U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

MSc

2.º
CICLO

FCUP
ICBAS
2019

U.PORTO

Uncovering Heparan Sulfate ProteoGlycans biosynthetic pathways and functional relevance in gastric cancer

Catarina Marques

FC

U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

U.PORTO
INSTITUTO DE CIÊNCIAS
BIOMÉDICAS ABEL SALAZAR
UNIVERSIDADE DO PORTO

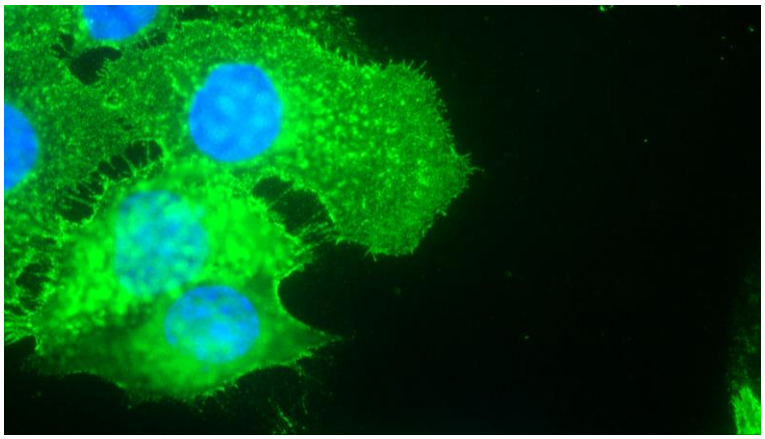
MSc

Uncovering Heparan Sulfate ProteoGlycans biosynthetic pathways and functional relevance in gastric cancer

Catarina Marques

Dissertação de Mestrado apresentada à
Faculdade de Ciências da Universidade do Porto, Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto
Bioquímica

2019



Immunofluorescence labeling of Heparan Sulfate chains in MKN74 gastric cancer cells deficient in the glycosyltransferase EXTL2.

Uncovering Heparan Sulfate ProteoGlycans biosynthetic pathways and functional relevance in gastric cancer

Catarina Marques

Mestrado em Bioquímica

Departamento de Química e Bioquímica

2019

Orientador

Ana Magalhães; Investigadora Júnior, Instituto de Patologia e Imunologia Molecular da Universidade do Porto – IPATIMUP/ Instituto de Investigação e Inovação em Saúde - i3S

Coorientador

Celso A. Reis; Professor Auxiliar Convidado, Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto – ICBAS; Investigador Coordenador, Instituto de Patologia e Imunologia Molecular da Universidade do Porto – IPATIMUP/ Instituto de Investigação e Inovação em Saúde - i3S

U.PORTO

INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR
UNIVERSIDADE DO PORTO

U.PORTO

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, 22 / 10 / 2019

Agradecimentos

Ao longo deste último ano de trabalho aprendi bastante, não só do ponto de vista científico, mas também à cerca de mim mesma e dos objetivos que pretendo alcançar. A tese que apresento é o produto de toda esta aprendizagem, feita não só de novas noções, mas principalmente de pessoas.

Gostaria de começar por agradecer à minha orientadora e co-orientador. À Dra. Ana Magalhães, por me ter permitido contribuir para o desenvolvimento deste projeto e pela confiança que depositou em mim, tanto no início desta jornada, como ao longo de todo o processo. Obrigada por toda a partilha e pelo apoio inesgotável, sem os quais chegar aqui teria sido impossível. Por via do exemplo ensinou-me a "fazer ciência", algo que levarei comigo para futuras etapas. Ao Dr. Celso A. Reis, agradeço por me ter dado oportunidade de realizar a minha dissertação de mestrado no grupo de investigação "Glycobiology in Cancer", bem como por todo o apoio e contributo científico, que foram essenciais à concretização desta dissertação.

Valorizo também todos os elementos do grupo "Glycobiology in Cancer" com quem aprendi bastante. Um especial agradecimento à Juliana Poças e à Isabel Faria-Ramos, pelo conhecimento e experiência que me transmitiram, por todo o tempo que disponibilizaram para me ensinar e pela generosidade. Possuem uma força e perseverança contagiantes e graças a isso pude crescer bastante com o nosso trabalho em equipa. Agradeço também à Dra. Catarina Gomes por ter fornecido os modelos celulares que foram fundamentas neste trabalho e ao Henrique Duarte pelo conhecimento que partilhou a nível de sinalização celular. Obrigada à Mariana Ferreira pelo apoio num período inicial de adaptação, o seu contributo foi sem dúvida fundamental.

E porque somos parte das partes que atravessam o nosso caminho, agradeço também à Lima, à Rafa, à Sia, à Telma, à Catarina, à Mariana e à Ju, por me aceitaram e ouvirem, pelas batalhas partilhadas, algumas derrotas, muitas vitórias e festejos. Num instante os anos passaram, e estas amigas que trago no coração mantiveram-se.

Obrigado também à minha família, em especial à minha mãe, por acreditar sempre em mim, por batalhar sempre ao meu lado, por me acompanhar nas sessões de estudos que entram madrugada dentro, resumindo, por me dar asas para voar e um ninho para voltar.

E por último, mas não menos importante, quero agradecer ao Leo, por me acompanhar sempre, mesmo nas viagens mais inóspitas, por ser o meu maior ouvinte e por continuar a "empurrar-me" em direção à melhor versão de mim própria.

Note

The work presented in this thesis has resulted in the following scientific contributions:

Poster presentations:

C Marques, I Faria-Ramos, J Poças, H Duarte, C Gomes, CA Reis, A Magalhães "Uncovering Heparan Sulfate ProteoGlycans biosynthetic pathways and functional relevance in gastric cancer". 13th International Meeting of the Portuguese Carbohydrate Group – GLUPOR 13 at i3S, Porto (3-5 September 2019).

J Poças, **C Marques**, I Faria-Ramos, C Gomes, F Pinto, M Ferreira, L Lima, A Polónia, LP Afonso, JA Ferreira, LL Santos, CA Reis, A Magalhães "Cell surface Heparan Sulfate ProteoGlycans as leading maestros of cancer cell behaviour". 13th International Meeting of the Portuguese Carbohydrate Group – GLUPOR 13 at i3S, Porto (3-5 September 2019).

J Poças, D Freitas, M Balmaña, **C Marques**, I Faria-Ramos, F Pinto, C Gomes, CA Reis, A Magalhães "Cell surface glycoconjugates as key players in gastric cancer". EMBO Practical Course: "Extracellular Vesicles: from biology to biomedical application" at Heidelberg, Germany (2-7 June 2019).

Oral presentations:

J Poças, **C Marques**, I Faria-Ramos, C Gomes, F Pinto, M Ferreira, L Lima, A Polónia, LP Afonso, JA Ferreira, LL Santos, CA Reis, A Magalhães "Cell surface heparan sulfate proteoglycans as novel key players in gastric cancer". Oral Communication at the 26th Porto Cancer Meeting at i3S (2 and 3 May 2019).

Manuscript in preparation:

C Marques, I Faria-Ramos, J Poças, H Duarte, C Gomes, CA Reis, A Magalhães "Heparan Sulfate ProteoGlycans biosynthetic pathways and functional implications in gastric cancer" (in preparation).

Resumo

A glicosilação tem um papel determinante tanto na fisiologia celular, como em processos patológicos. Uma característica comum a diferentes patologias malignas diz respeito, precisamente, à expressão aberrante de proteoglicanos principais da matriz extracelular e da superfície celular, como é o caso dos Proteoglicanos de Heparano Sulfato.

A capacidade de ligação dos Proteoglicanos de Heparano Sulfato a ligandos biologicamente ativos (fatores de crescimento, quimiocinas, citocinas, proteínas de matriz extracelular e recetores transmembranares) está intrinsecamente relacionada com o seu papel na modulação de eventos de sinalização celular e, consequentemente, nos mecanismos de regulação inerentes ao crescimento celular, proliferação, invasão e adesão, formação de gradientes extracelulares e tráfico de e para a membrana. Estas características conferem-lhes capacidade de controlar variados eventos celulares malignos, promovendo a aquisição de características carcinogénicas, incluindo aumento da proliferação, crescimento de tumores, invasão, metástase e angiogénese.

Neste projeto abordámos os mecanismos de biossíntese de Heparano Sulfato e implicações destes mesmos mecanismos na sinalização e comportamento das células de cancro gástrico, recorrendo a linhas celulares modificadas pela metodologia de CRISPR-Cas9. Avaliámos o papel de duas glicosiltransferases, a glicosiltransferase tipo exostina 2 e 3 (EXTL2 e EXTL3) na via de biossíntese do Heparano Sulfato. Através da caracterização bioquímica destes modelos, demonstramos, pela primeira vez em modelos de cancro gástrico, o papel da EXTL3 em iniciar a síntese de cadeias de Heparano Sulfato. Para além disto, revelámos a EXTL2 como uma enzima que regula negativamente a biossíntese de HS.

Adicionalmente, estabelecemos um ensaio de Biotinilação e Isolamento do Proteoma de Superfície Celular e demonstramos o impacto da modulação de HS na expressão do Recetor de Fatores de Crescimento Epidermal (EGFR) a nível da superfície celular, sendo estes Recetores de Tirosina Cinase frequentemente sobreexpressos em tumores gástricos.

Foram também avaliadas as implicações destas alterações no comportamento de células tumorais gástricas. Realizámos ensaios de migração para determinar a capacidade migratória das células em contacto com diferentes componentes da matriz extracelular, tendo sido utilizadas lâminas revestidas ou com fibronectina ou com colagénio IV. Também realizamos ensaios de invasão de modo a estudar a capacidade das células degradarem uma matriz de hidrogel enriquecida em componentes de matriz extracelular e inferir sobre

o fenótipo invasivo das células. Estes dois ensaios revelaram o papel determinante das cadeias de glicosaminoglicano de Heparano Sulfato na promoção de migração e invasão de células tumorais gástricas.

No cômputo geral, os nossos resultados demonstraram os papeis da EXTL2 e EXTL3 na regulação dos mecanismos de biossíntese de glicosaminoglicanos. Para além disto, revelaram também a importância das cadeias de Heparano Sulfato na regulação de Recetores Tirosina Cinase de superfície celular, bem como nas características metastáticas das células tumorais gástricas. Estes resultados evidenciam o potencial clínico da maquinaria biossintética de Heparano Sulfato no desenvolvimento de novas ferramentas relevantes para o diagnóstico precoce e tratamento eficiente do cancro gástrico, destacando-se a necessidade de continuar a explorar as vias de biossíntese dos glicosaminoglicanos, particularmente o impacto das glicosiltransferases na estrutura dos glicanos, juntamente com sua desregulação nas patologias de cancro.

Palavras-chave: Glicosilação, Proteoglicanos, Glicosaminoglicanos, Glicosiltransferase tipo exostina 2, Glicosiltransferase tipo exostina 3, Heparano Sulfato, Cancro gástrico.

Abstract

Glycosylation is known to have an essential role in cellular physiology as well as in pathological functions. A common feature amongst several malignant pathologies concerns the aberrant expression of major extracellular matrix (ECM) and cell surface proteoglycans, such as Heparan Sulfate ProteoGlycans (HSPGs).

HSPGs ability to bind to biological active ligands (growth factors, chemokines, cytokines, ECM proteins and transmembrane receptors) is intrinsically related to their main role in signaling modulation events and, consequently, in regulatory mechanisms related to cell growth, proliferation, invasion and adhesion, formation of extracellular gradients and membrane trafficking. These features allow them to take over a range of malignant cell events, promoting the acquisition of carcinogenic hallmark capabilities, including increased proliferation, tumor growth, invasion and metastasis, and angiogenesis.

In the present work, we addressed HS biosynthesis mechanisms and its implications in gastric cancer cell signaling and behavior, by resorting to CRISPR-Cas9 glycoengineered cell models. We have evaluated the roles of two glycosyltransferases, Exostosin Like Glycosyltransferase 2 and 3 (EXTL2 and EXTL3) in dictating the HS biosynthetic pathway. Upon biochemical characterization of these models, we demonstrated, for the first time in gastric models, the role of EXTL3 in initiating the synthesis of HS chains, and furthermore revealed EXTL2 as a negative regulator of HS biosynthesis.

In addition, we performed a Cell Surface Proteome Biotinylation and Isolation assay and demonstrated the impact of HS modulation in the cell surface expression of Epidermal Growth Factor Receptor (EGFR), a major Receptor Tyrosine Kinase (RTK) commonly overexpressed in gastric malignancies.

We further assessed the implications of these major alterations in tumor cells behavior. Wound Healing Migration assays were performed to evaluate cells migration capabilities when exposed to different ECM components, which was achieved by resorting to Fibronectin and Collagen IV coated slides. We have also conducted Matrigel Invasion assays to study cells ability to degrade an ECM-based hydrogel and to infer about cells invasive phenotype. Both these assays revealed the determinant role of HS GAG chains in the promotion of gastric tumor cells migration and invasion.

Overall, our results demonstrate the roles of EXTL2 and EXTL3 in regulating glycosaminoglycans (GAG) biosynthetic mechanisms, as well as the relevance of HS GAG chains to fine-tune of cell surface RTKs and metastatic features of gastric tumor cells.

Moreover, these results evidence the clinical potential of HS biosynthetic machinery for the development of new relevant tools for early diagnosis and efficient treatment of gastric cancer malignancies, highlighting the need to further explore GAGs biosynthetic pathways, particularly the impact of glycosyltransferases in the structure of these glycans, along with its deregulation in cancer pathologies.

Keywords: Glycosylation, Proteoglycans, Glycosaminoglycans, Exostosin Like 2 glycosyltransferase, Exostosin Like 3 glycosyltransferase, Heparan Sulfate, Gastric Cancer.

Table of Contents

List of Figures

List of Abbreviations

Asn – Asparagine
BSA – Bovine Serum Albumin
CA 19-9 – Cancer Related antigen 19-9
CA 72-4 – Cancer Related antigen 72-4
CEA – Carcinoembriogenic antigen
CHO – Chinese Hamster Ovary
CRISPR – Clustered Regularly Interspaced Short Palindromic Repeat DNA Sequences
CS – Chondroitin Sulfate
DS – Dermatan Sulfate
ECM – Extracellular Matrix
EGF – Epidermal Growth Factor
EGFR – Epidermal Growth Factor Receptor
EXT – Exostosin
EXTL2 – Exostosin Like 2
FGF2 – basic Fibroblast Growth Factor
FGFR – Fibroblast Growth Factor Receptor
Fuc – Fucose
FUT – Fucose Transporter
GAG – Glycosaminoglycan
Gal – Galactose
GalNAc – N-Acetylgalactosamine
GalNacT – N-Acetylgalactosaminyl Transferase
Glc – Glucose
GlcA – Glucuronic Acid
GlcATI – Glucuronyltransferase I
GlcNAc – N-Acetylglucosamine
GnT-V – β 1,6-N-acetylglucosaminyltransferase V
GPC – Glypican
GPI – Glycosylphosphatidylinositol
HA – Hyaluronic Acid
HB-EGF – Heparin-binding EGF-like Growth Factor
HER2 – Human Epidermal Growth Factor Receptor 2
HGF – Hepatocyte Growth Factor
HS – Heparan Sulfate
HSPG – Heparan Sulfate Proteoglycan
HS2ST – Heparan Sulfate 2-*O*-sulfotransferases
HS3ST – Heparan Sulfate 3-*O*-sulfotransferases
HS6ST – Heparan Sulfate 6-*O*-sulfotransferases
IdoA – Iduronic Acid
KS – Keratan Sulfate
Man – Mannose
MAPK – Mitogen-Activated Protein Kinase
NDST – N-deacetylase/N-sulfotransferases
PBS – Phosphate Buffered Saline
PD-1 – Programmed cell death protein 1
PDGF – Platelet-Derived Growth Factors

PD-L1 – Programmed death-ligand 1
RNA – Ribonucleic Acid
RTK – Receptor Tyrosine Kinase
SA – Sialic Acid
SDC – Syndecan
SDS-PAGE – Sulfate Polyacrylamide Gel Electrophoresis
Ser – Serine
sgRNA – single guide Ribonucleic Acid
SLeA – Sialyl Lewis^{a}
SLeX – Sialyl Lewis^{X}
T-antigen – Thomsen-Friedenreich antigen
TGF – Transforming Growth Factor
Thr – Threonine
UDP – Uridine Diphosphate
VEGF – Vascular Endothelial Growth Factors
VEGFR2 – Vascular Endothelial Growth Factor Receptor 2
Xyl – Xylose
XYLP – 2-Phosphoxylose Phosphatase
XYLT – Xylosyltransferase

1. Introduction

1.1. Gastric Cancer

Gastric cancer represents a serious global health issue as the fifth most frequent cancer (excluding non-melanoma skin cancer), with a crushing number of 1 033 701 estimated new cases in 2018, and as the second most common cause of death by cancer worldwide. In Portugal it follows a similar pattern, ranking as the sixth most common cancer and the third most common cause of death by cancer (1).

This pathology shows a huge divergence in terms of incidence amongst genders. In 2018 was estimated that approximately 1 in each 24 women and 1 in each 14 men were newly diagnosed with stomach cancer, and according to International Agency for Research on Cancer data it is estimated that these numbers will grow even further in the next 20 years, changing to 1 in each 15 women (68.3% growth) and 1 in each 8 men (71.1% growth) (2).

Gastric carcinoma is mainly defined as a malignant epithelial tumor of the stomach mucosa with glandular differentiation (3). Nevertheless, it is a very heterogeneous disease that can take different histological features and genotypes, whose initiation and progression can be triggered by multifactorial genetic, epigenetic and environmental events (4).

1.1.1. Classification

In order to classify the different forms of the gastric cancer pathology, different systems have been proposed over the years.

Gastric tumors can be classified according to their anatomic site, which distinguishes between “true” gastric cancer (non-cardia), in which the tumor epicenter is in the mid and distal stomach, and gastro-esophageal-junction cancers (cardia), in which the tumor epicenter is in the upper part of the stomach contiguous to the esophagus. This differentiation is extremely relevant since cardia and non-cardia subtypes are associated with different demographic trends, causes, clinical characteristics and treatment (5, 6).

From the histological perspective, gastric adenocarcinoma is the most common type, comprising approximately 90-95% of the malignant cases. Mesenchymal,

lymphoproliferative and neuroendocrine tumors make up the remaining 5-10% less common gastric cancer variants (4, 7).

Gastric adenocarcinomas can be further classified into intestinal, diffuse, mixed and unclassified types according to Laurén classification, which was proposed by Laurén in 1964 and is now one of the most widely used in diagnosis, clinical trials and in gastric cancer research (8).

Intestinal carcinomas arise from a process of dedifferentiation of the intestinal epithelium (intestinal metaplasia), forming glandular-like structures that range from well differentiated, containing well-matured goblet cells and expressing enzymes and mucins usually expressed in intestinal epithelium, to poorly differentiated, where tumors assume a more anaplastic profile. Atrophic gastritis and intestinal metaplasia are lesions that are commonly associated with this subtype (9).

Diffuse gastric adenocarcinomas lack a well-structured glandular differentiation or gland formation at all, being frequently observed as a mass of “mucocellular” cells, similar to epithelial cells of normal gastric foveolar epithelium, poorly cohesive, growing and infiltrating in mucosa and gastric wall (3, 9).

Differentiated forms of gastric adenocarcinoma are usually classified as the intestinal subtype while the diffuse subtype corresponds to poor differentiated tumors (9).

When the tissue appears to be equally occupied by both diffuse and intestinal subtype characteristics it is classified as a mixed carcinoma, and when it is undifferentiated to the point where it does not fit any of the descriptions above mentioned it is classified as indeterminate (3).

In histological terms it is also important to have in mind the range of differentiation morphologic features of adenocarcinomas. A well differentiated adenocarcinoma possesses well-structured metaplastic intestinal epithelium-like glands. In a poorly differentiated one we can observe highly irregular glands, whose structures are difficult to recognize, or single cells that can be either isolated or arranged in clusters with mucins or acinar structures (3).

The World Health Organization also proposed an alternative classification system that weights in the different histological features and acknowledges the predominant histological patterns in each case, subdividing gastric cancer into papillary, tubular, mucinous and signet-ring cell carcinomas (poorly cohesive) (3).

In 2014, as part of The Cancer Genome Atlas project, was also proposed a new molecular classification, seen as a valuable tool for improved morpho-molecular classification of gastric tumors and also for the development of new therapies. It considers molecular and genomic features of gastric cancer and divides gastric cancer into four different subtypes: microsatellite instable, Epstein-Barr-virus-positive, chromosomal instable and genomically stable gastric cancers (10, 11).

1.1.2. Etiology

As it was previously stated, gastric cancer is a multifactorial disease whose etiology can be impacted by both genetic and environmental factors. It is important to keep in mind that gastric mucosa epithelium cells are constantly exposed to several mutagenic agents. Reactive oxygen species, resulting from normal physiological processes or from defensive cell mechanisms against pathogens, gastroesophageal reflux disease, *Helicobacter pylori* infection, nutritional and smoking habits, genetic and socioeconomic factors, are considered some of the major risk factors for gastric cancer development (12).

For example, in terms of dietary habits, which are a modifiable risk factor, high intake of salt and smoked food had been suggested to play an important role in gastric carcinogenesis. Salt is postulated, in high doses, to damage gastric mucosa, inducing atrophic gastritis and thereby increasing the risk of gastric cancer and smoked food has a high polycyclic aromatic hydrocarbon content, whose higher intake is associated with higher gastric cancer rates (12). Low intake of fruits and vegetables was previously also associated with higher risks of developing gastric cancer. Fruit and vegetables are rich sources of vitamin C (ascorbate), folate, carotenoids, and phytochemicals that are hypothesized to act in a protective manner against gastric cancer (13, 14). However, several studies regarding some of the mentioned factors have shown to be inconsistent, and sometimes even contradictory (14, 15).

What remains evident, for the last 30 years, is the involvement of the pathogen *H. pylori* in the development of gastric cancer, to the point of it being classified as a “Group 1 human carcinogen” by the International Agency for Research on Cancer (3, 15).

H. pylori is a micro-aerophilic spiral-shaped Gram-negative bacterium known to be capable of colonize the host stomach from a very young age (16). *H. pylori* infection is considered one of the most common bacterial infection in the world, being reported to infect near to 50% of the world’s human population (17).

H. pylori causes progressive damage and structural changes in the gastric epithelium, triggering gastric inflammation. This condition is also known as gastritis and its long-term progression may result in multifocal atrophic gastritis, where it can be observed loss of gastric glands and the appearance of fibrotic tissue. This disease can progress to intestinal metaplasia, with the gastric epithelium being replaced by intestinal cells, and it latter can lead to dysplasia, a malignant pre-neoplastic lesion characterized by abnormal nuclear morphology and abnormal tissue architecture (13, 18).

This cascade of events frequently leads to the development of malignant tumors, that in general, show an intestinal subtype phenotype, whereas the diffuse subtype tends to be more often associated with genetic mutations (19).

The induction of gastric cancer by *H. pylori* infection is dependent on stomach colonization by highly virulent pathogen strains, with higher invasion and colonizing abilities. These include cag pathogenicity island positive and vacuolating cytotoxin A positive strains. However it is also dependent on the hosts genetic susceptibility, concerning more specifically its inflammatory response to the infection (15).

1.1.3. Diagnosis and treatment

With the increasing knowledge concerning molecular basis of gastric carcinogenesis, there has been increasing interest in studying and developing several target agents for gastric cancer diagnosis and treatment. However there are limitations in developing and applying new molecular-targeted therapies due to the intricate molecular and genetic events associated with this carcinogenesis process (20).

Currently, there are three approved targeted therapies for gastric malignancies, trastuzumab, ramucirumab and pembrolizumab, targeting Human epidermal growth factor receptor 2 (HER2/ c-erbB2 / ERBB2), Vascular Endothelial Growth Factor Receptor 2 (VEGFR2) and Programmed cell death-ligand 1 (PD-L1), respectively (21).

HER2 was the first molecular biomarker available for gastric cancer patients in clinical practice (22). HER2 belongs to the epidermal growth factor (EGF) Receptor Tyrosine kinase (RTK) family and, unlike the other members of the EGF family, its kinase activity is independent of the ligand binding, requiring instead the formation of heterodimers with HER3 or HER1, driving proliferative, survival, invasive, and metabolic functions (23, 24).

Amplification of the HER2 gene and overexpression of its product in gastric cancer patients has been correlated with more aggressive phenotypes and higher frequencies of tumor recurrence (25). One of the most efficient and most frequently used targeted agent in clinics is trastuzumab, the first agent approved in the first-line treatment of advanced gastric cancer patients overexpressing this proto-oncogene (22). Trastuzumab inhibits HER2-mediated signaling and it was shown that in combination with drugs commonly used in chemotherapy, such as capecitabine or 5-FU and cisplatin, it improves tumor response (20, 26).

VEGFR 2 is a transmembrane RTK, often expressed in tumor cells, involved in the process of angiogenesis, which is known to be a major hallmark of cancer. This receptor is activated by binding of Vascular Endothelial Growth Factor A (VEGFA), a soluble factor secreted by cancer cells in the state of hypoxia. As a consequence, it induces proliferation, migration and survival of neighboring endothelial cells, leading to the formation of tumor-associated blood vessels, contributing to tumor growth (21). Over the last years the expression of this receptor has been correlated with poorer prognosis in gastric cancer (27) and a recombinant monoclonal antibody, Ramucirumab, was already approved to be used in the treatment of patients with advanced progressing gastric adenocarcinoma (28).

Programmed cell death protein 1 (PD-1) is a receptor expressed at the surface of T and B cells, and when bound to one of its main targeted ligands, PD-L1, expressed at the surface of either tumor cells or antigen-presenting cells in several malignancies, suppresses T cell-mediated immune responses (29). This PD-1/PDL-1 is often pointed out as a main rout of immune evasion by tumor cells, contributing to tumor progression and poor patient outcomes in several malignancies, including gastric cancer (30).

PD-L1 expression was demonstrated to be positively positive correlated with gastric malignancy clinical features, such as tumor size, lymph node metastasis, invasion, and overall patents survival time (31-33). Taking this into consideration, several lines of research are being conducted to develop anti-tumor therapeutic agents that specifically target this PD-L1/PD-1 path. Pembrolizumab and Nivolumab are some of the PD-1/PD-L blocking agents tested in clinical trials as possible therapies for advanced gastric cancer, as alternatives to conventional therapy (34, 35).

Regardless of the targeted therapies that emerge, surgery has always played a crucial role in the treatment of this pathology, with gastrectomy procedures showing a strong realistic chance for the cure of advanced gastric cancer. These types of surgeries are commonly performed to remove the cancer and part or all of the stomach and some nearby lymph nodes (lymphadenectomy), depending on the type and stage of stomach cancer (36).

