NEW TUMOUR MARKERS FOR PANCREATIC CANCER BASED ON THE ALTERED GLYCOSYLATION OF SERUM GLYCOPROTEINS

Adrià Duran Sidera

Universitat de Girona

DOCTORAL THESIS

New tumour markers for pancreatic cancer based on the altered glycosylation of serum glycoproteins

Adrià Duran Sidera

2023

Universitat de Girona

DOCTORAL THESIS

New tumour markers for pancreatic cancer based on the altered glycosylation of serum glycoproteins

Adrià Duran Sidera

2023

Doctorate programme in Molecular Biology, Biomedicine and Health

Thesis supervisors

Dr. Rosa Peracaula Miró   Dr. Esther Llop Escorihuela

Tutor

Dr. Rosa Peracaula Miró

This thesis is submitted in fulfilment of the requirements to obtain the doctoral degree from the Universitat de Girona

Universitat
de Girona

Dr. Rosa Pearacaula Miró and Dr. Esther Llop Escorihuela, form Universitat de Girona,

DECLARE:

That the work entitled New tumour markers for pancreatic cancer based on the altered glycosylation of serum glycoproteins, presented by Adrià Duran Sidera for the obtention of the title of Doctor, has been realised under our direction and meets the requirements to be eligible for an International Mention.

And, for the record and all intents and purposes, we sign this document.

Dr. Rosa Peracaula Miró   Dr Esther Llop Escorihuela

Girona, 2023

Agraïments

Sembla que es comença a veure l'arribada d'aquesta marató que és la tesi doctoral, i això em fa pensar en totes les persones que m'han ajudat i donat el seu suport per tal de poder arribar fins aquest punt. Ja sabeu que normalment soc de poques paraules, espero amb les següents poder reflectir tot l'agraïment que sento.

En primer lloc vull donar les gràcies a les meves directores de tesis per tot el suport, aprenentatge i oportunitats que he rebut durant aquests anys. A l'Esther, per estar sempre al meu costat des d'aquell dia en què li vaig demanar per fer el TFG, i per tot el que m'has ensenyat tant dins com fora del laboratori; gràcies per introduir-me en aquest meravellós món de la investigació. I a la Rosa, per tota l'experiència, el coneixement i la constància per tirar endavant els projectes, per ajudar-me a creure en mi mateix en els moments més difícils i per trobar sempre el costat bo a tot allò que hem fet. També em vull recordar de la resta de membres de Bioquímica del càncer. A en Rafa, per tots els comentaris dels lab meetings i ensenyar-me que sempre es pot anar un pas més enllà. I a la Montse, per la seva proximitat, el seu somriure i fer-me sentir acompanyat en tantes hores al laboratori.

I com no, infinites gràcies als meus compis de doctorat bioquímics! A en Pedro, per acollir-me tan bé quan vaig arribar. Per tots els llisats, westerns i IPs al labo-ratori, i per l'alegria i bogeria que portaves al despatx i a qualsevol cosa que féssim. A la Laura, per estar sempre al meu costat -literalment-, per tot el que hem com-partit i per haver-me ajudat a créixer tant professionalment com personal. A l'Anna, que per moltes coses que escrigui no serien suficients. Gràcies per totes les xerrades, confidències, putivueltas i moments viscuts, per tots els somriures i també les llàgrimes. I a l'Àlex, amb qui vaig tenir la sort de coincidir durant el grau i hem pogut compartir tants anys junts. Gràcies per la nostra amistat, per tenir sempre les paraules adequades, per poder-nos entendre amb un sol gest, i per tot el que ens hem divertit. Tinc molt clar que aquesta etapa no hauria sigut el mateix sense vosaltres.

També vull tenir un petit record per tothom que ha passat per l'àrea de bioquímica durant aquesta etapa, començant per en Marlon i continuant per totes les estudiants de TFM, TFG i pràctiques en empresa, que d'una manera o altra (des d'un simple hola cada matí fins incomptables hores al laboratori) m'han acompanyat al llarg de la tesi. I a les doctorandes de més recent incorporació, espero que tot us vagi molt bé.

Moltes gràcies també a la resta de doctorandes del departament, que estaven reunides a "la granja" quan vaig arribar però que ja han colonitzat tot el LEAR. Per totes les pauses i cafès, per tenir sempre una bona conversa i una distracció a l'hora de dinar, per ser allà en els bons moments i en els no tan bons, i per tot el que hem fet fora de la universitat: sopars, aniversaris, festes, esquiades, calçotades, dies de platja o trobades sota qualsevol excusa.

I would like to specially acknowledge Dr. Radka Saldova to receive me into NIBRT's Glycoscience group and to teach me so much glycoscience. Thank you Hayden, Kieran, Laura, Ana Lucia and Maitrayee for all the fun during these three months in Dublin, in and out of the lab. And many thanks to Craig and Zoe, who welcomed me into their home and made all things much easier.

I per acabar, gràcies a la família, papa, mama i Dani. Per donar-me suport en tot el que faig, per animar-me sempre a seguir endavant i per confiar en mi. Tot és més fàcil al vostre costat. Us estimo molt!

Publication resulting from this thesis

Financial support

List of abbreviations

Abbreviation	Description
2-AB	2-aminobenzamide
AAL	*Aleuria aurantia* lectin
ABS	*Arthobacter ureafaciens* sialidase
ACN	Acetonitrile
AFP	α-Fetoprotein
AGP	α1-acid glycoprotein
Asn / N	Asparagine
Asp / D	Aspartic acid
ATCC	America Type Culture Collection
AUC	Area under the curve
BSA	Bovine serum albumin
CA125	Carbohydrate antigen 125
CA19-9	Carbohydrate antigen 19-9
CEA	Carcinoembryonic antigen
Cf-MSLN	Core fucosylated mesothelin
ChP	Chronic pancreatitis
CMP	Cytidine-5'-monophosphate
CTC	Circulating tumour cell
ctDNA	Circulating tumour DNA
DMEM	Dulbecco's modified Eagle's Medium
DNA	Deoxyribonucleic acid
ELISA	Enzyme-linked immunosorbent assay
ELLA	Enzyme-linked lectin assay
EMEM	Eagle's Minimum Essential Medium
EMT	Epithelial-to-mesenchymal transition
ER	Endoplasmic reticulum
ESI	Electrospray ionisation
FA	Formic acid
FBS	Fetal bovine serum
FDA	Food and drugs administration
Fuc	Fucose
Gal	Galactose
GalNAc	*N*-acetilgalactosamine
Glc	Glucose

GlcNAc	*N*-acetilglucosamine
GPI anchor	Glycosylphosphatidylinositol anchor
GTP	Guanosine-5'-triphosphate
GU	Glucose unit
Hex	Hexose
HILIC	Hydrophilic interaction liquid chromatography
HPLC	High-performance liquid chromatography
IPMN	Intraductal papillary mucinous neoplasm
LC	Liquid chromatography
LCA	*Lens culinaris* agglutinin
Le^a	Lewis a
Le^b	Lewis b
Le^x	Lewis x
Le^y	Lewis y
LOD	Limit of detection
LOQ	Limit of quantification
MAL-I	*Maackia amurensis* I lectin
MAL-II	*Maackia amurensis* II lectin
MALDI	Matrix assisted laser desorption/ionisation
Man	Mannose
MCN	Mucinous cystic neoplasm
MFAP4	Microfibril associated protein 4
MPF	Megakaryocyte-potentiating factor
MS	Mass spectrometry
MSLN	Mesothelin
MUC	Mucin
NAcHex	*N*-acetylhexosamine
NAN1	*Streptococcus pneumoniae* neuraminidase
NCCN	National Comprehensive Cancer Network
NCI	National cancer institute
Neu5Ac	*N*-acetylneuraminic acid
Neu5Gc	*N*-glycoylneuraminic acid
NIH	US National Institute of Health
NMR	Nuclear magnetic resonance
NPV	Negative predictive value
PaC	Pancreatic cancer

PanIN	Pancreatic intraepithelial neoplasia
PBS	Phosphate-buffered saline
PDAC	Pancreatic ductal adenocarcinoma
PHA-E	*Phaseolus vulgaris* erythroagglutinin
PHA-L	*Phaseolus vulgaris* leucoagglutinin
PhoSL	*Pholiota squarrosa* lectin
PNA	Peanut agglutinin
PNGaseF	Peptide:*N*-glycosydase F
poly-LacNAc	Poly-*N*-acetyllactosamine
PPV	Positive predictive value
PSA	Prostate specific antigen
PVDF	Polyvinylidene fluoride
QTof	Quadrupole time-of-flight
REG	Regenerating-islet derived proteins
REG1A	Regenerating-islet derived protein 1 alpha
REG1B	Regenerating-islet derived protein 1 beta
rMSLN	Recombinant MSLN
RNA	Ribonucleic acid
RNase 1	Ribonuclease 1
ROC curve	Receiver operating characteristic curve
RPMI	Roswell Park Memorial Institute medium
rREG1A	Recombinant REG1A
rREG1B	Recombinant REG1B
SA	Sialic acid
SDS-PAGE	Sodium dodecyl sulfaet polyacrylamide gel electrophoresis
Ser / S	Serine
Siglecs	Sialic acid-binding immunoglobulin-like lectins
sLe^a	Sialil-Lewis a
sLe^x	Sialil-Lewis x
SNA	*Sambucus nigra* agglutinin
SNFG	Symbol nomenclature for glycans
SRM	Selected reaction monitoring
ST	Sialyltransferase
STn	Sialyl-Thomsen-nouveau antigen
T antigen	Thomsen-Friedenreich antigen
TFA	Trifluoroacetic acid

Thr / T	Threonine
Tn antigen	Thomsen-nouveau antigen
TNM	Tumour node metastasis
Tof	Time-of-flight
TQ	Triple quadrupole
UEA	*Ulex europaeus* agglutinin
UICC	International Union Against Cancer
UPLC	Ultra-performance liquid chromatography
UTP	Uridine-5'-triphosphate
VVL	*Vicia villosa* lectin
WB	Western blotting

List of figures

List of tables

Table of contents

Summary

Pancreatic cancer (PaC) is the deadliest of all tumours, with a 5-year survival below 12% and a mortality/incidence ratio of 94.5%. One of the main reasons behind this dismal prognosis is the diagnosis of the disease at late stages, usually when metastasis have already occurred and no effective therapies are available. Nowadays, there is no accurate tumour marker for the detection of PaC. Only the use of the carbohydrate antigen 19-9 (CA19-9) is clinically validated for the management and recurrence evaluation of the disease, but lacks specificity to be used as a diagnostic test. For this reason, the discovery of novel biomarkers able to detect PaC in early stages, when current therapies are still effective, is of utmost significance for the research and medical communities.

