the study of the glycoproteome of the cells stands as an indispensable resource in identifying promising targets for diagnosis or therapy.

4. REFERENCES

1. Azevedo R, Peixoto A, Gaiteiro C, Fernandes E, Neves M, Lima L, et al. Over forty years of bladder cancer glycobiology: Where do glycans stand facing precision oncology? Oncotarget. 2017;8(53):91734-64.
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021;71(3):209-49.
3. Diniz F, Azevedo M, Sousa F, Osório H, Campos D, Sampaio P, et al. Polymeric nanoparticles targeting Sialyl-Tn in gastric cancer: A live tracking under flow conditions. Materials Today Bio. 2022;16:100417.
4. William W, Stella AVN, Sean C, Manuel R-J, Manon S, Ernst JK, et al. Recent advances in the detection and management of early gastric cancer and its precursors. Frontline Gastroenterology. 2021;12(4):322.
5. Ilic M, Ilic I. Epidemiology of stomach cancer. World J Gastroenterol. 2022;28(12):1187-203.
6. Karimi P, Islami F, Anandasabapathy S, Freedman ND, Kamangar F. Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention. Cancer Epidemiol Biomarkers Prev. 2014;23(5):700-13.
7. Lee YC, Chiang TH, Chou CK, Tu YK, Liao WC, Wu MS, et al. Association Between Helicobacter pylori Eradication and Gastric Cancer Incidence: A Systematic Review and Meta-analysis. Gastroenterology. 2016;150(5):1113-24.e5.
8. Thrift AP, Nguyen TH. Gastric Cancer Epidemiology. Gastrointestinal Endoscopy Clinics of North America. 2021;31(3):425-39.
9. Machlowska J, Baj J, Sitarz M, Maciejewski R, Sitarz R. Gastric Cancer: Epidemiology, Risk Factors, Classification, Genomic Characteristics and Treatment Strategies. Int J Mol Sci. 2020;21(11).
10. Duarte HO, Freitas D, Gomes C, Gomes J, Magalhães A, Reis CA. Mucin-Type O-Glycosylation in Gastric Carcinogenesis. Biomolecules. 2016;6(3).
11. Petryszyn P, Chapelle N, Matysiak-Budnik T. Gastric Cancer: Where Are We Heading? Digestive Diseases. 2020;38(4):280-5.
12. Correa P. Gastric cancer: overview. Gastroenterol Clin North Am. 2013;42(2):211-7.
13. Lynch HT, Grady W, Suriano G, Huntsman D. Gastric cancer: new genetic developments. J Surg Oncol. 2005;90(3):114-33; discussion 33.

14. Sexton RE, Al Hallak MN, Diab M, Azmi AS. Gastric cancer: a comprehensive review of current and future treatment strategies. Cancer Metastasis Rev. 2020;39(4):1179-203.
15. Rawla P, Barsouk A. Epidemiology of gastric cancer: global trends, risk factors and prevention. Prz Gastroenterol. 2019;14(1):26-38.
16. den Hoed CM, Kuipers EJ. Gastric Cancer: How Can We Reduce the Incidence of this Disease? Curr Gastroenterol Rep. 2016;18(7):34.
17. Thrift AP, El-Serag HB. Burden of Gastric Cancer. Clin Gastroenterol Hepatol. 2020;18(3):534-42.
18. Smyth EC, Nilsson M, Grabsch HI, van Grieken NCT, Lordick F. Gastric cancer. The Lancet. 2020;396(10251):635-48.
19. World Health Organization. Estimated number of new cases in 2020, Portugal, both sexes, all ages International Agency for Research on Cancer; 2020 [Available from: https://gco.iarc.fr/today/online-analysis-pie?v=2020&mode=cancer&mode_population=continents&population=900&populations=620&key=total&sex=0&cancer=39&type=0&statistic=5&prevalence=0&population_group=0&ages_group%5B%5D=0&ages_group%5B%5D=17&nb_items=7&group_cancer=1&include_nmsc=1&include_nmsc_other=1&half_pie=0&donut=0.
20. Ajani JA, Lee J, Sano T, Janjigian YY, Fan D, Song S. Gastric adenocarcinoma. Nature Reviews Disease Primers. 2017;3(1):17036.
21. Businello G, Galuppini F, Fassan M. The impact of recent next generation sequencing and the need for a new classification in gastric cancer. Best Practice & Research Clinical Gastroenterology. 2021;50-51:101730.
22. Lauren P. 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. 1965;64:31-49.
23. Berlth F, Bollschweiler E, Drebber U, Hoelscher AH, Moenig S. Pathohistological classification systems in gastric cancer: diagnostic relevance and prognostic value. World J Gastroenterol. 2014;20(19):5679-84.
24. Johnston FM, Beckman M. Updates on Management of Gastric Cancer. Curr Oncol Rep. 2019;21(8):67.
25. Lai JF, Xu WN, Noh SH, Lu WQ. Effect of World Health Organization (WHO) Histological Classification on Predicting Lymph Node Metastasis and Recurrence in Early Gastric Cancer. Med Sci Monit. 2016;22:3147-53.
26. Hu B, El Hajj N, Sittler S, Lammert N, Barnes R, Meloni-Ehrig A. Gastric cancer: Classification, histology and application of molecular pathology. J Gastrointest Oncol. 2012;3(3):251-61.

27. Sano T, Coit DG, Kim HH, Roviello F, Kassab P, Wittekind C, et al. Proposal of a new stage grouping of gastric cancer for TNM classification: International Gastric Cancer Association staging project. Gastric Cancer. 2017;20(2):217-25.
28. Lordick F, Carneiro F, Cascinu S, Fleitas T, Haustermans K, Piessen G, et al. Gastric cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. Annals of Oncology. 2022;33(10):1005-20.
29. Seevaratnam R, Cardoso R, McGregor C, Lourenco L, Mahar A, Sutradhar R, et al. How useful is preoperative imaging for tumor, node, metastasis (TNM) staging of gastric cancer? A meta-analysis. Gastric Cancer. 2012;15(1):3-18.
30. Tsendsuren T, Jun SM, Mian XH. Usefulness of endoscopic ultrasonography in preoperative TNM staging of gastric cancer. World J Gastroenterol. 2006;12(1):43-7.
31. Zhang M, Ding C, Xu L, Ou B, Feng S, Wang G, et al. Comparison of a Tumor-Ratio-Metastasis Staging System and the 8th AJCC TNM Staging System for Gastric Cancer. Front Oncol. 2021;11:595421.
32. Ajani JA, Lee J, Sano T, Janjigian YY, Fan D, Song S. Gastric adenocarcinoma. Nat Rev Dis Primers. 2017;3:17036.
33. The Cancer Genome Atlas Research N, et al. Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 2014;513(7517):202-9.
34. Sohn BH, Hwang JE, Jang HJ, Lee HS, Oh SC, Shim JJ, et al. Clinical Significance of Four Molecular Subtypes of Gastric Cancer Identified by The Cancer Genome Atlas Project. Clin Cancer Res. 2017;23(15):4441-9.
35. Röcken C. Molecular classification of gastric cancer. Expert Rev Mol Diagn. 2017;17(3):293-301.
36. Bhat G. Retrospective study of oxaliplatin, leucovarin and 5 fluoruracil regimen in patients with advanced gastric cancer with poor performance status: A study at a tertiary center of South India. South Asian J Cancer. 2018;7(4):223-5.
37. Takei S, Kawazoe A, Shitara K. The New Era of Immunotherapy in Gastric Cancer. Cancers (Basel). 2022;14(4).
38. Reily C, Stewart TJ, Renfrow MB, Novak J. Glycosylation in health and disease. Nat Rev Nephrol. 2019;15(6):346-66.
39. Thomas D, Rathinavel AK, Radhakrishnan P. Altered glycosylation in cancer: A promising target for biomarkers and therapeutics. Biochim Biophys Acta Rev Cancer. 2021;1875(1):188464.
40. Spiro RG. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. Glycobiology. 2002;12(4):43r-56r.

41. Schjoldager KT, Narimatsu Y, Joshi HJ, Clausen H. Global view of human protein glycosylation pathways and functions. Nature Reviews Molecular Cell Biology. 2020;21(12):729-49.
42. Ferreira JA, Magalhães A, Gomes J, Peixoto A, Gaiteiro C, Fernandes E, et al. Protein glycosylation in gastric and colorectal cancers: Toward cancer detection and targeted therapeutics. Cancer Letters. 2017;387:32-45.
43. Moremen KW, Tiemeyer M, Nairn AV. Vertebrate protein glycosylation: diversity, synthesis and function. Nature Reviews Molecular Cell Biology. 2012;13(7):448-62.
44. Smith BAH, Bertozzi CR. The clinical impact of glycobiology: targeting selectins, Siglecs and mammalian glycans. Nature Reviews Drug Discovery. 2021;20(3):217-43.
45. Chia J, Goh G, Bard F. Short O-GalNAc glycans: regulation and role in tumor development and clinical perspectives. Biochimica et Biophysica Acta (BBA) - General Subjects. 2016;1860(8):1623-39.
46. Aebersold R, Agar JN, Amster IJ, Baker MS, Bertozzi CR, Boja ES, et al. How many human proteoforms are there? Nat Chem Biol. 2018;14(3):206-14.
47. Magalhães A, Duarte HO, Reis CA. The role of O-glycosylation in human disease. Molecular Aspects of Medicine. 2021;79:100964.
48. Tran DT, Ten Hagen KG. Mucin-type O-glycosylation during development. J Biol Chem. 2013;288(10):6921-9.
49. Ohtsubo K, Marth JD. Glycosylation in Cellular Mechanisms of Health and Disease. Cell. 2006;126(5):855-67.
50. Lin B, Qing X, Liao J, Zhuo K. Role of Protein Glycosylation in Host-Pathogen Interaction. Cells. 2020;9(4).
51. Thompson AJ, de Vries RP, Paulson JC. Virus recognition of glycan receptors. Curr Opin Virol. 2019;34:117-29.
52. Rodrigues JG, Balmaña M, Macedo JA, Poças J, Fernandes Â, de-Freitas-Junior JCM, et al. Glycosylation in cancer: Selected roles in tumour progression, immune modulation and metastasis. Cell Immunol. 2018;333:46-57.
53. Marth JD, Grewal PK. Mammalian glycosylation in immunity. Nature Reviews Immunology. 2008;8(11):874-87.
54. Ugonotti J, Chatterjee S, Thaysen-Andersen M. Structural and functional diversity of neutrophil glycosylation in innate immunity and related disorders. Mol Aspects Med. 2021;79:100882.
55. Varki A. Biological roles of glycans. Glycobiology. 2017;27(1):3-49.
56. Oliveira-Ferrer L, Legler K, Milde-Langosch K. Role of protein glycosylation in cancer metastasis. Semin Cancer Biol. 2017;44:141-52.

57. Radovani B, Gudelj I. N-Glycosylation and Inflammation; the Not-So-Sweet Relation. Front Immunol. 2022;13:893365.
58. Stanley P, Moremen KW, Lewis NE, Taniguchi N, Aebi M. N-Glycans. In: Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al., editors. Essentials of Glycobiology. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press
Copyright © 2022 The Consortium of Glycobiology Editors, La Jolla, California; published by Cold Spring Harbor Laboratory Press; doi:10.1101/glycobiology.4e.9. All rights reserved.; 2022. p. 103-16.
59. Bieberich E. Synthesis, Processing, and Function of N-glycans in N-glycoproteins. Adv Neurobiol. 2014;9:47-70.
60. Fu C, Zhao H, Wang Y, Cai H, Xiao Y, Zeng Y, et al. Tumor-associated antigens: Tn antigen, sTn antigen, and T antigen. Hla. 2016;88(6):275-86.
61. González-Ramírez AM, Grosso AS, Yang Z, Compañón I, Coelho H, Narimatsu Y, et al. Structural basis for the synthesis of the core 1 structure by C1GalT1. Nat Commun. 2022;13(1):2398.
62. Brockhausen I, Wandall HH, Hagen KGT, Stanley P. O-GalNAc Glycans. In: Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al., editors. Essentials of Glycobiology. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press
Copyright © 2022 The Consortium of Glycobiology Editors, La Jolla, California; published by Cold Spring Harbor Laboratory Press; doi:10.1101/glycobiology.4e.10. All rights reserved.; 2022. p. 117-28.
63. 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.
64. Ju T, Aryal RP, Kudelka MR, Wang Y, Cummings RD. The Cosmc connection to the Tn antigen in cancer. Cancer Biomark. 2014;14(1):63-81.
65. Santos SN, Junqueira MS, Francisco G, Vilanova M, Magalhães A, Dias Baruffi M, et al. O-glycan sialylation alters galectin-3 subcellular localization and decreases chemotherapy sensitivity in gastric cancer. Oncotarget. 2016;7(50):83570-87.
66. Yeh J-C, Ong E, Fukuda M. Molecular Cloning and Expression of a Novel N-Acetylglucosaminyltransferase That Forms Core 2, Core 4, and I Branches. Journal of Biological Chemistry. 1999;274(5):3215-21.
67. Gill DJ, Clausen H, Bard F. Location, location, location: new insights into O-GalNAc protein glycosylation. Trends in Cell Biology. 2011;21(3):149-58.
68. 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. Journal of Proteome Research. 2013;12(3):1454-66.

69. Mereiter S, Balmaña M, Campos D, Gomes J, Reis CA. Glycosylation in the Era of Cancer-Targeted Therapy: Where Are We Heading? Cancer Cell. 2019;36(1):6-16.
70. Pinho SS, Reis CA. Glycosylation in cancer: mechanisms and clinical implications. Nature Reviews Cancer. 2015;15(9):540-55.
71. Delannoy C, Huang C, Coddeville B, Chen JY, Mouajjah D, Groux-Degroote S, et al. Mycobacterium bovis BCG infection alters the macrophage N-glycome. Mol Omics. 2020;16(4):345-54.
72. Groux-Degroote S, Cavdarli S, Uchimura K, Allain F, Delannoy P. Glycosylation changes in inflammatory diseases. Adv Protein Chem Struct Biol. 2020;119:111-56.
73. Paprocka J, Jezela-Stanek A, Tylki-Szymańska A, Grunewald S. Congenital Disorders of Glycosylation from a Neurological Perspective. Brain Sci. 2021;11(1).
74. Ondruskova N, Cechova A, Hansikova H, Honzik T, Jaeken J. Congenital disorders of glycosylation: Still "hot" in 2020. Biochim Biophys Acta Gen Subj. 2021;1865(1):129751.
75. Ferreira JA, Videira PA, Lima L, Pereira S, Silva M, Carrascal M, et al. Overexpression of tumour-associated carbohydrate antigen sialyl-Tn in advanced bladder tumours. Mol Oncol. 2013;7(3):719-31.
76. Fernandes E, Soares J, Cotton S, Peixoto A, Ferreira D, Freitas R, et al. Esophageal, gastric and colorectal cancers: Looking beyond classical serological biomarkers towards glycoproteomics-assisted precision oncology. Theranostics. 2020;10(11):4903-28.
77. Peixoto A, Relvas-Santos M, Azevedo R, Santos LL, Ferreira JA. Protein Glycosylation and Tumor Microenvironment Alterations Driving Cancer Hallmarks. Front Oncol. 2019;9:380.
78. Ferreira JA, Peixoto A, Neves M, Gaiteiro C, Reis CA, Assaraf YG, et al. Mechanisms of cisplatin resistance and targeting of cancer stem cells: Adding glycosylation to the equation. Drug Resist Updat. 2016;24:34-54.
79. Julien S, Videira PA, Delannoy P. Sialyl-tn in cancer: (how) did we miss the target? Biomolecules. 2012;2(4):435-66.
80. Stowell SR, Ju T, Cummings RD. Protein glycosylation in cancer. Annu Rev Pathol. 2015;10:473-510.
81. Häuselmann I, Borsig L. Altered tumor-cell glycosylation promotes metastasis. Front Oncol. 2014;4:28.
82. Freitas R, Peixoto A, Ferreira E, Miranda A, Santos LL, Ferreira JA. Immunomodulatory glycomedicine: Introducing next generation cancer glycovaccines. Biotechnology Advances. 2023;65:108144.