1.2. Cancer Microenvironment and Extracellular Matrix

The extracellular matrix (ECM) composes an extremely well-organized macromolecular network that provides a three-dimensional scaffold onto which the cells are embedded. Its precise composition and structure depend on the tissue where these matrices are established, however their main common constituents include collagens, laminins, fibronectin, glycoproteins and proteoglycans/glycosaminoglycans (37).

Collagen is a major component of the ECM, there are at least 28 types of collagens, each type being composed by polypeptide alpha chains homotrimers and heterotrimers, folded into triple helices (38). This triple helix motif potentiates collagen capacity of interacting with cell surface receptors, other proteins, glycosaminoglycans (GAGs) and nucleic acids (39).

In addition to providing binding sites for other macromolecules and cells, collagen is also crucial for the physical integrity of ECM, influencing the tissue architecture, shape and organization (40).

Proteoglycans constitute another abundant and important structural and functional ECM building block. These biomacromolecules consist of a core protein attached to long linear heteropolysaccharide chains, called GAG chains. These GAG chains are composed of varied and modified repeating disaccharides unites, assuming an immense variety of carbohydrate structures that can be distributed along the core protein in different manners, giving rise to different types of proteoglycans (37, 39).

These GAG chains have the capacity of interacting with several other ECM molecules and cells, influencing both ECM structure and cell functional properties, such as signaling, adhesion, proliferation, migration, apoptosis and differentiation events, which confers them high relevance in normal cells function and in pathologies. In fact, their biosynthesis was demonstrated to be altered during ECM remodeling in disease events (37, 39).

Another important constituent at ECM are laminins. Laminins compose a family of 20 glycoproteins, assembled in a cross-linked network together with collagen type IV in ECM basement membranes (40). They are formed by different isoforms of α, β and γ chains that can be combined to form 16 different laminin heterotrimers (39). Laminins can act as cell-adhesive molecules, binding to cells' surface receptors, such as integrins and dystroglycan, and to other components of the matrices, like nidogens, perlecans and collagens. By interacting with such biologic ligands, laminins can take leading roles in the maintenance of

the ECM structure and in cellular main events of embryonic development and organogenesis, partially through modulation of cell differentiation and migration (41).

Fibronectin is described as a “biological glue” in charge of the attachment and migration of cells (40). It is another ubiquitous ECM glycoprotein that is assembled into a fibrillar matrix in an intricate process mediated via cell interactions (42). This component is secreted as a dimer, however it is found as an insoluble multimer in the ECM. Each fibronectin subunit contains three modules of repeating units and several binding motifs responsible for the interaction of fibronectins with itself and cell surface receptors, such as integrins, collagen and gelatin (39, 41).

The different ECM constituents can be organized into two different types of structures: basement membrane and interstitial matrix.

Basement membranes are cell-adherent extracellular matrices that underlie all epithelia, sitting above the connective tissue of the organs. They are mainly composed by laminins, collagen IV, nidogens, the heparan sulfate proteoglycans (HSPGs) perlecan and agrin, and fibronectin, this latter one conferring tensile strength to the overlying cells (43, 44).

Interstitial (stromal) ECM surrounds and encapsulates most tissues. Its composing molecules, such as collagen fibrils, elastin, secreted proteoglycans, hyaluronan, and matricellular proteins, interact with each other to create a dynamic and complex three-dimensional network (37, 43).

Cells take part in the synthesis of the ECM, by controlling the spatial organization of the ECM fibril assembly upon ECM receptors-cytoskeleton interactions, and by producing and secreting matrix molecules upon specific external stimuli (45). Growth factors, cytokines and chemokines can bind to ECM molecules, depositing in this multi-molecular network. Later they can be set free and be quickly diffused in an adequate interval of time for important developmental and physiological events (37).

These biological scaffolds not only offer physical support to tissue cells but are also involved in the regulation of biochemical and biomechanical signaling, which is critical in a myriad of cellular vital processes, such as growth, migration, differentiation, survival, homeostasis and morphogenesis (46). This cell-ECM interaction occurs through cell surface receptors with high affinity for matrix macromolecules, including integrins, discoidin domain receptors and cell surface transmembrane proteoglycans, such as syndecans and CD44, allowing cells to integrate ECM signals that will dictate their fate (47).

The extreme relevance of the ECM is evident not only from its role in normal physiologic processes but also from the several pathologies that can emerge from abnormalities in ECM components, which can be set as potential targets for clinical diagnosis and treatment. For example, during tumorigenesis there are obvious alterations in the ECM, including formation of stiffer fibrotic stroma, increased deposition of macromolecular components and release of proteolytic active enzymes that induce aberrant ECM remodulation and contribute to cancer progression (37). In fact, there has been growing emerging evidence that points to ECM as an important influencer of cancer cell plasticity (48).

During cancer progression the phenomenon of epithelial-mesenchymal transition correlates with increased motility, enhanced survival and higher expression of matrix metalloproteinases, which are proteolytic enzymes capable of cleaving ECM components, that can either be linked to the cell surface membrane or be secreted and bind to integrins, heparan sulfate proteoglycans or collagen type IV, remodeling matrices. All of this contributes to increased invasiveness and metastasis of the tumoral cells. This phenomenon is potentiated by the rupture of the basement membrane, leading to the “leakage” of epithelial to the mesenchyma, followed by migration to the surrounding stroma and invasion. Ultimately these cells acquire the capacity of activate pro tumorigenic stromal cells, which in their turn can remodel the ECM, contributing to the development of a refined tumor microenvironment (48, 49) .

Among the components of the ECM, glycosaminoglycans seem to play major roles during cancer progression, hence the increasing need of studying the complex and highly regulated events related to glycosylation.

1.3. Cellular Glycosylation

Glycosylation is an abundantly occurring co- and posttranslational modification in cells (50, 51). Through this enzymatic process, in which glycosyltransferases and glycosidases assume leading roles, saccharides are covalently attached to other saccharides, lipids or proteins, through glyosidic linkages, producing a wide variety of glycoconjugates, namely glycolipids, glycoproteins and proteoglycans (52).

Vertebrates, more specifically, mammals, are known to possess a vast glycome repertoire, englobing an enormous variety of glycan structures as a result of tightly regulated synthesis mechanisms carried on by multiple enzymes in different conserved biosynthetic pathways (53).

There are ten different monosaccharides synthesized from sugar and other precursors provided by dietary intake, used as building blocks for this glycosylation reactions (fucose (Fuc), galactose (Gal), glucose (Glc), N-acetylgalactosamine (GalNAc), N-acetylglucosamine (GlcNAc), glucuronic acid (GlcA), iduronic acid (IdoA), mannose (Man), sialic acid (SA) and xylose (Xyl)), and the variety of produced glycans is dictated in the cell secretory pathways, with the glycans becoming more oligomeric and branched in the Golgi apparatus (53, 54).

Besides monosaccharide composition and branching structures, glycan diversity has its origin also in the different linkage possibilities between monosaccharides, in anomeric states, in other possible substitutions (such as sulfation) and in the linkage to their aglycone part (protein or lipid) (55).

Protein glycosylation requires interaction between nucleotide sugar transporters, glycosyltransferases and glycosidases, which operate in the endoplasmic reticulum and the Golgi apparatus, in a well-choreographed manner (56). Being important to mention that the level of protein glycosylation is intimately dependent on the frequency of glycosylation sites in the protein sequence, and the expression and activity of the glycosylation enzymes in cells and tissues (52).

Glycoproteins and proteoglycans diverge from each other by a matter of carbohydrate moiety structure and content, mostly (57).

Glycoproteins are formed by one, or more, short and branched glycan chains covalently attached to the polypeptide backbone of a certain protein. This class englobes *N*- and *O*-linked glycosides. *N*-glycans involve the carbohydrate attachment to a nitrogen atom on amine molecules of arginine or asparagine residues side chain, specifically in an Asn-X-Ser/Thr motif, whereas *O*-glycans are usually attached to the oxygen atom on hydroxyl groups of threonine, serine or tyrosine residues (53, 57, 58). There are other glycoproteins that can also be linked to the outer leaflet of the plasma membrane through a glycosylphosphatidylinositol (GPI) anchor, namely the GPI-anchored proteins (55).

Proteoglycans are composed by GAG chains also linked to specific serine and threonine residues of a core protein (except Hyaluronic acid). However, unlike glycoprotein glycan chains, GAGs constitute long linear hydrophilic and negatively charged polymers of repeating disaccharide units that can be extended to the extracellular space (57).

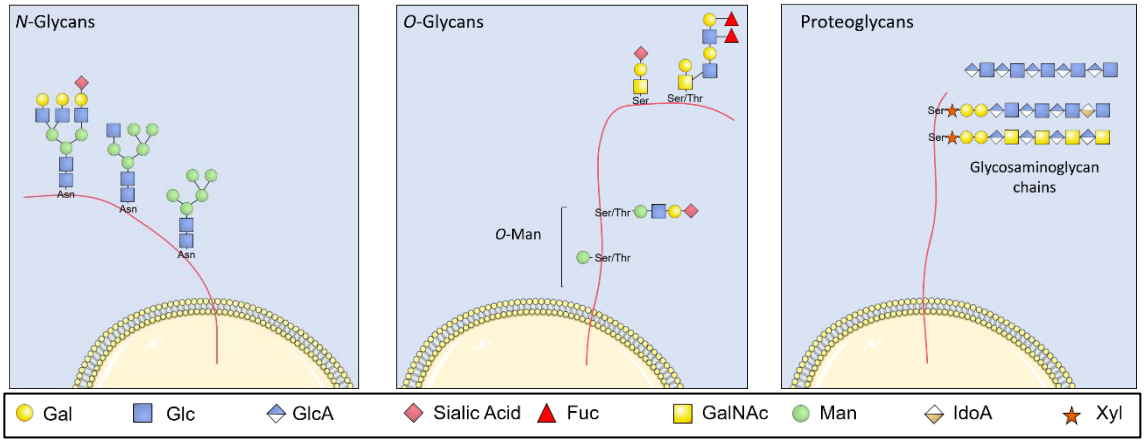


Figure 1. Glycoproteins and proteoglycans common structures. Glycoproteins comprise shorter and branched glycan chains, whereas proteoglycans are formed by longer and linear polysaccharide chains.

Other class of glycans that also compose the outer leaflet of the cell plasma membrane are the glycosphingolipids, which are ceramide-linked glycans with diverse structures and, consequently, functions. Their structures can be further decorated with terminal sialic acids (55).

All the mentioned carbohydrate moieties enhance some of the physicochemical properties of the aglycone portion they are attached to, intervening in their cellular localization, protein quality control and turnover mechanisms, and ligand interaction, amongst other (58).

More specifically, through their capacity of binding to “self” lectins, glycans promote cell-cell adhesion, and by binding to “nonself” lectins, they can recognize organisms like bacteria, yeast and invertebrates, inducing immunologic defenses activation. Glycosylation is also a key element in receptors activation and signaling modulation, having been previously demonstrated that glycans interfere in promoting ligands and their targeted receptors interaction, by binding to both, and also by interfering in the cell-surface glycoproteins endocytosis mechanisms, which, in a way, controls receptors expression (53, 59, 60).

With this said, it is easy to comprehend how mammal glycans take an active role in several molecular and cellular mechanisms intrinsically correlated with health and disease.

1.4. Aberrant Glycosylation in Cancer

Glycosylated proteins are mainly bound to cell plasma membranes, as part of the glycocalyx, or secreted as constituents of the ECM, and they are determinant in processes related to cancer development and progression, such as cell-cell adhesion, cell-matrix adhesion, cells metabolism, intercellular and intracellular signaling events, inflammation and immune surveillance (55).

In tumor cells the normal glycosylation patterns are altered, leading to aberrant glycan expression. These changes seem to go hand-in-hand with the overall carcinogenesis process, to the point of being suggested to be associated with the acquisition of hallmark capabilities, like proliferation, growth suppressors evasion, apoptosis resistance, invasion, metastasis and angiogenesis induction (51, 52).

Several factors were pointed out to be in the root of these alterations, namely: (1) changes in the expression levels of glycosyltransferases, due to dysregulated transcription and/or chaperon function; (2) altered glycosidase activity; (3) abnormal localization of glycosyltransferases and/or glycosidases in the Golgi apparatus; (4) alterations in the peptide backbone tertiary conformation and in the nascent glycan chain; (5) variability in terms of acceptor substrates and abundance of the sugar nucleotide donors and cofactors (55).

Amongst the most frequently occurring glycosylation modifications in cancer cells transformation are: sialylation, fucosylation, *O*-glycan truncation and *N*-linked glycan branching (55).

Sialylation of carbohydrates, catalyzed by sialyltransferases, affects the half-lives of circulating glycoproteins and plays an important role in cell communication and interaction abilities, cell adhesion and signaling events (51). Sialyl-Lewis X constitutes an example of a relevant sialylated structure, whose increased levels indicate poor survival. It interacts with selectins, which are cell adhesion molecules, and promotes the adhesion of cancer cells to the endothelium, therefore contributing to the development of metastasis (55).

Fucosylation relates to the addition of Fuc residues to glycans, catalyzed by several fucosyltransferases (FUT), it contributes to altered cell adhesion and impacts tumor growth rates (51). It can occur either at the terminal end of the oligosaccharide, giving rise, for example, to Sialyl-Lewis X (SLeX) antigen, or at the core level, where there is addition of α1,6-Fuc to the innermost GlcNAc residue of *N*-glycans, through FUT8 activity (55).

Alterations in sialylation or fucosylation in several different types of malignancies correlate, in a general manner, with tumor burden, invasion and metastasis. In terms of aberrant sialylation this was already demonstrated to occur for example in bladder (61), ovarian (62) and colorectal cancers (63). Altered fucosylation was also detected in colorectal cancer (64) as well as in gastric (65) and breast cancer (66), correlating with more aggressive phenotypes and poor outcomes.

The overexpression of truncated *O*-glycans constitutes another relevant glycocalyx feature observed in cancer cells. *O*-GalNAc glycans, or mucin type *O*-glycans, are the main target for this aberrant glycosylation. They are usually bound to membranes or in secreted glycoproteins and their biosynthesis is initiated by polypeptide GalNAc transferases (GalNAcTs) activity, whose expression had shown to be often altered in cancer, leading to aberrant/incomplete *O*-glycosylation. This leads to the abnormal increase in the expression of Thomsen-Friedenreich antigen (T-antigen) and the monosaccharide GalNAc (Tn antigen), as well as their sialylated forms, whose levels are usually low or absent in normal healthy tissues. Therefore, these *O*-truncated glycans are proposed to have high potential as biomarkers for cancer prognosis, having been usually associated with poor cancer patients' outcome (55, 56, 67).

The carcinogenesis process is also frequently accompanied by an increased activity of the GlcNAc transferases β 1,6-N-acetylglucosaminyltransferase V (GnT-V), leading to an increased expression of complex β1, 6-branched *N*-linked glycans. These can be further altered, elongated and ultimately terminated with SA and Fuc residues, increasing their capacity of binding to galectins, and therefore increasing cell motility, contributing to cancer cells metastasis (55, 68).

1.4.1. Alterations of Glycosylation in Gastric Cancer

In gastric cancer specifically, the glycosylation of cells from the stomach surface has been shown to suffer major alterations in the carcinogenesis course, in fact aberrant epithelial glycosylation can be detected either in pre-malignant lesions as well as in initial stages of gastric carcinoma (56, 69).

Tn and sialylated Tn antigens, which are rarely expressed in normal adult cells, were shown to be exceedingly expressed in gastric malignancies, at the cell surface level, having been demonstrated a correlation between increased sialylated Tn antigen expression and increased wall and lymphatic invasiveness of the tumor (70). More aggressive phenotypes,

and poor patients' prognosis, were also associated with increased sialylated Tn antigen, which is hypothesized to induce single cell detachment from the primary tumor and to increase cell-ECM adhesion, migration and invasiveness (71, 72).

SLeX and SLeA tetrasaccharides were also associated to gastric cancer. It was reported that upon infection by specific pathogenic *H. pylori* strains there is an increased expression of β1, 3-GlcNAc transferase, which is required for the synthesis of both sialylated antigens, and by being specific ligands for E- and P-selectins in endothelial cells, they promote the adhesion of tumor cells to the endothelium, increasing metastasis (56, 69, 73).

Another feature commonly detected in gastric cancer cells concerns the overexpression of GnT-V and consequent E-cadherin mislocalization and dysregulation. E-cadherin is described as a transmembrane glycoprotein that takes part in cell-cell adhesion. It presents four potential *N*-glycosylation sites and upon GnT-V overexpression occurs an abnormal addition of β 1, 6-GlcNAc-branched *N*-glycans to E-cadherin, which leads to its wrong assembly and dysfunctional activity, compromising cell cohesion and downstream signaling pathways, increasing tumor invasiveness and metastasis capacities (55, 74-77).

1.5. Glycans as Biomarkers in Cancer

It is important to consider the importance that aberrant glycosylation potentially holds for the development of new non-invasive detection and prognosis tools, and for the targeted treatment of cancer in general, and gastric cancer in particular.

Nowadays there are several biomarkers widely used in patients for diverse cancer pathologies. In short, biomarkers stand for objectively measures and quantifiable characteristics of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention (78). The main goal of using a certain biomarker is, through a simple blood test, or through analyses of other body fluids, to determine the risk of developing a certain disease, to perform early diagnosis and quick staging of a malignancy or to monitor disease progression (79).

In the past years, the increasing technologic progression in the fields of glycomics and proteomics led to the realization that the changes in the extent and nature of protein glycosylation, as well as the altered levels of blood glycoproteins, detected in the cancer progression process, could be of great significance to identify new sensitive and specific cancer biomarkers (80, 81).

In gastric cancer, the most commonly used tumor biomarkers are the cancer related antigen 72-4 (CA 72-4), the carcinoembryonic antigen (CEA) and the cancer related antigen 19-9 (CA 19-9) (82).

Currently, the most widely used serum-based marker for gastric cancer is the glycoprotein antigen CA 19-9, which recognizes and binds to the cancer-related sialylated terminal structure, SLeA, in *O*-glycoproteins, such as mucins, and in glycolipids (56, 79, 81).

Preoperative serologic levels of CA 19-9 may indicate the stage of the gastric cancer disease and it appears to be a predictive factor for gastric cancer recurrence and possibly higher risk of metastasis (82, 83).

This tumor marker also seems to be useful in the context of pancreatic carcinoma, predicting its outcome and patients response to adjuvant chemotherapy (84), of bladder cancer, as prognostic factor for mortality (85), and of colorectal cancer, as a poor prognostic factor (86).

CA 72.4 serological assay detects increased expression of the sialylated *O*-glycan, STn antigen, with clinical applications in gastric cancer, but also in colorectal tumors (56). According to previous studies, the serologic levels of this marker appear to be correlated with tumor stage and presence of lymph node involvement in gastric carcinoma (87).

CEA is a glycoprotein normally expressed in the gastrointestinal tract of embryos and adults, and by being an intercellular adhesion molecule, it also promotes the aggregation of human gastrointestinal cancer cells (88). Preoperative high levels of CEA correlate with the stage of the disease, higher risk of recurrence and locoregional relapse, being considered a useful predictive factor for gastric cancer (82, 89). Some studies also revealed this biomarker as a potentially indicative of more aggressive types of colorectal cancer (86).

Unfortunately, these three biomarkers show limitations that prevent them from being used as diagnosis tumor markers in clinic, including low organ specificity and sensitivity, making them unsuitable for early and reliable gastric cancer detection (56, 82).

The need for new biomarkers, the constant development of new technologies and the increasing knowledge in the fields of glycomics and glycoproteomics, enlighten the importance of studying the biosynthesis mechanisms of glycoconjugates as well as how they are altered in disease.

The most frequently described aberrant glycosylation in gastric cancer concerns glycoproteins. Less described is the altered glycosylation in proteoglycans, even though

they are major components of the ECM, known to be altered in several cancer pathologies, contributing to cancer progression and more aggressive outcomes.

Syndecans (SDCs), which are cell surface HSPGs, are a major example of altered proteoglycans in malignancies. In previous studies, the loss of SDC1 expression, detected in colonic epithelial cells in colorectal adenocarcinomas, was associated with more aggressive phenotypes (90). On the other hand, in breast cancer, SDC2 was shown to be upregulated at the epithelial level and suggested to have a key role in maintaining an invasive phenotype (91).

Besides its deregulated expression in different cancer types, further data points towards the involvement of one member of the SDC family, namely SDC4, in gastric carcinogenesis events. This proteoglycan is expressed in low levels in the healthy gastric mucosa of noninfected individuals, however, upon infection by highly pathogenic *H. pylori* strains, known to be associated with the development of gastric malignancies, its expression in the gastric mucosa epithelium was shown to be up-regulated (92).

Having in consideration the roles of proteoglycans in cancer development and progression, this study will mainly focus in determining the functional biological implications of proteoglycans in gastric cancer, and how its altered expression relates with the cancer cells ability to invade and metastasize. Having in view its potential use as biomarker for an efficient and early detection of gastric cancer.

1.6. Proteoglycans Molecular Features and Expression in Cancer

Proteoglycans are abundant cell surface and extracellular matrix components associated not only to normal tissue functions, but also to cells responses to injury and disease. They can also exist at the intracellular level and play an essential role in mediating interactions between cells and the environment. More specifically, they regulate the distribution of the extracellular signaling molecules and modulate signaling events occurring at the cell surface, which determines cell proliferation, adhesion and motility properties. Proteoglycans also participate in endocytosis and vesicular trafficking, and when secreted, proteoglycans can also enroll in the matrix structural functions (93).

1.6.1. Glycosaminoglycans and Proteoglycans Composition

Proteoglycans GAGs are long, negatively charged, unbranched, linear polysaccharides composed by repeating disaccharides units. And according to the content of these chains, GAGs can be classified as Hyaluronan (HA), Chondroitin Sulfate (CS), Dermatan Sulfate (DS), Heparin, Heparan Sulfate (HS) and Keratan Sulfate (KS) (37, 94).

HA consist of repeating disaccharides units of GlcNAc and GlcA residues, which exceptionally are not sulfated nor covalently bound to proteins. CS GAGs consist of repeating disaccharides units of GalNAc and GlcA residues. DS represents an altered form of CS, in which GlcA residues are replaced by IdoA, and lastly KS is formed by sulfated poly-*N*-acetyllactosamine chains (94).

Heparin and HS are considered the most complex subtypes of proteoglycans. They have linear polysaccharide chains composed by glucosamine molecules, that can either be *N*-acetylated or *N*- or *O*-sulfated, and uronic acid residues that can either be GlcA or IdoA (95). Even though HS appears to be less sulfated than heparin, both PGs are predominantly distinguished based on the protein being expressed and their cellular distribution. Unlike heparin, which only occurs in mast cells populating connective tissue, HS is synthesized, expressed and secreted by all mammalian tissue cells (96).

Proteoglycans can also be classified into five different groups according to their cellular and subcellular location, overall homology and function: (1) Intracellular proteoglycans (Serglycin), (2) cell surface proteoglycans (SDC, Condroitin Sulfate Proteoglycan 4/ Neuron glia antigen-2 (CSPG4/NG2), Betaglycan/TGFβ type III receptor; Phosphacan/receptor-type protein tyrosine phosphatase β; Glypican/GPI-anchored (GPC) proteoglycans), (3) pericellular and basement membrane proteoglycans (Perlecan; Agrin; Colagens XVIII and XV), (4) extracellular proteoglycans (Hyaluronan- and lectin binding proteoglycans; Aggrecan; Versican; Neurocan and Brevican) and (5) small leucine-rich proteoglycans/SLRPs (class I, II, III, IV and V), which are abundant in ECMs (97).

Proteoglycans GAG chains vary also in terms of molecular weight and level of monosaccharides sulfation which, in turn, is intimately associated with the tissue in which they are occurring (98). It is also important to mention that polysaccharide chains composition, structure of the protein core and the distribution of the proteoglycan in the organism, all determine proteoglycans biological activities (94). Hence why deregulated expression of proteoglycans and structurally altered GAG chains seem to affect cell properties and functions, playing a key role in pathology (98).

1.6.2. Altered Proteoglycans Expression in Cancer

One common feature detected amongst several cancers is the overexpression of proteoglycans. For example, serglycans, which can occur both at the intracellular and extracellular level, are important cell-cell interaction modulators, contributing also to extravasation, colonization and growth of metastatic cells. In mammary and nasopharyngeal carcinoma patients the detection of abnormally high levels of serglycans are usually associated to poor prognosis (74).

CSPG4/NG2, located at cell surface level, operates in pericyte recruitment and vascular morphogenesis and its upregulation indicates more adverse outcomes for hepatocellular and head and neck carcinoma patients (74).

As it was described before, HS and HSPGs, such as SDCs, were shown to be majorly deregulated in a diverse series of solid tumors and hematological malignancies, mainly due to their interference in cancer initiation, progression and cell signaling, having been the target of increasing scrutiny by the scientific community in recent years.

1.7. Heparan Sulfate ProteoGlycans Molecular Features and Expression in Cancer

1.7.1. Heparan Sulfate ProteoGlycans Structure

HS GAGs are long linear polysaccharide chains. The portion directly attached to the protein is the “tetrasaccharide linker”, which is composed by a residue of Xyl, directly attached to the core protein, two of Gal residues and one GlcA residue. This region is followed by repeating disaccharide units of glucosamine and uronic acid residues. The β-D-glucosamine residues can be either N-sulfated (GlcNS) or N-acetylated (GlcNAc), both of which can suffer 6-O-sulfation (GlcNS(6S) and GlcNAc(6S)). GlcNS and GlcNS (6S) can also be further 3-O-sulfated (GlcNS(3S) and GlcNS(3,6S)). The uronic acid residues, α-L-IdoA and β-D-GlcA, can further be 2-O-sulfated (Ido(2S) and Glc(2S)) (95).

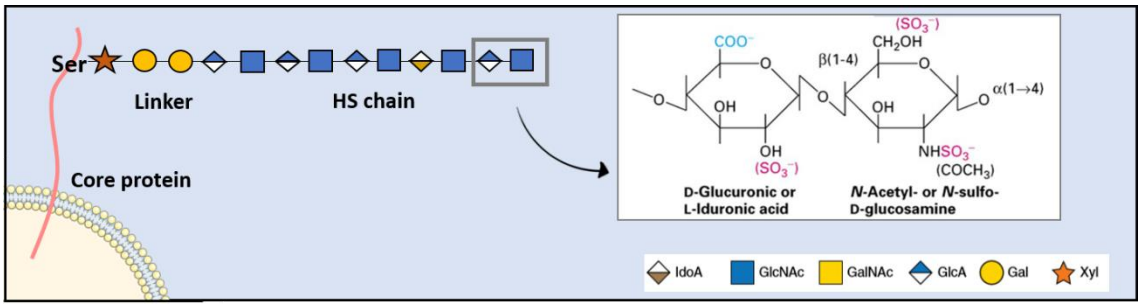


Figure 2. HSPGs structural features. HS GAG chains are divided into two regions, the "tetrasaccharide linker", and the following long linear glycosaminoglycan chain. This later is composed by repeating disaccharide units, namely glucosamine and uronic acid residues that are frequently subjected to several modification reactions.