Altered glycosylation, a common feature of cancer, stands as a potential source for developing new tumour markers. However, a specific glycan signature to diagnose PaC has not been described. As glycosylation is one of the main post-translational modification of proteins, we hypothesise that the combinatorial analysis of tumour-associated glycan structures on overexpressed PaC proteins could outperform the sensitivity and specificity of current methodologies. In this regard, we have studied the glycosylation pattern of two *neo-*/over-expressed proteins in PaC, mesothelin (MSLN) and regenerating islet-derived protein 1 (REG1), and we have assessed the potential of their glycoforms as PaC tumour markers.

On one hand, MSLN expression was assessed by western blotting (WB) on PaC cell lines protein lysates and conditioned media. MSLN was observed in the seven studied cell lines, with higher amounts on the secreted media. MSLN N-glycosylation was characterised in those cell lines with higher expression (Capan-2, AsPC-1 and HPAF-II). PNGaseF digestion, peptide mapping and total glycoprotein mass analysis showed that the three MSLN N-glycan sites were occupied. The structural characterisation of fluorescently labelled MSLN N-glycans by UPLC (in a HILIC column) combined with several exoglycosidase digestions revealed the expression of complex type sialylated structures, most of which were core fucosylated and highly branched. Next, the study of immunopurified MSLN N-glycan

determinants using lectins indicated the potential to use the lectins AAL (fucose recognition), PhoSL (core fucose recognition), PHA-E (bisected GlcNAc recognition) and PHA-L (ß1,6-antenna recognition) to detect specific glycan moieties on MSLN.

MSLN expression was also assessed in pancreatic tissues corresponding to tumours or healthy adjacent structures. MSLN expression was detected in 77.4% of tumours, while just one non-tumour sample (10%) expressed the protein. Then, we developed a methodology for the quantification of specific MSLN glycoforms in biological samples, as tissue lysates or blood serum. The method consisted in the quantification of MSLN protein levels by an enzyme-linked immunosorbent assay (ELISA) followed by the identification of specific core fucosylated MSLN (Cf-MSLN) glycoforms in an enzyme-linked lectin assay (ELLA) using the PhoSL lectin. Application of this methodology in a cohort of patients including PaC, chronic pancreatitis and healthy individuals showed a significant decrease in Cf-MSLN abundancy in PaC samples. Moreover, combination of this biomarker with CA19-9 enhanced its potential to be used as a diagnostic tumour marker, yielding a sensitivity and specificity of 93.3% and 81.2%, respectively.

On the other hand, REG1 expression was also assessed by WB on PaC cell lines and pancreatic tissues. While no expression was observed in cultured PaC cell lines, 90.9% of non-tumour tissues and 60.5-65.8% of PaCs expressed REG1. The analysis of REG1 *O*-glycosylation through WB with lectins on immunopurified REG1 showed that, while the lectin PNA (galactose recognition) could only detect control samples, the lectins SNA (α2,6-sialic acid recognition) and VVL (GalNAc recognition) were only reactive in tumour samples. However, the development of an ELLA assay using these lectins as detection probes to quantify REG1 glycoforms was not possible due to the lack of analytical sensitivity. To overcome this limitation, probably caused by the use of lectins, we performed an in-depth analysis of REG1 glycoforms by mass spectrometry (MS). MALDI-Tof evaluation of the total glycoprotein and its *O*-glycopeptide obtained after trypsin digestion showed up to 16 different structures, which included the Tn antigen, core 1, core

2 and core 4 *O*-glycans. Next, we developed a LC-MS application consisting of a hydrophilic interaction liquid chromatography (HILIC) separation of glycopeptides followed by their detection in a triple quadrupole under a selected reaction monitoring (SRM) data acquisition to detect the most relevant glycoforms in complex biological samples. This workflow has already been successful in analysing a control and a PaC serum sample, and has shown the increased expression of sialyl-T and sialyl-Tn antigens on REG1A in malignant conditions. However, these are still preliminary results, and validation in a cohort of patients is required.

To sum up, we have been able to perform a comprehensive analysis of MSLN and REG1 glycosylation in PaC. In addition, we have developed new strategies to quantify their tumour-associated glycoforms. While the reduction of Cf-MSLN in PaC patients holds a promising potential to detect PaC, the quantification of specific REG1 glycoforms still needs to be tested in more serum samples. Overall, this work has provided candidate MSLN and REG1 glycoforms to be used as novel PaC biomarkers, which will require further validation in large cohorts of patients.

Resum

El càncer de pàncrees (PaC) és el més mortífer de tots els tumors, amb una supervivència als 5 anys inferior al 12% i una relació mortalitat/incidència del 94.5%. Una de les principals raons darrera aquesta mala prognosis és la diagnosis de la malaltia en estadis avançats, normalment quan ha aparegut metàstasi i no hi ha teràpies efectives disponibles. Actualment, no hi ha cap marcador tumoral prou precís per detectar el PaC. Únicament l'ús de l'antigen carbohidrat 19-9 (CA19-9) està validat clínicament pel seguiment i recurrència de la malaltia, però li falta especificitat per ser utilitzat com un test diagnòstic. Per aquesta raó, el descobriment de nous biomarcadors capaços de detectar el PaC en estadis primerencs, quan les teràpies actuals encara son efectives, és de gran importància per la comunitat mèdica i investigadora.

La glicosilació alterada, una característica comú del càncer, és una font potencial per desenvolupar nous marcadors tumorals. Tot i així, no s'ha descrit un perfil glucídic específic per diagnosticar el PaC. Ja que la glicosilació és una de les principals modificacions post-traduccionals de les proteïnes, la nostra hipòtesis és que l'anàlisi combinat d'estructures glucídiques associades a tumor en proteïnes sobreexpressades en PaC podriar superar la sensibilitat i especificitat del les metodologies actuals. En aquest sentit, hem estudiat el patró de glicosilació de dues proteïnes sobreexpressades en PaC, la mesotelina (MSLN) i la proteïna regeneradora 1 (REG1), i hem avaluat el potencial de les seves glicoformes com a marcadors tumorals pel PaC.

D'una banda, l'expressió de la MSLN va ser avaluada per western blotting (WB) en lisats proteics i medis condicionats de línies cel·lulars de PaC. La MSLN va ser observada a les set línies cel·lulars estudiades, en major quantitat al medi secretat. La N-glicosialció de la MSLN va ser caracteritzada en les línies cel·lulars amb major expressió (Capan-2, AsPC-1 i HPAF-II). La digestió amb PNGasaF, el mapeig de glicopèptids i l'anàlisi de la massa de la proteïna total van mostrar que els tres llocs d'N-glicosilació de al MSLN estaven ocupats. La caracterització estruc-

tural dels *N*-glicans de la MSLN marcats fluorescentment per UPLC (en una columna HILIC) combinada amb vàries digestions amb exoglicosidases va revelar l'expressió d'estructures complexes sialilades, la majoria d'elles amb fucosa core i altament ramificades. A continuació, l'estudi dels determinants glucídics de la MSLN va mostrar el potencial ús de les lectines AAL (reconeix fucosa), PhoSL (reconeix fucosa core), PHA-E (reconeix el *bisected* GlcNAc) i PHA-L (reconeix l'antena ß1,6) per detectar glúcids específics en MSLN.

L'expressió de la MSLN també es va avaluar en teixits de tumor de pàncrees i estructures sanes adjacents. Es va detectar expressió de MSLN en el 77.4% dels tumors, mentre que només una mostra no tumoral (10%) expressava la proteïna. A continuació, vam desenvolupar una metodologia per quantificar glicoformes específiques de MSLN en mostres biològiques complexes, com lisats de teixits o mostres de sèrum sanguini. El mètode consistia en la quantificació dels nivells proteics de MSLN per ELISA seguit per la identificació de les glicoformes de MSLN amb fucosa core (Cf-MSLN) amb un assaig de lectina lligada a enzim (ELLA) utilitzant la lectina PhoSL. L'ús d'aquesta metodologia en una cohort de pacients que incloïa individus amb PaC, pancreatitis crònica i individus sans va mostrar una disminució significant en l'abundància de Cf-MSLN en les mostres de PaC. A més, la combinació d'aquest biomarcador amb el CA19-9 va millorar el seu potencial per ser utilitzat com un marcador tumoral de diagnosi, amb una sensibilitat i especificitat del 93.3% i 81.2%, respectivament.

D'altra banda, l'expressió de REG1 també va ser determinada per WB en línies cel·lulars de PaC i teixits pancreàtics. Mentre que no es va observar expressió en línies cel·lulars de PaC, es va veure REG1 en el 90.9% dels teixits no tumorals i el 60.5-65.8% dels teixits de PaC. L'anàlisi de l'*O*-glicosilació de REG1 mitjançant WB amb lectines sobre REG1 immunopurificada va mostrar que, mentre la lectina PNA (reconeix galactosa) només podia reconèixer mostres control, les lectines SNA (reconeix àcid siàlic en enllaç α2,6) i VVL (reconeix GalNAc) només reaccionaven amb mostres tumorals. Tot i així, el desenvolupament d'un assaig ELLA utilitzant aquestes lectines com a sonda de detecció no va ser possible per falta de

sensibilitat analítica. Per superar aquesta limitació, probablement causada per l'ús de lectines, vam fer un anàlisi a fons de les glicoformes de REG1 per espectrometria de masses (MS). L'avaluació per MALDI-Tof de la glicoproteïna i del seu *O*-glicopèptid obtingut per digestió amb tripsina va mostrar fins a 16 estructures diferents, que incloïen l'antigen Tn, i *O*-glycans de core 1, core 2 i core 4. A continuació, vam desenvolupar una aplicació LC-MS que consistia en la separació de glicopèptids en una cromatografia líquida d'interacció hidrofílica (HILIC) seguit de la seva detecció en un triple quadrupol utilitzant el monitoreig de reaccions seleccionades (SRM) com a mètode d'adquirir les dades per detectarles glicoformes més rellevants en mostres biològiques complexes. Aquest flux de treball ja ha permès analitzar una mostra de sèrum control i una de PaC amb èxit, i ha mostrat l'augment d'expressió dels antígens sialil-T i sialil-Tn en la REG1A de condicions malignes. Tot i així, aquests resultats encara son preliminars, i la seva validació en una cohort de pacients és necessària.