83. Macedo-da-Silva J, Santiago VF, Rosa-Fernandes L, Marinho CRF, Palmisano G. Protein glycosylation in extracellular vesicles: Structural characterization and biological functions. Molecular Immunology. 2021;135:226-46.
84. David L, Nesland JM, Clausen H, Carneiro F, Sobrinho-Simões M. Simple mucin-type carbohydrate antigens (Tn, sialosyl-Tn and T) in gastric mucosa, carcinomas and metastases. APMIS Suppl. 1992;27:162-72.
85. Chung YS, Yamashita Y, Kato Y, Nakata B, Sawada T, Sowa M. Prognostic significance of T antigen expression in patients with gastric carcinoma. Cancer. 1996;77(9):1768-73.
86. 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(4):1443-55.
87. Hofmann BT, Schlüter L, Lange P, Mercanoglu B, Ewald F, Fölster A, et al. COSMC knockdown mediated aberrant O-glycosylation promotes oncogenic properties in pancreatic cancer. Mol Cancer. 2015;14:109.
88. Ju T, Lanneau GS, Gautam T, Wang Y, Xia B, Stowell SR, et al. Human Tumor Antigens Tn and Sialyl Tn Arise from Mutations in Cosmc. Cancer Research. 2008;68(6):1636-46.
89. Trinchera M, Aronica A, Dall'Olio F. Selectin Ligands Sialyl-Lewis a and Sialyl-Lewis x in Gastrointestinal Cancers. Biology (Basel). 2017;6(1).
90. Flores AR, Lemos I, Rema A, Taulescu M, Seixas F, Reis CA, et al. Tn and Sialyl-Tn antigens in canine gastric tissues. Veterinary and Comparative Oncology. 2020;18(4):615-25.
91. Marcos NT, Pinho S, Grandela C, Cruz A, Samyn-Petit Bnd, 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 Research. 2004;64(19):7050-7.
92. Victorzon M, Nordling S, Nilsson O, Roberts PJ, Haglund C. Sialyl Tn antigen is an independent predictor of outcome in patients with gastric cancer. International Journal of Cancer. 1996;65(3):295-300.
93. Tamura F, Sato Y, Hirakawa M, Yoshida M, Ono M, Osuga T, et al. RNAi-mediated gene silencing of ST6GalNAc I suppresses the metastatic potential in gastric cancer cells. Gastric Cancer. 2016;19(1):85-97.
94. Freitas D, Campos D, Gomes J, Pinto F, Macedo JA, Matos R, et al. O-glycans truncation modulates gastric cancer cell signaling and transcription leading to a more aggressive phenotype. EBioMedicine. 2019;40:349-62.

95. Neves M, Azevedo R, Lima L, Oliveira MI, Peixoto A, Ferreira D, et al. Exploring sialyl-Tn expression in microfluidic-isolated circulating tumour cells: A novel biomarker and an analytical tool for precision oncology applications. New Biotechnology. 2019;49:77-87.
96. 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 Letters. 2007;249(2):157-70.
97. Cotton S, Azevedo R, Gaiteiro C, Ferreira D, Lima L, Peixoto A, et al. Targeted O-glycoproteomics explored increased sialylation and identified MUC16 as a poor prognosis biomarker in advanced-stage bladder tumours. Mol Oncol. 2017;11(8):895-912.
98. Ozaki H, Matsuzaki H, Ando H, Kaji H, Nakanishi H, Ikehara Y, et al. Enhancement of metastatic ability by ectopic expression of ST6GalNAcI on a gastric cancer cell line in a mouse model. Clin Exp Metastasis. 2012;29(3):229-38.
99. Oliveira T, Thaysen-Andersen M, Packer NH, Kolarich D. The Hitchhiker's guide to glycoproteomics. Biochem Soc Trans. 2021;49(4):1643-62.
100. Thomas DR, Scott NE. Glycoproteomics: growing up fast. Current Opinion in Structural Biology. 2021;68:18-25.
101. Pan S, Chen R, Aebersold R, Brentnall TA. Mass Spectrometry Based Glycoproteomics—From a Proteomics Perspective *. Molecular & Cellular Proteomics. 2011;10(1).
102. Bagdonaite I, Malaker SA, Polasky DA, Riley NM, Schjoldager K, Vakhrushev SY, et al. Glycoproteomics. Nature Reviews Methods Primers. 2022;2(1):48.
103. Di Marco F, Blöchl C, Esser-Skala W, Schäpertöns V, Zhang T, Wuhrer M, et al. Glycoproteomics of a Single Protein: Revealing Tens of Thousands of Myozyme Glycoforms by Hybrid HPLC-MS Approaches. Mol Cell Proteomics. 2023;22(9):100622.
104. Corfield A. Eukaryotic protein glycosylation: a primer for histochemists and cell biologists. Histochem Cell Biol. 2017;147(2):119-47.
105. Campos D, Freitas D, Gomes J, Magalhães 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.
106. Rudd P, Karlsson NG, Khoo KH, Packer NH. Glycomics and Glycoproteomics. In: Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al., editors. Essentials of Glycobiology. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press
 2015. p. 653-66.

CHAPTER II

AIM & SCOPES

Gastric cancer (GC), a prevalent malignancy globally, remains the fourth leading cause of cancer-related deaths (1). A substantial percentage of patients with GC are diagnosed in advanced disease stages, posing an obstacle to the availability of effective therapies. The management of the disease is a challenge due to lack of efficient molecular tools to assist clinical decisions and few effective targeted therapies (2). The premature stop in the *O*-glycosylation pathway is a well described hallmark of many cancers, including gastric carcinomas, leading to the overexpression of short-chain *O*-glycans, such as Tn and STn antigens. Those glycans drive oncogenic features of cancer cells, modifying their ability to proliferate, migrate, and invade (3-6). Therefore, the characterization of the glycoproteome in cancer cells has gained remarkable significance in discovering novel targets relevant to therapy, thus enhancing clinical strategies (4, 7, 8). Accordingly, **Chapter III** aimed to understand the functional implications of the overexpression of truncated *O*-glycans, namely Tn and STn antigens, in a gastric cancer cellular model. It addresses the following specific objectives:

- **I.** Establish a glycoengineered gastric cancer cellular model overexpressing the Tn and STn antigens;
- **II.** Characterize the genotypic profile of the glycoengineered cell line;
- **III.** Characterize the phenotype of the glycoengineered model;
- **IV.** Determine the functional implications of Tn and STn overexpression on gastric cancer cells.

Based on the previously established cell line, arises the need to study its glycoproteome to understand the alterations induced by glycosylation. Accordingly, **Chapter IV** focused on developing a standardized protocol for isolating membrane-associated proteins from cellular content to enable targeted glycoproteomic studies. It addresses the following specific aim:

- **I.** Determine the optimal conditions for cellular labelling with biotin, envisaging a robust protocol for glycoprotein annotation.

The chapters follow a research paper format comprised by an abstract, introduction, materials and methods, results, discussion, and conclusions. **Chapter V** addresses the overall conclusions of the work delves into possible future perspectives.

REFERENCES

1. Machlowska J, Baj J, Sitarz M, Maciejewski R, Sitarz R. Gastric Cancer: Epidemiology, Risk Factors, Classification, Genomic Characteristics and Treatment Strategies. Int J Mol Sci. 2020;21(11).
2. Carrasco-Garcia E, García-Puga M, Arevalo S, Matheu A. Towards precision medicine: linking genetic and cellular heterogeneity in gastric cancer. Ther Adv Med Oncol. 2018;10:1758835918794628.
3. Ferreira JA, Magalhães A, Gomes J, Peixoto A, Gaiteiro C, Fernandes E, et al. Protein glycosylation in gastric and colorectal cancers: Toward cancer detection and targeted therapeutics. Cancer Letters. 2017;387:32-45.
4. Campos D, Freitas D, Gomes J, Magalhães 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.
5. 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.
6. Freitas D, Campos D, Gomes J, Pinto F, Macedo JA, Matos R, et al. O-glycans truncation modulates gastric cancer cell signaling and transcription leading to a more aggressive phenotype. EBioMedicine. 2019;40:349-62.
7. Peixoto A, Ferreira D, Azevedo R, Freitas R, Fernandes E, Relvas-Santos M, et al. Glycoproteomics identifies HOMER3 as a potentially targetable biomarker triggered by hypoxia and glucose deprivation in bladder cancer. J Exp Clin Cancer Res. 2021;40(1):191.
8. Fernandes E, Freitas R, Ferreira D, Soares J, Azevedo R, Gaiteiro C, et al. Nucleolin-Sle A Glycoforms as E-Selectin Ligands and Potentially Targetable Biomarkers at the Cell Surface of Gastric Cancer Cells. Cancers. 2020;12(4):861.

CHAPTER III

ESTABLISHMENT OF A GLYCOENGINEERED GASTRIC CANCER CELL LINE LIBRARY

Establishment of a Glycoengineered Gastric Cancer Cell Line Library

Martina Dourado-Gonçalves[1,2], Dylan Ferreira[1,2,3], Marta Relvas-Santos[1,2,3,4] Paula Paulo[5], Lúcio Lara Santos[1,2,6,7,8,9], Andreia Peixoto[1,3], José Alexandre Ferreira[1,2,6,7,9]

[1]Experimental Pathology and Therapeutics Group, IPO Porto Research Center (CI-IPOP), Portuguese Oncology Institute (IPO Porto), 4200-072 Porto, Portugal; [2]Institute of Biomedical Sciences Abel Salazar (ICBAS), University of Porto, 4050-013; [3]Institute for Research and Innovation in Health (i3S), University of Porto, 4200-135 Porto, Portugal / Institute for Biomedical Engineering (INEB), University of Porto, 4200-135 Porto, Portugal; [4]REQUIMTE-LAQV, Department of Chemistry and Biochemistry, Faculty of Sciences of the University of Porto, 4169-007 Porto, Portugal; [5]Cancer Genetics Group, IPO Porto Research Center (CI-IPOP), Portuguese Oncology Institute, 4200-072 Porto, Portugal; [6]Health School of University Fernando Pessoa, 4249-004 Porto, Portugal; [7]GlycoMatters Biotech, 4500-162 Espinho, Portugal; [8]Department of Surgical Oncology, Portuguese Oncology Institute (IPO Porto), 4200-072 Porto, Portugal; [9]Porto Comprehensive Cancer Center (P.ccc), 4200-072 Porto, Portugal.

Corresponding author:

José Alexandre Ferreira (jose.a.ferreira@ipoporto.min-saude.pt)

Experimental Pathology and Therapeutics Group, Research Centre, Portuguese Oncology Institute of Porto, R. Dr. António Bernardino de Almeida 62, 4200-072 Porto, Portugal; Tel. +351 225084000 (ext. 5111).

ABSTRACT

Gastric cancer (GC) is currently managed through chemotherapy, radiotherapy, and surgery, which often fail to prevent disease progression and relapse. Targeted therapies and immunotherapy have provided significant, however insufficient improvements, in part due to lack of cancer specific signatures. In this sense, exploiting post-transcriptional modifications such as glycosylation appear as a strategy to identify novel molecular signatures. Aberrant protein glycosylation is a hallmark of cancer, contributing significantly to carcinogenesis, cancer progression, and metastasis characterized by the overexpression of short-chain *O*-glycans such as Tn and Sialyl-Tn (STn) antigens, which are mostly absent in the corresponding healthy tissues. Truncated *O*-glycans are prevalent in more aggressive and advanced gastric cancer tumours, being closely linked to tumour progression and unfavourable prognoses. As such, targeting specific glycans or glycoproteoforms may complement conventional cancer detection and treatment, leading to better theragnostic settings. In this sense, this work aimed to develop and characterize a GC cell line overexpressing Tn and STn antigens and determine their functional implications on cell aggressiveness. Accordingly, intestinal-type AGS gastric cancer cell line was glycoengineered using CRISPR-Cas9 technology to knock-out the *C1GALT1* gene, leading to premature stop in the *O*-glycosylation pathway and consequently overexpress Tn and STn antigens. The cell model was then validated through Indel Detection by Amplicon Analysis and Sanger sequencing. Furthermore, the model was phenotypically validated by immunofluorescence and flow cytometry analysis. The assessment of functional characterization was conducted through *in vitro* proliferation, migration, invasion, and colony formation assays. We observed that Tn and STn overexpressing cells exhibited reduced migratory ability while displaying increased capacity to invade and form colonies in an anchorage-independent manner. In summary, these results strongly indicate that truncation of *O*-glycans directly triggers oncogenic features related to cell invasion and resistance to anoikis. As such, defining the glycoengineered cells' glycoproteome might help envisage future molecular targets for early diagnosis and new therapies for gastric cancer.

Keywords: Gastric cancer (GC); Glycosylation; *O*-glycans; Sialyl-Tn antigen (STn); Tn antigen

1. INTRODUCTION

Gastric cancer (GC) is a significant global health concern, ranking as the fifth most incident cancer and the fourth leading cause of cancer-related deaths worldwide (1, 2). This disease, with a multifactorial and heterogeneous nature, poses challenges in developing effective treatments (3).

Alterations in protein glycosylation are among the primary molecular events that occur alongside oncogenic transformations (4, 5). Glycan biosynthesis is influenced by several factors, including altered activity of glycosyltransferases and glycosidases, as well as the availability of sugar donors, cofactors, and acceptor substrates (6, 7).

Dysregulation of the typical *O*-glycophenotype is associated with the progression of cancer, facilitating the amplification of more aggressive cell properties (8). The premature stop of the *O*-glycosylation synthesis is a well-known feature of malignant transformation leading to the presence of truncated *O*-glycans such Tn, sialyl-Tn, T, and sialyl-T antigens (9, 10). Compromised C1GalT1 glycosyltransferase activity, which is responsible for T antigen synthesis, caused by mutations or hypermethylation of its molecular chaperone COSMC promoter, leads to the overexpression of the Tn antigen (11, 12). Moreover, overexpression of ST6GALNAC1 glycosyltransferase, which originates STn antigen, is associated with a highly cancer-specific nature (13, 14). Tn and STn antigens are overexpressed in several tumours, including gastric, bladder and breast carcinomas, while absent or marginally expressed in the corresponding healthy tissues (15-18). High levels of these glycans have been correlated with more aggressive tumour behaviour, increased invasiveness, and a higher probability of metastasis, being associated with a worse prognosis (19-21). In gastric tumours the overexpression of Tn and STn antigens has also been associated with increased cellular invasion and decreasing cell-cell adhesion, ultimately instigating oncogenic characteristics in cancer cells (12). This correlation emphasizes the possible clinical importance of Tn and STn antigens as prognostic markers in GC, highlighting their relevance in assessing disease progression and guiding therapeutic strategies.

Envisaging the study of the functional implications of these aberrant *O*-glycans on gastric cancer aggressiveness, this chapter describes the establishment and characterization of a glycoengineered gastric cancer cell line overexpressing the Tn and STn antigens.

2. MATERIALS AND METHODS

2.1 Cell Culture Conditions

The AGS cell line, derived from human gastric adenocarcinoma, was acquired from the American Type Culture Collection (ATCC). Cell lines were cultured at 37°C in a 5% CO2 humidified atmosphere and atmospheric oxygen, with RPMI Medium 1640 (1X) + GlutaMAXTM-I (Gibco™, Thermo Fisher Scientific), supplemented with 10% heat-inactivated FBS (Gibco™, Thermo Fisher Scientific), and 1% penicillin-streptomycin (10,000Units/mL penicillin; 10,000mg/mL streptomycin; Gibco, Thermo Fisher Scientific).

2.2 *O*-glycomics

The *O*-glycome of the AGS wild-type cell line was characterized through the Cellular *O*-glycome Reporter/Amplification method, as described by Fernandes, et al (22). Briefly, benzyl 2-acetamido-2-deoxy-α-D-galactopyranoside (Sigma-Aldrich), which is a Tn mimetic, was peracetylated and administered to semi-confluent cells to a final concentration of 150μM in culture medium. After 24 h of incubation the secreted benzyl-*O*-glycans were isolated from the conditioned media by filtration with a 10kDa centrifugal filter (Amicon Ultra-4; Merck KGaA), followed by solid-phase extraction in Sep-Pak 3 cc C18 cartridges (Waters). Bn-*O*-glycans were then permethylated and analysed through nanoLC-ESI-MS/MS as previously described by us (22).

2.3 RT-PCR for Gene Expression Evaluation

Real-time polymerase chain reaction (RT-PCR) was used to quantify the expression level of the genes involved in the first steps of *O*-glycan biosynthesis (*C1GALT1* and *ST6GALNAC1*). Total RNA from cultured AGS cells was isolated using GRS Total RNA Kit (Grisp Research Solutions), according to the manufacturer's instructions. The purity and quantity of RNA extracts were estimated based on the A260/A280 using a Nanodrop (ND1000, Nano Drop Technologies Inc. Wilmington, DE, USA). RNA samples were further converted into complementary DNA (cDNA) using random primers from the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystem). RT-PCR amplification of cDNA samples was performed in a Mycycler Thermal cycler (BioRad): 25°C for 10 minutes, 37°C for 120 minutes, and reverse transcriptase inactivation at 85°C for 5 minutes. The relative expression of C1GalT1 and ST6GalNAC1 glycosyltransferases was determined using TaqMan Gene Expression Assays (*C1GALT1*: Hs00750511_s1; *ST6GALNAC1*: Hs00300842_m1) in a StepOne Real-Time PCR System (Applied Biosystems), with the

following thermal cycling conditions: 10 minutes at 95°C followed by 45 cycles of 15 seconds at 95°C and 1 minute at 60°C. Hypoxanthine-guanine phosphoribosyltransferase (HPRT: Hs99999909_m1) was used for normalization. All samples were run in duplicate and relative mRNA gene expression was calculated with the $2^{-\Delta Ct}$ formula.

2.4 Genetic Editing of Gastric Cancer Cell Line

The AGS cell line was genetically edited through CRISPR/Cas9 technology using a validated plasmid vector to deliver the Cas9 and the single guide RNA (GTAAAGCAGGGCTACATGAG). The genetic modification yielded specific double-strand breaks (DSBs) at the *C1GALT1* gene *in vitro*, resulting in the subsequent knock-out of the *C1GALT1* gene. Lipofectamine™ CRISPRMAX™ Transfection Reagent (Thermo Fisher Scientific) was employed in accordance with the instructions provided by the manufacturer. Complexes were formed in serum-free media (Opti-MEM™ I Reduced Serum Medium; Gibco) and were directly introduced to the cells in the culture medium, followed by a 24-hour incubation.