HS chains typically consist of 50 to 250 disaccharide units (20–100 kDa) and occur naturally in all cells with an enormously varied degree of sulfation, chain length and composition, depending on its biological origin (99).

Within HS varied disaccharide sequences, the different repeating units are specifically organized and divided into 3 main domains, NA-domain, NA/NS-domain and NS-domain, starting off with repeating (GlcA-(1→4)-GlcNAc) disaccharides forming the NA domain, followed by highly sulfated units (IdoA-(1→4)-GlcNS) composing the NS domain. Between these two main domains there is an intermediate NA/NS domain, consisting of a mixture of GlcNAc and GlcNS units (95).

1.7.2. Heparan Sulfate Biosynthetic Pathways

HS biosynthesis occurs in the Endoplasmic Reticulum and Golgi apparatus and is divided into three main stages: Linker assembly, HS initiation and elongation and HS modifications.

1.7.2.1. Linker Assembly

GAG chains are attached to their core proteins through specific linkers, glucuronosyl-galactosyl-galactosyl-xylosyl tetrasaccharides, covalently bound to serine residues (GlcA-β1-3Gal-β1-3Gal-β1-4Xyl-β1-O-Ser) (94, 100).

Xylosyltransferase 1 and 2 (XYLT1 and XYLT2) initiate the synthesis and linkage of this tetrasaccharide to a core protein by using UDP-Xyl as donor. It is not defined a specific

consensus sequence in core proteins for this xylosylation, but it was observed the presence of a glycine residue in the carboxy-terminal site of the targeted serine residue and two acidic amino acids, located nearby, on one or both sites, of the serine residue (94).

This is followed by the addition of a Gal residue, by galactosyltransferase-I/β4-galactosyltransferase 7 (β4Gal-T7), and subsequent transient phosphorylation of the Xyl residue at Carbon 2. This phosphorylation reaction is catalyzed by the kinase FAM20B, and it is described as an essential event for the following reactions of assembly, since it increases the activity of galactosyltransferase-II/β3-galactosyltransferase 6 (β3Gal-T6), which will then transfer a second residue of Gal (100, 101).

The formation of this tetrasaccharide is completed with the addition of a GlcA residue by glucuronyltransferase I (GlcATI), and the simultaneous dephosphorylation of the Xyl residue by 2-Phosphoxylose Phosphatase (XYLP) (102).

Through experimental analyses, performed on a Chinese hamster ovary (CHO) cell line, it was observed that abrogation of XYLT1, XYLT2, β4Gal-T7 and β3Gal-T6 totally shut down GAG synthesis, whilst the KO of the gene coding GlcATI and FAM20B only reduce its synthesis. PXYLP KO showed no effect on the synthesis of GAG chains (100). However, in a different report, by Koike T. *et al*., it was determined that the silencing of XYLP in HeLa cells led to a significant decrease in the levels of both HS and CS GAG chains (102). This indicates that the dephosphorylation of xylose residues, tightly regulated by XYLP and GlcAT-I, might actually be indispensable for the efficient maturation of the linkage region tetrasaccharide (102).

Altogether, this shows the relative importance of each enzyme in the HS synthesis process, and how potential different mechanisms overcome possible defects in the GAG synthesis pathways in different types of cells.

The above described step is common for all PGs containing heparin/HS or CS/DS GAG chains. What determines the synthesis of a specific GAG chain, HS in this case, are the following steps.

1.7.2.2. Heparan Sulfate Initiation and Elongation

The synthesis of HS chains is initiated upon the addition of one GlcNAc residue to the linkage tetrasaccharide. For a long time, it was hypothesized that two specific glycosyltransferases, belonging to the Exostosin (EXT) family, Exostosin Like 2 (EXTL2)

and EXTL3, catalyzed this reaction, however, most recently, this was called into question (96, 103).

EXTL3 is highly efficient when catalyzing the transfer of this first GlcNAc residue to the linker tetrasaccharide, and this gives rise to posterior elongation of the HS chains. This elongation comprises a series of polymerizing reactions carried out by Golgi-located transferase enzymes, namely EXT1 and EXT2, which transfer both GlcNAc and GlcA residues (96, 100).

It was demonstrated that upon KO of either EXTL3, EXT1 or/and EXT2, CHO cells became incapable of synthesizing HS chains, indicating the value of these enzymes in this specific pathway (100).

In recent years, it was also considered an alternative HS biosynthetic pathway mainly governed by EXTL2 activity. It is though that an exacerbated formation of the phosphorylated linkage region, by increased FAM20B activity, or/and an attenuated dephosphorylation of this same linkage region, by decreased XYLP activity, leads to the accumulation of phosphorylated tetrasaccharides (4GlcUA β1–3Gal β1–3Gal β1-4Xyl(2-O-phosphate)) (Fig. 3) (102).

Structural analyses have shown that contrarily to EXTL3, which can only act on dephosphorylated linker tetrasaccharides, EXTL2 immediately transfers a GlcNAc residue to the phosphorylated linkage tetrasaccharide. This induces a premature GAG chain termination, with both EXT1 and EXT2 unable to polymerize the resulting phosphorylated pentasaccharide (GalNAc α1 - 4GlcUA β1–3Gal β1–3Gal β1-4Xyl(2-O-phosphate)) (104).

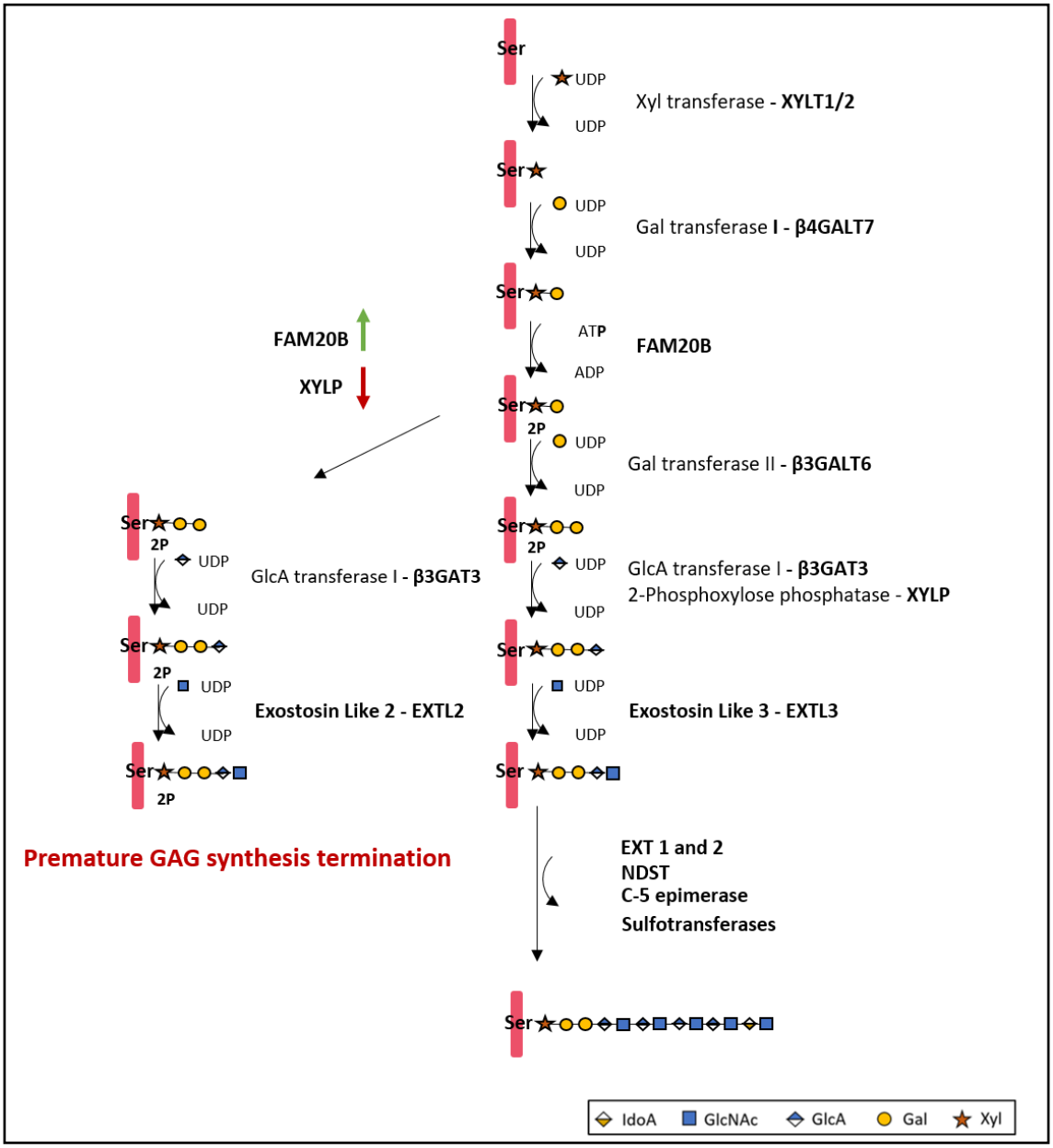


Figure 3. HS biosynthetic pathway. HS biosynthesis comprises three main events: tetrasaccharide linker formation, HS initiation and elongation, which is greatly dependent on EXTL2 and EXTL3 activity, and lastly, HS modification reactions, which underly HS GAG chain great diversity**.**

1.7.2.3. Heparan Sulfate modification

After being assembled, the copolymer of GlcNAcα4GlcAβ4 suffers a series of modification reactions catalyzed by several different enzymes.

In order to these reactions to occur, it is required the removal of N-Acetyl groups - deacetylation - from GlcNAc units, followed by the quick addition of sulfate groups - sulfation - to the free amino groups, catalyzed by N-deacetylase/N-sulfotransferases (NDST1-4), forming $GlcNSO_3$ residues (94, 96).

It was shown that KO of the genes coding for the isoenzymes NDST1 or NDST2 had no effect in the sulfation of HS disaccharides, however when it was performed KO of both genes simultaneous, it totally impaired the addition of the sulfate groups, suggesting that these are essential for the sulfation reactions, yet they can compensate each other's activity (100).

Following NDST action, subsequent modifications occur, including GlcA C5-epimerization and *O*-sulfation reactions. C-5 epimerase epimerizes GlcA residues to IdoA, followed by 2-*O*-sulfation of some IdoA units, catalyzed by 2-*O*-sulfotransferases encoded by Hs2st1. This same enzyme also 2-*O*-sulfates some GlcA units, and the 2-*O*-sulfation of GlcA or IdoA blocks further epimerization reactions (94, 96, 100).

Next, there is the addition of sulfate groups by 6-*O*-sulfotransferases (HS6sST1-3) to specific glucosamine residues and a rarer HS modification might also occur, namely the addition of a 3-*O*-sulfate group by 3-*O*-sulfotransferases (HS3ST1-6) to sulfated sugar units and uronic acid epimers (94).

It was demonstrated that KO of the genes coding for C-5 epimerase and 2-*O*-sulfotransferase led to absence of 2-*O*-sulfated disaccharides, and KO of the gene coding for 6-O-sulfotransferases eliminated 6-*O*-sulfated HS (100).

GAG chains' length, epimerization, acetylation and sulfation patterns seem to be mainly determined by the organization and specificity of the biosynthetic enzymes, availability of precursors, and flux through the Golgi apparatus, by means of mechanisms that remain still unknown (99).

Up until now, only 17 different HS proteoglycans have been described. Most of them are either located ate the cell surface membrane, associated through a transmembrane domain (SDCs 1–4, CD44v3, neuropilin, betaglycan) or a GPI anchor (GPCs 1–6), or are secreted into the ECM (collagen XVIII, agrin, and perlecan), rarely occurring as free units (99).

These variety of structures that HS chains can assume provide them with an immense amount of possible biological activities. However HSPGs core proteins are not simply carrying these HS chains, in the sense that they can also lead very important roles in the cell (95, 99).

1.7.3. Heparan Sulfate ProteoGlycans Biological Roles

Cell surface HSPGs are regulated in a very dynamic manner. When being synthesized HS chains can acquire different structures according to the cells needs and in response to extracellular stimuli. They are then internalized and turned over very quickly, faded to a relatively short half live period (95). Usually HS biosynthesis, modification and degradation, determines cells capacity of binding to extracellular ligands (105).

HSPGs are well known for their capacity to interact with several biological active ligands, including growth factors, chemokines, cytokines, morphogens, blood coagulation factors and their inhibitors, extracellular structural proteins, such as collagens, fibronectin, and vitronectin, several enzymes involved in different biochemical pathways, single-transmembrane signaling receptors and cell adhesion proteins (98, 99).

HS chains high negative charge, which comes from sulfate and carboxyl groups, prompts HSPGs to interact, in a non-covalent ionic manner, with several proteins, more specifically, positively charged lysine and arginine residues, although nonionic interactions may also contribute to this binding (106). It is also important to mention that, in most cases, there is not a specific HS sequence required for the binding of proteins to HSPGs. Proteins appear to bind to common structural motifs composing the NS domains, with the overall charge assuming a much more important role rather than a specific HS sequence (96).

This binding capacity allows HS chains to function as signaling co-receptors and involves HSPGs in several regulation mechanisms related to cellular growth and proliferation, cell adhesion and invasion, formation of extracellular gradients, membrane trafficking, blood coagulation inhibition, angiogenesis and viral invasion, amongst others (Fig. 4) (93, 95).

HSPGs act as co-receptors for cell-cell signaling by binding to several ligands in a HS chain dependent way and promoting interactions with their target receptors, behaving sort of as a scaffold for protein–protein interactions (93). FGF (fibroblast growth factor) was the first growth factor demonstrated to be dependent on HS to bind to its receptor (FGFR) (93, 107). This interaction is essential, not only to avoid FGF degradation (108), but also, to lower the concentration of FGF required to activate FGFRs, initiate signaling and increase the response period (109). Ever since then, more molecules, including growth factors, cytokines and morphogens, were discovered to act in similar HS dependent way.

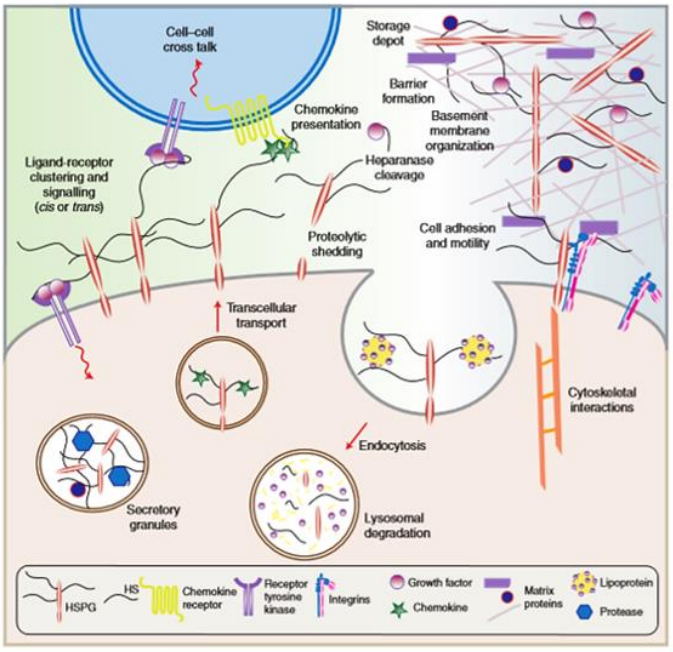


Figure 4. HSPGs major roles in cells. HSPGs bind to several biological active ligands and take over a range of important cell events related to cell growth, proliferation, invasion and adhesion, and membrane trafficking. Retrieved from Sarrazin, S., W. C. Lamanna and J. D. Esko (2011). "Heparan sulfate proteoglycans." Cold Spring Harb Perspect Biol 3(7))

It is important to highlight the Hedgehog, BMP and Wnt morphogens, whose signaling pathways are known to be impaired by the loss of EXT function (93, 110, 111). These three morphogens have major relevance in cell differentiation and tissue development, which depends on the formation of gradients of morphogens and growth factors (99). By binding to such molecules, through HS chains, HSPGs retain them on specific places at the cell surface, allowing an extracellular gradient formation. Besides influencing the distribution of extracellular signaling molecules in the ECM, interfering with their spreading through different tissues, HSPGs might also contribute to their stability and endocytosis (93).

Another example of the relevance of protein-HS chains interaction concerns the role of chemokines in inflammation. By binding to chemokines in the extracellular environment, HSPGs localize these chemotactic cytokines to the sites of inflammation where they are produced and then mediate leukocyte adhesion. Additionally, they also promote the formation of chemokine gradients that signal and guide cells migration (112).

Some HSPGs intervene in integrin signaling pathways in order to control cell adhesion and motility. Particularly SDCs are hypothesized to interfere with integrin responses, by interacting with cell surface molecules. For example, SDC1 senses matrix ligands, even at low concentrations, and stimulates αvβ3 integrin signaling pathways through its ectodomain, regulating cell attachment and spreading (93, 113).

As main constituents of the ECM, HSPGs can also interact with other basement membrane components in order to contribute to the support and resistance to mechanical stress that these matrices provide to the epithelium that lies over them (109).

Membrane HSPGs have also a major role in the uptake of macromolecules in mammalian cells. It is hypothesized that they can either interact with different biologic ligands, induce conformational changes, and present them to endocytic receptors, following predominantly lipid-raft-dependent pathways or, suffer physiological internalization for future degradation in lysosomes, and co-internalize with extracellular ligands bound to their HS chains, which will also be degraded (105, 114).

Lipid metabolism, for example, heavily relies on this type of mechanisms, being reported the pivotal role of SDC1 in the hepatic clearance of triglyceride-rich lipoproteins, by binding to triglycerides via HS chains and mediating its uptake by liver cells (115).

Furthermore, HSPGs also play a relevant role in the internalization of growth factors, whose activity may require not only interaction with target receptors, but also further internalization. SDC4, for example, was revealed to be crucial in the uptake of basic Fibroblast Growth Factor (FGF2) by the cells. It behaves as a FGF2 receptor and upon interaction with this growth factor, SDC4 oligomerizes and follows endocytosis, promoting the internalization of the total complex, SDC4-FGF2 (116).

Previous studies have also identified SDC as major key player in the biogenesis of extracellular vesicles. These small vesicles possess a similar composition to that of the cells that secrete them, mainly in terms of proteins, lipids, messenger RNAs and non-coding RNAs. They are of extreme relevance in the cells process of communication with the surrounding environment, being involved both in physiological processes, such as inflammation, coagulation, angiogenesis, and waste management, as well as in physiopathological processes, such as tumor progression, neurodegeneration and cardiovascular diseases (117).

It is described that SDCs are involved in the biogenesis pathway of these vesicles, more precisely, of the exosomes, which englobes a class of secreted vesicles characterized by their very small size. Moreover, they are also determinant in the signaling cargo of these vesicles, as SDCs HS bound receptors are also internalized to these vesicles as they are being formed (receptor trafficking) (118).

1.7.4. Heparan Sulfate ProteoGlycans Expression and Regulation in Cancer

Several HSPGs properties, such as their ability to bind proteins, implies them in regulation processes of cancer cells. In fact both HS and HSPGs can take over a range of malignant cell events, including cell transformation and proliferation, tumor growth, invasion, metastasis and angiogenesis (98).

HS is a key element in cancer cell proliferation events, intervening in altered signaling by binding to growth factor receptors, promoting their dimerization and consequent activation, leading to overstimulation of downstream signaling cascades (119). A great example of such activity relies on the overexpression of the HSPGs GPC1 and SDC1 in breast cancer, where they enhance the interaction of growth factors, such as FGF2, Heparin-binding EGF-like growth factor (HB-EGF) and Hepatocyte growth factor (HGF), with RTKs, even at low ligand concentrations, promoting cell proliferation and consequent tumor growth (120, 121). In multiple myeloma cells, SDC1 showed similar co-receptor activity, having been established its capacity of interacting with HGF via HS chains, promoting enhanced activation of Met and consequent activation of the phosphatidylinositol 3-kinase/protein kinase B and RAS-Raf mitogen-activated protein kinase (MAPK) pathways, which are intrinsically related with cell proliferation and survival (122).

Tumor progression to malignant phenotypes and following dissemination are accompanied by the pathological development of new tumor blood vessels in an event called “angiogenic switch” (123). Angiogenesis, the multi-step process through which new blood vessels are formed from previously existing ones, is fundamental, not only to supply tumor cells needs for oxygen and nutrients, but also to the cell’s invasive capabilities and tumor propagation (124). As a matter of fact, cancer cells ability to penetrate blood vessels is preponderant in their metastatic spread, being usually followed by circulation through the intravascular stream and ultimately proliferation and establishment in other sites (125).

HS GAG chains, by binding to angiogenic growth factors, namely FGFs, platelet-derived growth factors (PDGFs) and vascular endothelial growth factors (VEGFs), involve HSPGs in extremely relevant steps in angiogenesis (126). Perlecan, for example, which was described as an abnormally abundant HSPG in the basement membranes of highly metastatic human melanoma tumor cells (127), is known to promote the binding of pro-angiogenic FGF2 to its receptors, and consequently increasing angiogenesis (128).

SDC1, which is overexpressed in multiple myeloma patients’ bone marrow derived endothelial cells, constitutes another HSPG determined to promote *in vitro* and *in vivo*

angiogenesis. This is hypothesized to be correlated with its ability to physically interact with VEGFR2, preventing its intracellular recycling, which in turn, leads to increased levels of this same receptor in the plasma membrane, prolonging its interaction with its target ligand (129).

Still regarding HSPGs ability to bind to different RTKs, recently, HS features were also implicated in altered epidermal growth factor receptor (EGFR) signaling in tumor cells.

EGFR, also known as HER1 or c-erbB1, is a transmembrane glycoprotein, composed by an extracellular receptor domain, a hydrophobic transmembrane region and an intracellular domain with tyrosine kinase function (130). This receptor belongs to the ErbB family of RTKs, which also include the members erbB2/HER2, erbB3/HER3, and erbB4/HER4, all sharing similar structural features (131).

EGFR interacts and is activated by multiple ligands, such as epidermal growth factor (EGF) and the transforming growth factor α (TGF-α). Interaction with these target molecules promotes a phenomenon of homo- or hetero-dimerization of EGFR, which triggers its internalization and consequent autophosphorylation of the tyrosine kinase domains. These later ones will be responsible for binding to signal transducers and activators stimulating several intracellular downstream signaling cascades (130). Amongst these signaling cascades, the MAPK pathway and the Phosphatidyl Inositol 3' Kinase and Akt pathway should be highlighted as the main ones, as they modulate several biological processes related to cell growth, proliferation, survival, invasion and migration (131).

This receptor is reported to be overexpressed in different tumor pathologies, including non-small-cell lung cancer and gastric cancer, where it is associated to an overall poor survival (132, 133). In this context, deregulation of EGFR pathways, by either overexpression or receptor ligand-independent activation (constitutive receptor activation), governs tumorigenic events.

Cole C. L. *et al.* revealed the critical role of HS 6-*O*-sulfation levels (determined by the expression of HS6STs) on the activation of EGFR modulated by HB-EGF, and consequent increase in the expression of angiogenetic cytokines (interleukin 6, interleukin 8 and FGF2) on ovarian tumor cells (134). Furthermore, SDC1 expression in colorectal carcinomas was correlated to EGFR immunohistochemical reactivity, and even though the mechanisms underlying the interactions between this HSPG and EGFR remain to be described, it is very likely that these 2 molecular markers are related in this context, possibly contributing to tumor growth (135).

Cell surface HSPGs ability to interact with ECM elements, via HS chains, is hypothesized to affect cell adhesion and invasiveness, impacting cancer cell metastatic behavior (119). Loss of SDC1 expression, in hepatocellular (136) and colorectal (137) carcinomas, was shown to relate to weaker cell-ECM interactions, resulting in more invasive phenotypes, with higher metastatic potential. On the other hand, SDC2 overexpression in breast (91), colon (138) and pancreatic (139) tumor cells, amongst others, is associated to altered cell morphology, focal adhesion formation, spreading, enhanced migration capabilities, and overall to a more aggressive tumor cell behavior and disease progression (140).

Besides the altered expression of HSPGs, above briefly described, the abnormal expression levels of HS biosynthetic enzymatic machinery is also described as a major event behind HS deregulation in cancer (126).

GAG chains' length, epimerization, acetylation and sulfation patterns, determined by biosynthetic mechanisms, contribute to HS specificity when regulating signaling events. Comparative studies demonstrated that the expression of specific genes, coding for HS biosynthetic machinery, is deregulated in several types of cancer, when compared to normal cells, weighing in on its role on carcinogenic events (141). Major evidences of altered expression of genes involved in HS chains' polymerization, sulfation and epimerization reactions (including XYLT1, XYLT2, EXT1, NDST1, Glc Epimerase, HS2ST1, HS6ST1/3 and HS3ST6) were detected on colorectal cancer samples, these changes being also linked to metastatic features of the tumors (142, 143). Moreover, Karibe T. *et al.* showed that hypermethylation of EXTL3 promoter and consequent downregulation of this same gene induced the partial loss of HS expression in mucinous colorectal cancer lesions, which might be related to growth and proliferation advantages of these cells (144, 145).

Regarding HS biosynthetic enzymes in breast cancer, the transcription of major glycosyltransferases (namely EXT1, EXT2, EXTL2 and EXTL3) exhibited significative alterations in Estrogen Receptor positive and triple negative breast cancer cell lines (146). As for enzymes that catalyze modification reactions, these also presented deregulated expression levels on both metastatic and non-metastatic breast cancer tissue samples (147). In lung cancer, HS3ST2 hypermethylation, and consequent deregulation, was associated with lung tumorigenesis and poor overall patient survival, possibly resulting from the altered HSPGs ability to interact with a several proteins participating in cell growth and adhesion (148).

Chen *et al*. have shown that hypermethylation in HS 6-O-endosulfatase SULF1 promoter region down-regulates its expression in gastric cancer cell lines and tissue

samples (149). This enzyme removes 6-O-sulfate groups from HS GAG chains and has been determined to be a negative extracellular regulator of the Wnt signaling pathway in gastric cancer, inhibiting tumor cells growth, migration and invasion by interfering with of HSPGs' affinity for extracellular ligands. This could explain how the decreased levels of HS 6-O-endosulfatase lead to progression of gastric cancer (150). Conversely the sulfatase SULF2 was described to be overexpressed in hepatocellular carcinomas and associated to worse prognosis (151). It was hypothesized that SULF2 mediates desulfation of cell surface GPC3 HS chains that consequently loss their affinity to Wnt. This enables Wnt to bind to target Frizzled receptors, leading to activation of downstream Wnt/β-catenin signaling cascades and consequent increased cell proliferation and tumor growth (152).