En resum, hem sigut capaços de fer un anàlisi integral de la glicosilació de la MSLN i REG1 en PaC. A més, hem desenvolupat noves estratègies per quantificar les seves glicoformes associades a tumor. Mentre que la reducció de la Cf-MSLN en pacients de PaC té un potencial prometedor per detectar el PaC, la quantificació de glicoformes espepecífiques en REG1 encara necessita ser provada en més mostres de sèrum. En general, aquest treball ha proporcionat glicoformes de MSLN i REG1 com a candidates per ser utilitzades com nous biomarcadors pel PaC, cosa que requerirà futures validacions en grans cohorts de pacients.

xx

Resumen

El cáncer de páncreas (PaC) es el más mortífero de todos los tumores, con una supervivencia a 5 años inferior al 12% y una relación mortalidad/incidencia del 94.5%. Una de las principales razones detrás de esta mala prognosis es la diagnosis de la enfermedad en estadios tardíos, normalmente cuando ya ha aparecido metástasis y no hay terapias efectivas disponibles. Actualmente, no hay ningún marcador tumoral lo suficientemente preciso para detectar el PaC. Solo el uso del antígeno carbohidrato 19-9 (CA19-9) está validado clínicamente para el seguimiento y recurrencia de la enfermedad, pero le falta especificidad para ser utilizado como test diagnóstico. Por esta razón, el descubrimiento de nuevos biomarcadores capaces de detectar el PaC en estadios tempranos, cuando las terapias actuales todavía son efectivas, es de gran importancia para la comunidad médica e investigadora.

La glicosilación alterada, una característica común del cáncer, es una fuente potencial para desarrollar nuevos marcadores tumorales. Sin embargo, no se ha descrito ningún perfil glucídico específico para diagnosticar el PaC. Puesto que la glicosilación es una de las principales modificaciones post-traduccionales de las proteínas, nuestra hipótesis es que el análisis combinado de estructuras glucídicas asociadas a tumor en proteínas sobreexpresadas en PaC podría superar la sensibilidad y especificidad de las metodologías actuales. En este sentido, hemos estudiado el patrón de glicosilación de dos proteínas sobreexpresadas en PaC, la mesotelina (MSLN) y la proteína regeneradora 1 (REG1), y hemos evaluado el potencial de sus glicoformas como marcadores tumorales para el PaC.

Por un lado, la expresión de la MSLN fue evaluada por western blotting (WB) en lisados proteicos y medios condicionados de líneas celulares de PaC. La MSLN fue observada en las siete líneas celulares estudiadas, en mayor cantidad en el medio secretado. La N-glicosilación de la MSLN fue caracterizada en las líneas celulares con mayor expresión (Capan-2, AsPC-1 y HPAF-II). La digestión con PNGasaF, el mapeo de glicopéptidos y el análisis de la masa de la proteína total mostraron que los tres lugares de N-glicosilación de la MSLN estaban ocupados.

La caracterización estructural de los *N*-glicanos de la MSLN marcados fluorescentemente por UPLC (en una columna HILIC) combinada con varias digestiones con exoglicosidasas reveló la expresión de estructuras complejas sialiladas, la mayoría de ellas con fucosa core y altamente ramificadas. A continuación, el estudio de los determinantes glucídicos de la MSLN mostró el potencial uso de las lectinas AAL (reconoce fucosa), PhoSL (reconoce fucosa core), PHA-E (reconoce el bisected GlcNAc) y PHA-L (reconoce la antena ß1,6) para detectar glúcidos específicos en MSLN.

La expresión de la MSLN también se evaluó en tejidos de tumor de páncreas y estructuras sanas adyacentes. Se detectó expresión de MSLN en el 77.4% de los tumores, mientras que sólo una muestra no tumoral (10%) expresaba la proteína. A continuación, desarrollamos una metodología para cuantificar glicoformas específicas de MSLN en muestras biológicas complejas, como lisados de tejidos o muestras de suero sanguíneo. El método consistió en la cuantificación de los niveles proteicos de MSLN por ELISA seguido de la identificación de las glicoformas de MSLN con fucosa core (Cf-MSLN) con un ensayo de lectina ligada a enzima (ELLA) utilizando la lectina PhoSL. El uso de esta metodología en una cohorte de pacientes que incluía a pacientes con PaC, pancreatitis crónica e individuos sanos mostró una disminución significativa en la abundancia de Cf-MSLN en las muestras de PaC. Además, la combinación de este biomarcador con el CA19-9 mejoró su potencial para ser utilizado como marcador tumoral de diagnosis, con una sensibilidad y especificidad del 93.3% y 81.2%, respectivamente.

Por otra parte, la expresión de REG1 también fue determinada por WB en líneas celulares de PaC y tejidos pancreáticos. Mientras que no se observó expresión en líneas celulares de PaC, se vio REG1 en el 90.9% de los tejidos no tumorales y el 60.5-65.8% de los tejidos de PaC. El análisis de la *O*-glicosilación de REG1 mediante WB con lectinas sobre REG1 inmunopurificada mostró que, mientras la lectina PNA (reconoce galactosa) sólo podía reconocer muestras control, las lectinas SNA (reconoce ácido siálico en enlace α2,6) y VVL (reconoce GalNAc) solo reaccionaban con las muestras tumorales. Sin embargo, el desarrollo de un ensayo

ELLA utilizando estas lectinas como sonda de detección no fue posible por falta de sensibilidad analítica. Para superar esta limitación, probablemente causada por el uso de lectinas, realizamos un análisis a fondo de las glicoformas de REG1 por espectrometría de masas (MS). La evaluación por MALDI-Tof de la glicoproteína y de su *O*-glicopéptido obtenido por digestión con tripsina mostró hasta 16 estructuras diferentes, que incluían el antígeno Tn, y *O*-glycanes de core 1, core 2 y core 4. A continuación desarrollamos una aplicación de LC-MS que consistía en la separación en de glicopéptidos en una cromatografía líquida de interacción hidrofílica (HILIC) seguido de su detección en un triple cuadrupolo utilizando el monitoreo de reacciones seleccionadas (SRM) como método de adquirir los datos para detectar las glicoformas más relevantes en muestras biológicas complejas. Este flujo de trabajo ya ha permitido analizar una muestra de suero control y una de PaC con éxito, mostrando el aumento de expresión de los antígenos sialil-T y sialil-Tn en la REG1A de condiciones malignas. Sin embargo, estos resultados todavía son preliminares, y su validación en una cohorte de pacientes es necesaria.

En resumen, hemos sido capaces de realizar un análisis integral de la glicosilación de la MSLN y la REG1 en PaC. Además, hemos desarrollado nuevas estrategias para cuantificar sus glicoformas asociadas a tumor. Mientras que la reducción de la Cf-MSLN en pacientes de PaC tiene un potencial prometedor para detectar el PaC, la cuantificación de glicoformas específicas en REG1 todavía necesita ser probada en más muestras de suero. En general, este trabajo ha proporcionado glicoformas de MSLN y REG1 como candidatas para ser utilizadas como nuevos biomarcadores para el PaC, lo que requerirá futuras validaciones en grandes cohortes de pacientes.

INTRODUCTION

Duran Sidera, A.

Cancer

The term cancer comprises a group of more than a hundred diseases characterised by deregulated and sustained cell growth that causes the formation of cell masses named tumours, which often cause organs malfunction and, in the worst cases, death [1,2]. In addition, these cancerous cells can spread through the bloodstream or lymphatic vessels to distant organs and form new tumours in a process called metastasis [3,4]. In a more detailed manner, the carcinogenic process normally begins with the accumulation of genetic alterations in oncogenes and tumour-suppressor genes, those in control of the cell cycle regulation and deoxyribonucleic acid (DNA) repair [5]. As a result, small cell masses start to grow. When reaching a certain size, solid tumours require new vascularization to obtain nutrients for their survival. At this point, new blood vessels are produced through angiogenesis, a process driven by hypoxia in which anaerobic metabolism, cell survival and invasion are induced [6]. Cells continue acquiring new capacities through mutations accumulation, and some are able to escape the primary tumour and invade nearby tissues. An epithelial-to-mesenchymal transition (EMT) is the most usual process by which these cells detach from adjacent cells and achieve motility and migratory abilities [7]. However, recent evidence supports alternative modes of dissemination as clustered-based migration, in which cells do not completely lose adhesion nor express mesenchymal markers [8]. Finally, the tumour can disseminate to other tissues as a result of the ability of malignant cells to intravasate into capillaries or lymphatic vessels surrounding the tumour, and to circulate through them avoiding the immune system. In the end, these cells can extravasate in a process similar to that used by leukocytes [9], and may start proliferating in a new organ to create a secondary tumour. The evasion of the immune system and the promotion of angiogenesis will be crucial in the growth of the metastatic tumour [10].