2.5 *C1GALT1* Mutation Analysis

Single clones were obtained by serial dilution in 96 well plates. The GRS Genomic DNA Kit (GRISP Research Solutions) was used to extract and purify DNA according to the manufacturer's instructions. Sanger sequencing was used primarily to screen for mutations in the *C1GALT1* gene. Primers (**Table 1**) were designed in this context using the Primer-BLAST design tool from the National Center for Biotechnology Information (NCBI). A PCR reaction setup was prepared according to **Table 2** for PCR product amplification. A thermocycler was used for an initial denaturation step of 15 minutes at 95°C, followed by 35 cycles of denaturation at 95°C for 30 seconds, annealing at 59°C for 30 seconds, and extension at 72°C for 45 seconds. A final 9-minute extension step at 72°C was also included. Electrophoresis in a 2% (w/v) agarose gel stained with SYBR-safe DNA gel stain (Thermo Fisher Scientific) confirmed the amplification. The PCR product was purified using a 1:2 mixture of exonuclease I (20µg/µl) (Thermo Fisher Scientific) and Fast Thermosensitive Alkaline Phosphatase (1µg/µl) (Thermo Fisher Scientific). Purification was carried out in a thermocycler at 37°C for 50 minutes to ensure proper enzyme activity, followed by 15 minutes at 85°C to inactivate the enzymes. For this reaction, 2µl of ExoSap mix was added to 5µl of PCR product, allowing unincorporated primers and nucleotides to be degraded.

Table 3 shows how the DNA amplicons were sequenced using the AppliedBiosystems® (Life Technologies) BigDye® Terminator v3.1 Cycle Sequencing Kit. The thermocycler PCR protocol included a 10-minute denaturation stage at 96°C, followed by 35 cycles of denaturation at 96°C for 10 seconds, annealing at 52°C for 5 seconds, and extension at 60°C for 6 minutes. The sequencing product was purified using Illustra Sephadex® G-50 fine columns (GE Healthcare, Life Sciences) according to the manufacturer's instructions to reduce potential impurities. Finally, samples were eluted in 15µl of deionized formamide (Applied Biosystems) to stabilize the DNA. The sequencing was done using a Thermo Fischer Scientific 3500 Genetic Analyzer, and the CRISPR editing results were examined with Synthego's ICE.

Indel Detection by Amplicon Analysis (IDAA) was also carried out to identify the size and frequency of insertions and deletions produced by the transfection. The nuclease target site was amplified using genomic PCR with three primers: a locus-specific set of forward and reverse primers and a universal primer (FamFwd) having the same sequence as an extension of the Fwd and tagged with 6-FAM to make the amplicons fluorescent. Selected amplified products were diluted 1:20; 1µl of diluted amplified product was mixed with Hi-Formamide (Applied Biosystems) and appropriate dye size standard (GeneScan 600 LIZ or 500 ROX, Applied Biosystems). The mixture was analyzed in a fragment analyzer (3500 Genetic Analyzer, Applied Biosystems), revealing the size of the amplicon (bp), and thereby the inflicted indels, as well as their frequency in relative fluorescence units (RFU). IDAA was performed with an ABI 3130 instrument. Two AGS *C1GALT1* KO clones with distinct out-of-frame indel alterations were selected. A single clone with silent mutations provided the phenotypic control cell line.

Table 1. Primer sequences used for sequencing the *C1GALT1* gene.

C1GALT1	**Primer sequence (5' → 3')**
Primer Forward	CCGGCCCTCAAAACCTAGAG
Primer Reverse	TGCATCTCCCCAGTGCTAAG

Table 2. Amplifications PCR reactions setup.

Component	20µl -rxn
ddH_2O	To 20µl
10X PCR Buffer	2µl
10 Mm dNTP Mix	1µl
25 mM $MgCl_2$	2µl
10 µM forward primer	1µl
10 µM reverse primer	1µl
Template DNA	0.5µl
AmpliTaq Gold DNA Polymerase (5U/µl)	0.2µl

Table 3. BigDye™ Terminator v3.1 Cycle Sequencing Reaction.

Component	10 µl -rxn
BigDye® Mix	0.5µl
5X Sequencing Buffer	3.4µl
Primer	0.5µl
ddH_2O	4.78µl
PCR product	1µl (75ng)

2.6 Flow Cytometry for *O*-linked Glycans

Cells were detached using Versene 1x solution (Gibco; Life Technologies) and fixed with 2% paraformaldehyde (PFA; Sigma-Aldrich) for 15 minutes. Tn antigen was determined using 20µg/mL fluorescein-labelled *Vicia Villosa* (VVA) lectin (Vector Laboratories) for 1h in PBS 2% FBS, under mild agitation, at room temperature, whereas the STn antigen was assessed using the same lectin at the same conditions, after digestion with 70mU/10^6 cells α-Neuraminidase from *Clostridium perfringens* (Sigma-Aldrich), in sodium acetate buffer, at 37ºC overnight, under mild agitation. T antigen was determined with 20µg/mL fluorescein-labelled *Peanut Agglutinin* (PNA) lectin (Vector Laboratories), and ST antigen was determined by comparing the affinity of the PNA before and after overnight sialidase digestion. In addition, cancer cells were screened for *N*-acetylglucosamine residues using 20µg/mL *Griffonia Simplicifolia* Lectin II (GSL II) (Vector Laboratories). C-type lectin GSL II was incubated for 1h in 10mM HEPES, 0.15M NaCl, 0.1mM $CaCl_2$, pH=7.5 buffer, under mild agitation at room temperature. Under the above-mentioned conditions, GSL II lectin detection was also performed after PNGaseF (250mU/10^6 cells, in PBS 1x) enzymatic digestion to exclude *N*-glycan-associated *N*-acetylglucosamine

residues contribution. Sample acquisition and data analysis were performed through the FC500 Beckman Coulter flow cytometer.

2.7 Immunofluorescence for *O*-linked Glycans Detection

To evaluate short *O*-glycan expression, AGS glycoengineered cell models were cultured at low density in µ-Slide 8 Well (Ibidi) and fixed with 2% PFA (Sigma-Aldrich), following immunofluorescent staining as described in the **section 2.6**. STn and ST were evaluated in parallel with samples digested with 70mU/10^6 cells α-Neuraminidase from Clostridium perfringens (Sigma-Aldrich). Nuclei were marked with 2,3x10^{-3}µg/µL 4',6-Diamidino-2-phenylindole dihydrochloride (DAPI, Thermo Fisher Scientific) for 10 minutes at room temperature. Immunofluorescence images were acquired on a Leica DMI6000 FFW microscope using Las X software (Leica).

2.8 Cell Proliferation Assay

AGS and AGS glycoengineered cell models were submitted to the colorimetric Cell Proliferation ELISA (Roche) assay to evaluate their proliferation capacity. This method, which is based on the measurement of the incorporation of bromodeoxyuridine (BrdU) into newly synthesized DNA of proliferative cells, was performed according to the manufacturer's instructions. All experiments included 1% Triton-X in a complete cell culture medium as a cell death positive control. The results were monitored at 450nm using an iMARK™ microplate reader (Bio-Rad).

2.9 Cell Migration Assay

The migratory capacity of the AGS glycoengineered cells was assessed using Invasion Chambers (Corning). The 8-µm diameter pore-size filters were placed in a 24-well plate (Falcon®). The cells were then seeded on the upper side of the filter at a density of 25x10^3 cells/well. After 24 h at 37°C, filters were fixed in 4% paraformaldehyde and then stained with DAPI, for 10 minutes. Non-migrating cells, present on the upper side, were then removed, to facilitate analysis. Cells that had migrated to the underside of the filters were visualized through an IX51 fluorescence microscope (Olympus). When the DAPI-stained nuclei pass through the filter pores, the cells that had migrated were assessed in at least 12 microscopic areas (20x objective).

2.10 Cell Invasion Assay

The invasive capacity of the AGS glycoengineered cells was assessed using Matrigel Invasion Chambers (Corning) as described in Ferreira, JA. *et al*. (23). Briefly, 8-µm diameter pore size filters were placed in a 24-well plate, coated with 75µL of Matrigel (50µg/mL) diluted in RPMI Medium 1640 (1X) + GlutaMAXTM-I (Gibco™, Thermo Fisher Scientific), and were incubated for 1h at 37°C. The cells were then seeded on the upper side of the Matrigel-coated filter at a density of $25x10^3$ cells/well. After 24 h at 37°C, filters were fixed in 4% paraformaldehyde and non-invading cells, present on the upper side, were completely removed, to facilitate analysis. Cells that had invaded the underside of the filters were mounted in VECTASHIELD® Antifade Mounting Medium with DAPI (Vector Laboratories) and visualized through an IX51 fluorescence microscope (Olympus). When the DAPI-stained nuclei pass through the filter pores, the cells that had migrated were assessed in at least 12 microscopic areas (20x objective).

2.11 Colony Formation Assay

The anchorage-independent growth of AGS glycoengineered cells was determined using the soft agar colony formation assay. A 0.5% low melting point agarose (Lonza) solution in the complete medium was used as the bottom layer in 6-well flat bottom plates. A top layer of 0.3% agarose containing $5x10^4$ cells was then seeded and covered with a culture medium. Cells were maintained in standard growth conditions for 15 days. Colonies were then fixed with 10% neutral buffered formalin solution (Sigma-Aldrich) for 30 minutes and stained with 0,1% (w/v) crystal violet (Sigma-Aldrich) for 60 minutes. Colonies were photographed using an SZX2-ILLT stereomicroscope (Olympus) with an SC180 camera (Olympus) and automatically counted using the open-source software ImageJ (Fiji package). Only colonies with 40-2000 pixels were considered.

2.12 Statistical Analysis

Kruskal-Wallis with Dunn's multiple comparisons test was used to compare the groups in the Colony Forming Assay and One-way ANOVA with Tukey's multiple comparisons test was used to compare groups for Proliferation, Invasion and Migration assays. All experiments were performed in triplicate. The results are presented as mean ± standard deviation rate for each experiment. The graphing and statistical analysis were performed in GraphPad Prism Software version (GraphPad 8.4.2). Statistical significance was determined as $p < 0.05$.

3. RESULTS AND DISCUSSION

Aberrant protein glycosylation is one of the key molecular events in oncogenic transformations of the gastric tract (7). This posttranslational modification is a well-known hallmark of cancer due to its contribution to carcinogenesis, cancer progression, and metastasis (24-26). Expression of truncated *O*-glycans, namely Tn and STn antigens, is associated with tumour aggressiveness and poorer prognosis in several tumours, including GC (7, 12, 14, 27).

3.1 Genetic glycoengineering of AGS gastric cancer cell line

Short chain *O*-glycans have a tumour microenvironment dependent nature, being associated with specific immunological factors and nutrient restriction condition (16, 28). As such, the stable and robust expression of this glycans *in vitro* settings needs to be preset by cancer cell glycoengineering. We started by selecting AGS cell line from an inventory of gastric cancer cell lines, since AGS is representative of intestinal-type gastric cancer, which is the most common form of GC associated with poor prognosis. This cell line exhibited markers of GC aggressiveness such as loss of E-cadherin expression, however it maintained a mild aggressive profile *in vitro* (12, 29). The glycomic profile of this cell line has not revealed any glycosidic epitopes commonly associated with GC aggressiveness, including STn (22). As such, this cell line represents a suitable model for STn overexpression studies.

AGS will-type cells were analysed by *O*-glycomics nanoLC-MS/MS, using the hydrophobic peracetylated Benzyl-a-GalNAc mimetic (**Figure 1**). Accordingly, this cells mainly expressed T antigens (m/z 572.306) and fucosylated T antigens (m/z 746.395), corresponding, respectively, to 23.4% and 15.8% of all glycans expressed (**Figure S1**). Lower levels of mono-sialylated structures of T antigen (m/z 933.480), core 3 and extended core 3 structures (m/z 613.332; 817.432), as well as extended core 2 O-glycans (*m/z* 991.521; 1021.532; 1195.662; 1369.710) were also observed. These results suggest that the precursor Tn antigen is mainly extended to core 1 structures rather than STn, which is consistent with previous observations made by our group (22). Taking this into account, the transcript levels of *C1GALT1*, T antigen synthase, and *ST6GALNAC1*, the sialyl-Tn biosynthetic enzyme, were analysed by RT-PCR in AGS wild-type cells to determine alterations in glycosyltransferase expression that could indicate competitive advantage for the Tn substrate. This analysis revealed considerably high *C1GALT1* transcript levels compared to ST6GalNAC1, favouring T antigens biosynthesis over sialylated forms of core1

(**Figure 1B**). These findings are in accordance with the literature, which reports high levels of *C1GALT1* expression in the AGS cell line (30, 31).

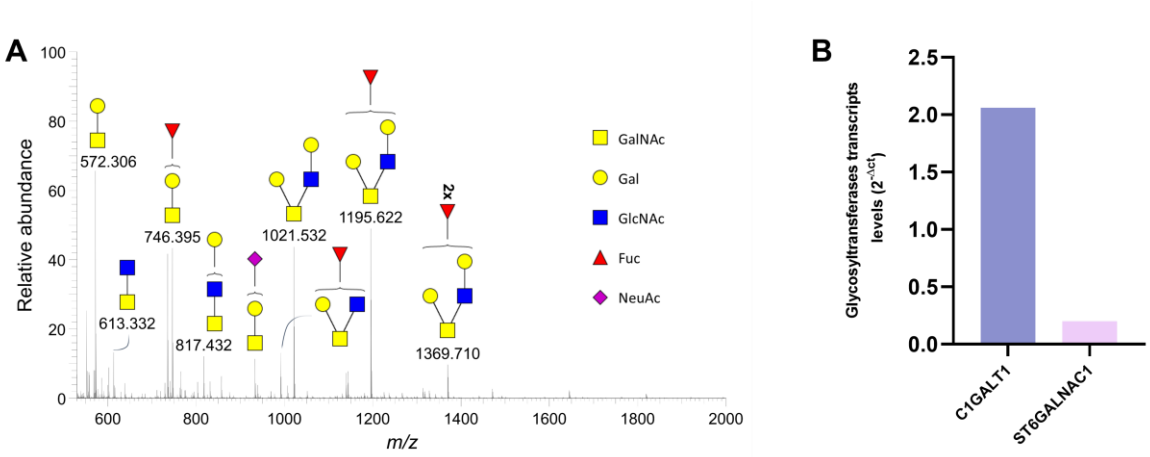

Figure 1. Characterization of the AGS wild-type cell line glycome and glycosyltransferases activity. A) *O*-glycomics analysis of AGS wild-type cells using Cellular *O*-Glycome Reporter/Amplification and nanoLC-MS/MS. Wild-type AGS cells mainly express T antigens (*m/z* 572.306) and fucosylated T antigens (*m/z* 746.395); Mono-sialylated T antigens (*m/z* 933.480), core 3 (*m/z* 613.332), extended core 3 (*m/z* 817.432), and extended core 2 *O*-glycans (*m/z* 991.521; 1021.532; 1195.662;1369.710) were also identified. **B)** Transcript levels of *C1GALT1* and *ST6GALNAC1* glycosyltransferases in the AGS WT cell line. *C1GALT1* transcript levels are considerably higher than those of *ST6GALNAC1* glycosyltransferase in the AGS gastric cancer cell line.

Accordingly, the intestinal-type AGS cell line was glycoengineered resorting to CRISPR-Cas9 technology, using validated guide RNAs (32), to knock-out the *C1GALT1* gene, hampering *O*-glycans extension to core 1 structures. Two cell clones carrying biallelic frameshift mutations of *C1GALT1* were elected for subsequent experiments, and a wild-type clone carrying silent mutations was elected as a control. The induction of mutations was confirmed by Indel Detection by Amplicon Analysis (IDAA) and Sanger sequencing (**Figure 2**), following the phenotypic characterization of the cell line (**Figure 3**).