1.7.5. Heparan Sulfate and Heparan Sulfate ProteoGlycans as Major Targets for Cancer Therapy

The altered expression of HSPGs and enzymes involved in the biosynthesis and posterior modifications of HS appears to have a huge role in the cancer related HS deregulation, which suggests their great potential as anti-cancer drug targets for a more personalized therapy.

Recent studies emphasized the promising roles of HS mimetics in cancer therapy, some of which are currently under clinical trials. The success of these compounds relies majorly on their capacity of slowing cancer progression, angiogenesis and overall metastatic dissemination (153). Antibodies drug conjugates have also been an emerging tool with elevated potential in cancer treatment. In recent studies, after being demonstrated that increased levels of GPC2 in neuroblastoma are associated with worse overall survival rates, researchers were able to develop a GPC2-directed antibody drug, D3-GPC2-PBD, highly cytotoxic to GPC2 expressing neuroblastoma cells (154).

As it was mentioned previously, enzymes involved in HS synthesis and following modification reactions are also considered valuable targets for therapy, a great example being sulfatases, which appear to be good drug targets. Taking as an example the altered expression levels of SULF2 in hepatocellular carcinomas and considering the limited effectiveness of chemotherapy for hepatocellular carcinoma patients, 2, 4-Disulfophenyl-N-tert-butylnitrone was proposed as a new efficient antitumor agent. This compound was already tested for human safety, and it has been demonstrated to be capable of decreasing SULF2 activity in Huh7 human liver cancer cells, inhibiting cell proliferation, viability, and migration, and to block xenograft tumors growth in nude mice (155).

In summary, the roles of HS at normal physiological environments, but most importantly in cancer microenvironment, highlight these molecules, as well as HSPG and HS-biosynthetic enzymes, as crucial intervenients in tumorigenic progression. Altogether, this shows how relevant this field of investigation is, as well as the potential it holds in clinical terms. The imminent and increasing need for new highly specific and efficient biomarkers and selective drugs prompts a more in-depth research of these promising tools, inciting further exploitation of HS main biosynthesis events and the effects of its deregulation in cancer related features.

2. General Aims

In the present work, we aim to uncover HSPGs biosynthetic pathways, more specifically the synthesis of HS GAG chains, to learn their role in tumorigenic cell signaling and functional relevance in gastric cancer.

To achieve this general aim, we focused on the following specific objectives:

1. To investigate the mechanisms behind HS biosynthesis in gastric cancer cell models, namely by addressing the impact of EXTL2 and 3, two different glycosyltransferases responsible for dictating and initiating the HS biosynthetic pathway.

2. To determine the implications of EXTL2 and EXTL3 altered expression in cancer related HS deregulation, evaluating the role of glycocalyx HS in modulating the cell surface expression of RTK (EGFR) and revealing its function in tumor cells behavior, namely migration and invasion.

To address the **1st objective**, we have selected the gastric adenocarcinoma cell line MKN74, and three clones from two CRISPR-Cas9 KO cell models, namely EXTL2 KO and EXTL3 KO cells, and conducted the following tasks:

Task 1.1. Characterize the selected gastric cell models, regarding morphological features and cell viability, by performing Phalloidin staining and Annexin V analysis, respectively.

Task 1.2. Characterize the selected gastric cell models, regarding their HS overall content, by performing Western Blot, Flow cytometry and Immunofluorescence analyses.

Task 1.3. Determine the impact of EXTL2 and EXTL3 specifically on the number of synthesized HS chains, by resorting to GAGs Heparinase Digestion.

Task 1.4. Explore the role of EXTL2 and EXTL3 in CS synthesis and sulfation profiles, by performing Western Blot analysis.

To address the **2nd objective** we have selected the gastric adenocarcinoma cell line MKN74, and one clone of each CRISPR-Cas9 KO cell models, namely EXTL2 KO and EXTL3 KO cells, and conducted the following tasks:

Task 2.1. Assess EGFR cell surface expression and activation on the HS modulated gastric cell models, by optimizing a Cell Surface Proteome Biotinylation and Isolation Assay and by performing Western Blot analysis.

Task 2.2. Evaluate the impact of HS modulation in the migration capabilities of the cells, by performing a Wound Healing Migration Assay. The influence of different ECM components (Fibronectin and Collagen IV) on the cells mobility was also considered.

Task 2.3. Evaluate the impact of HS modulation in the invasion capabilities of the cells, by performing a Matrigel Invasion Assay.

3. Materials and Methods

3.1. Antibodies

The antibodies and dilutions used for the analysis described below are listed in Table 1.

Table 1. List of antibodies and conditions used for Western Blot (WB), Immunofluorescence (IF) and Flow Cytometry (FACS) assays.

Primary Antibodies	**Applications and Dilutions**			**Source**
	WB	IF	FACS	
HS (10E4)	1:350	1:200	1:600	Amsbio
HS stub region (3G10)	1:500			Amsbio
Chondroitin 4-Sulfate (2H6)	1:300			Amsbio
Chondroitin 6-Sulfate (Mo-225)	1:300			Amsbio
EGF Receptor (D38B1)	1:1000			Cell Signalling Technology
p-EGF Receptor (D75A)	1:700			Cell Signalling Technology
Cytochrome C (A-8)	1:100			Santa Cruz Biotechnology
α-Tubulin (DM1A)	1:10000			Sigma-Aldrich
Secondary Antibodies	**Applications and Dilutions**			**Source**
	WB	IF	FACS	
Peroxidase-conjugated α rabbit IgG	1:25000			Jackson ImmunoResearch Laboratories, Inc
Peroxidase-conjugated α mouse IgG	1:5000			Jackson ImmunoResearch Laboratories, Inc
Peroxidase-conjugated α mouse IgG1	1:8000			Jackson ImmunoResearch Laboratories, Inc
Peroxidase-conjugated α mouse IgM	1:5000			Jackson ImmunoResearch Laboratories, Inc
Rabbit α-mouse Ig – FITC		1:70		DAKO
Alexa Fluor® 647 α mouse IgM			1:400	Jackson ImmunoResearch Laboratories, Inc

3.2. Cell Culture and Brightfield microscopy

In this study we resorted to the human gastric adenocarcinoma cell line MKN74, purchased from the Japanese Cancer Research Cell Bank (Tsukuba). Cells were cultured in RPMI–1640 (Gibco) culture medium, supplemented with 10% (v/v) fetal bovine serum (Biowest), at 37°C, under 5% (v/v) CO_2 conditions.

Besides the MKN74 WT cells, we also kept in culture knockouts of either EXTL2 or EXTL3 previously established in the lab through the CRISPR-Cas9 methodology (100).

Brightfield microscopy images of MKN74 WT, EXTL2 KO and EXTL3 KO cells, 70-75% confluent, were obtained by resorting to the Leica DMi1 Inverted Microscope (Wetzlar, Germany). Cells were observed with a 50x magnification.

3.3. Annexin V Viability Assay

To evaluate MKN74 WT and KO cells viability, we initially seeded $1{,}5x10^5$ cells per well on 6 well plates (Corning Incorporated Costar). After being kept in culture at 37°C under 5% (v/v) CO_2 conditions for 48h, cells were trypsinized (Biowest), resuspended on the medium they were cultured in and centrifuged at 1200 rpm for 5 minutes. The supernatant was then rejected, and the cell pellet was resuspended again in fresh medium. Cells were washed twice in PBS, followed by another two sets of washes with Annexin V Binding Buffer (BioLegend). Each wash was followed by centrifugation at 1200rpm for 5 minutes. The resulting cell pellet was resuspended and incubated with Annexin V-FITC (BioLegend) diluted in a ratio of 1:40 in Annexin V Binding Buffer for 15 minutes at room temperature. Cells were then filtered to flow cytometry tubes and data was acquired via FACscanCantoII and further analysed with the FlowJo Software.

3.4. Western Blotting

The levels of Heparan Sulfate, digested Heparan Sulfate, Chondroitin 4-Sulfate and Chondroitin 6-Sulfate and phosphorylated EGFR were measured by Western Blot (WB) analysis. To obtain the total protein lysates, ~ 100% confluent cells, cultured on $75cm^2$ flasks, were scraped and lysed with lysis buffer 17 (R&D Systems) supplemented with 1 mM sodium orthovanadate (Sigma-Aldrich), 1 mM phenylmethanesulfonyl fluoride (Sigma-Aldrich) and cOmplete™ protease inhibitor cocktail (Roche). The protein concentration of

these lysates was determined using the DC protein assay (BioRad). Specific amounts of the total protein lysates (30 µg for HS and digested HS analysis, 25 µg for C4S and C6S analysis and 15-30 µg for p-EGFR analysis) were incubated with 4x Laemmli Sample Buffer (BioRad), supplemented with 10% β-mercaptoethanol, at 98 ◦C for 5 minutes. This step leads to the denaturation of the protein content, causing the cleavage of intra and inter-chain disulfide bonds as well as the loss of secondary and tertiary conformational structures. Furthermore, it confers negative charges to the unfolded proteins, which will be crucial to their migration during sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The prepared protein samples were loaded and run in precast 4-10% precast polyacrylamide gels (BioRad) and then blotted to a nitrocellulose membrane (GE Healthcare Life Sciences).

The membranes were then blocked 1 hour at room temperature, with either 5% milk (Molico) or 5% of Bovine Serum Albumin (BSA) (Sigma-Aldrich) in 0.5% Tween® 20 (Sigma-Aldrich) in Phosphate Buffered Saline (PBS), to prevent non-specific binding of antibodies. They were then incubated overnight at 4°C with primary antibodies (Table 1), which were diluted in the blocking solution. This was followed by three rounds of washes, 10 minutes each, with PBS 0,5% tween® 20 and a new incubation for 1 h at room temperature with horseradish peroxidase-conjugated secondary antibodies, also diluted in the blocking solution. Three new washes with PBS 0,5% tween® 20 were performed, membranes were developed, and the protein bands visualized with ECL chemiluminescent western blotting detection reagent and films (GE Healthcare Life Sciences).

3.5. Flow Cytometry Assay

Cell surface levels of Heparan Sulfate were assessed through Flow Cytometry assay. First, cells previously cultured in 75cm^2 flasks, were detached with versene (Gibco), to maintain the cell membrane intact, resuspended in fresh medium and centrifuged at 1200 rpm for 5 minutes. The supernatant was then rejected, and the cell pellet was resuspended again in fresh medium. Cells (1×10^6 cells/ml) were washed 2 times in 1% BSA in PBS solution, each wash was followed by centrifugation at 1200rpm for 5 minutes. The remaining cell pellet was resuspended and incubated 1h at room temperature with the primary antibody, 10E4 (Table 1), in 1% BSA in PBS and then washed again with 1% BSA in PBS solution and incubated with APC fluorochrome-conjugate anti-mouse IgM. Lastly, cells were washed again and resuspended in 1% BSA in PBS solution and filtered to flow cytometry

tubes for analysis. Data was acquired via FACscanCantoII and further analysed with the FlowJo Software.

3.6. Fluorescence Microscopy

MKN74 cell surface and total Heparan Sulfate levels were evaluated through Immunofluorescence staining. 70-80% confluent cells seeded and grown on 12-well microscope glass slides (IBIDI) were washed twice with PBS, and then fixed on ice either for 20 minutes with 4% paraformaldehyde (Alfa Aesar) or for 10 minutes with methanol (Thermo Fisher Scientific). Cells were washed again with PBS, blocked with Rabbit Serum (Dako) diluted in a ratio of 1:5 in 10% BSA in PBS, and incubated with the primary antibody, 10E4, in 5% BSA in PBS, overnight at 4°C. This was followed by incubation with rabbit α-mouse Ig – FITC conjugated secondary antibody for 30 minutes, at room temperature. Finally, cells were incubated with 100 µg/ml DAPI (4′, 6-diamidino-2-phenylindole), a blue-fluorescent DNA stain, and diluted 1:100 in PBS and mounted in VectaShield mounting medium (Vector Laboratories).

MKN74 cells cytoskeleton was also evaluated by Phalloidin staining analysis. Cells were fixed on 4% paraformaldehyde and this was followed by an additional step of cell membrane permeabilization. In this step cells were initially incubated with NH_4Cl 50mM for 10 minutes, to quench aldehyde groups, and then washed with PBS. This was followed by incubation with 0.1% (v/v) Triton X-100 (Sigma-Aldrich) in PBS for 5 minutes at room temperature, and subsequent new three washes with PBS. After permeabilization, cells were blocked with Goat Serum (Dako) diluted in a ratio of 1:5 in 10% BSA in PBS and incubated with Alexa Fluor™ 568 Phalloidin (Thermo Fisher Scientific), diluted in 1:40 in 5% BSA in PBS, for 30 minutes at room temperature. Cells were washed again in PBS and were then incubated with DAPI, in similar conditions as the ones previously described.

Microscope images of the stained cells were obtained resorting to the Zeiss Axio Imager Z1, Axiocam MR ver3.0 and Axiovision 4.8 Software (Carl Zeiss, Germany).

3.7. Glycosaminoglycan Enzymatic Digestion Assay

Heparan Sulfate glycosaminoglycan chains digestion was performed incubating 500 µg of protein lysates of each sample with 7,8 µl Heparinase I ((Hep I) EC 4.2.2.7; 2.5 mU/ml) and 7,8 µl Heparinase III ((Hep III) EC 4.2.2.8; 2.5 mU/ml) (Sigma Aldrich) in TBS,

supplemented with Calcium Acetate (2mg/mL), overnight, at 37ºC with continuous mixing (Fig. 10). 25µg of the resulting digested samples were then analyzed through WB analysis, as the levels of resulting stub regions were labelled with 3G10 antibody and the staining was quantified by densitometry (Table 1).

3.8. Dot Blotting

Dot Blot analysis were performed for antibody validation and to evaluate the effectiveness of the glycosaminoglycan enzymatic digestion. 1 and 5 µg of both digested and non-digested protein lysates were directly spotted onto nitrocellulose membranes, which were then blocked and probed with the primary and secondary antibodies previously described for WB, following the same conditions and incubation periods (Table 1).

3.9. Cell Surface Proteome Biotinylation and Isolation

To isolate cell surface proteins, MKN74 cells ($1,5x10^5$ cells/ml) were cultured in 6-well plates and cell surface-labelling with a cell-impermeant amine-reactive biotinylation reagent was performed as follows.

Cells ~100% confluent were pre-incubated on ice for 10 minutes and maintained on ice during the subsequent steps. Cells were washed three times with ice-cold PBS^+ (supplemented with 0.1 mM $CaCl_2$, 1 mM $MgCl_2$) adjusted to pH 8 and were then incubated with 1mM Sulfo-NHS-SS-Biotin (Pierce; Thermo Scienific) in PBS^+ for 30 minutes to label cell surface proteins. Biotin molecules have high affinity for amine groups present on the lateral chains of amino acids, like lysine, moreover, their negative charge prevents them from penetrating the cells membrane.

Cells were washed again three times with ice-cold PBS^+, after which free biotin was quenched by incubating cells with 0,1M glycine in PBS^+ for 10 minutes. A new round of PBS^+ washes was performed followed by cell scrapping and lysis. Lysates were then centrifuged at 18 000g for 10 minutes at 4°C and the soluble fraction was collected.

To capture cell surface proteins, 150 µg of biotinylated protein extracts of each clone were incubated with 80 µl of streptavidin-agarose beads (Thermo Scienific), overnight at 4ºC with continuous mixing. Streptavidin is a polymer with high affinity to biotin molecules, which justifies the use of these beads in the capture process.

The non-bound/non-biotinylated proteins were first collected on the supernatant upon centrifugation at 2500g for 5min. The remaining pellets, that is, the protein-beads complexes, were washed four times with PBS containing 1% (v/v) Triton X-100, each wash intercalated with centrifugation at 2500g for 2 minutes and removal of the supernatant. To recover the bound protein from the streptavidin-agarose beads, biotin-labeled proteins were eluted using 80µl of Laemmli Buffer for immunoblotting heated at 95°C for 10 min (Fig.13). Streptavidin pull-downs (a volume of 15µl) were separated by SDS-PAGE electrophoresis and immunoblotted for EGFR and cytochrome c (used as intracellular fraction control) and quantified by densitometry (Table 1).

3.10. Wound Healing Assay

To evaluate the migration capacities of the gastric adenocarcinoma cells *in vitro*, we performed a wound-healing migration assay in which $5{,}3x10^5$ cells/ml were seeded in each side of a silicon insert, previously adhered to a well of a µ-Slide 8 Well Collagen IV or Fibronectin coated (IBIDI). These cells were maintained in culture for 24h, at 37◦C in 5% CO_2 atmosphere conditions. The inserts were then removed, the cells were washed with RPMI–1640 culture medium, supplemented with 10% FBS, and fresh supplemented medium was added to each well (Fig. 16 A). Time-lapse microscopy was performed using Leica DMI6000 (Wetzlar) and three bright field images per well/condition were obtained during 24h in intervals of 10 minutes, capturing the wound healing process. To evaluate the rate of the wound healing 13 time-points were considered (0h, 2h, 4h, 6h, 8h, 10h, 12h, 14h, 16h, 18h, 20h, 22h and 24h) and the total area of the wound in each of those time points was measured resorting to ImageJ software.

3.11. Matrigel Invasion Assay

In vitro invasion assays of gastric adenocarcinoma cells were performed resorting to a 24-well plate of BD BioCoat Matrigel™ Invasion Chambers (BD Biosciences). $2x10^5$ cells were initially seeded and incubated in RPMI–1640 medium in the upper chamber for 24h at 37°C in 5% CO_2 atmospheric conditions. The lower side of the well contained only RPMI–1640 medium supplemented with 10% FBS, used as a chemoattractant (Fig. 17 A).

The Matrigel coated chambers were then washed seven times in PBS, the non-invasive cells adhered on the inner side of the chamber were removed with a cotton swab, previously

wet with PBS, and the remaining cells in the inner and outer side of the chamber (non-invasive and invasive cells, respectively) were fixed in ice-cold methanol for 10 minutes and air dried. The chambers were once again washed with PBS, the matrigel coated membranes were removed and mounted in glass coverslips with Vectashield mounting medium with DAPI (Vector Laboratories). Microscope images of the stained nuclei were obtained resorting to the Zeiss Axio Imager Z1, Axiocam MR ver3.0 and Axiovision 4.8 Software (Carl Zeiss, Germany) and the total number of invasive nuclei was counted manually.

3.12. Statistical Analysis

Statistical analysis was carried out using GraphPad Prism 6 and statistical significance was considered when P values were <0.05 (** means p<0.01).

In Western blot analysis, band density values of the different clones were normalized to the band density values of the WT, which were defined as a unit value. Averages of the normalized results were calculated, and the results were presented as average + SD.

Migration assay results were depicted as the average values of the % of closing wound + SD. The % of closing wound was calculated by subtracting to the area of the open wound in the first time point (t=0h), the area determined in each time point, followed by normalization of the resulting values to the wound area determined for the first time point (t=0) (only the first 8 time points were considered, as the wound was closed after the first 14h in all assays). Two-way ANOVA tests from independent replicates were used to calculate the statistical significance in an interval of 95% confidence level (n≥4).

Invasion assay results were depicted as the average values of the fold changes of the nº of invasive cells + SD. The nº of invasive cells of the KO models were normalized to the greatest value of invasive cells of the WT condition, which was defined as a unit value (the number of the remaining WT invasive cells were also normalized to this higher value). Significance was calculated, in an interval of 95% confidence level, using the unpaired Student's t test with Welch's correction (n=6).

4. Results

4.1. Characterization of EXTL2 and EXTL3 CRISPR-Cas9 KO cell models

EXTL2 and EXTL3 glycosyltransferases have been described as two key regulators of the synthesis of HS GAG chains. The role of EXTL2 under pathological conditions, where GAG biosynthesis is crucial, has already been described in mammal models, such as models of liver disease (156). Furthermore, in recent studies, it was developed a cell library of genetically modified cells, resorting to CRISPR-Cas9 combinatorial knockout (KO) and knock-in (KI) methodologies, and the GAG biosynthetic capacities of these cells were profiled. This helped to reveal the role of different enzymes, including EXTL2 and EXTL3, on the synthesis of HS (100).

Therefore, and taking into consideration the previously described contributions of HS and HSPGs in several malignant features, we considered of major relevance to evaluate the role of these two glycosyltransferases, EXTL2 and EXTL3, in gastric cancer models.

We selected and maintained in culture cells from the MKN74 cell line, derived from a moderately differentiated tubular gastric adenocarcinoma, classified as intestinal type, with an epithelial-like morphology. These cells grow in a monolayer pavement-like pattern and are described to possess a high number of microvilli with filaments and an abundant glycocalyx (157).

Using the CRISPR-Cas9 methodology, two different glycoengineered cell models were previously stablished in our lab, namely MKN74 EXTL2 KO and MKN74 EXTL3 KO cell lines, using the gRNA published in (100).

CRISPR (clustered regularly interspaced short palindromic repeat DNA sequences) methodology refers to a widely used and highly efficient genome editing tool, based on a natural system used by bacteria to protect themselves from infection by viruses. Particularly CRISPR-Cas9 relies on a Cas9 protein-RNA complex, formed by the nuclease Cas9 (CRISPR-associated) that recognizes Photospacer-Adjacent Motifs (PAM) on DNA. This nuclease is usually paired up with two RNA sequences: the crRNA, corresponding to guide RNA, which contains a sequence complementary to the target sequence, and the trans-activating crRNA (tracrRNA), which is paired to crRNA, and forms the active complex. These two sequences can also be fused into a single chimeric RNA named single guide RNA (sgRNA) (158).

This complex can be reprogrammed to target specific DNA sequences to perform both *in vitro* and *in vivo* genome editing. In summary, upon recognizing and binding to the PAM sequence, Cas9 unzips the DNA strands, allowing the guide RNA to bind to its target sequence. Upon this match, Cas9 induces a DNA double strand break, and as a response, cells try to repair this break through either an error prone mechanism – Non-homology end joining – that often leads to insertion or deletion mutations that randomly disrupt the gene of interest, or via homology-directed repair, which is more precise (158).

Out of each KO cell model obtained through this methodology, were selected three representative clones to be further analyzed, and we started out by analyzing their morphologic features.

We could detect morphology variations between the WT and the different clones as there were certain features more prominent in certain KO cell lines (Fig. 5). For example, EXTL2 KO cells exhibited more frequently a large number of giant cells, with decreased nucleus to cytoplasm ratio, which was not as frequent neither in WT nor in EXTL3 KO cells. In the other hand, in EXTL3 KO cells we detected cytoplasmic filaments shared between neighboring cells, and sometimes between a specific cell, stablished amongst a large population of cells, and another cell, more distant and isolated (Fig. 5).

We cannot state that these features were exclusive to each KO cell line, nevertheless they were more commonly observed in the mentioned clones.

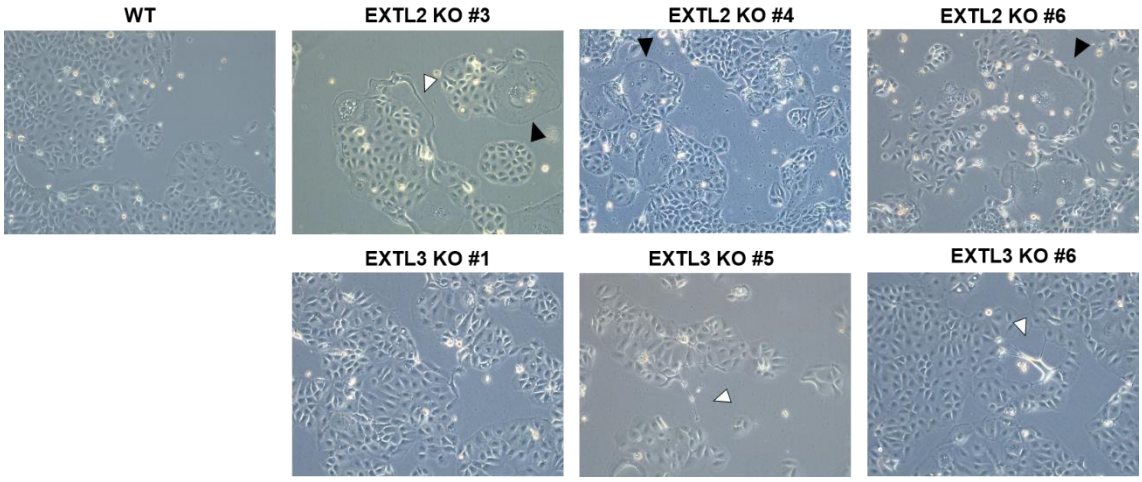


Figure 5. Brightfield microscopy of MKN74 WT and glycoengineered cell models evidenced different morphologic features. Cells were 70-75% confluent (50x magnification). Black arrows indicate giant cells, with decreased nucleus to cytoplasm ratio. White arrows indicate cytoplasmic filaments.

We further investigated the morphologic features of the EXTL2 and EXTL3 CRISPR-Cas9 KO gastric cell models, by assessing the expression of a major cytoskeletal protein, namely F-actin. Two clones of each KO cell line were selected and labelling of this target was performed in 4% paraformaldehyde (4-PFA) fixed cells. Prior to the labelling, cells were permeabilized with a non-ionic detergent, Triton X-100.

As depicted in Fig. 6, the cytoskeletal actin networks showed similar patterns in both cell models and in the WT cells. Noteworthy, the observations made via brightfield microscopy (Fig. 5) were further corroborated. EXTL2 deficient cells exhibited a larger cytoplasm area, while in the absence of EXTL3, cells presented numerous filamentous protrusions stablishing cell to cell contacts (Fig. 6).