Cancer types are usually classified by their tissue of origin and cell type, being most of their characteristics different in each case. However, about twenty years ago, Hanahan & Weinberg listed six common capabilities cells may gather to become malignant, known as the hallmarks of cancer [11]. Two new capabilities and two

enabling characteristics were included years later, and have been enlarged up to fourteen hallmarks recently [12]. Some of these capacities include sustained proliferative signalling, replicative immortality, cell death resistance, invasion and metastasis activation or cell metabolism deregulation, among others (Figure 1).
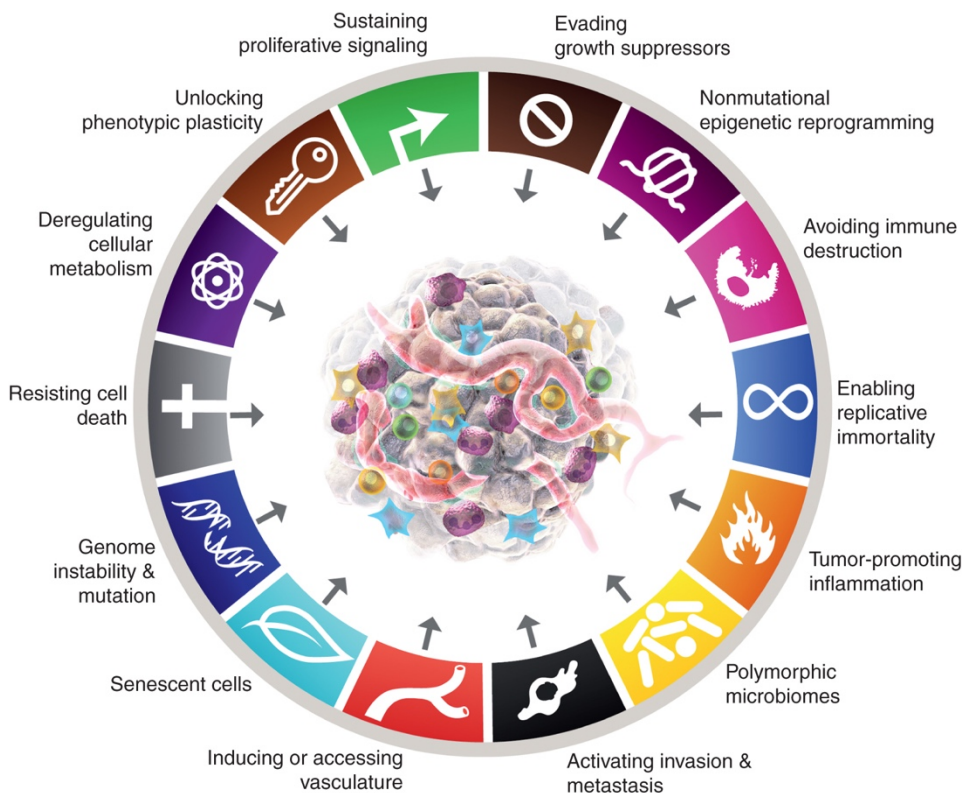
Figure 1. Hallmarks of cancer. Functional capabilities and enabling characteristics acquired by tumour cells in order to form malignant tumours. Extracted and modified from Hanahan, 2022 [12].

With about one in five deaths worldwide, cancer was the second cause of death in 2020 just behind cardiovascular diseases in both Europe and the US [13–15]. This turns the disease into a global sanitary issue that must be of concern for the medical and research community. In 2020, 19.3 million new cases and 9.9 million deaths occurred worldwide, which are predicted to increase until 28.4 million diagnosed cancers in 2040 [16]. Prostate, breast, lung and colorectal cancers are the principal diagnosed cancers in the United States [17]. These tumours also occupy the top positions in the mortality causes ranking, together with pancreatic malignancies rising to the second place. The situation is similar in Europe, with 4 million

new cases and 1.9 million deaths. The same malignancies stay on the first positions for incidence, but regarding mortality pancreatic neoplasms are found in the fourth position, just above prostate cancer [18].

Pancreatic cancer

Epidemiology and problematics

Pancreatic cancer (PaC) comprises several malignancies, which can be classified according to the cell type of origin: exocrine or neuroendocrine cells. Most of diagnosed PaC's, about 90%, correspond to pancreatic ductal adenocarcinomas (PDAC), an exocrine-based tumour with ductal epithelium phenotype [19–21]. For this reason, the general term PaC usually refers to PDAC, as will be in this thesis. Other exocrine tumours include acinar cell carcinomas, intraductal papillary-mucinous neoplasms and mucinous cystic neoplasms. The remaining neoplasms, corresponding to neuroendocrine tumours, can be distinguished between functional (produce hormones) and non-functional (do not produce hormones), being the latter the most common type [22]; or by their cell differentiation degree [23,24].
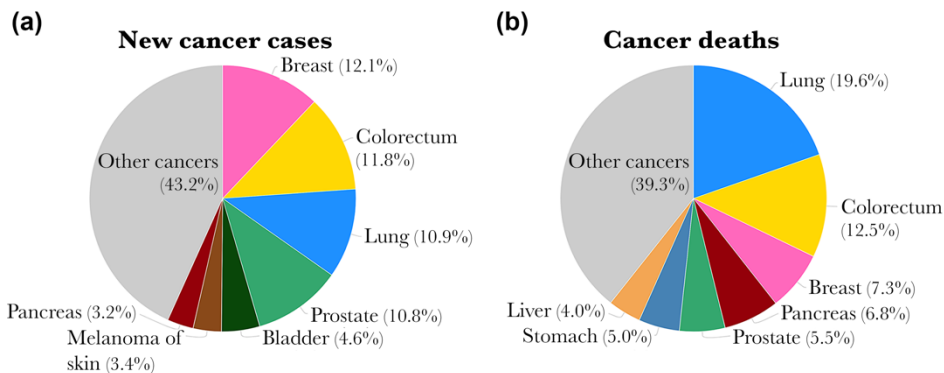
Figure 2. Cancer statistics for the European population in 2020. Localisation and percentage of **(a)** diagnosed new cases and **(b)** cancer deaths in Europe in 2020. Modified from GLOBOCAN 2020, Global Cancer Observatory: https://gco.iarc.fr [25].

PaC ranks seventh in cancer incidence by site in Europe (Figure 2.a) [18]. However, it is the fourth case of death (Figure 2.b), presenting a mortality/incidence ratio of 94.5%. This raise in the ranking is due to a dismal 5-year survival rate of

just 12%, the lowest for all cancer types. In addition, about half PaC cases are diagnosed at distant stages, when survival is of just 3%, and only 15% PaCs are detected at a localized phase, which present a 44% survival at 5 years [17]. Thus, PaC management presents a dual problematic: it lacks proper tools for an early diagnosis, when current treatments could be more efficient, as well as novel therapeutic options to conveniently fight the disease [26,27].

Precursor lesions and disease development

The human pancreas is an organ with a dual role, with exocrine and endocrine functions. The exocrine part, responsible of secreting digestive enzymes into the gut, is made up of acinar cells and duct cells, which produce the enzymes and sodium bicarbonate constituting the pancreatic juice, respectively. On the other hand, the endocrine gland consists in five types of secretory islet cells which secrete peptide hormones into the bloodstream to maintain glucose homeostasis [28]. Therefore, these secretory functions require a delicate regulation, as inappropriate activation or inactivation can produce severe health impacts [29].

Anatomically, the pancreas is a retroperitoneal organ of about 15-25 cm in length and 100-150g of weight which produces and pours around a kilogram of pancreatic juice daily into the duodenum. The organ is strongly vascularized by several major arteries, is highly innervated, and is connected with other abdominal organs as the stomach, duodenum, colon and spleen. Despite being basically divided in three parts (head, body and tail), it can be considered as four structurally distinct components: the exocrine pancreas, the endocrine pancreas, the blood vessels and the extracellular space. The exocrine portion, which accounts for almost the 80% of the total volume, is a blind-ended ductal system made up of several acini grouped into lobules and connected by tubules. Highly orientated acinar cells linked by gap junctions and with their apical membrane faced towards the central lumen form each acinus. These cells secrete α-amylases, lipases and proteases, which flow through ductal epithelial cells canaliculus and are finally discharged into the duodenum [28]. Along with the digestive enzymes, the pancreatic juice is enriched with sodium bicarbonate secreted by the duct cells, which neutralize the

duodenum content and provides an optimal pH for pancreatic digestive enzymes [30].

The endocrine pancreas is morphologically distinct from the exocrine part and is composed of highly vascularized spherical clusters of cells known as the islets of Langerhans. Accounting for the 1-2% of the pancreatic volume, these islets are composed by five major cell types that secrete distinct hormones. Glucagon, insulin, somatostatin, pancreatic polypeptide and ghrelin are produced by α-cells, ß-cells, δ-cells, F-cells and ε-cells, respectively [28]. Due to the lack of basal membranes, the islets are found within the exocrine acini, so that its secretory products can alter the acinar function [30]. In this regard, several exocrine diseases could be launched by endocrine malfunction, and vice versa [31–34].

The accepted PaC progression model defines a multi-step process originated by non-malignant lesions that finally become malignant through the acquisition of several genetic mutations (Figure 3). Morphologically, after an acinar-to-ductal metaplasia event (in which acinar cells promote to a more ductal-like phenotype) [35], pancreatic intraepithelial neoplasia (PanIN) occurs, followed by increased desmoplasia, *in situ* carcinoma and finally invasive adenocarcinoma [36,37]. In addition to PanINs, which are the most common precursor lesions, PaC can arise from other precursor lesions including intraductal papillary mucinous neoplasms (IPMNs) and mucinous cystic neoplasms (MCN) [38–40].

A PanIN is a non-invasive microscopic (<5 mm) lesion in the smaller pancreatic ducts which presents architectural atypia [39,41]. Depending on its cytological and histological abnormality degree, these lesions can be classified into three groups [39]. Low-grade lesions with minimal atypia compose PanIN-1 lesions, which are further divided in PanIN-1A (flat epithelium) and PanIN-1B (papillary epithelium). PanIN-2 show moderate cytological and architectural atypia with frequent papillae and nuclear alterations. Finally, high-grade PanINs (PanIN-3), also referred as *in situ* carcinoma, show severe atypia and dysplasia, usually with papillary morphology. However, it must be noted that none of these lesions trespass the

basement membrane, and thus an extra step is still necessary for them to become invasive.
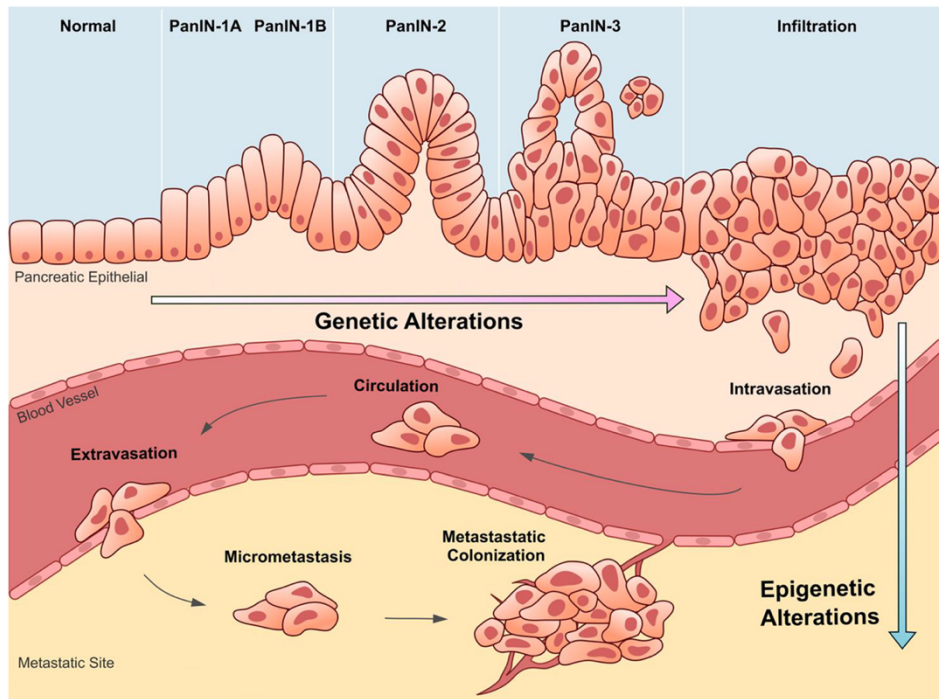
Figure 3. PaC progression model. Development of PaC from normal epithelia to invasive and metastatic cancer, progressing through different PanIN grades thanks to the accumulation of different genetic mutations that finally, together with other epigenetic alterations, let the tumour invade distant tissues. Extracted and modified from Wang *et al.*, 2021 [42].