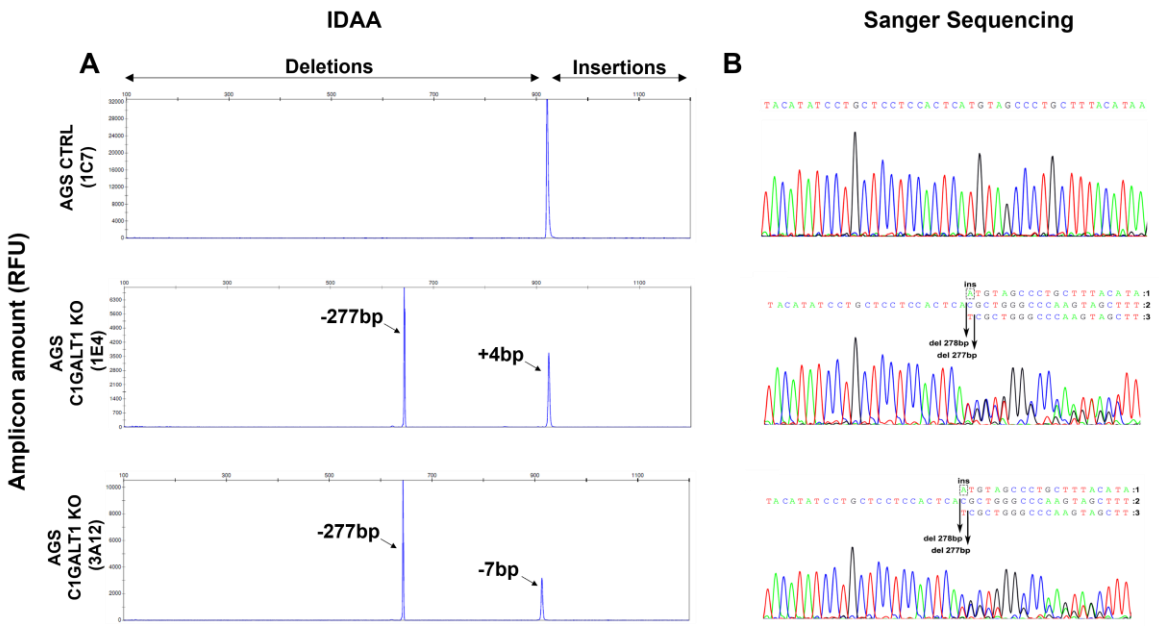

Figure 2. Genomic profiling of glycoengineered AGS *C1GALT1* KO models. AGS cells were genetically modified to knock-out the *C1GALT1* gene, resulting in the selection of two clones (1E4 and 3A12) for proof-of-concept experiments. IDAA **(A)** and Sanger Sequencing **(B)** allowed the characterization of the indel mutations induced in the glycoengineered cells. Electropherograms of Sanger Sequencing with the reverse primer are shown, along with the corresponding sequencing readings (1, 2 and 3). Accordingly, clones 1E4 and 3A12 are characterized by three different DNA sequences, with the following observed variants and predicted consequences: 1: c.620dup [p.(Met207IlefsTer19)], where the duplication of the nucleotide 620 results in the replacement of Met 207 by Ile, leading to a frame-shift introducing a stop codon 19 a.a ahead; 2: c.343_620del [p.(Cys115GlufsTer18)], resulting in a 277bp deletion and the replacement of Cys 115 by Glu, leading to a frame-shift introducing a stop codon 18 a.a. ahead; 3: c.342_618del [p.(Cys115Ter)], where a 278bp deletion results in the replacement of Cys 115 by a stop codon.

C1GALT1 KO cells were then screened for *O*-glycan expression by flow cytometry (**Figure 3A**) and immunocytochemistry (**Figure 3B**) combining glycan-specific lectins and enzymatic digestions to expose the glycans of interest. The Tn antigen content was determined using VVA lectin, whereas T antigen levels were assessed using PNA lectin. Before VVA and PNA staining, neuraminidase digestion was performed to assess the expression levels of sialylated forms of Tn (STn) and T antigens (ST), particularly STn and ST antigens. Moreover, core 3 expression was indirectly assessed using GSL II lectin, targeting GlcNAc residues at the nonreducing end of glycan chains. To mitigate potential cross-reactivity with N-glycans, GSL II was determined following PNGaseF digestion.

Accordingly, *C1GALT1* KO cells exhibited higher expression of Tn antigen than the control (**Figure 3A** and **3B**). However, *C1GALT1* KO clones demonstrated a reduced T antigen expression, directly attributed to T-synthase knock-out mutation (**Figure 3A** and

3B). The availability of Tn epitopes due to the knock-out led to an overexpression of STn antigen. Moreover, there were no significant changes in the expression of core 3 (**Figure 3A**) when comparing the clones with the control. Overall, *C1GALT1* gene knock-out (**Figure 3C**) enabled the establishment of a glycoengineered cell line exhibiting consistent Tn and STn expression. Further glycophenotype optimization will be pursued in future studies, including the elimination of competing pathways as core 3 biosynthesis, promoted by β3GnT. *ST6GALNAC1* knock-in could also further enhance STn levels.

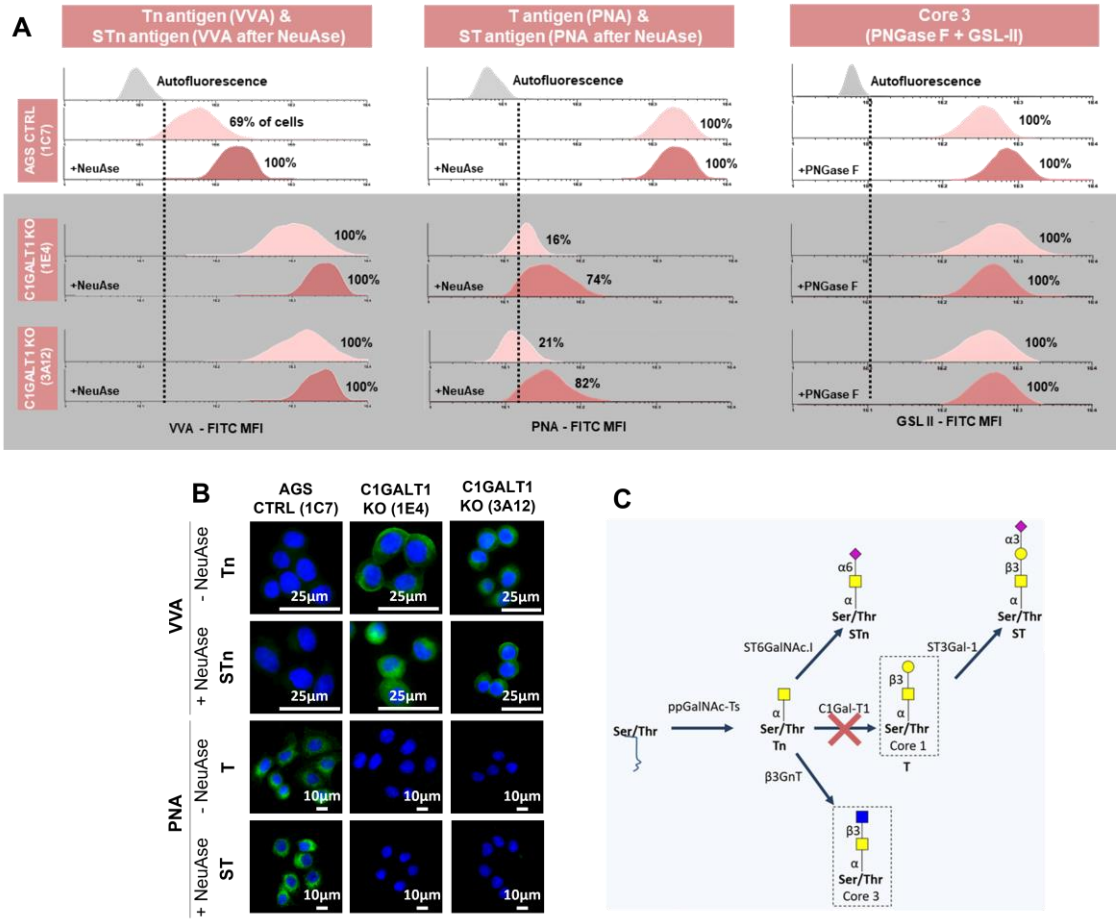

Figure 3. Phenotypic characterization of the glycoengineered AGS cell line. A) Flow cytometry assessment demonstrates the reduction of *O*-glycosylation extension beyond the core 1 structure, resulting from the knock-out of the *C1GALT1* gene. Tn and T antigens were identified directly through VVA and PNA lectins. STn and ST antigens were detected indirectly by employing VVA and PNA lectins following NeuAse treatment of the cells. No significant changes in the levels of core 3 expression were detected. **B)** The findings from immunocytochemistry align with the results obtained through flow cytometry. The *C1GALT1* knock-out clones exhibit an absence of T and ST antigens while overexpressing Tn and STn antigens, contrasting to the wild-type control cells. **C)** Schematic representation of Tn, STn and T antigens and core 3 biosynthesis pathways.

3.2 Functional Impact of STn Overexpression in AGS Cells

To clarify the functional role of Tn and STn expression in GC, the Tn and STn overexpressing cell line was characterized according to aggressiveness features as invasion, migration, anchorage-independent growth, and changes in proliferation.

The Tn and STn overexpressing cell line did not show considerable differences in proliferation rates compared to control (**Figure 4A**). These findings are in accordance with previous studies were Tn overexpressing cells baring *C1GALT1C1* mutations also maintained proliferative rates compared to wild-type controls (12). Both *C1GALT1* and *C1GALT1C1* knock-outs are translated by the same phenotypical alterations, since, *C1GALT1C1* is a T synthase chaperone, aiding its activity. As such, both inactivating mutations lead to Tn antigen accumulation (12).

The Tn and STn overexpressing cells migration capacity was also addressed in transwell assays. *C1GALT1* knock-out cells demonstrated significantly decreased migratory ability than control cells (**Figure 4B**). Our findings do not support the findings of a previous study conducted by other research team, which showed that the AGS cell line, which expressed truncated *O*-glycans, did not exhibit any significant differences in migration compared to the control group (12). Interestingly, AGS cells arise from cohesive gastric tumours, suggesting increased cell-cell adhesion events and possibly a less mobile phenotype.

The invasive capacity of Tn and STn overexpressing cells was also tested using Matrigel coated cell culture inserts. Notably, the inhibition of the C1GalT1 pathway led to a significant increased invasive capacity of cells compared to wild-type control (**Figure 4D**). These results are in line with other groups reports of elevated invasive capacity of GC cell lines overexpressing truncated *O*-glycans (12). Future studies will expand the glycoengineered cell line library to other gastric cancer cell lines to further interrogate the impact of Tn and STn overexpression in migration and invasion. Other migration assays using matrix proteins coating could better replicate the tumour microenvironment faced by cells. Similarly, other extracellular matrixes could be used to infer on GC cell lines invasion capacity. Moreover, the metalloproteinases expression profile of GC cell lines will be determined to clarify the dichotomy between migration and invasion processes, since increased invasive capacity is often associated with enhanced migratory capability in cancer cells, which seems to not be the case in AGS cells (33-35).

Finally, we looked into the adhesion-independent proliferation of Tn and STn overexpressing cells using colony forming assays (**Figure 4C**). This strategy gives us an appropriate indication of cell resistance to anoikis. Of note, glycoengineered cells

overexpressing truncated *O*-glycans showed an increased proliferative capacity and resistance to cell death in anchorage-independent cell culture conditions (**Figure S2**).

Overall, Tn and STn antigens play significant roles in driving the aggressiveness of GC cells, promoting invasion and resistance to anoikis. Moreover short-chain *O*-glycans impact on invasion and cell death seem to be ubiquitous, with similar findings reported in bladder cancer (28). This suggests that truncated *O*-glycans endow GC cells with an increased ability to escape the primary tumour mass and disseminate. *In vivo* findings by other authors are in line with this hypothesis since *C1GALT1* loss accelerated gastric cancer progression and metastasis (36, 37).

Future studies will address the functional impact of Tn and STn in GC in a more comprehensive manner, resorting to the glycoproteomic characterization of glycoengineered cell lines. The glycoproteomics analysis of Tn and STn overexpressing cells will reveal relevant intermediates of cell aggressiveness by pinpointing hub proteins responsible for cancer progression. Moreover, studies focused on cell surface unique glycoproteomic signatures will reveal potential targets for intervention. **Chapter IV** will address the establishment of a robust protocol for the glycoproteomics analysis of the gastric cancer cell lines.

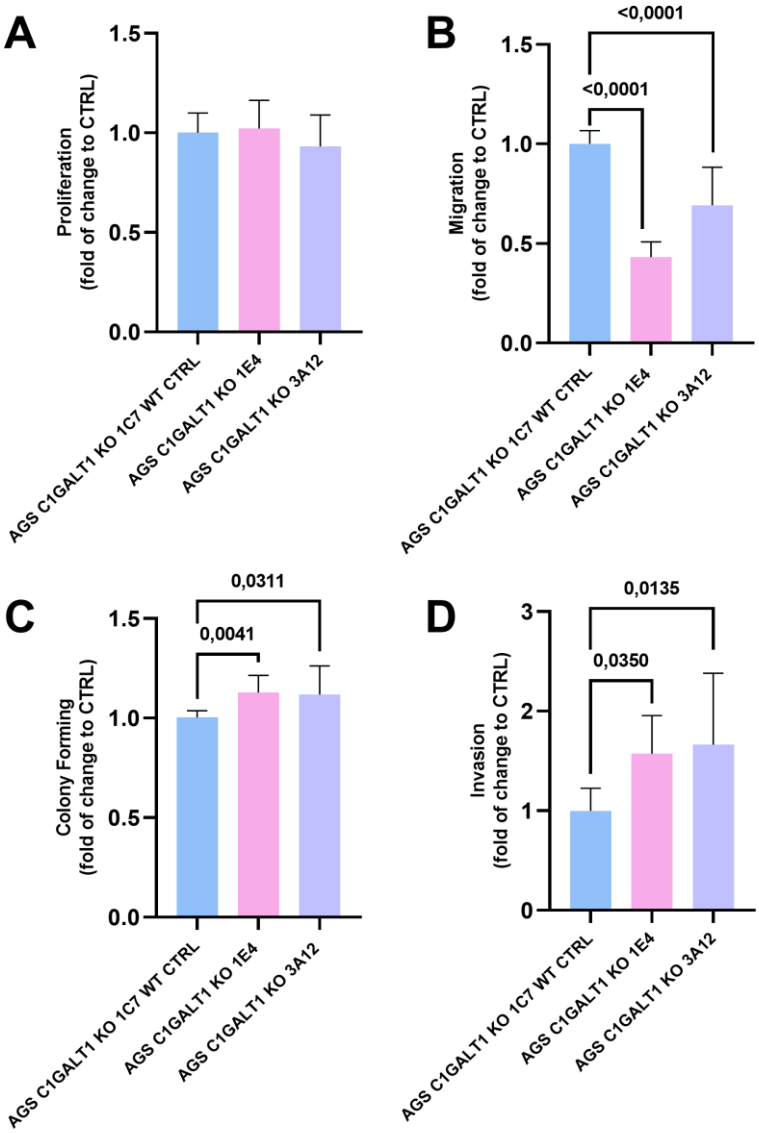

Figure 4. Functional characterization of the AGS C1GALT1 KO cells. A) Abrogation of extended glycans and Tn overexpression does not alter AGS cell's proliferation capacity compared to the control (results were obtained from three independent experimental replicates and expressed as the means ± SD analysed by ordinary one-way ANOVA with multiplicity-adjusted P values from Tukey's multiple comparisons test). On the other hand, it decreased their **(B)** migration capacity, while increasing their **(D)** invasion capacity, compared to the control (results were obtained from two independent experimental replicates and expressed as the means ± SD analysed by ordinary one-way ANOVA with multiplicity-adjusted P values from Tukey's multiple comparisons test). **C)** The colony formation capacity of the C1GALT1 KO cells increased when compared to the control (results were obtained from three independent experimental replicates and expressed as the means ± standard deviation rates analysed by Kruskal-Wallis one-way ANOVA with multiplicity-adjusted P values from Dunn's multiple comparisons test).

4. CONCLUDING REMARKS

Altered glycosylation is a distinctive characteristic of cancer, particularly the presence of truncated *O*-glycans namely Tn and STn antigens, which are associated with the aggressive nature of gastric cancer and the poor prognosis of patients (13, 38). This excessive glycan production is primarily linked to impaired *C1GALT1* activity, associated with *C1GALT1* mutations or the epigenetic suppression of its chaperone COSMC (*C1GALT1C1*). Additionally, overexpression of *ST6GALNAC1* glycosyltransferase has been observed in various cancers, including gastric cancer (14). The intricate interplay between cancer cell surface glycosylation and the tumour microenvironment presents challenges in establishing reliable and robust *in vitro* cell models for studies in glycobiology. As such, to better understand the role of Tn and STn antigens overexpression in GC cells, this chapter described the establishment of a glycoengineered cell line overexpressing immature *O*-glycans. Moreover, the Tn and STn overexpressing cell model provided additional support for the functional role of Tn and STn antigens expression in cancer cells proliferation, migration, invasion, and independent anchorage growth. The expression of those glycans on the cell surface was responsible for increasing GC cell's ability to invade and form colonies, resulting in cells with a more aggressive phenotype. These results are in line with previous studies showing COSMC knock-out affecting both STn and Tn levels, consequently associated with a higher capacity for cellular invasion, inducing an oncogenic phenotype (10, 39, 40). However, *C1GALT1* knock-out cells showed a decreased capacity to migrate. Overall, we disclosed the functional role of Tn and STn antigens in a glycoengineered gastric cancer cell line, that exhibited a more aggressive phenotype when these truncated *O*-glycans are overexpressed.

In future studies a novel glycoengineered cell model could be established to enhance STn levels. Furthermore, it would be interesting to consider the elimination of competitive pathways, such as core 3. Additionally, to understand the uncoupled association between migration and invasion capabilities of the Tn and STn overexpressing cells, it could be relevant to conduct cell-cell adhesion assays using atomic force microscopy (AFM) for single-cell force spectroscopy (SCFS) as well as gelatinase zymography assays to evaluate their metalloproteinases profile. A comprehensive research approach, encompassing both organoid and *in vivo* studies, is imperative to deepen our understanding of the potential implications and consequences of short-chain *O*-glycans on the overall aggressiveness of gastric cancer. This will allow us to explore the intricate complexities within the microenvironment and its influential factors.