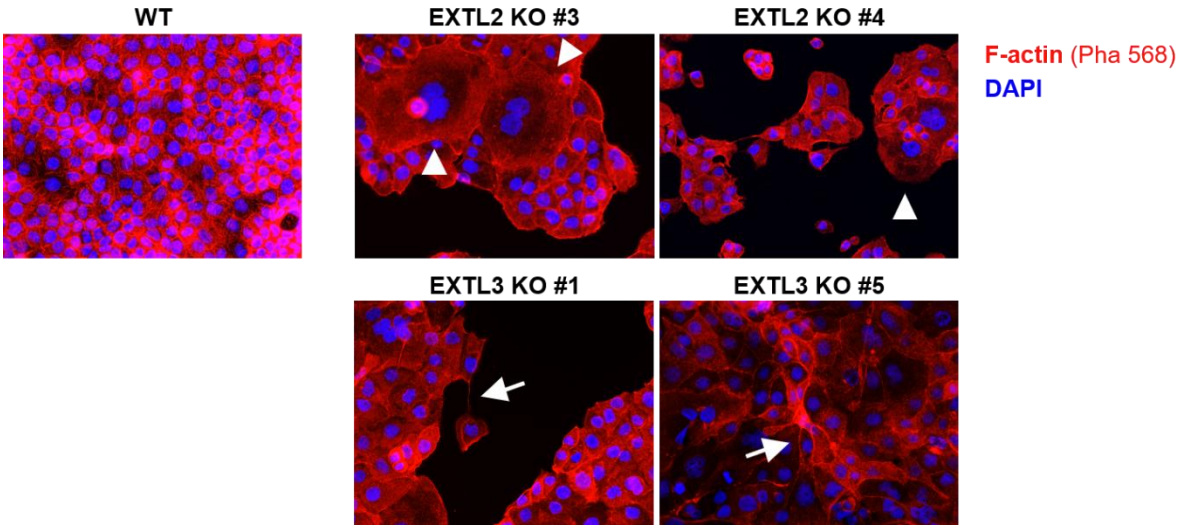


Figure 6. MKN74 EXTL2 KO and EXTL3 KO gastric cell models cytoskeletal features revealed by Phalloidin staining. 4-PFA fixed cells were permeabilized with Triton X-100 and stained for F-actin via Phalloidin 568 (Pha 568) (200x magnification). White arrowheads indicate giant cells, a common feature detected in cells lacking EXTL2. White arrows indicate cytoplasmic filamentous protrusions shared frequently between EXTL3 KO cells. Representative image of 2 independent biological experiments.

4.2. Influence of EXTL2 and EXTL3 CRISPR-Cas9 KO in cell viability

To determine the rate of apoptosis of the EXTL2 and EXTL3 CRISPR-Cas9 KO gastric cell models, we have performed a cellular viability assay, performing the 7-AAD/Annexin V-Fluorescein isothiocyanate (FITC) staining.

A common feature detected in the early steps of cellular apoptosis is the loss of plasma membrane asymmetry, which is accompanied by the translocation of the membrane

phospholipid phosphatidylserine (PS) from the inner to the outer leaflet of the plasma membrane. This event results in the exposition of this phospholipid to the external cellular environment. The Annexin V viability assay takes this major apoptotic event into consideration, as Annexin V is a protein with high binding affinity to PS. In this assay we used Annexin V conjugated with FITC, a derivative of fluorescein, as a probe for flow cytometry analysis of apoptotic cells (159).

We observed that the percentage of viable cells was, overall, higher than 70%, and as it is shown in the Dot Plots represented in Fig. 7, no major differences were detected between the WT and the different clones. This validates that CRISPR-Cas9 KO of either *EXTL2* or *EXTL3* genes did not interfere with cells mechanisms of survival.

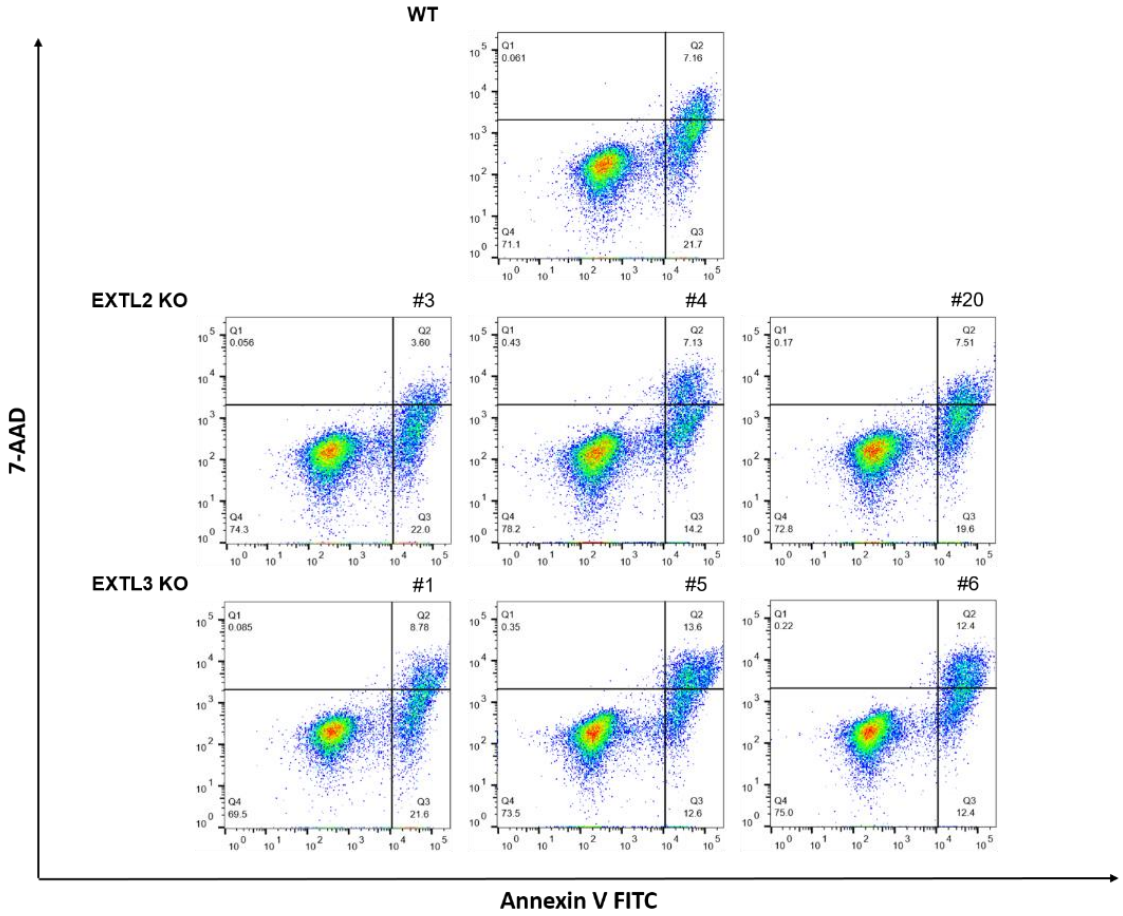


Figure 7. Cellular viability of MKN74 WT, EXTL2 and EXTL3 glycoengineered cell models. Viability of the cells was determined by 7-AAD/ Annexin V-FITC cell labelling. All KO clones presented similar viability rates in relation to the WT cells (Q1 = Cells in necrosis; Q2 = Cells in late apoptosis; Q3 = Cells in early apoptosis; Q4= Viable cells). Representative image of 3 experiments with independent biological replicates.

4.3. Evaluation of EXTL2 and EXTL3 impact in HS levels in gastric cancer cell models

We first evaluated the phenotype of our CRISPR-Cas9 glycoengineered cell models, MKN74 EXTL2 KO and EXTL3 KO, focusing on the end product of the biosynthetic pathway we were interfering with, namely HS GAG chains.

WB and Flow Cytometry analyses were performed, and when compared with the WT, EXTL2 KO cells showed evident increased levels of HS, both in terms of the total cell content (Fig. 8 A) and also at the cell surface level (Fig. 8 B). In Fig. 8 A, it stands out an intense HS signal in a variety of high molecular weights, corresponding to different lengthened HS chains attached to a variety of proteins of different molecular weights. This demonstrates that the abrogation of EXTL2 affects HS biosynthesis general mechanism, and not only a specific HSPG.

EXTL3 KO cells, on the other hand, exhibited total lack of HS labelling, which is in agreement with what is described in the literature for other models (Fig. 8 A and B) (100).

Noteworthy, even though a similar trend is observed through both technics, the augmented levels of HS chains in EXTL2 KO clones is more substantial in terms of total cell content (Fig. 8 A), and not as drastic in the cell surface (Fig. 8 B).

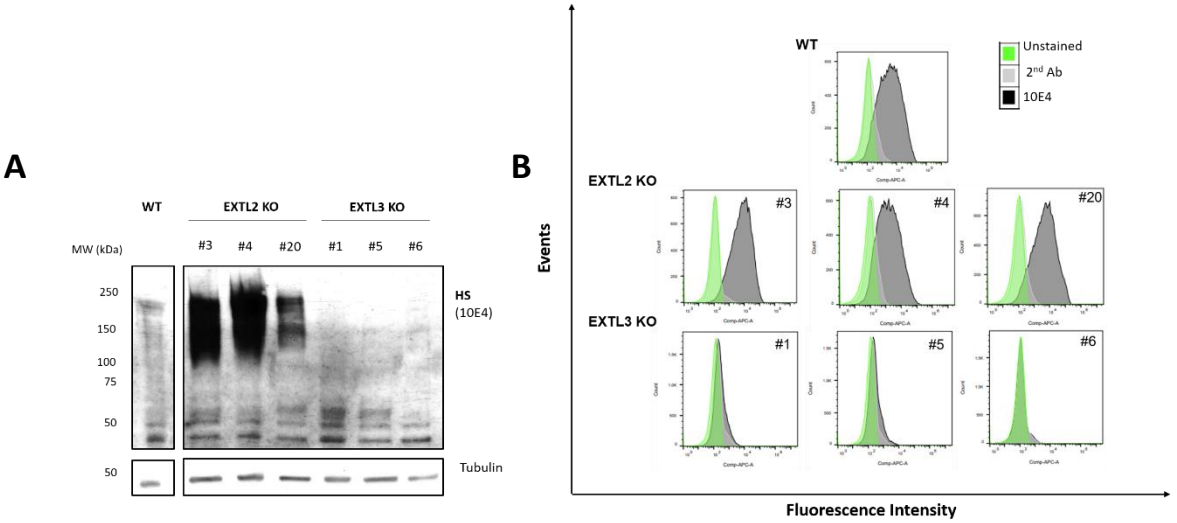


Figure 8. MKN74 EXTL2 and EXTL3 glycoengineered cell models presented altered synthesis of HS. A. Cells total content of HS was evaluated via WB analysis. EXTL2 KO cells showed increased levels of HS in comparison to the WT. In EXTL3 deficient cells was evident the absence of these GAG chains. α-Tubulin was used as loading control. Representative image of 3 experiments with independent biological replicates. **B.** Flow cytometry detection of cell surface levels of HS follows a similar trend as the one observed in terms of total cell content, though the increase of HS in EXTL2 KO clones was not as significant. Representative image of 3 experiments with independent biological replicates.

We also performed Immunofluorescence assays on cells either fixed with 4-PFA or methanol (MeOH). These further confirmed the results previously described, EXTL2 KO clones presented higher immunofluorescence staining intensity for HS chains, in comparison with the WT, while in EXTL3 deficient clones, it was evident the complete absence of these polysaccharides (Fig. 9 A and B).

Comparing the results obtained when executing the different types of cell fixation, we could observe improved HS labeling in MeOH fixed cells (Fig. 9 B). MeOH, contrarily to 4-PFA, permeabilizes the cell membrane, without further steps of permeabilization having to be implemented. This explains the higher HS signal intensity dispersed through the cells when using this specific fixative, further confirming the more drastic increase of HS levels specifically in the cytoplasm of EXTL2 KO cells.

We also must consider that PFA is an aldehyde-based fixative, which implies that some of its aldehyde groups can crosslink with target epitopes, leading to the loss of antigenicity of the target molecules. This can also explain the fainter HS signal in WT and EXTL2 KO clones fixed with 4-PFA.

Interestingly, EXTL2 KO cells did not only present higher abundance of HS scattered through the whole cell, but we also detected thin HS enriched protrusions directed towards the exterior of the cells and towards neighbouring cells (Fig. 9 B - 630x magnification). This was rarely detected on the WT and completely absent on the EXTL3 KO cells (Fig. 9 B - 630x magnification).

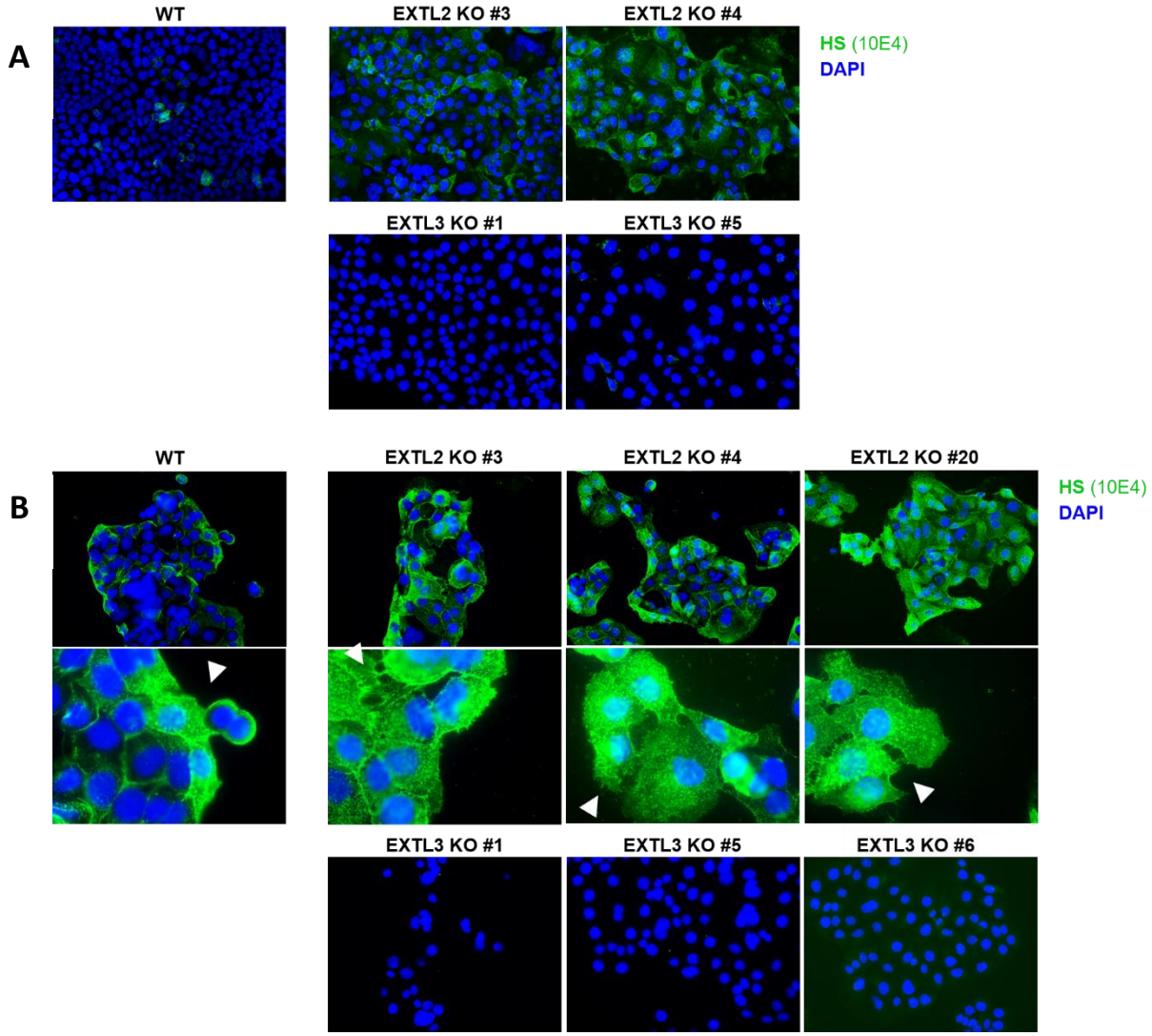


Figure 9. Immunofluorescence labeling of HS GAG chains in MKN74 WT and glycoengineered cell models. A. 4-PFA and **B**. MeOH fixed cells were immunolabelled with 10E4 and DAPI. EXTL2 KO clones showed intense labelling of HS, mainly in MeOH fixed cells. This was not observed in EXTL3 KO clones, which were HS-negative in both fixation methods (200x magnification). White arrows indicate HS enriched protrusions observed in higher ampliation pictures taken from the WT and EXTL2 KO clones (630x magnification). Representative image of 2 experiments with independent biological replicates.

4.4. Addressing the regulatory roles of EXTL2 and EXTL3 in the biosynthesis of HS chains in gastric cancer cells

We then questioned the effect that abrogating either EXTL2 or EXTL3 could take specifically on the number of synthesized HS chains. To evaluate this, we performed an enzymatic digestion assay, in which HS GAG chains were subjected to subsequent cleavage reactions catalysed by Heparinase I and III.

Heparinase I cleaves the glycosidic linkage between N-sulfated glucosamine ($GlcNSO_3$) and 2-O-sulfated iduronic acid ($IdoA,2SO_3$) residues. These residues are usually located

the farthest from the core protein of HSPGs and are quite rare in this type of GAG chains, being confined strictly in the sulfated domains (Fig. 10).

Heparinase III cleaves the glycosidic linkage between N-sulfated or N-acetylated glucosamine ($GlcNSO_3$ or GlcNAc) and GlcA residues, which are much more frequent in HS chains and are present in regions of low sulfation. This enzyme is the one cleaving the closest from the core protein of HSPGs (Fig. 10).

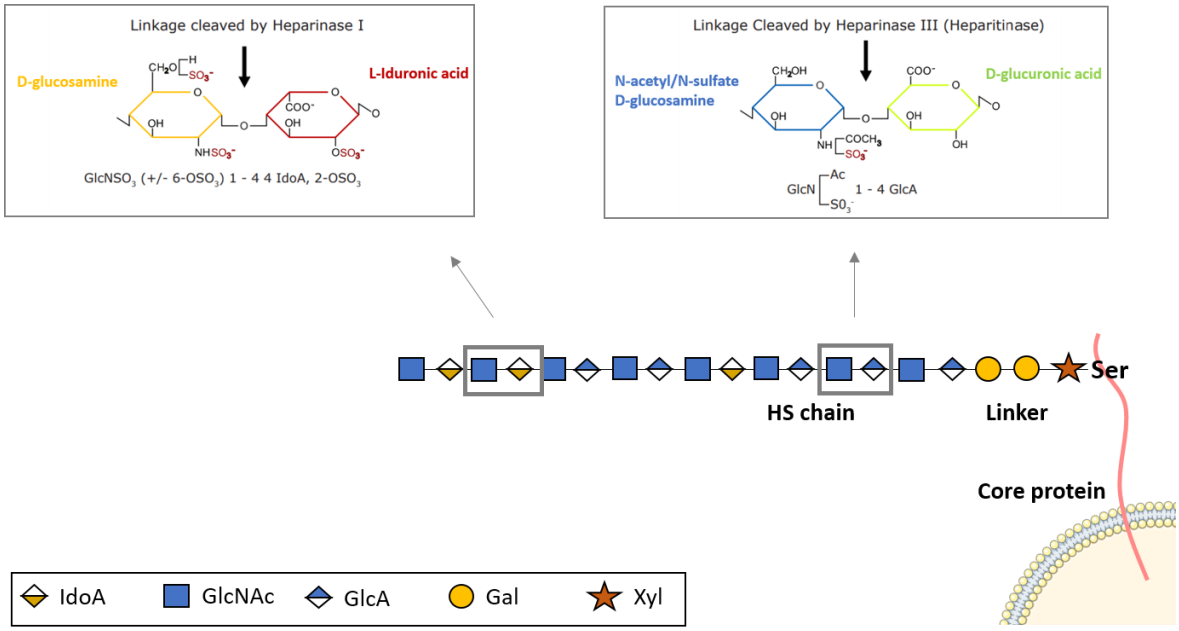


Figure 10. Schematic representation of HS GAGs enzymatic digestion by Heparinase I and Heparinase III, depending on the structural features of the sugar residues.

By performing Heparinase I and Heparinase III digestion of the HS GAGs contained in the protein lysates of our cell models, we exposed one specific unsaturated uronic acid saccharide residue at the extremity of each cleaved polysaccharide, here referred as the "HS stub region". The resulting quantity of HS stub regions was then quantified by Dot Blot and WB analysis. This experience allowed us to dismiss the length of HS chains and to directly infer about the number of HS chains present on each clone.

We first assessed the efficiency of the digestion, by performing Dot Blot analysis of both non-digested (-) and digested lysates (+). For this preliminary evaluation, we selected one clone of each KO cell model, and two different concentrations of lysates, of each selected clone, were tested.

As we can see in the Fig. 11 A, the levels of HS in the non-digested lysates followed a similar trend as the one previously detected, higher levels of labelled HS were observed in the EXTL2 KO clone, while the labelling of HS was less intense in the EXTL3 KO clone,

when compared with the WT. Upon digestion we observed a major reduction in the levels of HS in the WT, EXTL2 KO #3 and EXTL3 KO #5 lysates, as it would be expected from an efficient digestion.

We must consider that the Dot Blot methodology does not include the electrophoresis step, which means that the total protein content of the lysates is concentrated in one specific dot, giving rise to a higher number of unspecific interactions between the antibodies and the target molecules. This could explain the detected labelling for HS in the EXTL3 KO clone non-digested and digested lysates, previously absent in the WB analysis (Fig. 8 A).

Regarding the levels of HS stub regions, a faint signal was detected in all conditions of non-digested lysates. Considering that 3G10 antibody only interacts specifically with unsaturated glucuronic acid residues that result from Heparinase I and Heparinase III digestion, we attributed this to unspecific interactions (background staining).

Digested EXTL2 KO cell lysates presented increased levels of HS stub regions, further indicating a successful digestion. The levels of HS stub regions in EXTL3 KO digested lysates were almost undetectable and curiously, the same was observed for the WT (Fig. 11 A).

We then performed WB analysis to evaluate the levels of HS and HS stub regions on the non-digested and digested lysates of the WT and all clones of each KO cell line in study. As it can be seen in Fig.10 B, non-digested and digested lysates of the WT presented similar levels of HS, which was not previously noticed in the Dot Blot analysis. Meanwhile, the high levels of HS in EXTL2 KO cells were significantly decreased upon digestion, except for EXTL2 KO #20, in which this reduction was not as noticeable, similar to what was observed in the WT. In agreement, the levels of HS in the non-digested lysates of this clone, were also not as drastically increased as in the other two EXTL2 KO clones (#3 and #4) (Fig. 11 B). EXTL3 KO clones were 10E4 negative, presenting equally null levels of HS in digested and non-digested lysates, as expected (Fig. 11 B).

Regarding the levels of HS stub regions in the digested protein lysates, EXTL2 KO clones revealed intense 3G10 labelling in high molecular weights, around 150 kDa and 200 kDa, indicating a greater number of HS chains in these cells. In the digested protein lysates from the WT and EXTL3 KO clones, this was not observed (Fig. 11 C). The absence of HS stub regions in EXTL3 KO clones was expected, considering the lack of HS chains, previously determined (Fig. 8). However, the lack of HS stub regions in the WT was somehow unexpected, since these cells possess indeed HS GAG chains that should have been targeted by Heparinase I and Heparinase III (Fig. 8).

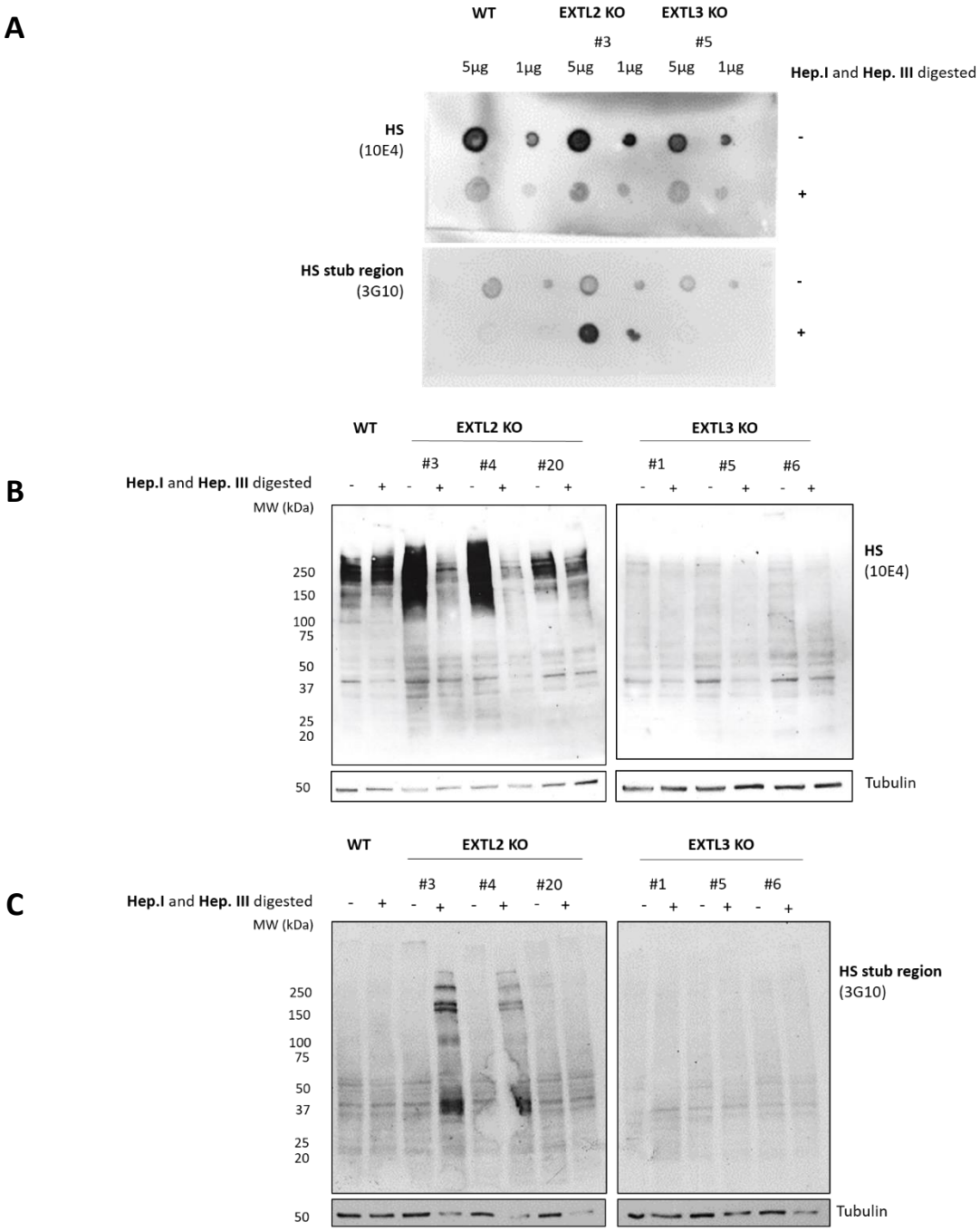


Figure 11. HS GAGs enzymatic digestion assay of MKN74 WT and glycoengineered cell models revealed increased synthesis of HS chains in the absence of EXTL2. A. Dot Blot analysis of the HS chains and HS stub regions present on non-digested (-) and digested (+) lysates from the WT and one representative clone of each EXTL2 KO and EXTL3 KO cell line **B.** WB analysis of HS levels and **C.** HS stub regions exposed by Heparinase digestion. Hep. I and Hep. III terminology was used to indicate Heparinase I and Heparinase III, respectively. α-Tubulin was used as loading control. Representative image of 2 experiments with technical replicates.