All these morphological changes are driven by some well-established genetic and epigenetic changes [39–42]. Early mutation in the Kirsten rat sarcoma (*KRAS*) gene is presumably the initiator of PaC [43]. Actually, more than 90% PaCs present activating alterations in this gene, usually through point mutations in codon 12 or 13. This activation translates to enhanced proliferation through activation of MAPK or PI3K transduction cascades, constituting a first step towards a more malignant phenotype [44,45]. Telomer shortening is another initial event in PaC transformation, which induces chromosome instability and promotes neoplastic progression in cells [46].

Progression to increased dysplasia and higher-grade PanIN is associated with the loss of mainly three tumour suppressor genes. Inactivation of *CDKN2A* is already detectable at PanIN-2, while loss of *TP53* and *SMAD4* is observed in PanIN-3 stages [47,48]. These alterations cause the deregulation of the cell cycle and apoptosis mechanisms, thus promoting cell division and impairing cell death. A number of other genetic alterations, with lower prevalence, are also altered in PaCs [49–51]. Finally, when becoming invasive, PaC cells experience a huge genomic instability that cause changes in structural variants and gene copy number affecting known oncogenes and tumour-suppressor genes. They also undergo chromosomal changes including polyploidisation and chromothripsis, which cause gains and losses in PaC driver genes [36].

PaC cell-of-origin is a controversial topic, and contrarily to the presented acinar-to-ductal metaplasia initiation and PanIN progression, several authors have described a ductal origin for the malignancy which results in different PaC subtypes [52]. Despite *KRAS* and *TP53* remain as the driving genes [53], different progression models and phenotypes can be observed [54,55].

Many other genes present point mutations with much lower prevalence. These can be aggregated into several molecular pathways including RAS signalling, regulation of the cell cycle, TGFβ signalling, JUN aminoterminal kinase signalling, integrin signalling, WNT–Notch signalling, Hedgehog signalling, apoptosis, DNA damage control, small GTPase-dependent signalling, invasion and homophilic cell adhesion and embryonic regulators of axon guidance [36]. All these possible genetic alterations derive in a great heterogeneity of tumours [56]. Moreover, differences in the mutation landscape of single cells generate a remarkable intratumour heterogeneity and the presence of several genetic (and transcriptomic) subpopulations [57]. Interestingly, secondary tumour subpopulations are represented in the original one, indicating a high mutational conservation between paired primary and metastatic tumours [58].

Opposed to DNA mutations, PaC subtypes have also been classified based on its transcriptional landscape. Classic and basal-like have emerged as two consensus

groups, with a third "hybrid" capturing those with overlapping features [59]. There is potential to differentiate these into further subclasses, and multiple types have potential to co-exist within individual tumours [60]. Actually, single-cell RNA sequencing platforms have shown up to seven coexisting clusters of neoplastic cell subpopulations [61]. Interestingly, Peng and collaborators [57] found two ductal cell types with different transcriptomic profiles, and also identified novel genes participating in the carcinogenic process as *EGLN3*, *MMP9* and *PLAU*. Notably, recent clinical trials have shown that different transcriptional subtypes present different responses to chemotherapy, and hence sorting of patients could derive in a better selection on the therapeutic strategy [62].

Several evolutionary models to explain progression from normal ductal epithelium to metastasis have been proposed, with special emphasis on the required time [36]. It is estimated that the time from the first somatic mutation until the origin of invasive cells is about 12 years, with 7 more years to the development of metastatic cells and 3 years from then until patient death [58]. Thus, according to this relatively large time intervals, valuable opportunities for intervention are present.

Apart from all these reported alterations in pancreatic ductal cells, PaC is also characterised by the presence of a dense stroma that occupies up to 80% of the tumour volume and constitutes a supportive microenvironment for the tumour growth [63]. It is mainly composed of cancer-associated fibroblasts, vascular structures and extracellular matrix rich in collagen, hyaluronan and fibronectin. This stroma may promote cell proliferation and acts as a barrier that impairs drug delivery and cell infiltration, restraining chemotherapy and other therapies effectiveness. On the other hand, PaC tumour microenvironment is also characterised by the presence of immune cells such as myeloid-derived suppressor cells, tumourassociated macrophages, neutrophils and regulatory T cells, which release cytokines and generate an immunosuppressive environment [59,64]. Hence, the tumour microenvironment also plays a key role in PaC progression and must be considered as a source of biomarkers and a target for therapy.

Risk factors and diagnosis

Several risk factors are associated with the development of PaC. As it generally happens in cancer, PaC risk is increased with age. Gender and ethnicity may also affect PaC risk. Increased incidence is observed in men respect women, and in black versus white individuals [65]. Diabetes is another factor to consider, as well as several genetic mutations (although family related PaC is uncommon). Genetic syndromes involved in PaC include mutations in *BRCA1/BRCA2*, *CDKN2A*, *PRSS1*, *MLH1*, *MSH2*, *STK11* and other not known genes [66,67].

Tobacco use and body overweight stand as the principal risk factors that can be modified. Smokers have 74% increased chance of undergoing PaC respect people who have never smoked. In a similar way, obese people are more likely to develop PaC. Chronic pancreatitis (ChP) also increases PaC risk. This factor, as obesity, can be prevented through a controlled and varied diet, alcohol evasion and physical activity. Human microflora is another aspect to be aware of, which can also be modifiable [66]. Finally, exposure to chemicals such as hydrocarbon compounds, pesticides, heavy metals and fine particles may also increase PaC risk and should thus be avoided [68].

PaC symptomatology is diffuse and mild, and by the time it is recognised, the tumour is usually very large and has even expanded. Jaundice is the most prevalent symptom [69,70], despite PaC is the less common cause of jaundice, which is mainly induced by hepatocellular diseases, intrahepatic cholestasis and extrahepatic biliary tract obstruction [71]. Other generic symptoms that can be related to PaC include abdominal or back pain, nausea, vomiting, poor appetite and weight loss [70]. However, all these aspects are really unspecific, and can hardly be clinically linked to PaC.

Due to the limited physical evidence, PaC diagnosis presents a great challenge. However, several tests can be performed to diagnose it, most of which rely on imaging methodologies. Ultrasonography is the easiest methodology, despite the pancreas can be difficult to visualise, especially for small tumours [72]. Improved sensitivity can be achieved by endoscopic ultrasonography, which is nowadays

considered the most sensitive imaging modality for PaC detection [73]. Upon PaC suspicion, computed tomography is the other preferred diagnostic tool, which can report the tumour and invasive lesions. Magnetic resonance might also be performed, despite it presents restrictions in distinguishing neoplastic from inflammatory lesions [74]. These techniques can be combined with other methodologies, and endoscopic ultrasonography-guided fine needle aspiration is a common method to obtain tissue samples, which have shown great performance in detecting PaC (90.8% sensitivity and 96.5% specificity) [75]. Unfortunately, tissue samples obtention is an invasive process (despite minimal in this methodology) that optimal diagnostic tools should avoid. Finally, although positron emission tomography may also be a useful tool, the often-used glucose analogue radiotracer presents low sensitivity and specificity for diagnosis, while it is effective for detecting recurrence and metastasis [76].

A main problematic in PaC diagnosis is the absence of blood-based biomarkers, which could be easily used as a screening methodology to detect the tumour in asymptomatic stages. The only FDA-approved biomarker nowadays is the carbohydrate antigen 19-9 (CA19-9), corresponding to the sialyl-Lewis a (sLe^a) antigen [77]. However, it is only recommended for therapy monitoring as it lacks accuracy for PaC diagnosis [78]. Thus, at the very end most diagnoses are performed through pancreatic biopsies, which can be percutaneous, endoscopic or surgical, and subsequent microscope observation [79]. Overall, the lack of clear symptoms and early diagnosis markers results in a late diagnosis, at which time the tumour has already grown and less effective treatments are available.

Once diagnosed, PaC can be classified in different stages according to their progression, which will require different actions from the medical point-of-view. This classification is based on the TNM system [80], which considers three key points: tumour size (T), lymph nodes invasion (N) and presence of metastasis (M). Once these parameters are established, pancreatic tumours can be sorted into stages I-IV as shown in Table 1 [81].

Table 1. Pancreatic cancer staging. Classification of PaC progression stage based on the TNM system according to the Cancer Staging Manual of the American Joint Committee on Cancer [82]. A brief description is included for each situation.

PaC stage	TNM classification	Description
IA	T1N0M0	Tumour confined in the pancreas, no bigger than 2 cm.
IB	T2N0M0	Tumour confined in the pancreas, with a size between 2 and 4 cm.
IIA	T3N0M0	Tumour confined in the pancreas, larger than 4 cm.
IIB	T1N1M0	The tumour is confined in the pancreas (any size) and has spread to 1 to 3 lymph nodes.
	T2N1M0	
	T3N1M0	
III	T1N2M0	The tumour is confined in the pancreas (any size) and has spread to more than 3 lymph nodes.
	T2N2M0	
	T3N2M0	
	T4NxM0	The cancer is growing outside the pancreas. Lymph nodes invasion may (or may not) occur.
IV	TxNxM1	The cancer has spread to distant organs (metastasis). It can be any size and any N.