5. REFERENCES

1. Ilic M, Ilic I. Epidemiology of stomach cancer. World J Gastroenterol. 2022;28(12):1187-203.
2. Lordick F, Carneiro F, Cascinu S, Fleitas T, Haustermans K, Piessen G, et al. Gastric cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. Annals of Oncology. 2022;33(10):1005-20.
3. Sexton RE, Al Hallak MN, Diab M, Azmi AS. Gastric cancer: a comprehensive review of current and future treatment strategies. Cancer Metastasis Rev. 2020;39(4):1179-203.
4. Pinho SS, Carvalho S, Marcos-Pinto R, Magalhães A, Oliveira C, Gu J, et al. Gastric cancer: adding glycosylation to the equation. Trends in Molecular Medicine. 2013;19(11):664-76.
5. Pinho SS, Reis CA. Glycosylation in cancer: mechanisms and clinical implications. Nature Reviews Cancer. 2015;15(9):540-55.
6. Dall'Olio F, Malagolini N, Trinchera M, Chiricolo M. Mechanisms of cancer-associated glycosylation changes. Front Biosci (Landmark Ed). 2012;17(2):670-99.
7. Ferreira JA, Magalhães A, Gomes J, Peixoto A, Gaiteiro C, Fernandes E, et al. Protein glycosylation in gastric and colorectal cancers: Toward cancer detection and targeted therapeutics. Cancer Letters. 2017;387:32-45.
8. Reily C, Stewart TJ, Renfrow MB, Novak J. Glycosylation in health and disease. Nat Rev Nephrol. 2019;15(6):346-66.
9. Julien S, Videira PA, Delannoy P. Sialyl-tn in cancer: (how) did we miss the target? Biomolecules. 2012;2(4):435-66.
10. 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.
11. Sun X, Zhan M, Sun X, Liu W, Meng X. C1GALT1 in health and disease. Oncol Lett. 2021;22(2):589.
12. Freitas D, Campos D, Gomes J, Pinto F, Macedo JA, Matos R, et al. O-glycans truncation modulates gastric cancer cell signaling and transcription leading to a more aggressive phenotype. EBioMedicine. 2019;40:349-62.
13. 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 Letters. 2007;249(2):157-70.

14. 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(4):1443-55.
15. Rømer TB, Aasted MKM, Dabelsteen S, Groen A, Schnabel J, Tan E, et al. Mapping of truncated O-glycans in cancers of epithelial and non-epithelial origin. Br J Cancer. 2021;125(9):1239-50.
16. Carrascal MA, Severino PF, Guadalupe Cabral M, Silva M, Ferreira JA, Calais F, et al. Sialyl Tn-expressing bladder cancer cells induce a tolerogenic phenotype in innate and adaptive immune cells. Mol Oncol. 2014;8(3):753-65.
17. Fu C, Zhao H, Wang Y, Cai H, Xiao Y, Zeng Y, et al. Tumor-associated antigens: Tn antigen, sTn antigen, and T antigen. Hla. 2016;88(6):275-86.
18. Ju T, Wang Y, Aryal RP, Lehoux SD, Ding X, Kudelka MR, et al. Tn and sialyl-Tn antigens, aberrant O-glycomics as human disease markers. PROTEOMICS – Clinical Applications. 2013;7(9-10):618-31.
19. Chia J, Goh G, Bard F. Short O-GalNAc glycans: regulation and role in tumor development and clinical perspectives. Biochimica et Biophysica Acta (BBA) - General Subjects. 2016;1860(8):1623-39.
20. Santos SN, Junqueira MS, Francisco G, Vilanova M, Magalhães A, Dias Baruffi M, et al. O-glycan sialylation alters galectin-3 subcellular localization and decreases chemotherapy sensitivity in gastric cancer. Oncotarget. 2016;7(50):83570-87.
21. Munkley J. The Role of Sialyl-Tn in Cancer. Int J Mol Sci. 2016;17(3):275.
22. Fernandes E, Freitas R, Ferreira D, Soares J, Azevedo R, Gaiteiro C, et al. Nucleolin-Sle A Glycoforms as E-Selectin Ligands and Potentially Targetable Biomarkers at the Cell Surface of Gastric Cancer Cells. Cancers. 2020;12(4):861.
23. Ferreira JA, Videira PA, Lima L, Pereira S, Silva M, Carrascal M, et al. Overexpression of tumour-associated carbohydrate antigen sialyl-Tn in advanced bladder tumours. Mol Oncol. 2013;7(3):719-31.
24. Peixoto A, Relvas-Santos M, Azevedo R, Santos LL, Ferreira JA. Protein Glycosylation and Tumor Microenvironment Alterations Driving Cancer Hallmarks. Front Oncol. 2019;9:380.
25. Wang M, Zhu J, Lubman DM, Gao C. Aberrant glycosylation and cancer biomarker discovery: a promising and thorny journey. Clin Chem Lab Med. 2019;57(4):407-16.
26. Oliveira-Ferrer L, Legler K, Milde-Langosch K. Role of protein glycosylation in cancer metastasis. Semin Cancer Biol. 2017;44:141-52.
27. Fernandes E, Soares J, Cotton S, Peixoto A, Ferreira D, Freitas R, et al. Esophageal, gastric and colorectal cancers: Looking beyond classical serological

biomarkers towards glycoproteomics-assisted precision oncology. Theranostics. 2020;10(11):4903-28.
28. Andreia P, Dylan F, Andreia M, Marta R-S, Rui F, Tim SV, et al. Multilevel Plasticity and Altered Glycosylation Drive Aggressiveness in Hypoxic and Glucose-Deprived Bladder Cancer Cells. bioRxiv. 2023:2023.10.21.561355.
29. Jawhari AU, Noda M, Farthing MJ, Pignatelli M. Abnormal expression and function of the E-cadherin-catenin complex in gastric carcinoma cell lines. Br J Cancer. 1999;80(3-4):322-30.
30. Lee P-C, Chen S-T, Kuo T-C, Lin T-C, Lin M-C, Huang J, et al. C1GALT1 is associated with poor survival and promotes soluble Ephrin A1-mediated cell migration through activation of EPHA2 in gastric cancer. Oncogene. 2020;39(13):2724-40.
31. Dong X, Chen C, Deng X, Liu Y, Duan Q, Peng Z, et al. A novel mechanism for C1GALT1 in the regulation of gastric cancer progression. Cell Biosci. 2021;11(1):166.
32. Narimatsu Y, Joshi HJ, Yang Z, Gomes C, Chen Y-H, Lorenzetti FC, et al. A validated gRNA library for CRISPR/Cas9 targeting of the human glycosyltransferase genome. Glycobiology. 2018;28(5):295-305.
33. Schaeffer D, Somarelli JA, Hanna G, Palmer GM, Garcia-Blanco MA. Cellular migration and invasion uncoupled: increased migration is not an inexorable consequence of epithelial-to-mesenchymal transition. Mol Cell Biol. 2014;34(18):3486-99.
34. Albuquerque C, Manguinhas R, Costa JG, Gil N, Codony-Servat J, Castro M, et al. A narrative review of the migration and invasion features of non-small cell lung cancer cells upon xenobiotic exposure: insights from in vitro studies. Transl Lung Cancer Res. 2021;10(6):2698-714.
35. Antón IM, Wandosell F, Vicente-Manzanares M. Chapter 15 - Cancer cell development, migratory response, and the role of the tumor microenvironment in invasion and metastasis. In: Schnoor M, Yin L-M, Sun SX, editors. Cell Movement in Health and Disease: Academic Press; 2022. p. 245-70.
36. Chugh S, Barkeer S, Rachagani S, Nimmakayala RK, Perumal N, Pothuraju R, et al. Disruption of C1galt1 Gene Promotes Development and Metastasis of Pancreatic Adenocarcinomas in Mice. Gastroenterology. 2018;155(5):1608-24.
37. Liu F, Fu J, Bergstrom K, Shan X, McDaniel JM, McGee S, et al. Core 1–derived mucin-type O-glycosylation protects against spontaneous gastritis and gastric cancer. Journal of Experimental Medicine. 2019;217(1).
38. David L, Nesland JM, Clausen H, Carneiro F, Sobrinho-Simões M. Simple mucin-type carbohydrate antigens (Tn, sialosyl-Tn and T) in gastric mucosa, carcinomas and metastases. APMIS Suppl. 1992;27:162-72.

39. Campos D, Freitas D, Gomes J, Magalhães 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.
40. Hofmann BT, Schlüter L, Lange P, Mercanoglu B, Ewald F, Fölster A, et al. COSMC knockdown mediated aberrant O-glycosylation promotes oncogenic properties in pancreatic cancer. Mol Cancer. 2015;14:109.

Establishment of a Glycoengineered Gastric Cancer Cell Line Library

Martina Dourado-Gonçalves, Dylan Ferreira, Marta Relvas-Santos, Paula Paulo, Lúcio Lara Santos, Andreia Peixoto, José Alexandre Ferreira

SUPPLEMENTARY INFORMATION

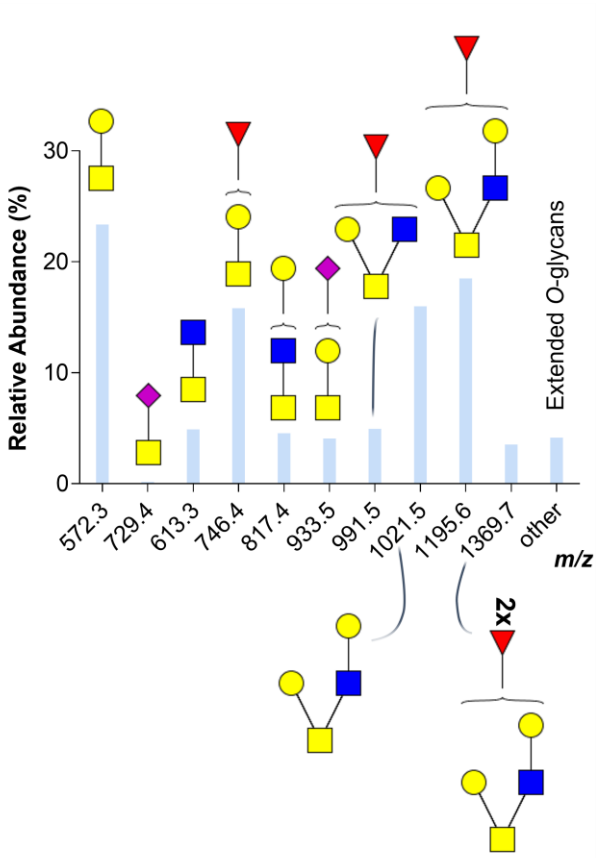

Figure S1. *O*-Glycomics profile of AGS GC cell line using *C*ellular *O*-Glycome Reporter/Amplification by nanoLC-MS/MS. The expression profile of wild-type AGS cells primarily consists of T antigen (*m/z* 572.306; 23.4%), fucosylated T antigen (*m/z* 746.395; 15.8%), extended core 2 (*m/z* 1021.5; 16%), and fucosylated core 2 *O*-glycans (*m/z* 1195.6; 18.5%). Additionally, lower levels of mono-sialylated T antigen (*m/z* 933.480; 4.1%), core 3 (*m/z* 613.332; 4.9%), extended core 3 (*m/z* 817.432; 4.6%), fucosylated core 2 (*m/z* 991.5; 5%), and di-fucosylated extended core 2 (*m/z* 1369.7; 3.5%) are also present. Nevertheless, the presence of STn (*m/z* 729.4; 0.2%) is minimal.

Figure S2. AGS *C1GALT1* KO cell model overexpressing truncated *O*-glycans showed an increased capacity to form colonies. Glycoengineered cells overexpressing truncated *O*-glycans showed an increased capacity to form colonies in anchorage-independent cell culture conditions.

CHAPTER IV

GLYCOPROTEOMICS CHARACTERIZATION OF GASTRIC CANCER CELLS

Glycoproteomics Characterization of Gastric Cancer Cells

Martina Dourado-Gonçalves[1,2], Rui Freitas[1,2,3], Marta Relvas-Santos[1,2,3,4], Dylan Ferreira[1,2,3], Lúcio Lara Santos[1,2,5,6,7,8], Andreia Peixoto[1,3], José Alexandre Ferreira[1,2,5,6,8]

[1]Experimental Pathology and Therapeutics Group, IPO Porto Research Center (CI-IPOP), Portuguese Oncology Institute (IPO Porto), 4200-072 Porto, Portugal; [2]Institute of Biomedical Sciences Abel Salazar (ICBAS), University of Porto, 4050-013; [3]Institute for Research and Innovation in Health (i3S), University of Porto, 4200-135 Porto, Portugal / Institute for Biomedical Engineering (INEB), University of Porto, 4200-135 Porto, Portugal; [4]REQUIMTE-LAQV, Department of Chemistry and Biochemistry, Faculty of Sciences of the University of Porto, 4169-007 Porto, Portugal; [5]Health School of University Fernando Pessoa, 4249-004 Porto, Portugal; [6]GlycoMatters Biotech, 4500-162 Espinho, Portugal; [7]Department of Surgical Oncology, Portuguese Oncology Institute (IPO Porto), 4200-072 Porto, Portugal; [8]Porto Comprehensive Cancer Center (P.ccc), 4200-072 Porto, Portugal.

Corresponding author:

José Alexandre Ferreira (jose.a.ferreira@ipoporto.min-saude.pt)

Experimental Pathology and Therapeutics Group, Research Centre, Portuguese Oncology Institute of Porto, R. Dr. António Bernardino de Almeida 62, 4200-072 Porto, Portugal; Tel. +351 225084000 (ext. 5111).

ABSTRACT

Glycoproteomics analysis is a fundamental tool within molecular-level biological research, yielding invaluable insights into the intricate molecular mechanisms influencing various physiological and pathological processes. The relevance of glycoproteomics is particularly pronounced in the field of oncology, revealing unique aberrantly glycosylated cancer specific molecular signatures across different tumours. These alterations in glycan structures play a central role in critical biological functions, however, glycoproteomic analysis face significant challenges, such as the complexity of biological samples, the vast array of proteins present, and the heterogeneity spawned by protein post-translational modifications. To overcome these barriers, the optimization of techniques is crucial. The objective of this study was to establish a standardized procedure for extracting membrane-bound proteins from cellular material in order to facilitate further glycoproteomic analysis.

Accordingly, two mechanical cellular extraction methods were tested, namely cell detachment with versene and cell scraping. Cell lysates were subjected to differential centrifugation to obtain an enriched membrane protein fraction, with further enrichment with neutravidin, to clear potential contaminants. Immunocytochemistry analyses determined a range of biotin concentrations to label the cell surface without internalization. Moreover, the degree of nuclear and cytoplasmic protein contaminants was determined by Western blot. Finally, glycoproteomic analysis were conducted using mass spectrometry, coupled with subsequent in-depth examination of the data employing advanced tools such as Proteome Discoverer and Panther software.

Therefore, a robust protocol based on biotin labelling followed by cell scraping and neutravidin enrichment was established. Moreover, the glycoproteomic analysis revealed a notable enrichment of membrane proteins, underscoring the effectiveness of the established protocol.

Overall, this work has enabled the development of a robust protocol for the isolation of membrane proteins, which will be useful for subsequent analysis of the cell glycoproteome.

Keywords: Glycoproteomics, Biotinylation, Neutravidin, *O*-glycans, Gastric Cancer

1. INTRODUCTION

Proteomic and glycoproteomic analyses of cell lines have emerged as pivotal tools for understanding cellular processes, revealing crucial insights into the molecular mechanisms that govern various physiological and pathological conditions (1). Comprehensive studies have provided insights of the intricate network of proteins and glycoproteins within a cell, highlighting their functional roles, interactions, and post-translational modifications (2). Glycosylation, the most frequent post-translational modification, plays a crucial role in modifying proteins, particularly those found in the cell surface. Alterations in glycosylation patterns have been implicated in the initiation, progression, and metastasis of tumours, rendering them pivotal factors to analyse within the scope of cancer biology (3-6).

One of the main challenges in glycoproteomic research lies in the dynamic range of protein abundance within a cell. Membrane proteins are known to be expressed at much lower levels than their soluble counterparts, making their detection and characterization challenging (7). However, mass spectrometry (MS) approaches, coupled with advanced separation methods, has allowed for comprehensive profiling of cellular membrane proteomes, thus introducing an additional layer of complexity, englobing the study of multiple omics fields (8, 9).

In glycoproteomics, biotinylation, a process involving the covalent binding of biotin molecules to proteins, has proven its utility in selectively labelling and isolating glycoproteins on the cell surface (10, 11). This technique targets proteins, establishing a covalent linkage with primary amines of amino acids. Biotinylated proteins can then be separated using an avidin-based capture/enrichment technique and subsequently analysed, often using mass spectrometry methods (12). This enables the identification and characterization of the proteins that have been labelled. Biotin labelling of cell surface proteins has been applied for *in vitro* and *in vivo* profiling of different cell lines and tissues (13-15). This process can be particularly useful in identifying drug targets and discovering new biomarkers, mostly in diseases like cancer where glycosylation abnormalities are significant (16, 17). Therefore, integrating proteomics and glycoproteomics not only advances our understanding of cancer biology but also unveils precise molecular signatures, contributing to the development of novel strategies for diagnosis and therapy (1, 6, 18).

Envisaging the development of standardized protocols for isolating membrane-associated proteins from cellular content, to enable future targeted glycoproteomic studies, this chapter describes the establishment of a robust protocol for membrane glycoprotein extraction.

2. MATERIALS AND METHODS

2.1. Cell Culture Conditions

The AGS cell line, derived from human gastric adenocarcinoma, was acquired from the American Type Culture Collection (ATCC), while the AGS *C1GALT1* KO cell lines were obtained as previously described in Chapter III. These cells were cultured at 37°C in a 5% CO2 humidified atmosphere and atmospheric oxygen, with RPMI Medium 1640 (1X) + GlutaMAXTM-I (Gibco™, Thermo Fisher Scientific), supplemented with 10% heat-inactivated FBS (Gibco™, Thermo Fisher Scientific), and 1 % penicillin-streptomycin (10,000 Units/mL penicillin; 10,000 mg/mL streptomycin; Gibco, Thermo Fisher Scientific).