4.5. Influence of EXTL2 and EXTL3 expression in the synthesis and sulfation profiles of CS in gastric cancer models

After observing major alterations in the levels of HS on both gastric cancer cell models we evaluated the influence of the abrogation of either EXTL2 or EXTL3 glycosyltransferases on the levels of CS GAG chains.

CS GAGs are long linear polysaccharide chains composed by repeating disaccharide units, these intercalated units correspond to GlcA and GalNAc residues.

CS chains share the same first stage of assembly with HS: the formation of the tetrasaccharide linker. The polymerization of this linker into mature CS GAG chains is then started out by CS N-acetylgalactosaminyl transferase 1 and 2 (CSGalNAcT-1 and -2), through the addition of a first residue of GalNAc to the dephosphorylated tetrasaccharide. CS further elongation is carried out by heterodimeric enzyme complexes of bi-functional chondroitin synthases 1, 2 and 3 (ChSy-1, -2 and -3), and chondroitin polymerizing factor (ChPF), which catalyze the intercalated transfer of GalNAc and GlcA residues (160).

Lastly, CS chains suffer several modification reactions by different CS sulfotransferases. There are different categories of CS chains that vary in terms of sulfation patterns. We mainly focused on Chondroitin Sulfate A, in which GalNAc residues are sulfated at the C4 position (Chondroitin 4-sulfate/C4S) and Chondroitin Sulfate D, in which GalNAc residues are sulfated at the C6 position (Chondroitin 6-sulfate/C6S), as these are the two most frequent modifications mentioned in the literature (106, 161) .

When evaluating the labelling of CS GAG chains, via WB, we only considered the bands present on the regions of high molecular weights, above 150 kDa, since it is very unlikely that the bands showing at lower molecular weights, in the intervals of 20-60 kDa, correspond to CS chains attached to a core protein in CSPGs (Fig. 12 A and C). Furthermore, even though CS chains alone are reported to weigh around 25kDa, it is also improbable that after the process of protein lysis those remain intact and in the soluble fraction.

As it can be noted in Fig. 12 A and B, there was an increase of C4S levels upon abrogation of either EXTL2 or EXTL3. Interestingly, when evaluating C6S levels, we observed a different tendency. EXTL3 deficient cells harbored similar levels of C6S, whereas in the absence of EXTL2, cells harbored lower levels of C6S, when compared to the WT (Fig. 12 C and D).

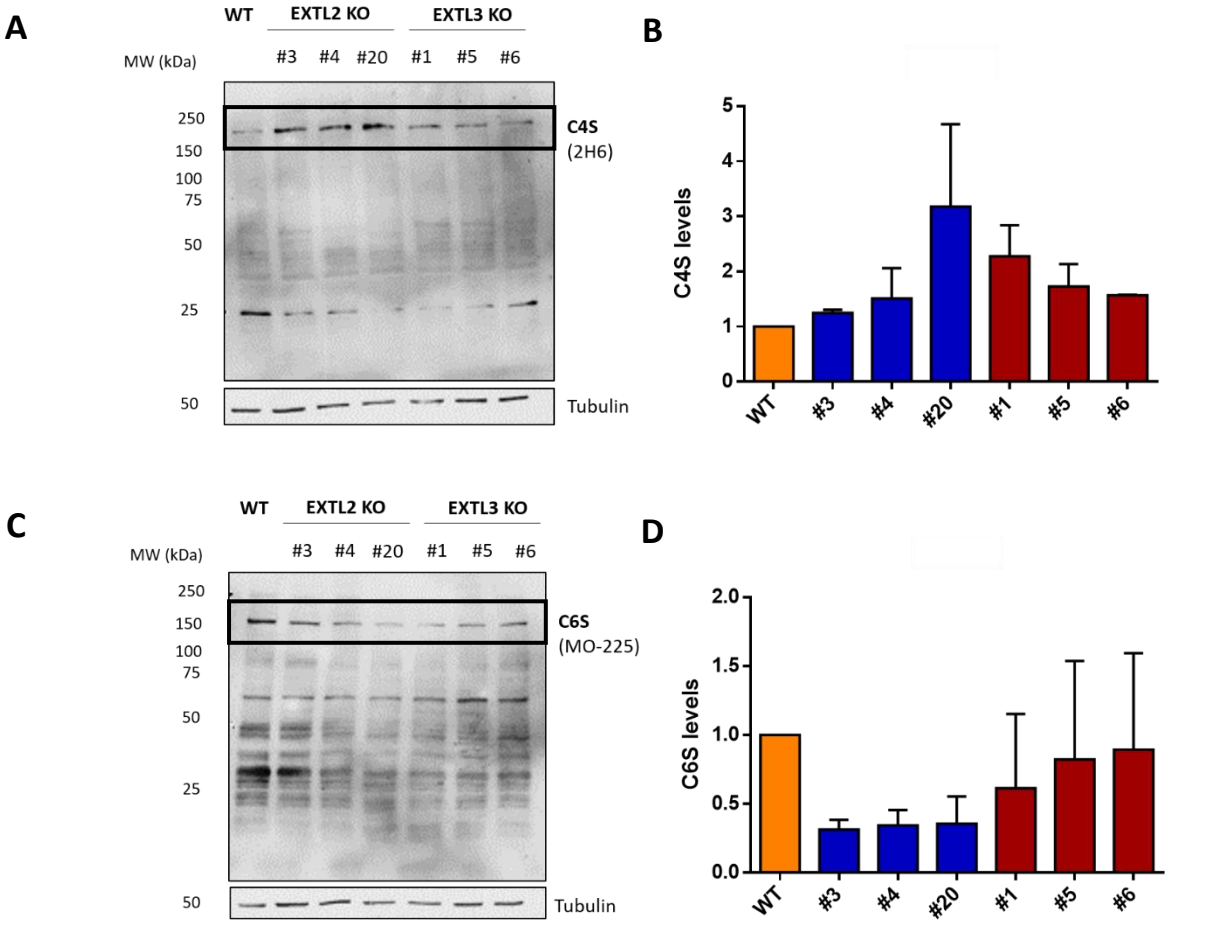


Figure 12. Altered synthesis of CS GAG chains in MKN74 glycoengineered cell models. A and **B.** WB analysis depicting the variations of C4S, **C** and **D.** C6S levels in EXTL2 KO and EXTL3 KO clones in comparison with the WT. α-Tubulin was used as loading control. Band densities were normalized for the loading control. Representative image of 2 experiments with independent biological replicates.

4.6. Analysis of Cell Surface EGFR by Biotinylation Assay

HSPGs assume leading roles as fundamental components of the cells glycocalyx. As it was described in the introduction, for a long time, the multiple roles of HS in modulating the activity of several growth factors, as well as its capacity of interacting with different RTKs, including EGFR, have been studied (134). This major cell surface receptor is known to be overexpressed in gastric cancer, and to our knowledge, the interactions it shares with HS, and consequent signaling and functional outcomes in gastric cancer malignancies, are yet to be described.

To stablish the impact of HS GAG chains in the cell surface expression levels of EGFR in epithelial gastric tumor cells, we optimized a biotinylation assay, through which we isolated the cell surface proteome. This assay was performed on the WT, and in one

representative clone of each KO cell model, and the levels of our targeted receptor on the resultant fractions were then quantified via WB analysis.

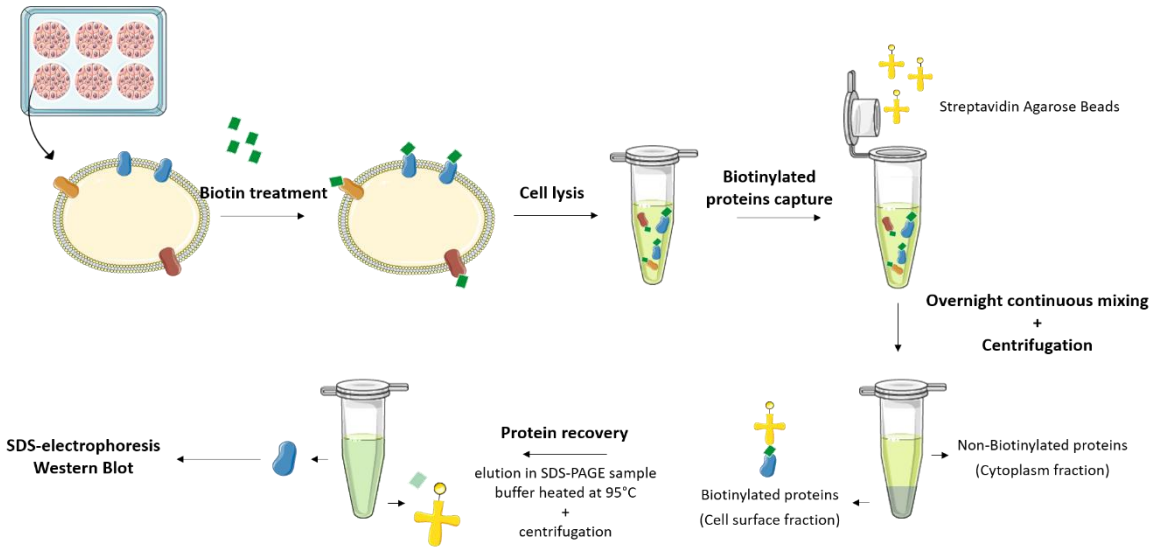


Figure 13. Schematic representation depicting cell-surface proteome biotinylation and isolation methodology.

We targeted cytochrome c, located in the mitochondria inner membrane, for the cytoplasm fraction, and as we show in Fig.14 A, there was no labeling of this complex in our cell surface fraction, demonstrating the efficiency of this isolation. When considering the cytoplasm fraction, we could detect the presence of cytochrome c, as it would be expected. We could also detect the expression of EGFR in the cytoplasm fraction. Interestingly, in this fraction we can distinguish two different bands labelled for EGFR, one rounding the 170 kDa and another one slightly lighter, closer to the 150 kDa, which was not detected in the cell surface fraction (Fig. 14 A).

The expression levels of EGFR were considerably lower in the cell surface fraction of the EXTL2 KO representative clone (Fig. 14 A and B). Interestingly, cells lacking HS GAG chains, from the EXTL3 KO cell line, exhibited values of EGFR cell surface expression similar to the WT (Fig. 14 A and B).

These results suggest that HS GAG chains might regulate the cell surface expression of EGFR in gastric tumor cells.

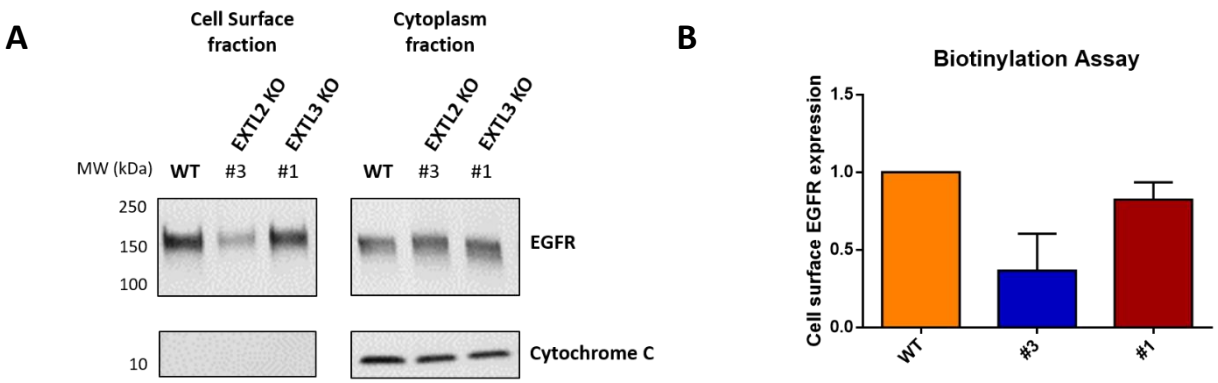


Figure 14. Cell Surface Proteome Biotinylation and Isolation Assay revealed the impact of HS modulation in EGFR cell surface expression. A. and **B.** WB analysis of total EGFR and cytochrome c of the isolated cell surface and cytoplasm fractions collected from the WT, EXTL2 KO and EXTL3 KO clones. Representative images of 2 experiments with technical replicates.

To understand the implications of the altered expression of EGFR in cell signaling events in EXTL2 KO cells, we evaluated the activation state of this receptor, by determining the phosphorylation status of its tyrosine-kinase domain, via WB analysis. This was performed on whole cell lysates and, as a positive control, we evaluated EGFR expression and activation in cell lysates from the gastric adenocarcinoma cell line NCI-N87, known to display high levels of activated-EGFR (p-EGFR).

No major differences, in the levels of p-EGFR, were detected between the MKN74 WT and the EXTL2 KO and EXTL3 KO clones (Fig. 15). Moreover, despite both cell lines presenting high levels of total EGFR, we detected substantially lower phosphorylation of the receptor in the MKN74 cells (WT and KO clones), when compared with the NCI-N87 cells (Fig. 15). In the later we can distinguish intense labelling of p-EGFR (150-180 kDa), absent in MKN74 cells.

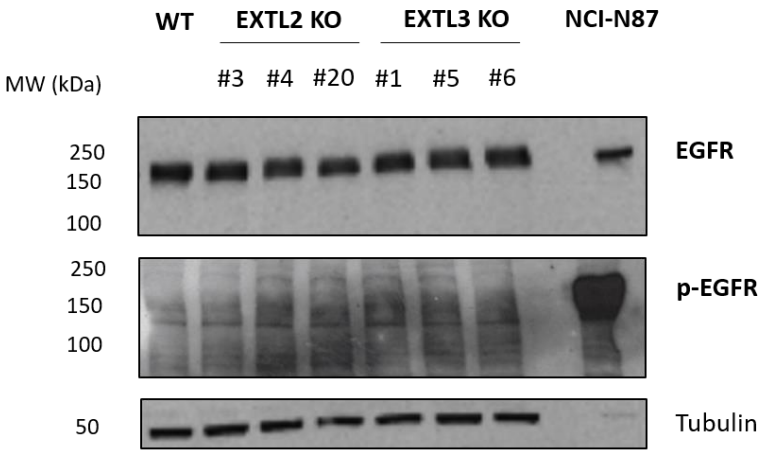


Figure 15. MKN74 EXTL2 and EXTL3 glycoengineered cell models displayed low EGFR activation levels similar to the WT cells. WB analysis of the total expression and activation of EGFR in two gastric cancer cell lines, MKN74 and NCI-N87, and in MKN74 EXTL2 KO and EXTL3 KO clones. NCI-N87 cell protein lysates were used as EGFR activation control. α-Tubulin was used as loading control. Representative image of 2 experiments with independent biological replicates.

4.7. Impact of HS modulation in gastric cancer cell motility features

Upon detecting major phenotypic alterations on our glycoengineered gastric cancer cell models, both in HS and CS content, and considering its possible impact on cell signaling features, as it interferes with EGFR cell surface expression, we then evaluated its consequent impact on the tumor cells behavior.

We started out by exploring cells' migration capabilities by resorting to a wound healing migration assay. This method implies the creation of a wound on confluent monolayered cells, to study the ability of the cells, siting on one of the edges of the created gap, to move forward, towards the cells that sit on the opposite edge, closing the "wound" as new cell-cell contacts are created (Fig. 16 A) (162).

Overall this assay mimics the behavior of cells during *in vivo* migration, and particularly in this project, it was used to learn the role of HS in the regulation of gastric tumor cells migration, accounting also with its interference in cells' interactions with ECM components and cell-cell interactions. The migration rates of the different models were assessed on different coated slides, the coatings being either fibronectin or collagen IV, two major ECM proteins.

EXTL2 KO cells were determined to migrate the fastest in both coatings, presenting higher rates of migration, meaning higher percentages of closed wounds in shorter periods of time, in comparison with the WT (Fig.16 C). Cells migration also varied between different coats. EXTL2KO cells migrated faster in collagen IV coated slides, as the wound was practically closed around the first 4 hours, while in fibronectin coated slides this only happened after 6 hours (Fig.16 B and C).

EXTL3 KO cells migrated at similar rates as the WT cells, in both coatings. In fibronectin coated slides, both WT and the EXTL3 KO cells reached 100% of closed wound only after 12 hours, taking twice as much time as the EXTL2 KO cells (Fig.16 C), while in the collagen IV coated slides the wounds were closed after 10 hours (Fig.16 C).

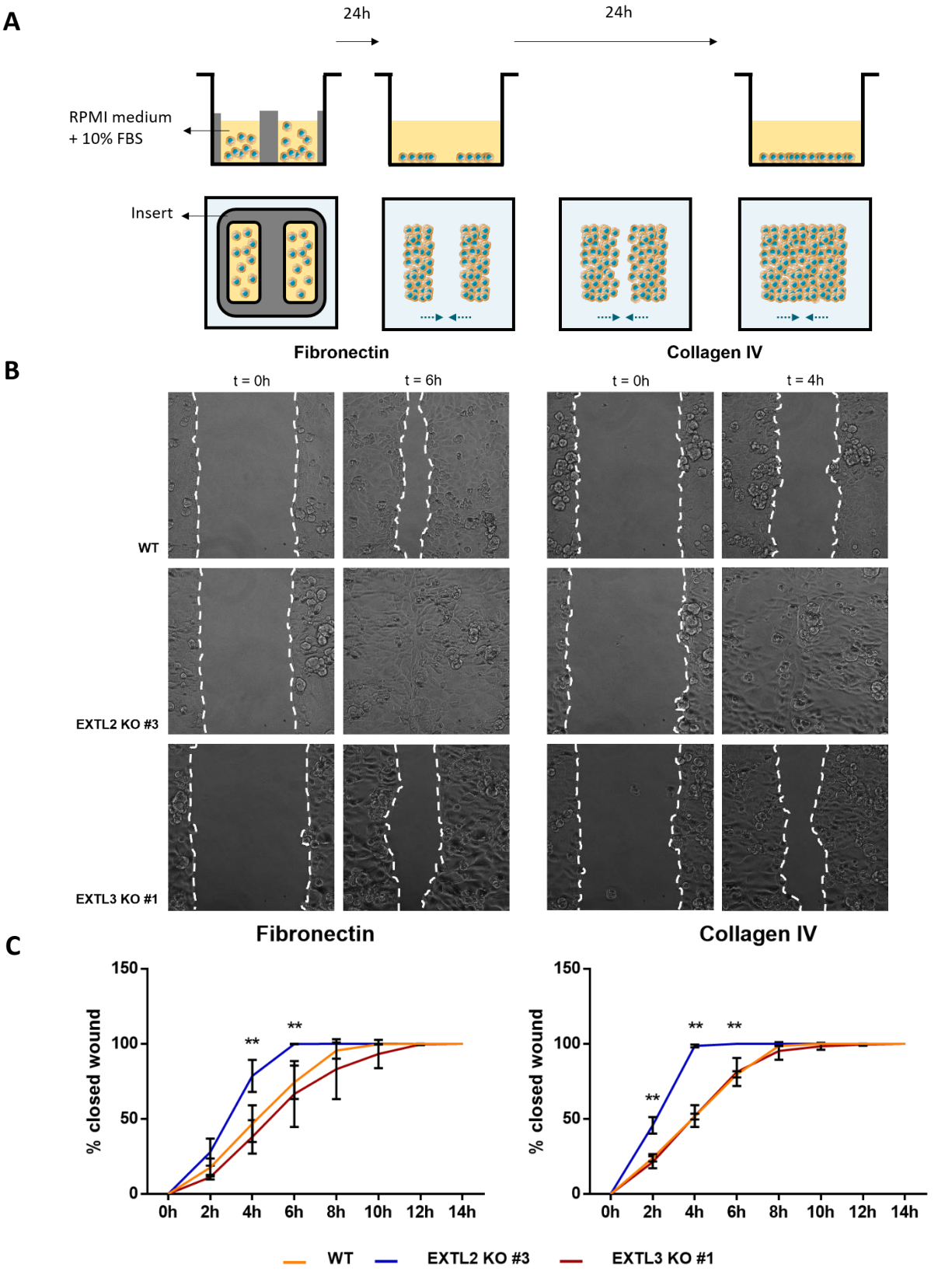


Figure 16. Migration Assay of MKN74 WT and glycoengineered cell models A. Schematic representation illustrating the wound healing migration assay methodology. **B.** Representative image of the wounded cell monolayer in fibronectin or collagen IV coated slides, at t=0h and t=6h or t=4h, respectively. **C.** Graphic representation of the rate of closing wounds evaluated for each condition, in either fibronectin or collagen IV coated slides, on the first 14 hours. Significance is shown relative to the WT and was calculated by two-way ANOVA, with error bars representing S.D. of 4 and 6 independent replicates, of the WT and the clones, respectively, over two experiments (**p<0.01).

We then assessed these cells' invasion capabilities. We selected the same representative clone of each KO cell line already tested in the biotinylation assay and performed a Matrigel invasion assay (Fig. 17 A). In this *in vitro* assay we evaluated cells' capacity of crossing Matrigel coated membranes, which are basement membrane like structures, enriched in ECM proteins, such as collagen IV, laminin, proteoglycans and several growth factors. This allowed to explore the functional role of cell surface HS in this process.

The EXTL2 KO clone, previously determined to have higher levels of HS GAG chains and consequent lower cell surface expression of EGFR, displayed increased invasion capability, showing a higher percentage of invaded cells (Fig. 17 C). This means that more cells were able to degrade the Matrigel and establish themselves in the outer layer. The difference between the EXTL3 KO cells and the WT was not statistically significant (Fig. 17 B).

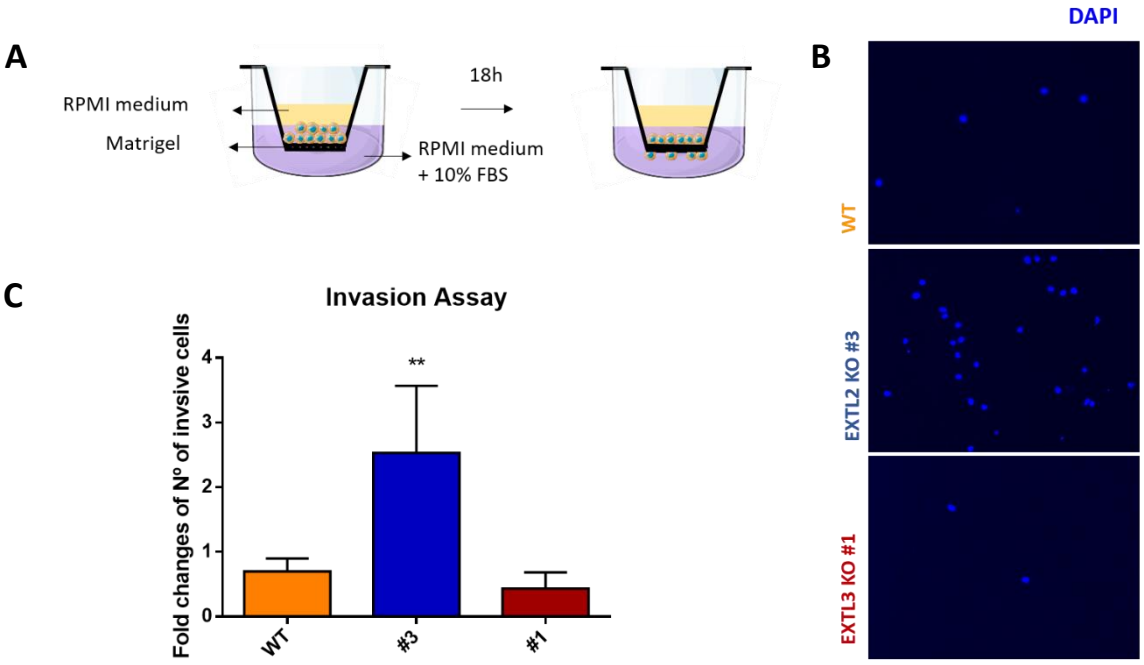


Figure 17. Invasion Assay of MKN74 WT and glycoengineered gastric cell models A. Schematic representation illustrating the Matrigel invasion assay methodology. **B**. Graphic representation of the % of invaded cells highlighted the altered invasion capabilities of the HS modulated gastric cell models. **C**. Representative microscope images of the DAPI stained nucleus of the invasive cells, counted for each condition (200x magnification). Representative images of 2 experiments with independent biological triplicates. Significance is shown relative to the WT and was calculated by Student's t test, with error bars representing S.D. of 6 independent replicates over two experiments (**p<0.01).

5. Discussion

HSPGs are described as fundamental molecules composing both cells' glycocalyx and the ECM environment. These are known to mediate an extended set of mechanisms implicated in development and normal cell physiology, contributing to homeostasis maintenance. Yet, they are also leading players in the context of malignant pathologies, involved in key oncogenic events (98).

HS biosynthetic pathways have been the subject of interest of recent studies, which have revealed the regulatory role of several enzymes that intervene in the different stages of assembly of HS (100-102, 104).

Based on the state of art, the first aim of this project was to gain knowledge about GAGs main biosynthetic pathways, as we took close interest in the roles of EXTL2 and EXTL3 in modulating and determining the HS profile of gastric tumor epithelial cells, which up until now remained unstudied in this specific context.

We resorted to glycoengineered gastric cell models, namely MKN74 EXTL2 KO and EXTL3 KO cells, and started by stablishing the viability of our models by performing an Annexin-V viability assay. All clones presented viability rates higher than 70%, similar to the ones detected for the WT (Fig. 7). This revealed that cells did not suffer substantial modifications regarding apoptotic features, upon CRISPR-Cas9 KO of either *EXTL2* or *EXTL3* genes. In addition, we think that the conditions to which cells are exposed when performing this assay, such as the use of trypsin to detach the cells from the plate they were seeded on, might account for the rates of early apoptosis ranging from 12-22%.

The HS total content of MKN74 WT, EXTL2 KO and EXTL3 KO clones was then determined via WB, FACS and IF approaches. According to our results in the absence of EXTL3, cells were unable to synthesize HS GAG chains, which were not detected neither in the cells' cytoplasm (Fig. 8 A and 9 B) nor in the cells' glycocalyx (Fig. 8 B and 9 A), agreeing to what is described in the literature (100).

We hypothesize that the abrogation of EXTL3, inhibits the transfer of the first residue of GlcNAc to the mature tetrasaccharide linker, and as consequence, it prevents further polymerization of this oligosaccharide, leading to a generalized premature HS GAG synthesis termination in EXTL3 KO cells (Fig. 18).