Additionally, other aspects can be evaluated to determine tumour progression and future outcomes, which include tumour grade and resectability. The former assesses how histologically similar to normal pancreatic tissue the tumour is. Grade 1 (or well-differentiated) refers to tumour tissues that look very like normal pancreatic cells, where duct-like structures can still be observed. These tend to grow slowly and are less likely to spread. On the other hand, grade 3 (or poorly differentiated) show abnormal tissues, with cells embedded in the stroma without any arrangement. Finally, grade 2 is found in between, showing abundant glands with papillary patterns [83]. Regarding resectability, tumours are classified by the capacity of being surgically removed in resectable, borderline resectable or unresectable [84].

Treatment

Several therapeutic strategies are nowadays available for the management of PaC, despite most of them lack effectivity. Actually, the National Comprehensive Cancer Network (NCCN) pancreatic adenocarcinoma panel suggests that enrolment in a clinical trial may be the best option for PaC treatment [85].

Surgical resection is the only option that offers a potential cure [86], and is usually combined with chemotherapy in the adjuvant or neo-adjuvant setting. Three principal strategies are available for such, the use of which will depend on the anatomical location of the tumour. Pancreaticoduodenectomy (Whipple's procedure) is the first choice for resectable and borderline resectable PaC [87]. It consists in the removal of the pancreas head, common bile duct and duodenum. It can also include gallbladder, lymph nodes and part of the stomach extirpation. Then, the remaining pancreas and bile duct are reattached to the small intestine, which is joined to the stomach. It is a rather complicated surgery that should be performed in high-volume centres, and despite mortality rates have decreased it still causes high morbidity [88]. Surgery is usually performed through an open technique, despite laparoscopic methodologies are available and result in better postoperative clinical outcomes [87].

Total or distal pancreatectomy are the remaining surgery options. Total pancreatectomy is still recommended in particular situations [89], but distal pancreatectomy is much preferred, either in the open, laparoscopic or robotic approach [90,91], being those minimally invasive methodologies more popular. Actually, novel surgery options for pancreatectomy that include vessel reconstruction constitute a group of candidates offering curative surgery to borderline resectable and locally advanced PaC [92].

However, as stated before, most PaCs are diagnosed at late stages, when the tumour is unresectable, and other strategies including systemic therapy are mandatory. The first-line therapy includes gemcitabine (mainly in combination with other chemotherapeutic agents) or FOLFIRINOX [85,93]. Gemcitabine has been the reference therapy drug for metastatic PaC since 1997, showing a clinical benefit response in 23.8% of patients and a median survival time of 5.6 months [94]. Combination of gemcitabine with different biological or cytotoxic agents has been tested, but increased overall survival has rarely been observed [95–97]. A meta-analysis including more than 10000 patients with advanced PaC showed signifi-

cant improvement on overall survival when using combined chemotherapy, despite toxicity was also increased [98]. Gemcitabine combination with nab-Paclitaxel has also shown significant improved clinical outcomes, again in expenses of increased toxicity [99]. This is the reason why combination is just accepted for good performance status patients [85].

FOLFIRINOX was developed as a regimen that combines folinic acid, 5-fluoroacil, irinotecan and oxaliplatin, which should not present overlapping toxic effects. It presents higher survival benefits compared to gemcitabine, but also increased toxicity [100]. However, the use of FOLFIRINOX versus gemcitabine plus nab-Paclitaxel improves survival while reducing post-treatment complications [101]. A modified version of FOLFIRINOX without 5-fluoroacil and including growth factors has demonstrated improved safety while maintaining efficacy [102,103], and constitute the first-line therapy preferred regiment together with classical FOLFIRINOX and gemcitabine plus nab-Paclitaxel [85]. These chemotherapeutic strategies are not just restricted to unresectable tumours, and both FOLFIRINOX and gemcitabine plus nab-Paclitaxel have demonstrated efficacy for resected malignancies in both the adjuvant [104–107] and neo-adjuvant [105,108–110] therapy.

Novel approaches relying on targeted therapies have also been investigated [111,112]. PaC targeted therapies have been directed towards growth factor receptors, K-ras pathway inhibition, angiogenesis, tumour-stroma interactions, cancer stem cells, DNA repair mechanisms and other targets [113,114]. Unfortunately, while some have proven promising results in preclinical stages, none of them have passed phase II/III trials [113]. Immunotherapy, which has shown really positive effects in other cancer types, has also been explored in PaC, but encouraging results have not been obtained so far, mainly due to the barrier formed by the tumour microenvironment [115]. The use of antibodies [116–118], antibody-drug conjugates [119–121], immunotoxins [122,123] or CAR-T cells [124] have not significantly improved chemotherapy clinical outcomes. Recently, the development of personalised RNA neoantigen vaccines have demonstrated T cells

stimulation and improved free relapse survival in patients with surgically resected tumours, but validation in a larger cohort is still needed to transfer the methodology to the clinical routine [125].

Glycobiology

Glycobiology comprises the study of the structure, biosynthesis, function, and biology of carbohydrates (also called glycans, saccharides, sugars, or sugar chains) and their derivatives. These carbohydrates, widely distributed in nature, constitute one of the major biomolecules of living organisms, with functions ranging from structural components to signalling effectors [126].

Monosaccharides are the basic structural units from which glycans are assembled. Chemically, simple monosaccharides are biomolecules with carbon, hydrogen and oxygen atoms, with an empiric formula of $C_x(H_2O)_n$ with an aldehyde or ketone group [126]. Monosaccharides are usually not found in their linear form but in a cyclic one, resulted from the reaction of a hydroxyl group with the carbonyl of the aldehyde or ketone to generate a hemiacetal group containing an anomeric carbon.

Despite several hundred monosaccharides have been described, only nine of them are common in vertebrates, which are classified in pentoses (xylose), hexoses (glucose (Glc), galactose (Gal), mannose (Man)), hexosamines (*N*-acetylglucosamine (GlcNAc), *N*-acetylgalactosamine (GalNAc)), deoxyhexoses (fucose (Fuc)), uronic acids (glucuronic acid) and nonulosonic acids (*N*-acetylneuraminic acid (Neu5Ac)) (Figure 4).

Monosaccharides link to each other through a glycosidic bond, an acetal linkage which occurs between the anomeric carbon of a monosaccharide and a hydroxyl group from another. This linkage let polysaccharides embrace different conformations, which are essential for their biological activity. Many different unions can occur between two monosaccharides, as the anomeric carbon form the first can be an α- or ß- stereoisomer while the hydroxyl groups of the other monosaccharide allow for various regioisomers [126]. The number of monosaccharides forming a mature structure can also vary, from short disaccharides to enormous structures (some models predict glycogen to have until 55000 glucose units [127]).

In addition, structural diversity is further broadened by the possibility of monosaccharides to be involved in more than two glycosidic linkages and assemble branched structures [126].
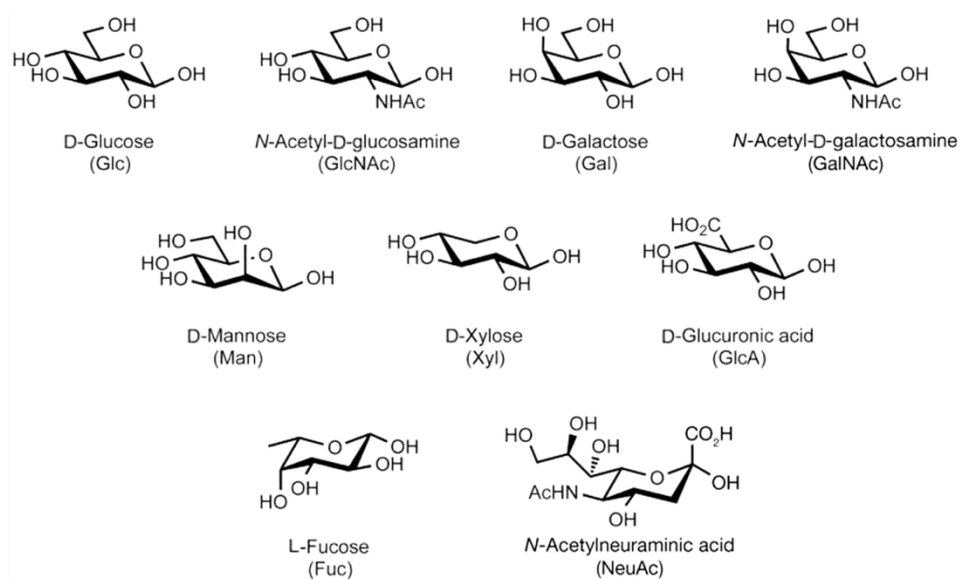
Figure 4. Most common monosaccharides in vertebrates. Common monosaccharides found in vertebrates. *N*-Acetylneuraminic acid is the most common form of sialic acid (SA). Extracted from Varki *et al.*, 2022 [126].

Glycan biosynthesis is a universal, complex and well-regulated process. Interestingly, it is not directly genome encoded, and relies on the expression of several enzymes, the bioavailability of glycan precursors, donors, transporters and growing structures, the steric impediments of the synthesised biomolecule and the location of all the involved agents [126]. The main effectors that modify glycan structures are enzymes that can be divided in two groups: those which add monosaccharides (glycosyltransferases) and those removing them (glycosidases). Both glycosyltransferases and glycosidases are extremely specific enzymes. Around 200 glycosyltransferases have been described, which can be further classified upon the monosaccharide that add [128]. For example, galactosyltransferases catalyse the addition of Gal, while sialyltransferases (STs) transfer a sialic acid (SA) monosaccharide. They are also specific for the monosaccharide to which the bound is

formed, and even to the specific carbon, making them linkage specific. The number of different glycosyltransferases is further increased by their redundancy, as various enzymes can catalyse the same reaction. For instance, up to twenty STs are described in mammals, which are classified in four groups depending on the linkage and acceptor. Six of them, ST3GalI-VI, are responsible for the same reaction: addition of a SA to a Gal in α2,3-linkage [129]. On the other hand, glycosidases hydrolyse the glycosidic bond to remove monosaccharides. Equally, glycosidases are both monosaccharide and linkage specific, despite some of them can catalyse a broad spectrum of different linkages [128].