2.2. Plasma Membrane Protein Enrichment & Glycoproteomics

In order to optimise the biotinylation protocol, two different methods of detaching the cells were used. In the first one (A), 4 flasks with $60x10^6$ cells (AGS *C1GALT1* KO) each were detached with Versene 1x solution (Gibco; Life Technologies). Subsequently, the cells were incubated with different concentrations of EZ-link™ Sulfo-NHS-SS-Biotin (Thermo Fisher Scientific) for 10 minutes at RT. The conditions used were non-biotinylated (control) and 33.3, 166.7, and 666.7 µg/mL of biotin. In the second method (B), the cells were incubated for 10 min at RT under previously described conditions while still attached to the flasks. After incubation with biotin, the cells were subjected to subcellular fractionation and membrane protein extraction. Cells from group A were resuspended in fractionation buffer (20 mM HEPES buffer (pH = 7.4), 10 mM KCl, 2 mM MgCl2, 1 mM EDTA, and 1 mM EGTA) on ice. Cells from group B were detached by scraping with the fractionation buffer on ice. Cell suspensions from both groups were passed through a 27G needle, left on ice for 20 min, and then centrifuged at 720 g for 5 min at 4°C to remove nuclei. The supernatants were transferred to new tubes and recentrifuged at 10 000 g for 5 min at 4°C to remove mitochondria. Samples were then transferred to polycarbonate centrifuge bottles with cap assemblies and centrifuged for 1 h at 100 000 g at 4°C. The pellets were recovered, resuspended in the fractionation buffer, and passed through a 25G needle before a new centrifugation for 1 h at 100 000 g at 4°C. The final pellets corresponding to membrane proteins were resuspended in an appropriate volume of RIPA buffer and proteinase inhibitors (1mg /mL aprotinin, 1 mg/mL leupeptin, 35 mg/mL PMSF, 10 mM NaF, 50 mM $NaVO_3$, 50 mg/mL $Na_4P_2O_7$), followed by a brief sonication (10 seconds) to homogenise the lysate. The protein content in each sample was estimated using a DC protein assay kit (Bio-Rad). The protein extracts were further enriched to membrane protein with neutravidin.

Briefly, Pierce™ High Capacity NeutrAvidin™ Agarose (Thermo Fisher Scientific) was added to Pierce Spin Columns (Thermo Fisher Scientific). The columns were washed three times with Dulbecco's Phosphate Buffered Saline (PBS 1x; Thermo Fisher Scientific) and then the samples were added to the columns and incubated for 1 h at RT. All samples were eluted with loading buffer, for 5 min at 100ºC, followed by a centrifugation at 14 000 g, 1 min at RT. The samples were further separated through SDS-PAGE employing precast polyacrylamide gels with a gradient of 4-20% (BioRad). The bands were excised from the gel and proteins were reduced with 5 mM dithiothreitol (DTT; Sigma Aldrich) for 40 min at 60°C, alkylated with 10 mM iodoacetamide (Sigma Aldrich) for 45 min in the dark and digested with 50 µL of chymotrypsin (Promega) at 37°C, overnight. Samples were analysed through nanoLC-ESI-MS/MS.

2.3. Fluorescence Immunocytochemistry

After the enrichment of the AGS glycoengineered cells with different concentrations of EZ-link™ Sulfo-NHS-SS-Biotin (Thermo Fisher Scientific) (section **2.2**), the biotin internalization by the gastric cancer cells was evaluated, using fluorescence immunocytochemistry. One million cells of each condition were further fixed with 2% paraformaldehyde (PFA; Sigma-Aldrich) for 15 min, following immunofluorescent staining with Streptavidin, Alexa Fluor™ 488 Conjugate (1:1000; Thermo Fisher Scientific) for 15 min, at RT, and CellMask™ staining. Immunofluorescence images were acquired on a Leica DMI6000 FFW microscope using Las X software (Leica).

2.4. Western Blot

Extracted proteins were separated through SDS-PAGE employing precast polyacrylamide gels with a gradient of 4-20% (BioRad). Subsequently, they were transferred onto a nitrocellulose membrane (GE Healthcare Life Sciences). After blocking with 5% bovine serum albumin (BSA), the membranes were blotted with β-actin (1.4 µg/mL, Sigma-Aldrich) for 1 h at room temperature (RT) and incubated with Peroxidase AffiniPure Goat Anti-Mouse IgG (1:70000, Jackson ImmunoResearch), as secondary antibody, for 30 minutes at RT. To evaluate nucleolin and B2M expression, the membranes were blocked with Carbo-free blocking solution 1x, blotted with anti-nucleolin antibody (1:1000; Abcam) and anti-beta 2 microglobulin antibody (1:2000; Abcam) for 1h at RT, and incubated with goat anti-rabbit IgG (H+L) HPR conjugate antibody (1:6000, Life Technologies), for 30 minutes at RT. For VVA analysis, the membranes, previously blocked with Carbo-free

blocking solution 1x, were blotted for the Tn antigen with biotinylated VVA lectin (1:5000, VectorLabs) for 1 h at room temperature and incubated in the Vectastain Elite ABC reagent, Peroxidase, R.T.U. The Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare Life Sciences) was used as developing reagent. Data analysis was performed through Image Lab Software (Bio-Rad) in a ChemiDoc XRS (Bio-Rad). β-actin and VVA signals were normalized to total protein content.

2.5. Proteomics Data Acquisition

A nanoLC system (Vanquish Neo UHPLC system) was coupled online to a QExactive Plus Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific) equipped with a nano-electrospray ion source (EASY-Spray source; Thermo Fisher Scientific). Eluent A was aqueous formic acid (0.1%), and eluent B was formic acid (0.1%) in acetonitrile (80%). Samples (10 µL) were injected directly into a trapping column (C18 PepMap Neo, 5µm particle size) and redirected to the analytical column (EASY-Spray C18 PepMap, 100 Å, 75 µm x 150 mm ID and 3 µm particle size) at a flow rate of 0.25µL/min. Column temperature was set at 35°C. Peptide separation occurred using the following gradient: 7 min (2.5% B to 12% B), 50 min (12% B to 46% B), 5 min (46% B to 99% B), 10 min (hold 99% B). Subsequently, the column was equilibrated with 2.5% eluent B for 13 min. The mass spectrometer was operated in the positive ion mode, over the *m/z* range 300-2000, with a spray voltage of 1.9 kV and a transfer capillary temperature of 275°C. Full MS settings were 140,000 resolution, AGC target 3×10^6, and maximum injection time 200 ms. The top 15 most intense ions were selected for high-energy collision dissociation (HCD). MS2 settings were normalized collision energy 30%, 17,500 resolution, AGC target 1×10^5, maximum injection time 100 ms, isolation window 4.0 *m/z*, and fixed first mass 110 *m/z*. The following data-dependent settings were used: minimum AGC target 7×10^3, peptide match preferred, dynamic exclusion 30 s, exclusion of unassigned charge and charge =1 and >8, and exclude isotopes on. Data were recorded with Xcalibur software version 4.5 (Thermo Fisher Scientific).

2.6. Data Analysis

Data were analysed using the Proteome Discoverer 3.0 software (Thermo Fisher Scientific), namely the SequestHT search engine and the Percolator algorithm for validation. Protein identification analysis was performed using the human proteome database, containing the reviewed canonical proteins, obtained from Swiss-Prot on 07/07/2023. Chymotrypsin was selected as the enzyme and up to three missed cleavage

sites were allowed, as well as a precursor ion mass tolerance of 10 ppm and 0.02 Da for roduct ions. Carbamidomethylcysteine was selected as a fixed modification. Oxidation of methionine (+15.995u), modification of serine and threonine with HexNac (+203.079u) and/or HexNacNeuAc (STn) (+494.175u) were selected as variable modifications. In addition, variable modification of lysine and protein N-terminus with CAMthiopropanoyl (+145.020u) was considered for biotinylated samples. Proteins with high and medium false discovery rate confidence were considered for downstream analysis. Celular component annotation was performed using Panther (19), considering the following go terms 0005886 for plasma membrane, 0005737 for cytoplasm and 0005634 for nucleus.

3. RESULTS & DISCUSSION

Gastric cancer is characterized by a profound deregulation of glycosylation pathways (20). This work is devoted to the study of one of the most common alterations of glycosylation in gastric tumours, namely Tn and STn antigens. Aberrant glycosylation, translated by the overexpression of these glycans, has been correlated with increased invasive capacity and resistance to cell death, however, through still unexplored mechanisms (21, 22). Glycoproteomics characterization of glycoengineered cell lines will help disclose particularly relevant molecular glycoproteomics signatures driving this behaviour. Great part of potentially glycosylated proteins of the human proteome are cell surface proteins (23). As such, there is a need to establish standardized protocols to isolate membrane associated proteins from the remaining cellular content for targeted glycoproteomic studies. To address this need, we proposed a differential cellular fractioning protocol coupled to a cell surface biotin-neutravidin enrichment method for membrane glycoproteins extraction.

Using AGS *C1GALT1* knock-out cells as a starting point, a reactive biotin ester was employed to label the extracellular domains of integral plasma membrane proteins through primary amines (**Figure 1A**). After labelling, two distinct mechanical cellular extraction methods were tested, namely cell detachment using a chelating agent (versene) or cell scraping (**Figure 1B**). Whole cell lysates were then subjected to differential ultracentrifugation to obtain an enriched membrane protein fraction. Based on previous studies, this enriched fraction still contains fair contamination with both cytoplasm and nuclear proteins. To clear potential contaminants from other cellular organelles, biotin labelled proteins were rescued from this enriched fraction by neutravidin coupling (**Figure 1B**).

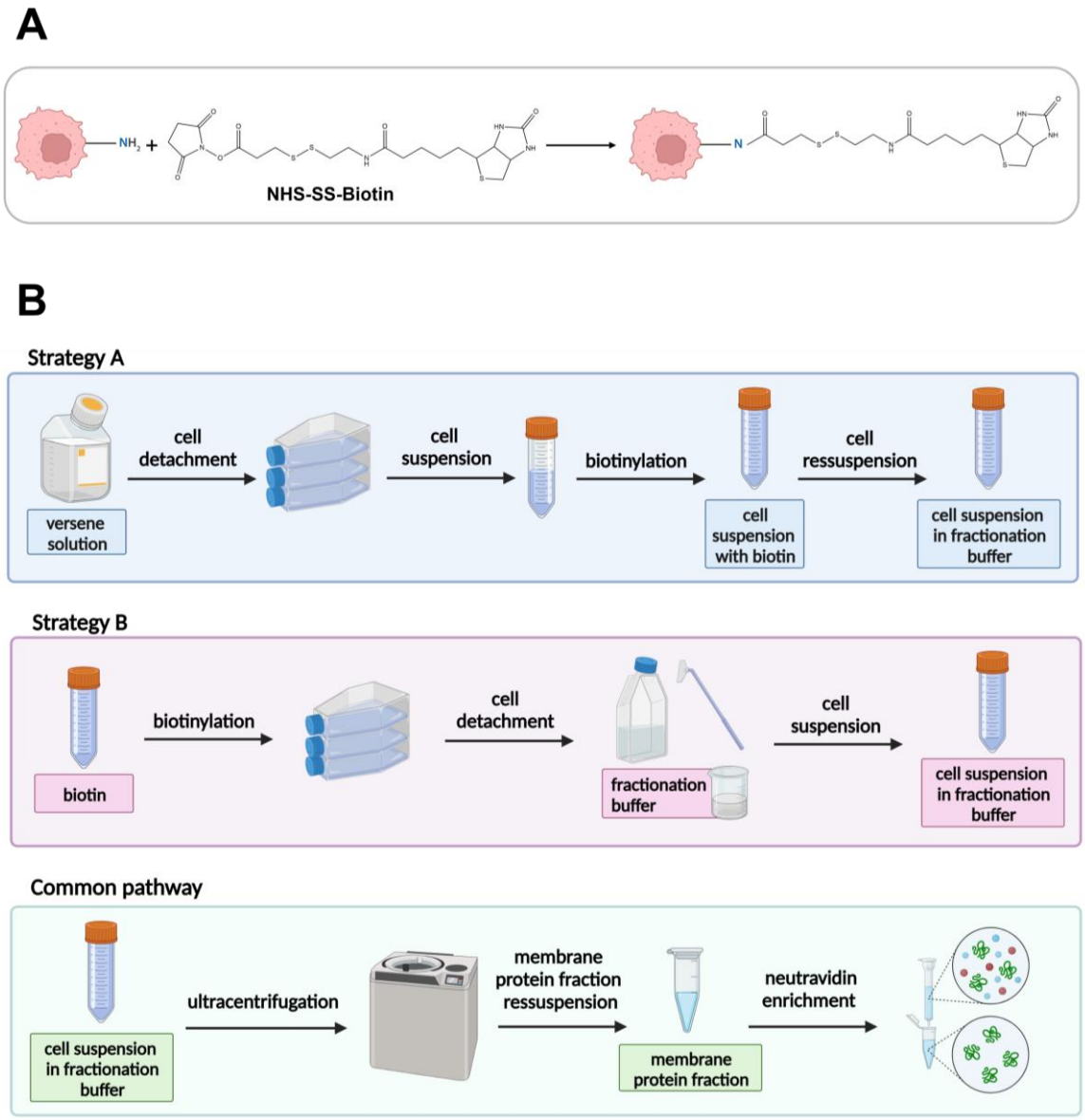

Figure 1. Representative schemes of biotin-neutravidin coupling and differential cellular fractionation protocol. A) Chemical reaction between biotin and proteins. In the presence of a suitable buffer, the NHS-activated biotin reacts with primary amines on the proteins, creating a stable amide bond between biotin and the protein, rendering biotinylated proteins. **B)** Schematic representation of biotinylated membrane proteins isolation from the AGS *C1GALT1* KO cell model. In strategy A, cells were detached using versene following biotinylation and cell resuspension in fractionation buffer for membrane protein isolation. Strategy B was employed firstly by cell biotinylation with further cell detachment by scraping and resuspension in the buffer. Then, the cell suspension obtained using both strategies was submitted to differential ultracentrifugation to obtain the membrane protein fraction. Finally, the isolated fraction was enriched with neutravidin for biotinylated protein isolation. Image created using Biorender.com (accessed on 19 October 2023).

We started by testing different concentrations of biotin (33.3, 166.7, and 666.7 µg/mL) to achieve the best labelling coverage of cell surface proteins without promoting biotin internalization into the cell. Immunocytochemistry analysis was used to assess the degree of biotin internalization throughout the range of concentrations. Of note, biotin was not internalized in any tested condition and cell surface protein staining was achieved, establishing a range of work concentrations that were explored in subsequent studies (**Figure 2**).

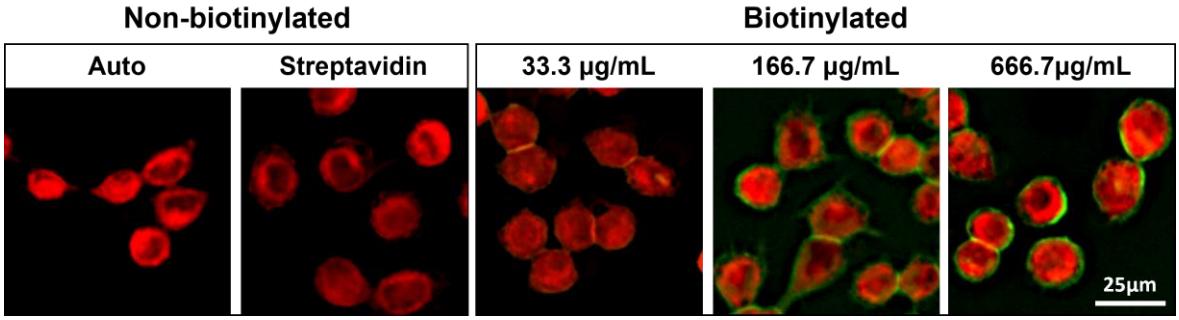

Figure 2. Immunocytochemistry analysis to evaluate biotin labelling in AGS *C1GALT1* KO cells. Biotin incorporation into the cytoplasm did not occur at any of the different concentrations tested. A concentration of 33.3 µg/mL is not sufficient to label membrane proteins. Moreover, the 666.7 µg/mL concentration shows optimal membrane protein labelling, making it the preferred condition for enrichment. The absence of labelling in the autofluorescence condition and in the condition marked only with streptavidin corroborates the previously described results, used as negative controls.

We then sought to determine the degree of nuclear and cytoplasmic protein contaminants in the final cell membrane protein fraction using both versene and scraping cell lysis methods. We used the protein nucleolin as a nuclear protein marker, β-actin as cytoplasmic protein marker, and B2M as a plasma membrane marker. Using versene chelating solution for cellular extraction we were not able to discern any enrichment of cell membrane components (**Figure 3A**), although, several extracted proteins were positive for Tn antigen as highlighted by VVA staining (**Figure 3B**). Using cell scraping as a cellular extraction method we were able to confirm an enrichment of B2M cell membrane marker over nuclear and cytoplasmic components across all ranges of biotin concentrations. Moreover, increased biotin concentration was translated by an increased cell membrane components extraction efficiency (**Figure 3D**). In addition, the extracted cell membrane proteins included an increased representativity of Tn glycosylated proteins as demonstrated by increased VVA staining (**Figure 3E**).