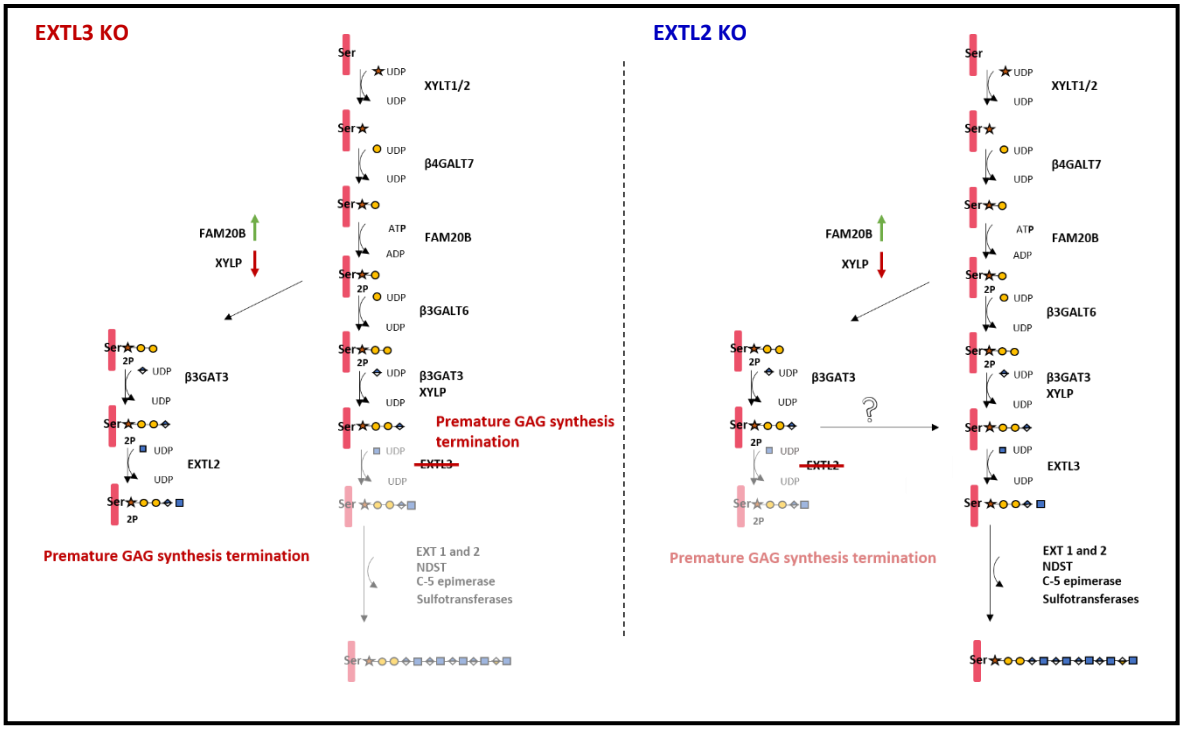


Figure 18. Schematic representation of HS biosynthetic pathways in EXTL2 and EXTL3 glycoengineered gastric cancer models.

On the other hand, EXTL2 abolished expression showed the opposite effect, as EXTL2 KO cells presented increased HS content, which could indicate the presence of longer HS chains and/or a higher number of HS GAG chains (Fig. 8).

To evaluate the number of HS chains in each cell model we performed Heparinase I and Heparinase III digestion of the cell lysates (Fig 10). This exposed one specific unsaturated uronic acid saccharide residue at the extremity of each HS stub region, posteriorly labelled by 3G10 antibody in Dot Blot and WB analysis.

We could not detect HS stub regions in MKN74 WT cells (Fig. 11 C). This has to be carefully analyzed and we must take into consideration the possible disparities in terms of HS chains composition between the WT and each EXTL2 KO clone. The major focus is on the possible different patterns of sulfation and epimerization, since we know that these two factors seem to play a major role in the recognition of these GAG chains by the heparinases performing the cleavage reactions. These possible differences could contribute to a greater resistance to the enzymatic digestion by WT cells when employing similar conditions to the ones used for the enzymatic digestion of EXTL2 KO cells, which was successfully performed.

In addition, we can also hypothesize that if the WT cells indeed possess noteworthy fewer HS chains, when compared to the EXTL2 KO cells, less 3G10 epitopes were exposed upon digestion, and consequently, these could be unnoticeable in the WB analysis (Fig. 11 C).

Altogether, this data suggests that in EXTL2 KO cells, the phosphorylated tetrasaccharide, that follows the EXTL2 pathway, might be dephosphorylated by an unidentified phosphatase, considering that XYLP was proven to be unable to perform this step (102). The resulting tetrasaccharide can then be polymerized by EXTL3, EXT1 and EXT2 that couldn't previously act on the phosphorylated form. Therefore, in EXTL2 KO cells, the immature forms of the tetrasaccharide, that in normal conditions would be polymerized by EXTL2, avoiding further elongation by EXT1 and EXT2, are also being converted into mature HS, contributing to the increased levels of these GAG chains (Fig. 18).

Furthermore, in the absence of EXTL2, cells did not only exhibit more HS chains in a general manner (Fig. 8 A and Fig. 9 B), but also specifically at the cell surface (Fig. 8 B and Fig. 9 A). This indicates that HS chains excessively synthesized in these cells, are also being translocated to the cells' membrane. Nevertheless, there is still more HS accumulated in the cells' cytoplasm than in the cells' membrane, which indicates that the translocation of HSPGs to the membrane of the cells is limited, preventing more HSPGs from reaching the cell surface.

The high levels of HS determined in our EXTL2 KO models are consistent with previous observations in different *in vitro* (100, 104) and *in vivo (163)* models. However, in a recent study by Sembajwe L.F. *et al*, the reduced expression of EXTL2, detected in MCF7 breast cancer cells, was correlated to decreased levels of HS cell surface expression (146). The disparity between these results and what we observed in our work might further indicate that glycosyltransferases act in a tissue-specific manner.

Altogether our results demonstrated, for the first time in gastric cancer models, the critical role of EXTL3 in dictating and initiating the synthesis of HS chains. Moreover, they revealed EXTL2 as a decisive factor in determining the number of HS chains on these cells, possibly assuming a preponderant role of a negative regulator on the HS GAGs biosynthetic pathway.

We then examined the effect of either EXTL2 or EXTL3 altered expression in the synthesis and sulfation profiles of CS chains.

Nadanaka S. *et al* have previously studied the association between EXTL2 expression and the levels of CS *in vivo*. They examined the disaccharide composition of CS chains isolated from Embryonic Fibroblasts, Livers and Brains of WT mice and EXTL2 KO mice, and observed a drastic increase in the overall content of CS, particularly C4S disaccharides, in the EXTL2 KO mice. Meanwhile the levels of C6S did not appear to vary significatively between the two conditions (104).

Our results followed a similar trend to what was previously demonstrated. C4S levels were augmented, not only in EXTL2 KO cells but also in EXTL3 KO cells (Fig. 12 A and B). However, the same was not detected in terms of C6S levels (Fig. 12 C and D), since we could clearly observe a drastic decay of C6S levels in the absence of EXTL2, while in EXTL3 KO cells these levels did not vary substantially.

Our results indicate that, in EXTL3 KO cells, there might be an abnormally high amount of unphosphorylated tetrasaccharide linkers being accumulated, as they cannot be further elongated. It is possible that part of these accumulated structures could follow the CS biosynthetic pathway, being firstly polymerized by CSGalNAcTI and CSGalNAcTII and further elongated into mature CS chains.

In EXTL2 KO cells, as we hypothesized, there are increased levels of phosphorylated tetrasaccharide linkers possibly being dephosphorylated, following then HS elongation. Our results demonstrate that these linkers might also follow the CS biosynthetic pathway. The mechanisms underlying this “biosynthetic shift” from HS to CS remain unknown as this data must be further validated and explored.

Considering this hypothesis, it is possible that the rates by which new CS chains are being synthesized, as result of a process that in normal conditions would not occur, have an effect also on CS modification reactions, as the enzymes that carry out these reactions approach their activity threshold. This could justify the lower and unaltered levels of C6S in the absence of EXTL2 and EXTL3, respectively.

After evaluating the phenotypic features of our models, we performed a biotinylation assay with the intent of studying in which manner the modulation of major GAGs synthesis, by EXTL2 and EXTL3, could affect EGFR cell surface expression, as this major RTK is upregulated in late stages of gastric adenocarcinoma patients (133).

Our WB results demonstrated that we were able to optimize this methodology, validated by the lack of cytochrome c staining in the cell surface fraction, whereas this complex was detected in the cytoplasm fraction (Fig. 14 A). In addition, we could also detect the expression of EGFR in this later fraction. This can be justified by cells extremely dynamic

environment. Receptors are constantly being synthesized and then transported across the cytoplasm to the cell membrane, and the opposite also occurs, there are mechanisms of internalization of these receptors for posterior degradation. In fact, the second EGFR band, labelled in the lower molecular weights (150 kDa) on the cytoplasm fractions, could represent immature forms of EGFRs recently synthesized.

The levels of EGFR expression at the cell surface of epithelial gastric tumor cells drastically decrease when the expression of EXTL2 was abolished, meanwhile no significant changes were observed in the absence of EXTL3, when compared to the WT (Fig. 14 A and B).

In regard to these results, we hypothesize that HS might potentiate the process of internalization of EGFR, which would also explain why there are no major variations, between the WT and all clones, in the levels of total EGFR in total cell lysates (Fig. 15). We also must consider that MKN74 WT cells appear to express low levels of total HS GAG chains, therefore the difference, in the content of HS, between these cells and the EXTL3 KO cells is not as drastic as it is between the WT and the EXTL2 KO cells (Fig. 8 A). This could explain the similar results obtained for EGFR cell surface expression both in the WT and in the EXTL3 KO cells (Fig. 14 B).

In accordance to this, it has also been previously reported the role of other PGs in the internalization and degradation of EGFR. Neill T. *et al* demonstrated, in a human squamous carcinoma cell line, that decorin, which is a CS/DS proteoglycan, is able to interact with cell surface EGFR and to promote its co-internalization via caveolae-mediated endocytosis, leading to its final degradation in late endosomes (164).

To determine the outcome of altered EGFR cell surface expression in cell signaling, we also determined the activation state of this receptor in our HS modulated cell models. We did not evaluate EGFR phosphorylation in biotinylated fractions since this methodology is not “phospho-friendly”, involving several steps of freezing and thawing, to which phospho groups are sensitive. The tyrosine-kinase domain of this receptor is attached to an intracellular juxtamembrane region, located at the cell membrane, therefore, using a cell surface fraction isolation assay to evaluate EGFR state of phosphorylation would be redundant.

When evaluating the total cell lysates, no variations were detected in the activation state of EGFR, comparing the WT with the EXTL2 KO and EXTL3 KO clones. Moreover, comparing these results with the levels of EGFR activation determined for other gastric

cancer cell line, NCI-N87, we learned that, regardless of total EGFR expression, the MKN74 cell line displays extremely low levels of EGFR constitutive activation (Fig. 15).

Lastly, we explored the consequences of HS modulation in gastric tumor cells behavior by performing migration and invasion assays, as these two functional features are closely related to the metastatic potential of tumor cells.

In the absence of EXTL2, and consequent abundance of HS chains, cells displayed high migration capabilities, being the fastest to close the wounds in both coatings (Fig. 16 B and C). Interestingly the pattern of migration was influenced by the content of these coatings since these cells migrated slightly faster when in contact with collagen IV coated slides. However, this is probably not related to the cells content in HS, since the WT and EXTL3 KO cells also migrated faster in the collagen IV coated slides.

A similar trend was observed in terms of invasive features, as EXTL2 KO cells presented severely increased percentages of Matrigel invasive cells, while EXTL3 KO cells did not differ significantly from the WT (Fig. 17 B).

The observations of the promoted migration and invasion in EXTL2 KO cells were intriguing, as these cells showed decreased EGFR cell surface expression, and low EGFR activation. However, we cannot disregard that in our EXTL2 KO models, the HS chains that are being majorly synthesized might interact and promote the activation of different RTKs, leading to the functional features that we observed in our models. Moreover, these results are in agreement with other literature descriptions depicting the functional role of HS and HSPGs in promoting adhesion and migration in different cell lines (91, 113, 139).

SDC4, for example, is a known focal adhesion key component, playing an important role in cells motility and migration capabilities (165). Focal adhesions are signaling multi-complexes known to establish the communication between the external environment and the cells, intervening also in events of cell cytoskeleton remodulation, which are essential to the process of migration, in response to extracellular cues (166). It has been shown that SDC4 interaction with fibronectin in these multi-complexes, promotes the activation of a downstream signaling cascade, involving stabilization of phosphatidylinositol 4, 5-bisphosphate, resulting in activation of PKCα and following regulation of a myriad of Rho-family GTPases. Consequently, this leads to promotion of events associated to cell motility and migration, such as lamellipodium extension, formation of membrane protrusions and contraction of the cell body (165).

In agreement to what we observed in our EXTL2 KO models, in studies performed in murine embryonic fibroblasts, it was determined that the role of SDC4 in the organization

of cells' cytoskeleton actin components and focal adhesion formation, requires multiple HS chains (at least 2 or 3 chains), whereas mutants forms of SDC4, with reduced number of HS chains (0 or 1 chain), presented aberrant cytoskeleton features and diminished focal adhesion formation (167).

In conclusion, our results demonstrate, for the first time in gastric tumor cell models, that EXTL3 is a master regulator of HS biosynthesis, while EXTL2 assumes the role of a negative regulator, leading to an alternative HS biosynthetic pathway. We have also determined the impact of HS modulation in cell surface expression of an RTK and cells' migration and invasion features, revealing HS negative effect in EGFR cell surface expression as well as its role in promoting enhanced motility and invasiveness of gastric carcinoma cells. Overall, our body of work supports the clinical potential of HS biosynthetic machinery, as new biomarkers and drug targets, for early diagnosis and efficient treatment of gastric cancer malignancies

6. Conclusion and Future Perspectives

The main objectives of this project were to understand the biosynthetic mechanisms underlying the synthesis of HS GAG chains, as well as to disclose the impact of HS in tumor cells main signaling events and, consequently, in their functional behavior, using as model gastric cancer.

We resorted to MKN74 EXTL2 KO and EXTL3 KO glycoengineered cell models and started out by evaluating their cellular viability and performing a biochemical characterization via WB, FACS and IF analysis. We were able to determine the part taken by EXTL3 in initiating the synthesis of HS GAG chains, which was revealed to be indispensable in this biosynthetic pathway. We also highlighted the key role of EXTL2 as a key negative regulator on this biosynthetic pathway in gastric cancer models.

Furthermore, we concluded that CS synthesis and sulfation profiles are affected by EXTL2 and EXTL3 altered expression in gastric cancer cells, which indicates an extremely broad regulatory role of these glycosyltransferases in the GAGs biosynthetic pathways.

We optimized a biotinylation assay, through which we isolated the cell surface proteome fraction of the cell models in study. This assay revealed that HS negatively regulates EGFR cell surface expression. We hypothesize that HS promotes the internalization of this receptor, which could have significant implication in the downstream signaling cascades, as less EGFR is exposed to ligand binding.

Lastly, we performed a Wound Healing assay and a Matrigel invasion assay in order to determine the implications of our glycoengineered models' phenotypic features in their behavior. We concluded that HS chains promote a more invasive and motile phenotype on gastric tumor cells.

Overall, our body of work highlighted EXTL2 and EXTL3, as two relevant glycosyltransferases with conflicting roles in the synthesis of GAGs in the context of gastric cancer, and HS, as a key component both in the modulation of RTKs cell surface expression and in tumor cells invasive and migration proficiencies.

In the future we are planning to evaluate the possible implications of EXTL2 and EXTL3 deregulation in the mRNA expression of other major enzymes intervenient in the GAGs biosynthetic pathways. This will allow us to validate our results and confirm if there are any compensatory upregulation mechanisms behind the variations we observed.

We also pretend to further study the impact of our models on CS glycosaminoglycan chains synthesis. Considering the lack of broader anti-CS antibodies that could allow to infer about general levels of all sulfated forms of CS GAG chains, we are planning to perform a CS GAGs enzymatic digestion assay, with Chondroitinase ABC, and then to examine directly the levels of resultant digested chains, disregarding its sulfation profiles.

Additionally, we plan to further evaluate the impact of these glycosyltransferases in GAGs biosynthesis, as we look forward to exploring the part that EXTL2 and EXTL3 take in determining GAGs saccharide overall composition. To achieve this specific aim, we pretend to conduct structural analyses of HS and CS GAGs isolated from our glycoengineered gastric models. These analyses will be crucial to assess the impact of these glycosyltransferases in the regulatory mechanisms related to GAGs' chain length, composition and modifications, primarily their sulfation profiles.

Taking into consideration the influence of HS modulation in EGFR cell surface expression, as well as the low levels of EGFR constitutive activation determined for the glycoengineered cell models, it would be interesting to assess the impact of this modulation in downstream signaling cascades by, for example, analyzing the expression and activation of molecular targets intervening in the MAPK pathway.

To further validate the results obtained in the migration and invasion assays, and considering the similar percentages of cell viability assessed for the WT and all clones, we also plan to evaluate the proliferation rate of the different cell models. This will allow to confirm that the migration or invasion rates, determined for EXTL2 KO and EXTL3 KO clones result from cells functional behavior, rather than their mitotic rate.

Additionally, and once again considering the results obtained in the functional analysis, as HS levels were associated with gastric tumor cells increased migration and invasion capabilities, we also want to study other possible RTKs, besides EGFR, whose expression and activation may be deregulated in our models, and to do this we are planning to use a Phospho-RTK Array Kit.

It would also be valuable to study EGFR downstream molecular targets Src and focal adhesion kinase (FAK), two intracellular RTKs whose increased expression/phosphorylation relates to more invasive phenotypes in tumor cells (168). Since MKN74 cells show low levels of activated EGFR, hard to detect via WB, maybe this could clarify the mechanisms underlying HS promotion of migration and invasion.

In conclusion, we consider that it is fundamental to understand the mechanisms underlying GAGs biosynthetic machinery, and the specific molecular outcomes of its

deregulation in cancer pathologies. Furthermore, HS biosynthetic enzymes might hold potential as predictive markers and drug targets, and its study could lead to the development new relevant tools to early diagnosis and efficient treatment of gastric cancer malignancies.

7. References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. 2018;68(6):394-424.
2. Cancer tomorrow: International Agency for Research on Cancer (IARC); 2018 [Available from: https://gco.iarc.fr/tomorrow/graphic-isotype?type=0&population=900&mode=population&sex=1&cancer=39&age_group=value&apc_male=0&apc_female=0.
3. Hamilton SR, Aaltonen LA, Organization WH, Cancer IAfRo. Pathology and Genetics of Tumours of the Digestive System: IARC Press; 2000.
4. Ajani JA, Lee J, Sano T, Janjigian YY, Fan D, Song S. Gastric adenocarcinoma. Nature reviews Disease primers. 2017;3:17036.
5. Van Cutsem E, Sagaert X, Topal B, Haustermans K, Prenen H. Gastric cancer. Lancet (London, England). 2016;388(10060):2654-64.
6. Mukaisho K-I, Nakayama T, Hagiwara T, Hattori T, Sugihara H. Two distinct etiologies of gastric cardia adenocarcinoma: interactions among pH, Helicobacter pylori, and bile acids. Frontiers in microbiology. 2015;6:412-.
7. Crew KD, Neugut AI. Epidemiology of gastric cancer. World journal of gastroenterology. 2006;12(3):354-62.
8. 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 pathologica et microbiologica Scandinavica. 1965;64:31-49.
9. Vauhkonen M, Vauhkonen H, Sipponen P. Pathology and molecular biology of gastric cancer. Best Pract Res Clin Gastroenterol. 2006;20(4):651-74.
10. Bass AJ, Thorsson V, Shmulevich I, Reynolds SM, Miller M, Bernard B, et al. Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 2014;513(7517):202-9.
11. Rocken C. Molecular classification of gastric cancer. Expert review of molecular diagnostics. 2017;17(3):293-301.
12. 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.
13. Zabaleta J. Multifactorial etiology of gastric cancer. Methods in molecular biology (Clifton, NJ). 2012;863:411-35.
14. Kelley JR, Duggan JM. Gastric cancer epidemiology and risk factors. Journal of clinical epidemiology. 2003;56(1):1-9.
15. Nagini S. Carcinoma of the stomach: A review of epidemiology, pathogenesis, molecular genetics and chemoprevention. World J Gastrointest Oncol. 2012;4(7):156-69.
16. Wen S, Moss SF. Helicobacter pylori virulence factors in gastric carcinogenesis. Cancer letters. 2009;282(1):1-8.
17. Nejati S, Karkhah A, Darvish H, Validi M, Ebrahimpour S, Nouri HR. Influence of Helicobacter pylori virulence factors CagA and VacA on pathogenesis of gastrointestinal disorders. Microbial pathogenesis. 2018;117:43-8.
18. Sugano K, Tack J, Kuipers EJ, Graham DY, El-Omar EM, Miura S, et al. Kyoto global consensus report on Helicobacter pylori gastritis. Gut. 2015;64(9):1353-67.
19. Gomceli I, Demiriz B, Tez M. Gastric carcinogenesis. World journal of gastroenterology. 2012;18(37):5164-70.
20. Song H, Zhu J, Lu D. Molecular-targeted first-line therapy for advanced gastric cancer. The Cochrane database of systematic reviews. 2016;7:Cd011461.
21. Duarte HO, Gomes J, Machado JC, Reis CA. Gastric cancer: Basic aspects. Helicobacter. 2018;23 Suppl 1:e12523.
22. Matsuoka T, Yashiro M. Biomarkers of gastric cancer: Current topics and future perspective. World J Gastroenterol. 2018;24(26):2818-32.
23. Gerson JN, Skariah S, Denlinger CS, Astsaturov I. Perspectives of HER2-targeting in gastric and esophageal cancer. Expert Opin Investig Drugs. 2017;26(5):531-40.
24. Moasser MM. The oncogene HER2: its signaling and transforming functions and its role in human cancer pathogenesis. Oncogene. 2007;26(45):6469-87.
25. Abrahao-Machado LF, Scapulatempo-Neto C. HER2 testing in gastric cancer: An update. World journal of gastroenterology. 2016;22(19):4619-25.

26. Proserpio I, Rausei S, Barzaghi S, Frattini F, Galli F, Iovino D, et al. Multimodal treatment of gastric cancer. World journal of gastrointestinal surgery. 2014;6(4):55-8.
27. Hirashima Y, Yamada Y, Matsubara J, Takahari D, Okita N, Takashima A, et al. Impact of vascular endothelial growth factor receptor 1, 2, and 3 expression on the outcome of patients with gastric cancer. Cancer science. 2009;100(2):310-5.
28. Fuchs CS, Tomasek J, Yong CJ, Dumitru F, Passalacqua R, Goswami C, et al. Ramucirumab monotherapy for previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma (REGARD): an international, randomised, multicentre, placebo-controlled, phase 3 trial. Lancet (London, England). 2014;383(9911):31-9.
29. Long J, Lin J, Wang A, Wu L, Zheng Y, Yang X, et al. PD-1/PD-L blockade in gastrointestinal cancers: lessons learned and the road toward precision immunotherapy. Journal of hematology & oncology. 2017;10(1):146.
30. Wang X, Teng F, Kong L, Yu J. PD-L1 expression in human cancers and its association with clinical outcomes. OncoTargets and therapy. 2016;9:5023-39.
31. Wu C, Zhu Y, Jiang J, Zhao J, Zhang XG, Xu N. Immunohistochemical localization of programmed death-1 ligand-1 (PD-L1) in gastric carcinoma and its clinical significance. Acta histochemica. 2006;108(1):19-24.
32. Boger C, Behrens HM, Mathiak M, Kruger S, Kalthoff H, Rocken C. PD-L1 is an independent prognostic predictor in gastric cancer of Western patients. Oncotarget. 2016;7(17):24269-83.
33. Amatatsu M, Arigami T, Uenosono Y, Yanagita S, Uchikado Y, Kijima Y, et al. Programmed death-ligand 1 is a promising blood marker for predicting tumor progression and prognosis in patients with gastric cancer. Cancer science. 2018;109(3):814-20.
34. Muro K, Chung HC, Shankaran V, Geva R, Catenacci D, Gupta S, et al. Pembrolizumab for patients with PD-L1-positive advanced gastric cancer (KEYNOTE-012): a multicentre, open-label, phase 1b trial. The Lancet Oncology. 2016;17(6):717-26.
35. Le DT, Bendell JC, Calvo E, Kim JW, Ascierto PA, Sharma P, et al. Safety and activity of nivolumab monotherapy in advanced and metastatic (A/M) gastric or gastroesophageal junction cancer (GC/GEC): Results from the CheckMate-032 study. Journal of Clinical Oncology. 2016;34(4_suppl):6-.
36. Orditura M, Galizia G, Sforza V, Gambardella V, Fabozzi A, Laterza MM, et al. Treatment of gastric cancer. World Journal of Gastroenterology. 2014;20(7):1635-49.
37. Theocharis AD, Skandalis SS, Gialeli C, Karamanos NK. Extracellular matrix structure. Adv Drug Deliv Rev. 2016;97:4-27.
38. Ricard-Blum S. The collagen family. Cold Spring Harbor perspectives in biology. 2011;3(1):a004978.
39. Mouw JK, Ou G, Weaver VM. Extracellular matrix assembly: a multiscale deconstruction. Nature reviews Molecular cell biology. 2014;15(12):771-85.
40. Yue B. Biology of the extracellular matrix: an overview. Journal of glaucoma. 2014;23(8 Suppl 1):S20-3.
41. Mecham RP. Overview of extracellular matrix. Current protocols in cell biology. 2012;Chapter 10:Unit 10.1.
42. Singh P, Carraher C, Schwarzbauer JE. Assembly of fibronectin extracellular matrix. Annu Rev Cell Dev Biol. 2010;26:397-419.
43. Kular JK, Basu S, Sharma RI. The extracellular matrix: Structure, composition, age-related differences, tools for analysis and applications for tissue engineering. Journal of tissue engineering. 2014;5:2041731414557112.
44. Pozzi A, Yurchenco PD, Iozzo RV. The nature and biology of basement membranes. Matrix biology : journal of the International Society for Matrix Biology. 2017;57-58:1-11.
45. Hynes RO. Stretching the boundaries of extracellular matrix research. Nature reviews Molecular cell biology. 2014;15(12):761-3.
46. Frantz C, Stewart KM, Weaver VM. The extracellular matrix at a glance. J Cell Sci. 2010;123(Pt 24):4195-200.
47. Multhaupt HA, Leitinger B, Gullberg D, Couchman JR. Extracellular matrix component signaling in cancer. Adv Drug Deliv Rev. 2016;97:28-40.
48. Poltavets V, Kochetkova M, Pitson SM, Samuel MS. The Role of the Extracellular Matrix and Its Molecular and Cellular Regulators in Cancer Cell Plasticity. Front Oncol. 2018;8:19.
49. Egeblad M, Werb Z. New functions for the matrix metalloproteinases in cancer progression. Nat Rev Cancer. 2002;2(3):161-74.
50. Spiro RG. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. Glycobiology. 2002;12(4):43R-56R.