The addition of monosaccharides require their activation before being transferred into the growing glycans. This activation is based on the phosphorylation of their anomeric carbon in the form of a nucleoside triphosphate (generally UTP or GTP) or CMP [126]. These will be recognized by glycosyltransferases and will be linked to the corresponding growing glycan.
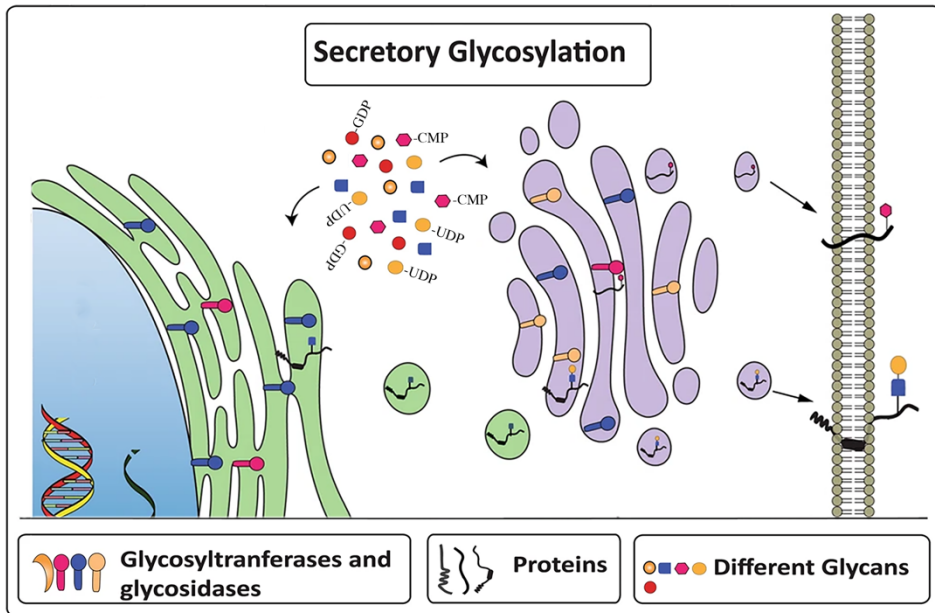
Figure 5. Overview of the glycosylation process. Schematic representation of the secretory glycosylation process, which is mostly restricted to the ER and Golgi. Activated monosaccharides are transported to these compartments, were glycosyltransferases and glycosidases, responsible of monosaccharides addition or digestion in the nascent glycan, will shape the final glycan structure. Extracted and modified from Mehboob *et al.*, 2021 [130].

Glycosylation is a process with a clear cellular organisation, which is mainly developed inside the endoplasmic reticulum (ER) and Golgi apparatus (Figure 5). Despite some structures can be preassembled outside this organelles before their translocation and minor glycosylation processes occur in the nucleus or cytosol, most glycosyltransferases are transmembrane proteins with the catalytic domain faced towards the ER or Golgi lumen, where the modification of glycans will occur during the secretory pathway. Thus, no specific instructions guide how glycans are synthesised, nor a template can be used for such. It will ultimately depend on which precursors, donors and enzymes meet inside the reticulum or Golgi. Hence, while genomic alteration of glycosyltransferases and glycosidases expression can vary these enzymes concentrations, the remaining aspects are more difficult to control [126].

Although glycans conform one of the four main biomolecule families, they are usually covalently linked to other biomolecules, forming glycoconjugates. Several categories can be found, such as glycoproteins, glycolipids, glycosaminoglycans, proteoglycans or lipopolysaccharides. The position of the glycan part in the other molecule, the number of added glycans and its structural variation make glycoconjugates highly heterogenic molecules, which may present vast amounts of isomers.

The biological functions of glycans are extensive and can be divided in three broad categories: structural contribution, energy metabolism and information carriers, being the last one subdivided in intrinsic or extrinsic [126] (Figure 6). Carbohydrates are the main source of cell energy and their role in the metabolism have already been widely reviewed [131]. Regarding their structural role, glycans provide physical structure and protection, modulate macromolecules solubility, membrane and extracellular matrix organization, and modulation of membrane receptors among others [132]. Actually, most cells present an outer-membrane layer rich in glycoconjugates known as the glycocalyx, responsible for physical expulsion of pathogens, protection from immune recognition or antiadhesive action. These glycoconjugates may also act as signal transductors by recognising bacterial and other parasites adhesins, toxins, agglutinins, molecular patterns and antigens, thus

modulating the immune response. Intrinsic recognition is also relevant for glyco-protein trafficking, intercellular signalling and adhesion or cell-matrix interactions [132].
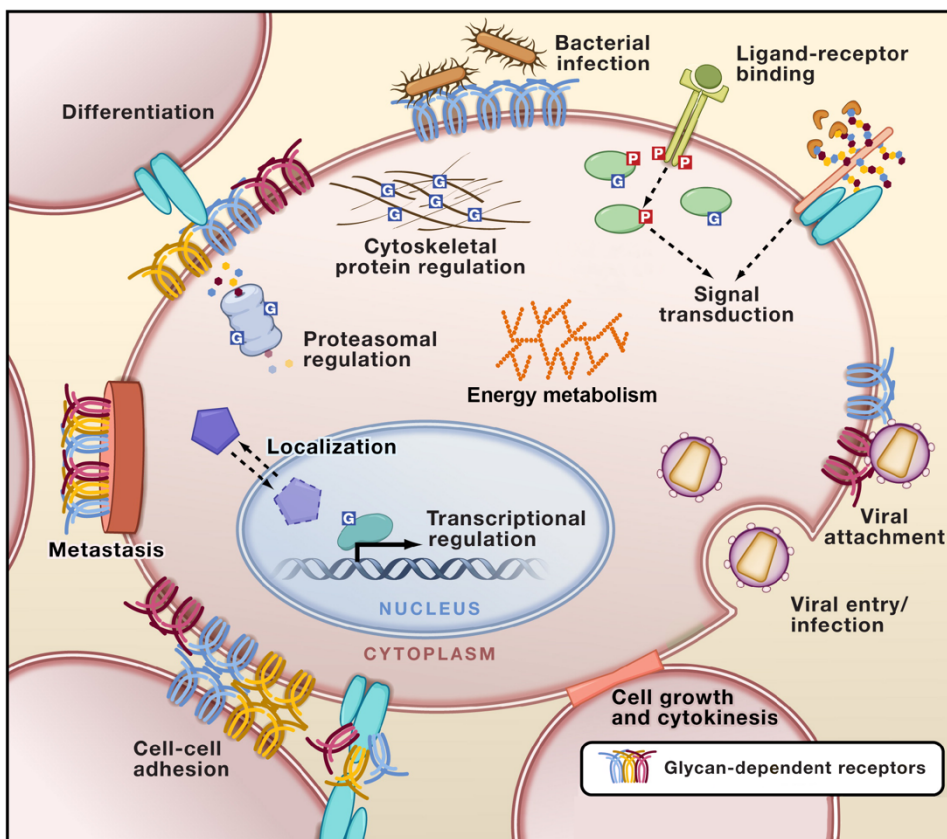
Figure 6. Biological functions of glycans. Representation of glycans main functions, which include structural, energetic and recognition activities. Extracted and modified from Hart *et al.*, 2010 [133].

Protein glycosylation

Glycosylation is one of the main post-translational modifications of proteins, which occurs in over 85% of secreted proteins [128]. Protein glycosylation is a site-specific mechanism that, in general, reflects the glycosyltransferase repertoire and glycosylation capacities of the producing cell, which can be affected by numerous cellular and environmental factors. Thus, different glycosylation patterns may occur on the proteins in a particular cell, which will result in protein isoforms

(glycoforms) with modified properties/functions. In short, attachment of glycans at different sites within a protein will result in macroheterogeneity, while different processing of glycan structures on a particular site will provide glycan microheterogeneity [134]. Thus, glycosylation broadens the proteome providing tenths to hundreds of glycoforms for a single protein.

Glycosylation may influence protein folding, quality control, stability, transport and function. Initiation of glycosylation in the ER guides protein folding and quality control. In addition, different glycans modify protein solubility, half-life time, and interaction to assemble complex protein structures. At a cellular level, glycans pay a crucial role in adhesion, receptor activation and cell-cell, cell-matrix and cell-pathogen interactions. Glycans not only alter the affinity of interactions involving proteins, as glucids themselves can be recognised by ligands and directly participate in several recognition mechanisms [126].

Given the extensive functions of glycoproteins, minimal changes may alter cell physiology. Actually, dysregulation of the glycan machinery, which finally results in altered glycosylation, has been determined in several diseases including cancer [135–137], neurological diseases such as Alzheimer's [138,139], autoimmune diseases [140,141], inflammatory diseases [142] and infectious diseases [143,144]. Thus, altered glycosylation raises as a source of potential therapeutic targets and biological markers [145].

According to the residue to which glycans are attached and the structure of the glycan itself, protein glycosylation can be sorted into various groups, being *N*-glycosylation and *O*-glycosylation the major types. Minor glycosylation processes include *C*-glycosylation, *S*-glycosylation, phosphoglycosylation and glypiation [146], but these will not be further discussed.

N-glycosylation

Protein *N*-glycosylation is characterised by the covalent attachment of oligosaccharides to the carboxamide on the side chain of asparagine (Asn) residues in polypeptide chains within the consensus sequence Asn-X-Ser/Thr, where X is any

aminoacid except proline [147]. Three main steps summarise this process: the formation of the oligosaccharide donor, the transfer to the nascent protein and the final processing of the glycan chain in the ER-Golgi (Figure 7).

The first step starts with the synthesis of a common precursor ($DolPPGlcNAc_2$-Man_5) at the cytoplasmatic side of the ER. It consists in the transfer of a GlcNAc onto a dolichol phosphate carrier, followed by the sequential addition of a second GlcNAc and five Man to build a branched heptasaccharide. Then, specific ER membrane proteins, flipases, promote the flip-flop of this precursor to the internal part of the ER, where it is further elongated by four Man and three Glc residues, yielding the oligosaccharide donor of 14 monosaccharides [126].