Overall, cell scraping coupled to biotin-neutravidin enrichment has allowed increased cell membrane extract purity, with fewer nuclear and cytoplasmic component contaminants.

Moreover, this methodology achieved an increased representativity of Tn glycosylated glycoproteins over the versene extraction method. Accordingly, mechanical scraping followed by biotin staining, differential ultracentrifugation and neutravidin coupling to biotin was the elected experimental procedure for subsequent studies.

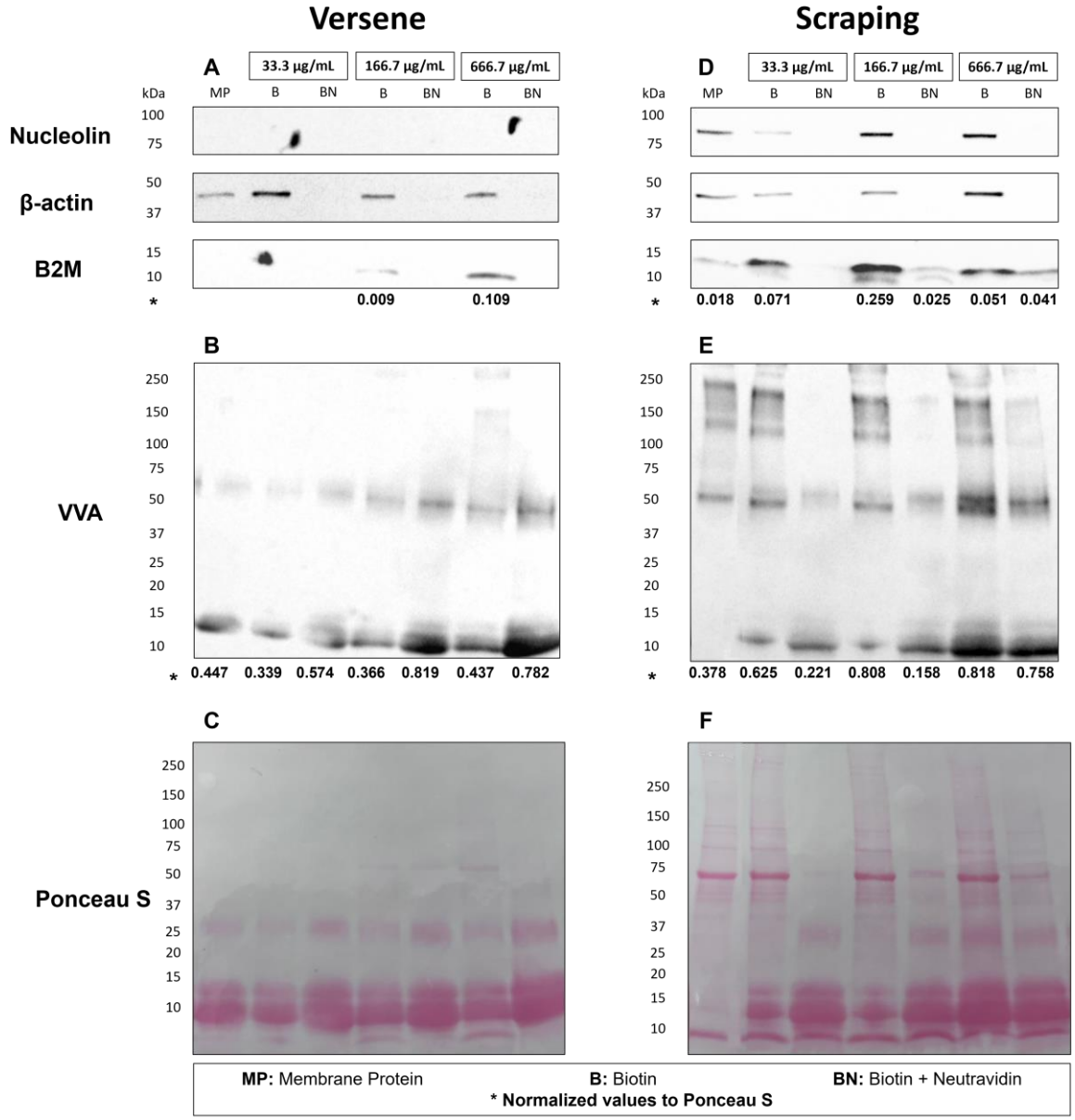

Figure 3. Evaluation of cellular fraction methods through Western blot analysis. A) Nucleolin, β- actin and B2M enrichment evaluation upon versene cellular extraction. **B)** VVA lectin staining of versene cellular extraction conditions. **C)** Ponceau S loading control of versene cellular extraction conditions. **D)** Nucleolin, β-actin and B2M enrichment evaluation upon scraping cellular extraction. **E)** VVA lectin staining of scraping cellular extraction conditions. F**)** Ponceau S loading control of scraping cellular extraction conditions.

After protocol establishment and validation, we have progressed to the glycoproteomic analysis of protein content in AGS *C1GALT1* knock-out cells. Membrane protein extraction was performed as above described, following band excision from SDS-PAGE gels (**Figure 4A**) and protein annotation using Proteome Discoverer software and the human proteome from the Swiss-Prot database. According to Proteome Discoverer assignments the biotin-neutravidin enrichment step has allowed the identification of an increased number of proteins (**Figure 4B**). Glycoproteomics data was integrated using Panther software to assign glycoproteins to specific cellular compartments (**Table S1** and **Table S2**). Go terms analysis highlighted a decreased nuclear protein content in the total identified proteins (44% to 14%, **Figure 4C**), using the proposed enrichment strategy. Moreover, an enrichment of the membrane proteins fraction was achieved (11% to 40%, **Figure 4C**). Of note, according to Panther analysis, the cytoplasmic proteins content was maintained regardless of the enrichment protocol (45% to 46%) for membrane proteins. This is mainly due to a data mining issue related to protein assignments to specific subcellular components. Specifically, proteins that had been identified in the cytoplasm and cellular membrane compartments will be automatically assigned to both compartments. As such, the 40% of proteins assigned to the cytoplasm fraction could be duplicated in the plasma membrane. The clarification of the exact number of membrane protein gains over cytoplasmic proteins would demand the manual removal of all proteins detected in the plasma membrane from the cytoplasmic proteins list (**Table S2**). Even though the automatic assignment of membrane proteins to the cytoplasm fraction is a relevant issue to be addressed, the sensitivity of glycoproteomics in detecting cytoplasmic contamination in the membrane protein fraction will be taken into account. Accordingly, an additional step of protein enrichment will be added to the proposed experimental protocol. Namely, after biotin-neutravidin enrichment, proteins will be subjected to an additional affinity chromatography based on VVA lectin to further enrich the protein fraction in plasma membrane glycosylated proteins.

Overall, this work has allowed the establishment and refinement of an enrichment protocol targeting plasma membrane proteins. Further protocol refinement will be pursued by addressing data mining issues concerning protein assignment to subcellular compartments and adding extra steps of affinity chromatography targeted to glycosylated cell surface proteins.

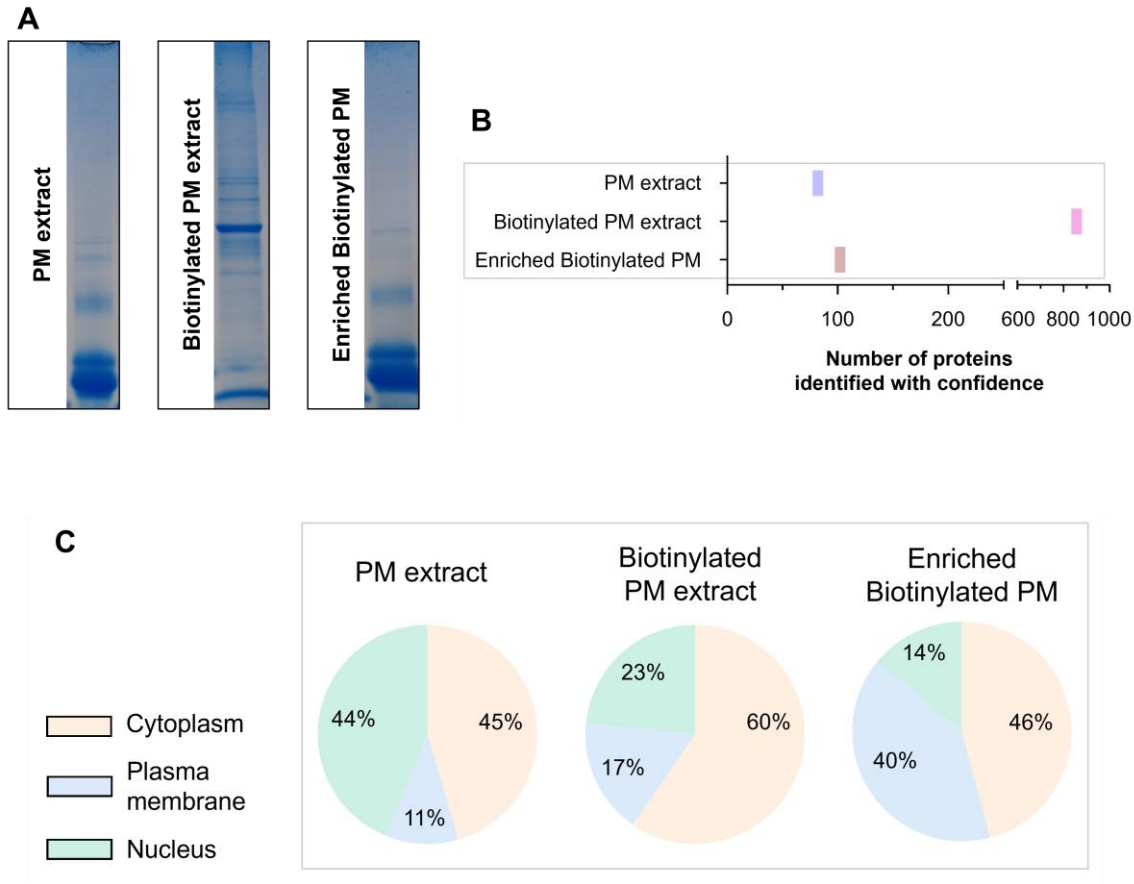

Figure 4. Glycoproteomic analysis of protein content in AGS *C1GALT1* knock-out cells. A) SDS-PAGE gel bands stained with PageBlue Protein Staining Solution, highlighting protein separation. **B)** Number of proteins identified with confidence, using Proteome Discoverer. The biotin-neutravidin enrichment step has allowed the identification of an increased number of proteins. **C)** Glycoproteomics data integrated using Panther software to assign glycoproteins to specific cellular compartments. Go terms analysis highlighted a decreased nuclear protein content in the total identified proteins after enrichment with neutravidin (44% to 14%), and an enrichment of the membrane proteins fraction (11% to 40%). PM stands for Plasma membrane protein.

4. CONCLUDING REMARKS

Glycobiology has emerged as a pivotal field in cancer research due to the intricate interplay between glycans and proteins, playing a crucial role in several aspects of cancer biology, including cell signalling, metastasis, and immune response modulation (4, 24, 25). Understanding the glycoproteomic profile holds huge potential for uncovering novel biomarkers and therapeutic targets, revolutionizing cancer diagnosis and treatment strategies (6, 26). Mass spectrometry coupled with other analytical techniques have promoted increased sensitivity and resolution in glycoprotein analysis (27, 28). Moreover, novel strategies for glycan labelling and enrichment have emerged, facilitating the isolation and characterization of glycoproteins from complex biological samples (9, 29). Nevertheless, it is imperative to develop uniform procedures for the separation of membrane-bound proteins from the rest of the cellular components in order to conduct focused investigations on glycoproteomics. As such, to address this need, this chapter describes the establishment of a robust differential fractioning protocol for membrane protein extraction. Accordingly, biotin staining followed by mechanical scraping, differential ultracentrifugation and neutravidin coupling to biotin was the elected experimental procedure, since it allowed increased cell membrane extract purity, with less nuclear and cytoplasmic components. Moreover, we progressed to glycoproteomic analysis, which revealed that there was a successful attainment of an enrichment of membrane proteins.

In future studies, an additional step of protein enrichment, namely affinity chromatography based on VVA lectin, will be included after biotin-neutravidin enhancement process in order to increase the protein component within the glycosylated proteins found in the plasma membrane. Further protocol refinement will be pursued by addressing data mining issues concerning protein assignment to subcellular compartments.

5. REFERENCES

1. Tan HT, Lee YH, Chung MC. Cancer proteomics. Mass Spectrom Rev. 2012;31(5):583-605.
2. Oliveira T, Thaysen-Andersen M, Packer NH, Kolarich D. The Hitchhiker's guide to glycoproteomics. Biochem Soc Trans. 2021;49(4):1643-62.
3. Fernandes E, Soares J, Cotton S, Peixoto A, Ferreira D, Freitas R, et al. Esophageal, gastric and colorectal cancers: Looking beyond classical serological biomarkers towards glycoproteomics-assisted precision oncology. Theranostics. 2020;10(11):4903-28.
4. Ferreira JA, Magalhães A, Gomes J, Peixoto A, Gaiteiro C, Fernandes E, et al. Protein glycosylation in gastric and colorectal cancers: Toward cancer detection and targeted therapeutics. Cancer Letters. 2017;387:32-45.
5. Campos D, Freitas D, Gomes J, Magalhães 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.
6. Abyadeh M, Meyfour A, Gupta V, Zabet Moghaddam M, Fitzhenry MJ, Shahbazian S, et al. Recent Advances of Functional Proteomics in Gastrointestinal Cancers- a Path towards the Identification of Candidate Diagnostic, Prognostic, and Therapeutic Molecular Biomarkers. Int J Mol Sci. 2020;21(22).
7. Leth-Larsen R, Lund RR, Ditzel HJ. Plasma membrane proteomics and its application in clinical cancer biomarker discovery. Mol Cell Proteomics. 2010;9(7):1369-82.
8. Mermelekas G, Zoidakis J. Mass spectrometry-based membrane proteomics in cancer biomarker discovery. Expert Rev Mol Diagn. 2014;14(5):549-63.
9. Bagdonaite I, Malaker SA, Polasky DA, Riley NM, Schjoldager K, Vakhrushev SY, et al. Glycoproteomics. Nature Reviews Methods Primers. 2022;2(1):48.
10. Elia G. Biotinylation reagents for the study of cell surface proteins. Proteomics. 2008;8(19):4012-24.
11. Li Y, Wang Y, Mao J, Yao Y, Wang K, Qiao Q, et al. Sensitive profiling of cell surface proteome by using an optimized biotinylation method. J Proteomics. 2019;196:33-41.
12. Ye X, Kaczmarczyk JA, Luke B, Saul RG, Whiteley GR, Nissley DV, et al. Chapter 24 - Cell surface protein enrichment for biomarker and drug target discovery using mass spectrometry-based proteomics. In: Issaq HJ, Veenstra TD, editors. Proteomic and Metabolomic Approaches to Biomarker Discovery (Second Edition). Boston: Academic Press; 2020. p. 409-20.
13. Katsantoni E. Protein Complexes and Target Genes Identification by in Vivo Biotinylation: The STAT5 Paradigm. Science Signaling. 2012;5(248):pt13-pt.

14. Hanke SA, Kerner A, Nadler WM, Trumpp A, Zhang Y, Rösli CP. The pivotal role of reactivity in the design of novel biotinylation reagents for the chemical-proteomics-based identification of vascular accessible biomarkers. Journal of Proteomics. 2016;141:57-66.
15. Vilen Z, Reeves AE, O'Leary TR, Joeh E, Kamasawa N, Huang ML. Cell Surface Engineering Enables Surfaceome Profiling. ACS Chemical Biology. 2023;18(4):701-10.
16. Smeekens JM, Chen W, Wu R. Mass Spectrometric Analysis of the Cell Surface N-Glycoproteome by Combining Metabolic Labeling and Click Chemistry. Journal of The American Society for Mass Spectrometry. 2015;26(4):604-14.
17. Conrotto P, Roesli C, Rybak J, Kischel P, Waltregny D, Neri D, et al. Identification of new accessible tumor antigens in human colon cancer by ex vivo protein biotinylation and comparative mass spectrometry analysis. International Journal of Cancer. 2008;123(12):2856-64.
18. Azevedo R, Peixoto A, Gaiteiro C, Fernandes E, Neves M, Lima L, et al. Over forty years of bladder cancer glycobiology: Where do glycans stand facing precision oncology? Oncotarget. 2017;8(53):91734-64.
19. Mi H, Muruganujan A, Huang X, Ebert D, Mills C, Guo X, et al. Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). Nat Protoc. 2019;14(3):703-21.
20. Pinho SS, Carvalho S, Marcos-Pinto R, Magalhães A, Oliveira C, Gu J, et al. Gastric cancer: adding glycosylation to the equation. Trends in Molecular Medicine. 2013;19(11):664-76.
21. Munkley J, Elliott DJ. Hallmarks of glycosylation in cancer. Oncotarget. 2016;7(23):35478-89.
22. Oliveira-Ferrer L, Legler K, Milde-Langosch K. Role of protein glycosylation in cancer metastasis. Semin Cancer Biol. 2017;44:141-52.
23. Takahashi M, Kizuka Y, Ohtsubo K, Gu J, Taniguchi N. Disease-associated glycans on cell surface proteins. Mol Aspects Med. 2016;51:56-70.
24. Josic D, Martinovic T, Pavelic K. Glycosylation and metastases. Electrophoresis. 2019;40(1):140-50.
25. Pinho SS, Reis CA. Glycosylation in cancer: mechanisms and clinical implications. Nature Reviews Cancer. 2015;15(9):540-55.
26. Costa AF, Campos D, Reis CA, Gomes C. Targeting Glycosylation: A New Road for Cancer Drug Discovery. Trends in Cancer. 2020;6(9):757-66.
27. Sun F, Suttapitugsakul S, Wu R. Systematic characterization of extracellular glycoproteins using mass spectrometry. Mass Spectrom Rev. 2023;42(2):519-45.