51. Vajaria BN, Patel PS. Glycosylation: a hallmark of cancer? Glycoconjugate journal. 2017;34(2):147-56.
52. Munkley J, Elliott DJ. Hallmarks of glycosylation in cancer. Oncotarget. 2016;7(23):35478-89.
53. Ohtsubo K, Marth JD. Glycosylation in Cellular Mechanisms of Health and Disease. Cell. 2006;126(5):855-67.
54. Moremen KW, Tiemeyer M, Nairn AV. Vertebrate protein glycosylation: diversity, synthesis and function. Nature reviews Molecular cell biology. 2012;13(7):448-62.
55. Pinho SS, Reis CA. Glycosylation in cancer: mechanisms and clinical implications. Nature reviews Cancer. 2015;15(9):540-55.
56. Ferreira JA, Magalhaes 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.
57. Tarbell JM, Cancel LM. The glycocalyx and its significance in human medicine. J Intern Med. 2016;280(1):97-113.
58. Joseph M, Trinh HM, Mitra AK. Chapter 7 - Peptide and Protein-Based Therapeutic Agents∗. In: Mitra AK, Cholkar K, Mandal A, editors. Emerging Nanotechnologies for Diagnostics, Drug Delivery and Medical Devices. Boston: Elsevier; 2017. p. 145-67.
59. Gomes C, Osório H, Pinto MT, Campos D, Oliveira MJ, Reis CA. Expression of ST3GAL4 leads to SLe(x) expression and induces c-Met activation and an invasive phenotype in gastric carcinoma cells. PloS one. 2013;8(6):e66737-e.
60. Mereiter S, Magalhaes A, Adamczyk B, Jin C, Almeida A, Drici L, et al. Glycomic analysis of gastric carcinoma cells discloses glycans as modulators of RON receptor tyrosine kinase activation in cancer. Biochimica et biophysica acta. 2016;1860(8):1795-808.
61. Videira PA, Correia M, Malagolini N, Crespo HJ, Ligeiro D, Calais FM, et al. ST3Gal.I sialyltransferase relevance in bladder cancer tissues and cell lines. BMC Cancer. 2009;9:357.
62. Wang PH, Lee WL, Juang CM, Yang YH, Lo WH, Lai CR, et al. Altered mRNA expressions of sialyltransferases in ovarian cancers. Gynecologic oncology. 2005;99(3):631-9.
63. Schneider F, Kemmner W, Haensch W, Franke G, Gretschel S, Karsten U, et al. Overexpression of sialyltransferase CMP-sialic acid:Galbeta1,3GalNAc-R alpha6-Sialyltransferase is related to poor patient survival in human colorectal carcinomas. Cancer Res. 2001;61(11):4605-11.
64. Petretti T, Kemmner W, Schulze B, Schlag PM. Altered mRNA expression of glycosyltransferases in human colorectal carcinomas and liver metastases. Gut. 2000;46(3):359-66.
65. Carvalho AS, Harduin-Lepers A, Magalhaes A, Machado E, Mendes N, Costa LT, et al. Differential expression of alpha-2,3-sialyltransferases and alpha-1,3/4-fucosyltransferases regulates the levels of sialyl Lewis a and sialyl Lewis x in gastrointestinal carcinoma cells. The international journal of biochemistry & cell biology. 2010;42(1):80-9.
66. Yan X, Lin Y, Liu S, Aziz F, Yan Q. Fucosyltransferase IV (FUT4) as an effective biomarker for the diagnosis of breast cancer. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie. 2015;70:299-304.
67. Mereiter S, Polom K, Williams C, Polonia A, Guergova-Kuras M, Karlsson NG, et al. The Thomsen-Friedenreich Antigen: A Highly Sensitive and Specific Predictor of Microsatellite Instability in Gastric Cancer. Journal of clinical medicine. 2018;7(9).
68. Kang HY, Wu QH, Sun AQ, Liu X, Fan YB, Deng XY. Cancer Cell Glycocalyx and Its Significance in Cancer Progression. Int J Mol Sci. 2018;19(9):23.
69. Magalhaes A, Marcos-Pinto R, Nairn AV, Dela Rosa M, Ferreira RM, Junqueira-Neto S, et al. Helicobacter pylori chronic infection and mucosal inflammation switches the human gastric glycosylation pathways. Biochimica et biophysica acta. 2015;1852(9):1928-39.
70. David L, Nesland JM, Clausen H, Carneiro F, Sobrinho-Simoes M. Simple mucin-type carbohydrate antigens (Tn, sialosyl-Tn and T) in gastric mucosa, carcinomas and metastases. APMIS Supplementum. 1992;27:162-72.
71. 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.
72. 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.

73. Marcos NT, Magalhaes A, Ferreira B, Oliveira MJ, Carvalho AS, Mendes N, et al. Helicobacter pylori induces beta3GnT5 in human gastric cell lines, modulating expression of the SabA ligand sialyl-Lewis x. The Journal of clinical investigation. 2008;118(6):2325-36.
74. Rodrigues JG, Balmana M, Macedo JA, Pocas J, Fernandes A, de-Freitas-Junior JCM, et al. Glycosylation in cancer: Selected roles in tumour progression, immune modulation and metastasis. Cellular immunology. 2018.
75. Pinho SS, Reis CA, Paredes J, Magalhaes AM, Ferreira AC, Figueiredo J, et al. The role of N-acetylglucosaminyltransferase III and V in the post-transcriptional modifications of E-cadherin. Human molecular genetics. 2009;18(14):2599-608.
76. Pinho SS, Figueiredo J, Cabral J, Carvalho S, Dourado J, Magalhaes A, et al. E-cadherin and adherens-junctions stability in gastric carcinoma: functional implications of glycosyltransferases involving N-glycan branching biosynthesis, N-acetylglucosaminyltransferases III and V. Biochimica et biophysica acta. 2013;1830(3):2690-700.
77. Ihara S, Miyoshi E, Ko JH, Murata K, Nakahara S, Honke K, et al. Prometastatic effect of N-acetylglucosaminyltransferase V is due to modification and stabilization of active matriptase by adding beta 1-6 GlcNAc branching. The Journal of biological chemistry. 2002;277(19):16960-7.
78. Strimbu K, Tavel JA. What are biomarkers? Current opinion in HIV and AIDS. 2010;5(6):463-6.
79. Adamczyk B, Tharmalingam T, Rudd PM. Glycans as cancer biomarkers. Biochimica et biophysica acta. 2012;1820(9):1347-53.
80. Sethi MK, Hancock WS, Fanayan S. Identifying N-Glycan Biomarkers in Colorectal Cancer by Mass Spectrometry. Accounts of chemical research. 2016;49(10):2099-106.
81. Drake PM, Cho W, Li B, Prakobphol A, Johansen E, Anderson NL, et al. Sweetening the pot: adding glycosylation to the biomarker discovery equation. Clinical chemistry. 2010;56(2):223-36.
82. Căinap C, Nagy V, Gherman A, Cetean S, Laszlo I, Constantin AM, et al. Classic tumor markers in gastric cancer. Current standards and limitations. Clujul medical (1957). 2015;88(2):111-5.
83. Duraker Nt, Çelik AN. The prognostic significance of preoperative serum CA 19-9 in patients with resectable gastric carcinoma: comparison with CEA. Journal of Surgical Oncology. 2001;76(4):266-71.
84. Humphris JL, Chang DK, Johns AL, Scarlett CJ, Pajic M, Jones MD, et al. The prognostic and predictive value of serum CA19.9 in pancreatic cancer. Annals of oncology : official journal of the European Society for Medical Oncology. 2012;23(7):1713-22.
85. Wang Q-h, Ji Z-g, Chen Z-g, Li H-z, Fan H, Fan X-r, et al. Serum CA 19-9 as a good prognostic biomarker in patients with bladder cancer. International Journal of Surgery. 2015;15:113-6.
86. Lech G, Słotwiński R, Słodkowski M, Krasnodębski IW. Colorectal cancer tumour markers and biomarkers: Recent therapeutic advances. World J Gastroenterol. 2016;22(5):1745-55.
87. Ychou M, Duffour J, Kramar A, Gourgou S, Grenier J. Clinical significance and prognostic value of CA72-4 compared with CEA and CA19-9 in patients with gastric cancer. Disease markers. 2000;16(3-4):105-10.
88. Jo JC, Ryu MH, Koo DH, Ryoo BY, Kim HJ, Kim TW, et al. Serum CA 19-9 as a prognostic factor in patients with metastatic gastric cancer. Asia-Pacific journal of clinical oncology. 2013;9(4):324-30.
89. Tachibana M, Takemoto Y, Nakashima Y, Kinugasa S, Kotoh T, Dhar DK, et al. Serum carcinoembryonic antigen as a prognostic factor in resectable gastric cancer. Journal of the American College of Surgeons. 1998;187(1):64-8.
90. Hashimoto Y, Skacel M, Adams JC. Association of loss of epithelial syndecan-1 with stage and local metastasis of colorectal adenocarcinomas: an immunohistochemical study of clinically annotated tumors. BMC Cancer. 2008;8:185-.
91. Lim HC, Couchman JR. Syndecan-2 regulation of morphology in breast carcinoma cells is dependent on RhoGTPases. Biochimica et biophysica acta. 2014;1840(8):2482-90.
92. Magalhaes A, Marcos NT, Carvalho AS, David L, Figueiredo C, Bastos J, et al. Helicobacter pylori cag pathogenicity island-positive strains induce syndecan-4 expression in gastric epithelial cells. FEMS immunology and medical microbiology. 2009;56(3):223-32.
93. Kirkpatrick CA, Selleck SB. Heparan sulfate proteoglycans at a glance. J Cell Sci. 2007;120(11):1829-32.
94. Lindahl U, Couchman J, Kimata K, Esko JD. Proteoglycans and Sulfated Glycosaminoglycans. In: rd, Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, et al., editors. Essentials of Glycobiology. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press

 2015. p. 207-21.
95. Rabenstein DL. Heparin and heparan sulfate: structure and function. Natural Product Reports. 2002;19(3):312-31.
96. Li JP, Kusche-Gullberg M. Heparan Sulfate: Biosynthesis, Structure, and Function. In: Jeon KW, editor. International Review of Cell and Molecular Biology, Vol 325. International Review of Cell and Molecular Biology. 325. San Diego: Elsevier Academic Press Inc; 2016. p. 215-73.
97. Iozzo RV, Schaefer L. Proteoglycan form and function: A comprehensive nomenclature of proteoglycans. Matrix Biol. 2015;42:11-55.
98. Afratis N, Gialeli C, Nikitovic D, Tsegenidis T, Karousou E, Theocharis AD, et al. Glycosaminoglycans: key players in cancer cell biology and treatment. The FEBS journal. 2012;279(7):1177-97.
99. Xu D, Esko JD. Demystifying heparan sulfate-protein interactions. Annu Rev Biochem. 2014;83:129-57.
100. Chen Y-H, Narimatsu Y, Clausen TM, Gomes C, Karlsson R, Steentoft C, et al. The GAGOme: a cell-based library of displayed glycosaminoglycans. Nature Methods. 2018;15(11):881-8.
101. Wen J, Xiao J, Rahdar M, Choudhury BP, Cui J, Taylor GS, et al. Xylose phosphorylation functions as a molecular switch to regulate proteoglycan biosynthesis. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(44):15723-8.
102. Koike T, Izumikawa T, Sato B, Kitagawa H. Identification of phosphatase that dephosphorylates xylose in the glycosaminoglycan-protein linkage region of proteoglycans. The Journal of biological chemistry. 2014;289(10):6695-708.
103. Kitagawa H, Shimakawa H, Sugahara K. The tumor suppressor EXT-like gene EXTL2 encodes an alpha1, 4-N-acetylhexosaminyltransferase that transfers N-acetylgalactosamine and N-acetylglucosamine to the common glycosaminoglycan-protein linkage region. The key enzyme for the chain initiation of heparan sulfate. The Journal of biological chemistry. 1999;274(20):13933-7.
104. Nadanaka S, Zhou S, Kagiyama S, Shoji N, Sugahara K, Sugihara K, et al. EXTL2, a member of the EXT family of tumor suppressors, controls glycosaminoglycan biosynthesis in a xylose kinase-dependent manner. The Journal of biological chemistry. 2013;288(13):9321-33.
105. Christianson HC, Belting M. Heparan sulfate proteoglycan as a cell-surface endocytosis receptor. Matrix biology : journal of the International Society for Matrix Biology. 2014;35:51-5.
106. Soares da Costa D, Reis RL, Pashkuleva I. Sulfation of Glycosaminoglycans and Its Implications in Human Health and Disorders. Annual review of biomedical engineering. 2017;19:1-26.
107. Yayon A, Klagsbrun M, Esko JD, Leder P, Ornitz DM. Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor. Cell. 1991;64(4):841-8.
108. Ruoslahti E, Yamaguchi Y. Proteoglycans as modulators of growth factor activities. Cell. 1991;64(5):867-9.
109. Bishop JR, Schuksz M, Esko JD. Heparan sulphate proteoglycans fine-tune mammalian physiology. Nature. 2007;446(7139):1030-7.
110. Takei Y, Ozawa Y, Sato M, Watanabe A, Tabata T. Three Drosophila EXT genes shape morphogen gradients through synthesis of heparan sulfate proteoglycans. Development (Cambridge, England). 2004;131(1):73-82.
111. Han C, Belenkaya TY, Khodoun M, Tauchi M, Lin X, Lin X. Distinct and collaborative roles of Drosophila EXT family proteins in morphogen signalling and gradient formation. Development (Cambridge, England). 2004;131(7):1563-75.
112. Dyer DP, Migliorini E, Salanga CL, Thakar D, Handel TM, Richter RP. Differential structural remodelling of heparan sulfate by chemokines: the role of chemokine oligomerization. Open biology. 2017;7(1).
113. Beauvais DM, Rapraeger AC. Syndecan-1-mediated cell spreading requires signaling by alphavbeta3 integrins in human breast carcinoma cells. Experimental cell research. 2003;286(2):219-32.
114. Sarrazin S, Lamanna WC, Esko JD. Heparan sulfate proteoglycans. Cold Spring Harbor perspectives in biology. 2011;3(7).
115. Stanford KI, Bishop JR, Foley EM, Gonzales JC, Niesman IR, Witztum JL, et al. Syndecan-1 is the primary heparan sulfate proteoglycan mediating hepatic clearance of triglyceride-rich lipoproteins in mice. The Journal of clinical investigation. 2009;119(11):3236-45.

116. Tkachenko E, Lutgens E, Stan RV, Simons M. Fibroblast growth factor 2 endocytosis in endothelial cells proceed via syndecan-4-dependent activation of Rac1 and a Cdc42-dependent macropinocytic pathway. J Cell Sci. 2004;117(Pt 15):3189-99.
117. Friand V, David G, Zimmermann P. Syntenin and syndecan in the biogenesis of exosomes. Biol Cell. 2015;107(10):331-41.
118. Baietti MF, Zhang Z, Mortier E, Melchior A, Degeest G, Geeraerts A, et al. Syndecan-syntenin-ALIX regulates the biogenesis of exosomes. Nature cell biology. 2012;14(7):677-85.
119. Knelson EH, Nee JC, Blobe GC. Heparan sulfate signaling in cancer. Trends in biochemical sciences. 2014;39(6):277-88.
120. Matsuda K, Maruyama H, Guo F, Kleeff J, Itakura J, Matsumoto Y, et al. Glypican-1 is overexpressed in human breast cancer and modulates the mitogenic effects of multiple heparin-binding growth factors in breast cancer cells. Cancer Res. 2001;61(14):5562-9.
121. Barbareschi M, Maisonneuve P, Aldovini D, Cangi MG, Pecciarini L, Angelo Mauri F, et al. High syndecan-1 expression in breast carcinoma is related to an aggressive phenotype and to poorer prognosis. Cancer. 2003;98(3):474-83.
122. Derksen PW, Keehnen RM, Evers LM, van Oers MH, Spaargaren M, Pals ST. Cell surface proteoglycan syndecan-1 mediates hepatocyte growth factor binding and promotes Met signaling in multiple myeloma. Blood. 2002;99(4):1405-10.
123. Bergers G, Benjamin LE. Tumorigenesis and the angiogenic switch. Nature reviews Cancer. 2003;3(6):401-10.
124. De Palma M, Biziato D, Petrova TV. Microenvironmental regulation of tumour angiogenesis. Nat Rev Cancer. 2017;17:457.
125. Nishida N, Yano H, Nishida T, Kamura T, Kojiro M. Angiogenesis in cancer. Vasc Health Risk Manag. 2006;2(3):213-9.
126. Nagarajan A, Malvi P, Wajapeyee N. Heparan Sulfate and Heparan Sulfate Proteoglycans in Cancer Initiation and Progression. Front Endocrinol. 2018;9:11.
127. Cohen IR, Murdoch AD, Naso MF, Marchetti D, Berd D, Iozzo RV. Abnormal expression of perlecan proteoglycan in metastatic melanomas. Cancer Res. 1994;54(22):5771-4.
128. Iozzo RV, Sanderson RD. Proteoglycans in cancer biology, tumour microenvironment and angiogenesis. Journal of cellular and molecular medicine. 2011;15(5):1013-31.
129. Lamorte S, Ferrero S, Aschero S, Monitillo L, Bussolati B, Omedè P, et al. Syndecan-1 promotes the angiogenic phenotype of multiple myeloma endothelial cells. Leukemia. 2011;26:1081.
130. Herbst RS. Review of epidermal growth factor receptor biology. International journal of radiation oncology, biology, physics. 2004;59(2 Suppl):21-6.
131. Scaltriti M, Baselga J. The epidermal growth factor receptor pathway: a model for targeted therapy. Clinical cancer research : an official journal of the American Association for Cancer Research. 2006;12(18):5268-72.
132. Veale D, Kerr N, Gibson GJ, Kelly PJ, Harris AL. The relationship of quantitative epidermal growth factor receptor expression in non-small cell lung cancer to long term survival. British journal of cancer. 1993;68(1):162-5.
133. Wang D, Wang B, Wang R, Zhang Z, Lin Y, Huang G, et al. High expression of EGFR predicts poor survival in patients with resected T3 stage gastric adenocarcinoma and promotes cancer cell survival. Oncology letters. 2017;13(5):3003-13.
134. Cole CL, Rushton G, Jayson GC, Avizienyte E. Ovarian cancer cell heparan sulfate 6-O-sulfotransferases regulate an angiogenic program induced by heparin-binding epidermal growth factor (EGF)-like growth factor/EGF receptor signaling. The Journal of biological chemistry. 2014;289(15):10488-501.
135. Kim SY, Choi EJ, Yun JA, Jung ES, Oh ST, Kim JG, et al. Syndecan-1 expression is associated with tumor size and EGFR expression in colorectal carcinoma: a clinicopathological study of 230 cases. International journal of medical sciences. 2015;12(2):92-9.
136. Matsumoto A, Ono M, Fujimoto Y, Gallo RL, Bernfield M, Kohgo Y. Reduced expression of syndecan-1 in human hepatocellular carcinoma with high metastatic potential. International journal of cancer. 1997;74(5):482-91.
137. Fujiya M, Watari J, Ashida T, Honda M, Tanabe H, Fujiki T, et al. Reduced expression of syndecan-1 affects metastatic potential and clinical outcome in patients with colorectal cancer. Jpn J Cancer Res. 2001;92(10):1074-81.
138. Park H, Kim Y, Lim Y, Han I, Oh ES. Syndecan-2 mediates adhesion and proliferation of colon carcinoma cells. The Journal of biological chemistry. 2002;277(33):29730-6.

139. De Oliveira T, Abiatari I, Raulefs S, Sauliunaite D, Erkan M, Kong B, et al. Syndecan-2 promotes perineural invasion and cooperates with K-ras to induce an invasive pancreatic cancer cell phenotype. Molecular cancer. 2012;11:19-.
140. Mytilinaiou M, Nikitovic D, Berdiaki A, Kostouras A, Papoutsidakis A, Tsatsakis AM, et al. Emerging roles of syndecan 2 in epithelial and mesenchymal cancer progression. IUBMB life. 2017;69(11):824-33.
141. Suhovskih AV, Domanitskaya NV, Tsidulko AY, Prudnikova TY, Kashuba VI, Grigorieva EV. Tissue-specificity of heparan sulfate biosynthetic machinery in cancer. Cell adhesion & migration. 2015;9(6):452-9.
142. Crespo A, Garcia-Suarez O, Fernandez-Vega I, Solis-Hernandez MP, Garcia B, Castanon S, et al. Heparan sulfate proteoglycans undergo differential expression alterations in left sided colorectal cancer, depending on their metastatic character. BMC Cancer. 2018;18:17.
143. Fernandez-Vega I, Garcia-Suarez O, Garcia B, Crespo A, Astudillo A, Quiros LM. Heparan sulfate proteoglycans undergo differential expression alterations in right sided colorectal cancer, depending on their metastatic character. BMC Cancer. 2015;15:742.
144. Karibe T, Fukui H, Sekikawa A, Shiratori K, Fujimori T. EXTL3 promoter methylation down-regulates EXTL3 and heparan sulphate expression in mucinous colorectal cancers. The Journal of pathology. 2008;216(1):32-42.
145. Arai T, Akiyama Y, Nagasaki H, Murase N, Okabe S, Ikeuchi T, et al. EXTL3/EXTR1 alterations in colorectal cancer cell lines. Int J Oncol. 1999;15(5):915-9.
146. Sembajwe LF, Katta K, Grønning M, Kusche-Gullberg M. The exostosin family of glycosyltransferases: mRNA expression profiles and heparan sulphate structure in human breast carcinoma cell lines. Bioscience reports. 2018;38(4).
147. Fernandez-Vega I, Garcia O, Crespo A, Castanon S, Menendez P, Astudillo A, et al. Specific genes involved in synthesis and editing of heparan sulfate proteoglycans show altered expression patterns in breast cancer. BMC Cancer. 2013;13:24.
148. Hwang JA, Kim Y, Hong SH, Lee J, Cho YG, Han JY, et al. Epigenetic inactivation of heparan sulfate (glucosamine) 3-O-sulfotransferase 2 in lung cancer and its role in tumorigenesis. PloS one. 2013;8(11):e79634.
149. Chen Z, Fan JQ, Li J, Li QS, Yan Z, Jia XK, et al. Promoter hypermethylation correlates with the Hsulf-1 silencing in human breast and gastric cancer. International journal of cancer. 2009;124(3):739-44.
150. Li J, Mo ML, Chen Z, Yang J, Li QS, Wang DJ, et al. HSulf-1 inhibits cell proliferation and invasion in human gastric cancer. Cancer science. 2011;102(10):1815-21.
151. Lai JP, Sandhu DS, Yu C, Han T, Moser CD, Jackson KK, et al. Sulfatase 2 up-regulates glypican 3, promotes fibroblast growth factor signaling, and decreases survival in hepatocellular carcinoma. Hepatology (Baltimore, Md). 2008;47(4):1211-22.
152. Lai JP, Oseini AM, Moser CD, Yu C, Elsawa SF, Hu C, et al. The oncogenic effect of sulfatase 2 in human hepatocellular carcinoma is mediated in part by glypican 3-dependent Wnt activation. Hepatology (Baltimore, Md). 2010;52(5):1680-9.
153. Lanzi C, Cassinelli G. Heparan Sulfate Mimetics in Cancer Therapy: The Challenge to Define Structural Determinants and the Relevance of Targets for Optimal Activity. Molecules. 2018;23(11):36.
154. Bosse KR, Raman P, Zhu Z, Lane M, Martinez D, Heitzeneder S, et al. Identification of GPC2 as an Oncoprotein and Candidate Immunotherapeutic Target in High-Risk Neuroblastoma. Cancer cell. 2017;32(3):295-309.e12.
155. Zheng X, Gai X, Han S, Moser CD, Hu C, Shire AM, et al. The human sulfatase 2 inhibitor 2,4-disulfonylphenyl-tert-butylnitrone (OKN-007) has an antitumor effect in hepatocellular carcinoma mediated via suppression of TGFB1/SMAD2 and Hedgehog/GLI1 signaling. Genes, chromosomes & cancer. 2013;52(3):225-36.
156. Nadanaka S, Kagiyama S, Kitagawa H. Roles of EXTL2, a member of the EXT family of tumour suppressors, in liver injury and regeneration processes. Biochem J. 2013;454(1):133-45.
157. Motoyama T, Hojo H, Watanabe H. Comparison of seven cell lines derived from human gastric carcinomas. Acta pathologica japonica. 1986;36(1):65-83.
158. Adli M. The CRISPR tool kit for genome editing and beyond. Nature communications. 2018;9(1):1911.
159. Hingorani R, Deng J, Elia J, McIntyre C, Mittar D. Detection of apoptosis using the BD annexin V FITC assay on the BD FACSVerse™ system. BD Biosciences, San Jose. 2011:1-12.
160. Mikami T, Kitagawa H. Biosynthesis and function of chondroitin sulfate. Biochimica et biophysica acta. 2013;1830(10):4719-33.

161. Nandini CD, Sugahara K. Role of the sulfation pattern of chondroitin sulfate in its biological activities and in the binding of growth factors. Advances in pharmacology (San Diego, Calif). 2006;53:253-79.
162. Liang CC, Park AY, Guan JL. In vitro scratch assay: a convenient and inexpensive method for analysis of cell migration in vitro. Nature protocols. 2007;2(2):329-33.
163. Purnomo E, Emoto N, Nugrahaningsih DA, Nakayama K, Yagi K, Heiden S, et al. Glycosaminoglycan overproduction in the aorta increases aortic calcification in murine chronic kidney disease. Journal of the American Heart Association. 2013;2(5):e000405.
164. Zhu JX, Goldoni S, Bix G, Owens RT, McQuillan DJ, Reed CC, et al. Decorin evokes protracted internalization and degradation of the epidermal growth factor receptor via caveolar endocytosis. The Journal of biological chemistry. 2005;280(37):32468-79.
165. Brooks R, Williamson R, Bass M. Syndecan-4 independently regulates multiple small GTPases to promote fibroblast migration during wound healing. Small GTPases. 2012;3(2):73-9.
166. Kim DH, Wirtz D. Focal adhesion size uniquely predicts cell migration. FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 2013;27(4):1351-61.
167. Gopal S, Bober A, Whiteford JR, Multhaupt HA, Yoneda A, Couchman JR. Heparan sulfate chain valency controls syndecan-4 function in cell adhesion. The Journal of biological chemistry. 2010;285(19):14247-58.
168. Bolós V, Gasent JM, López-Tarruella S, Grande E. The dual kinase complex FAK-Src as a promising therapeutic target in cancer. OncoTargets and therapy. 2010;3:83-97.