28. Chandler KB, Costello CE. Glycomics and glycoproteomics of membrane proteins and cell-surface receptors: Present trends and future opportunities. Electrophoresis. 2016;37(11):1407-19.
29. Delafield DG, Li L. Recent Advances in Analytical Approaches for Glycan and Glycopeptide Quantitation. Mol Cell Proteomics. 2021;20:100054.

Glycoproteomics Characterization of Gastric Cancer Cells

Martina Dourado-Gonçalves, Rui Freitas, Marta Relvas-Santos, Dylan Ferreira, Lúcio Lara Santos, Andreia Peixoto, José Alexandre Ferreira

SUPPLEMENTARY INFORMATION

Table S1. Cellular component identification in three subcellular compartment for membrane protein isolated by the versene approach followed by neutravidin enrichment.

Gene ID[1]	Protein ID[2]	Gene Name	Gene Symbol
		Cytoplasm	
12669	P21796	Voltage-dependent anion-selective channel protein 1	VDAC1
31332	Q6IQ22	Ras-related protein Rab-12	RAB12
20768	Q9BQE3	Tubulin alpha-1C chain;	TUBA1C
20776	Q9BUF5	Tubulin beta-6 chain	TUBB6
14902	Q9NRW1	Ras-related protein Rab-6B	RAB6B
33276	Q5D862	Filaggrin-2	FLG2
22215	Q9P2G1	Ankyrin repeat and IBR domain-containing protein 1	ANKIB1
9760	P62491	Ras-related protein Rab-11A	RAB11A
30269	Q96E17	Ras-related protein Rab-3C	RAB3C
9792	P51151	Ras-related protein Rab-9A	RAB9A
20846	Q86YZ3	Hornerin	HRNR
537	P07355	Annexin A2	ANXA2
9788	P51149	Ras-related protein Rab-7a	RAB7A
16940	Q96PD7	Diacylglycerol O-acyltransferase 2	DGAT2
6502	P08865	40S ribosomal protein SA	RPSA
30829	Q9BVA1	Tubulin beta-2B chain	TUBB2B
7720	P20929	Nebulin	NEB
4141	P04406	Glyceraldehyde-3-phosphate dehydrogenase	GAPDH
9767	O00194	Ras-related protein Rab-27B	RAB27B
20772	Q13509	Tubulin beta-3 chain	TUBB3
1473	P27824	Calnexin	CANX
4399	P63244	Receptor of activated protein C kinase 1	RACK1
9774	Q15286	Ras-related protein Rab-35	RAB35
170	P61158	Actin-related protein 3	ACTR3
20778	P07437	Tubulin beta chain	TUBB
9762	P51153	Ras-related protein Rab-13	RAB13
9411	P14314	Glucosidase 2 subunit beta	PRKCSH
9778	P20337	Ras-related protein Rab-3B	RAB3B
30273	Q92930	Ras-related protein Rab-8B	RAB8B
9777	P20336	Ras-related protein Rab-3A	RAB3A
		Plasma membrane	
31332	Q6IQ22	Ras-related protein Rab-12	RAB12
4933	P10321	HLA class I histocompatibility antigen, C alpha chain	HLA-C
400	Q13740	CD166 antigen	ALCAM
7373	P26038	Moesin	MSN
4965	P01893	Putative HLA class I histocompatibility antigen, alpha chain H	HLA-H
33276	Q5D862	Filaggrin-2	FLG2
30269	Q96E17	Ras-related protein Rab-3C	RAB3C
20846	Q86YZ3	Hornerin	HRNR
4931	P04439	HLA class I histocompatibility antigen, A alpha chain	HLA-A
537	P07355	Annexin A2	ANXA2
9774	Q15286	Ras-related protein Rab-35	RAB35
9762	P51153	Ras-related protein Rab-13	RAB13
6153	P05556	Integrin beta-1	ITGB1
9778	P20337	Ras-related protein Rab-3B	RAB3B
30273	Q92930	Ras-related protein Rab-8B	RAB8B
6137	P17301	Integrin alpha-2	ITGA2
9777	P20336	Ras-related protein Rab-3A	RAB3A
30041	Q6P1J6	Phospholipase B1, membrane-associated	PLB1
30422	Q8TF72	Protein Shroom3	SHROOM3
1116	P35613	Basigin	BSG
		Nucleus	
33276	Q5D862	Filaggrin-2	FLG2
5187	Q99873	Protein arginine N-methyltransferase 1	PRMT1
20846	Q86YZ3	Hornerin	HRNR
537	P07355	Annexin A2	ANXA2
9536	O14818	Proteasome subunit alpha type-7	PSMA7
4399	P63244	Receptor of activated protein C kinase 1	RACK1

[1] Gene database source: HUGO Gene Nomenclature Committee; [2] Protein database source: Uniprot

Table S2. Cellular component identification in three subcellular compartment for membrane protein isolated by the scrapping approach followed by neutravidin enrichment.

Gene ID[1]	Protein ID[2]	Gene Name	Gene Symbol
		Cytoplasm	
9761	Q15907	Ras-related protein Rab-11B	RAB11B
14188	O75366	Advillin	AVIL
24460	Q7L2H7	Eukaryotic translation initiation factor 3 subunit M	EIF3M
6656	Q9UIQ6	Leucyl-cystinyl aminopeptidase	LNPEP
4138	Q14697	Neutral alpha-glucosidase AB	GANAB
18311	Q9BS26	Endoplasmic reticulum resident protein 44	ERP44
1546	P50454	Serpin H1	SERPINH1
33276	Q5D862	Filaggrin-2	FLG2
10371	P05388	60S acidic ribosomal protein P0	RPLP0
9760	P62491	Ras-related protein Rab-11A	RAB11A
9783	P20339	Ras-related protein Rab-5	RAB5A
30269	Q96E17	Ras-related protein Rab-3C	RAB3C
20846	Q86YZ3	Hornerin	HRNR
5238	P11021	Endoplasmic reticulum chaperone BiP	HSPA5
5253	P07900	Heat shock protein HSP 90-alpha	HSP90AA1
537	P07355	Annexin A2	ANXA2
11186	Q99523	Sortilin	SORT1
6752	P20645	Cation-dependent mannose-6-phosphate receptor	M6PR
9784	P61020	Ras-related protein Rab-5B	RAB5B
4606	P30101	Protein disulfide-isomerase A3	PDIA3
7720	P20929	Nebulin	NEB
1473	P27824	Calnexin	CANX
4399	P63244	Receptor of activated protein C kinase 1	RACK1
2728	P39656	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit	DDOST
30167	P13667	Protein disulfide-isomerase A4	PDIA4
170	P61158	Actin-related protein 3	ACTR3
9778	P20337	Ras-related protein Rab-3B	RAB3B
36809	A0A8I5KQE6	Ribosomal protein SA pseudogene 58	RPSAP58
12028	P14625	Endoplasmin	HSP90B1
16986	Q12907	Vesicular integral-membrane protein VIP36	LMAN2
		Plasma membrane	
2665	P08174	Complement decay-accelerating factor	CD55
4933	P10321	HLA class I histocompatibility antigen, C alpha chain	HLA-C
400	Q13740	CD166 antigen	ALCAM
6656	Q9UIQ6	Leucyl-cystinyl aminopeptidase	LNPEP
4965	P01893	Putative HLA class I histocompatibility antigen, alpha chain H	HLA-H
4964	P17693	HLA class I histocompatibility antigen, alpha chain G	HLA-G
33276	Q5D862	Filaggrin-2	FLG2
6150	P06756	Integrin alpha-V	ITGAV
9783	P20339	Ras-related protein Rab-5A	RAB5A
4962	P13747	HLA class I histocompatibility antigen, alpha chain E	HLA-E
30269	Q96E17	Ras-related protein Rab-3C	RAB3C
20846	Q86YZ3	Hornerin	HRNR
2548	O60494	Cubilin	CUBN
5253	P07900	Heat shock protein HSP 90-alpha	HSP90AA1
537	P07355	Annexin A2	ANXA2
1756	Q12864	Cadherin-17	CDH17
9784	P61020	Ras-related protein Rab-5B	RAB5B
4932	P01889	HLA class I histocompatibility antigen, B alpha chain	HLA-B
12691	P15311	Ezrin	EZR
6470	P32004	Neural cell adhesion molecule L1	L1CAM
6153	P05556	Integrin beta-1	ITGB1
9778	P20337	Ras-related protein Rab-3B	RAB3B
11246	O43278	Kunitz-type protease inhibitor 1	SPINT1
11026	P08195	4F2 cell-surface antigen heavy chain	SLC3A2
6137	P17301	Integrin alpha-2	ITGA2
11763	P02786	Transferrin receptor protein 1	TFRC
188|	O14672	Disintegrin and metalloproteinase domain-containing protein 10	ADAM10

3236	P00533	Epidermal growth factor receptor	EGFR
6142	P23229	Integrin alpha-6	ITGA6
		Nucleus	
33276	Q5D862	Filaggrin-2	FLG2
20846	Q86YZ3	Hornerin	HRNR
6037	Q12905	Interleukin enhancer-binding factor 2	ILF2
5238	P11021	Endoplasmic reticulum chaperone BiP	HSPA5
5253	P07900	Heat shock protein HSP 90-alpha	HSP90AA1
537	P07355	Annexin A2	ANXA2
4399	P63244	Receptor of activated protein C kinase 1	RACK1
3213	P26641	Elongation factor 1-gamma	EEF1G
12833	P13010	X-ray repair cross-complementing protein 5	XRCC5
9583	P26599	Polypyrimidine tract-binding protein 1	PTBP1

[1] Gene database source: HUGO Gene Nomenclature Committee; [2] Protein database source: Uniprot

CHAPTER V

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

Despite the progress in targeted cancer therapy, the effectiveness of most treatments for GC remains limited, underscoring the critical need for safer and effective therapies (10, 11). Altered glycosylation resulting from alterations in cellular and metabolic processes results in atypical displays of aberrant glycans on cell membranes, contributing to initiate the malignant transformation of cells (1-3). One of the main mechanisms underlying tumour-associated alterations is the expression of truncated *O*-glycans, namely Tn and sialyl-Tn antigens (4-6). These glycans are detected in a wide range of solid tumours, including gastric cancer, exhibiting associations with metastasis and poor prognosis (7-9). Given that glycosylation is a well-documented hallmark of cancer, in order to improve the specificity and efficacy of current targeted techniques, strategies that take advantage of abnormal glycosylation of cancer cells can offer therapeutic options (12-14). Recent advancements in mass spectrometry and its integration with other analytical methodologies, have resulted in enhanced sensitivity and resolution in the investigation of the glycoproteome for the identification of cancer specific molecular signatures to be used as reliable therapeutic targets (15, 16). However, it is imperative to establish standardized protocols for the isolation of membrane proteins from other cellular constituents, ensuring the acquisition of dependable and consistent data about the glycoproteomic profiles.

Taking this insights in consideration, this work started by genetically modifying a gastric cancer cell line (AGS; intestinal type) disabling the *C1GALT1* gene, resulting in an overexpression of shortened *O*-glycans, namely Tn and STn. Following phenotypic characterization of the model, functional tests were conducted to discern the functional impact of the premature stop in *O*-glycosylation. Consequently, cells with overexpression of Tn and STn antigens exhibited a reduction in their migratory capacity, while displaying increased invasive potential and resistance to anoikis. These findings suggest that *O*-glycan truncation induces heightened aggressive and metastatic characteristics in GC cells.

Given the importance of developing robust and standardized protocols for isolating membrane proteins used in targeted glycoproteomic analysis, we have established a differential fractioning protocol for extracting biotinylated membrane proteins. Accordingly, we found that the method that allowed higher membrane protein yield with lower nuclear and cytoplasmatic contamination was the biotinylation of attached cells following scrapping detachment and neutravidin enrichment. Glycoproteomic analysis demonstrated that the developed strategy allowed a significant enrichment of the plasma membrane protein fraction. This methodology holds promise for facilitating future studies aiming to characterize the glycoproteome of the cells.

The work outlined in this dissertation represents an advancement in glycoproteomic studies in gastric cancer. Thus, it enabled the creation of a stable and thoroughly characterized cell line, addressing both phenotypic and functional aspects of short-chain *O*-

glycans overexpression, as well as the successful establishment of a protocol for isolating membrane proteins.

In the future, *in vitro* studies will be employed to disclosure the phenotypic and functional impact of ST6GALNAC1 knock-in and core 3 pathway knock-out in glycoengineered GC cells. Additionally, conducting cell-cell adhesion assays and evaluating the metalloproteinases profile of Tn and STn overexpressing cells will help to understand their migration and invasion capabilities. Moreover, a comprehensive approach, encompassing both organoid and *in vivo* studies, is crucial to deepen our understanding of the implications of short-chain *O*-glycans on GC aggressiveness within the microenvironment. Finally, the established methodology of membrane protein fraction isolation exhibits significant potential to drive future studies dedicated to characterizing the glycoproteome of cells. This research holds promise for identifying targets of interest, thereby enhancing theragnostic approaches in the scope of gastric cancer and ultimately improving the clinical management of the disease. Of note, refinement of the protocol for isolating membrane proteins will address data mining issues and test additional steps of affinity chromatography, targeting glycosylated cell surface proteins.

REFERENCES

1. Thomas D, Rathinavel AK, Radhakrishnan P. Altered glycosylation in cancer: A promising target for biomarkers and therapeutics. Biochim Biophys Acta Rev Cancer. 2021;1875(1):188464.
2. Mereiter S, Balmaña M, Campos D, Gomes J, Reis CA. Glycosylation in the Era of Cancer-Targeted Therapy: Where Are We Heading? Cancer Cell. 2019;36(1):6-16.
3. Stowell SR, Ju T, Cummings RD. Protein glycosylation in cancer. Annu Rev Pathol. 2015;10:473-510.
4. Freitas D, Campos D, Gomes J, Pinto F, Macedo JA, Matos R, et al. O-glycans truncation modulates gastric cancer cell signaling and transcription leading to a more aggressive phenotype. EBioMedicine. 2019;40:349-62.
5. Rømer TB, Aasted MKM, Dabelsteen S, Groen A, Schnabel J, Tan E, et al. Mapping of truncated O-glycans in cancers of epithelial and non-epithelial origin. Br J Cancer. 2021;125(9):1239-50.
6. Chia J, Goh G, Bard F. Short O-GalNAc glycans: regulation and role in tumor development and clinical perspectives. Biochimica et Biophysica Acta (BBA) - General Subjects. 2016;1860(8):1623-39.
7. Ju T, Wang Y, Aryal RP, Lehoux SD, Ding X, Kudelka MR, et al. Tn and sialyl-Tn antigens, aberrant O-glycomics as human disease markers. PROTEOMICS – Clinical Applications. 2013;7(9-10):618-31.
8. Lima L, Neves M, Oliveira MI, Dieguez L, Freitas R, Azevedo R, et al. Sialyl-Tn identifies muscle-invasive bladder cancer basal and luminal subtypes facing decreased survival, being expressed by circulating tumor cells and metastases. Urologic Oncology: Seminars and Original Investigations. 2017;35(12):675.e1-.e8.
9. Oliveira-Ferrer L, Legler K, Milde-Langosch K. Role of protein glycosylation in cancer metastasis. Semin Cancer Biol. 2017;44:141-52.
10. Guan W-L, He Y, Xu R-H. Gastric cancer treatment: recent progress and future perspectives. Journal of Hematology & Oncology. 2023;16(1):57.
11. Carrasco-Garcia E, García-Puga M, Arevalo S, Matheu A. Towards precision medicine: linking genetic and cellular heterogeneity in gastric cancer. Ther Adv Med Oncol. 2018;10:1758835918794628.
12. Abyadeh M, Meyfour A, Gupta V, Zabet Moghaddam M, Fitzhenry MJ, Shahbazian S, et al. Recent Advances of Functional Proteomics in Gastrointestinal Cancers- a Path towards the Identification of Candidate Diagnostic, Prognostic, and Therapeutic Molecular Biomarkers. Int J Mol Sci. 2020;21(22).

13. Azevedo R, Peixoto A, Gaiteiro C, Fernandes E, Neves M, Lima L, et al. Over forty years of bladder cancer glycobiology: Where do glycans stand facing precision oncology? Oncotarget. 2017;8(53):91734-64.
14. Campos D, Freitas D, Gomes J, Magalhães 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.
15. Sun F, Suttapitugsakul S, Wu R. Systematic characterization of extracellular glycoproteins using mass spectrometry. Mass Spectrom Rev. 2023;42(2):519-45.
16. Chandler KB, Costello CE. Glycomics and glycoproteomics of membrane proteins and cell-surface receptors: Present trends and future opportunities. Electrophoresis. 2016;37(11):1407-19.