The oligosaccharyltransferase complex is the responsible to transfer this assembled glycan of 14 monosaccharides to the Asn of the nascent polypeptide chain into the ER lumen, specifically in an Asn-X-Ser/Thr consensus sequence. However, not all *N*-glycosylation sequons are occupied within a protein. It is known that the aminoacids surrounding the consensus sequence and the secondary structure of the nascent polypeptide can modulate the affinity of the glycan precursor for the Asn, mainly due to steric hindrances. For instance, there is a greater preference for Asn-X-Thr over Asn-X-Ser sites, aromatic and aliphatic groups around the sequon increase glycosylation efficiency while it is decreased by basic groups [148]. *N*-glycosylation is supported by turns and bends secondary structures and is less favoured on buried or very exposed sequons.

Next, the processing of the 14-oligosaccharide glycan chain takes place, starting with the digestion of Glc and Man residues in the ER and then with the removal of more mannoses at the cis-Golgi, to finally form $GlcNAc_2Man_5$. However, some glycoproteins can escape this process and express high mannose *N*-glycans [126] (Figure 8). In the medial-Golgi, synthesis of complex structures is initiated by the addition of GlcNAc by MGAT1. At this point, evasion of Man digestion by MAN2A1 or MAN2A2 will give rise to hybrid structures (Figure 8), while total processing will digest two Man and add another GlcNAc, which will form the precursor for complex biantennary structures. Additional branches can be started
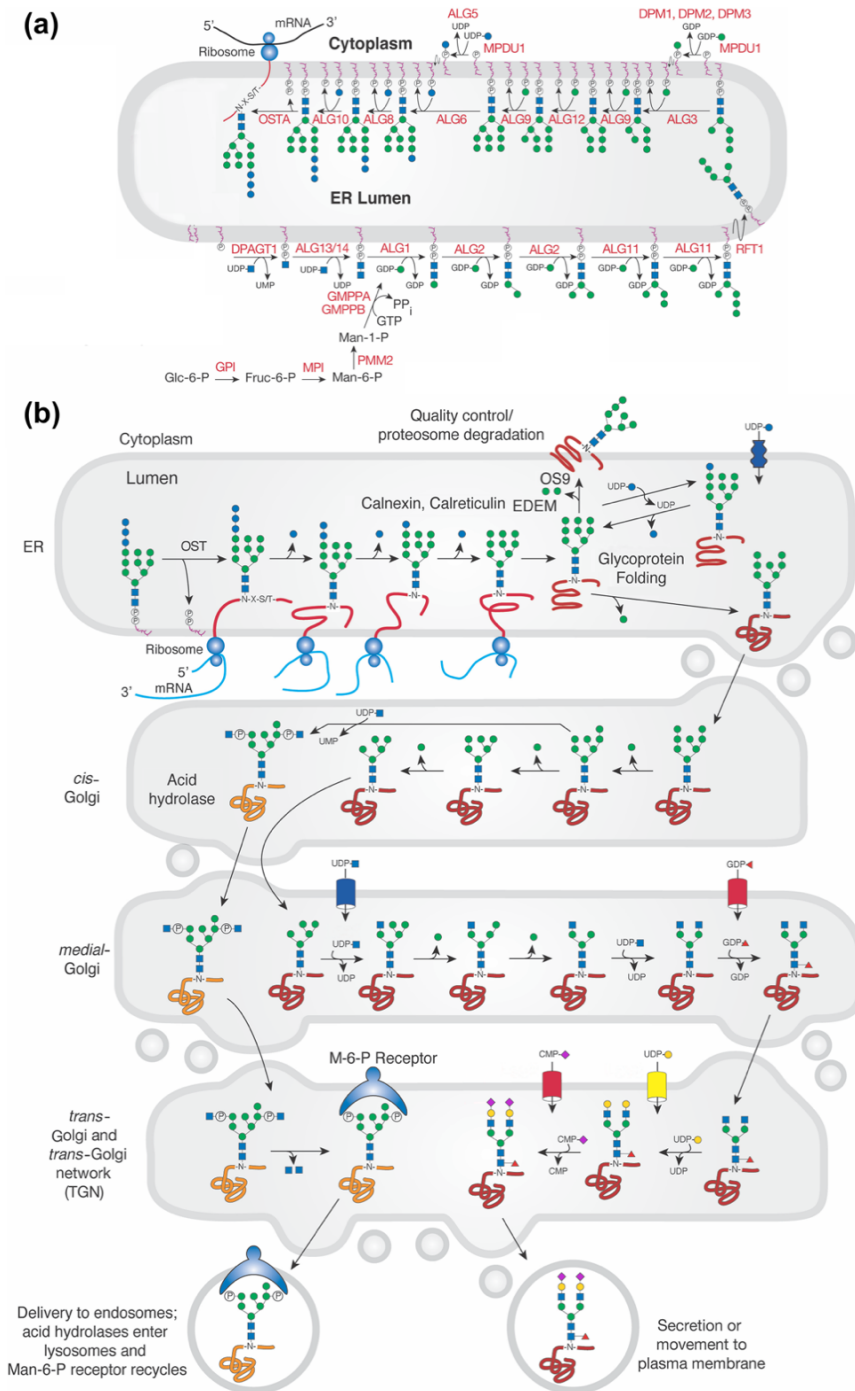
Figure 7. Protein *N*-glycosylation process. (a) Synthesis of the DolPPGlcNAc$_2$-Man$_5$ precursor in the membrane of the endoplasmic reticulum. **(b)** Processing and maturation of an *N*-glycan through the ER/Golgi compartments. Extracted and modified from Varki *et al.*, 2022 [126].

by addition of GlcNAc to Man by other *N*-acetylglucosaminyltransferases. Interestingly, addition of GlcNAc to the ß-Man of the core by MGAT3 will result in a bisecting GlcNAc, which is usually not further elongated.

Final maturation of complex *N*-glycans in the trans-Golgi comprises sugar addition to the *N*-glycan core, branch elongation and terminal capping of these branches. The main core modification is the addition of an α1,6-Fuc residue to the inner GlcNAc by the fucosyltransferase FUT8. Branch elongation is performed by addition of Gal to the initiated GlcNAc branches, usually forming a Galß1,4GlcNAc block which can be found in tandem repeats named poly-*N*-acetyllactosamine (poly-LacNAc). Common capping sugars include SA, Fuc, Gal and GlcNAc, usually in α-linkage and with the ability to bond to different hydroxyls on the attached residue. It must be noted that all *N*-glycan structures are characterised by a common pentasaccharide core (GlcNAc$_2$Man$_3$) (Figure 8), which is maintained along the polysaccharide processing.
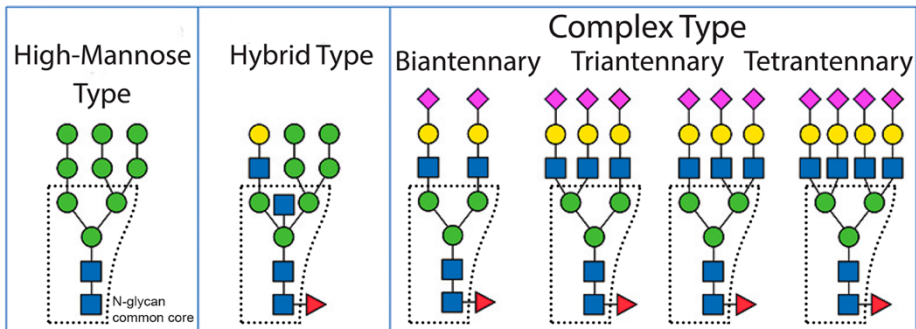
Figure 8. Types of *N*-glycans. Representation of the *N*-glycan main structures that can be synthesised along the glycosylation process, dividing *N*-glycans in high-mannose, hybrid or complex. Among complex *N*-glycans, a further distinction in bi, tri- and tetra-antennary structures can be observed. Remarkably, all *N*-glycans share a common core (dotted box). Glycans represented following the symbol nomenclature for glycans (SNFG) guidelines [149].

Many functions have already been described for carbohydrates. Among all them, *N*-glycans have a key role in folding, structure, activity, antigenicity and susceptibility to proteases of proteins [147].

O-glycosylation

Protein *O*-glycosylation comprises the covalent linkage of glycans to the hydroxyl group of Ser and Thr residues in polypeptides [126]. Several subtypes of *O*-glycans can be found depending on the first residue attached to the Ser/Thr, namely *O*-GalNAc, *O*-Fuc, *O*-Glc, *O*-GlcNAc, *O*-Gal and *O*-Man. However, in this thesis we will mainly focus on *O*-GalNAc glycans, which are the most abundant.

Unlike *N*-glycosylation, *O*-glycans are not processed from a common precursor but are constructed by adding monosaccharides one by one. Moreover, no consensus sequence is needed for the initiation of this process. Hence, the attachment of the first GalNAc to a Ser or Thr residue is a crucial step, which is mediated by a polypeptide GalNAc transferase [150]. Twenty isoenzymes of GalNAc transferases are described in humans [151], which select the glycosylation position based on peptide sequence motifs and the presence of previous *O*-glycosylation. Depending on the preference for different substrates, each GalNAc transferase will display different affinities and specificities, thus modulating *O*-glycosylation [152].

Once the first GalNAc is added, sequential transference of monosaccharides will result in oligosaccharides of between 1 and 20 residues. A single GalNAc (known as the Thomsen-nouveau (Tn) antigen) will rarely be found in normal glycoproteins, but it is usually overexpressed in tumour glycoproteins, suggesting that cancer somehow can stop *O*-glycan elongation. Addition of SA in α2,6-linkage by ST6GalNAc1 will form the sialyl-Tn antigen (STn), which cannot be further elongated, and is also vastly expressed in advanced tumours [126].

In general, four different *O*-glycan cores can be synthesised and additionally extended (Figure 9). Attachment of ß1,3-Gal generates core 1 structures (T antigen), which similarly to the Tn antigen can be decorated with SA or extended with other monosaccharides. Subsequent addition of ß1,6-GlcNAc to the inner GalNAc of core 1 results in core 2. These are the most common *O*-glycan cores found in glycoproteins. Synthesis of core 3 *O*-glycans seems to be restricted to mucus epithelia, and is driven by the addition of a ß1,3-GlcNAc to the Tn antigen. Then, the addition of ß1,6-GlcNAc to the inner GalNAc will result in a core 4 structure.

Synthesis of complex *O*-glycans through elongation of core structures is mediated by the addition of ß1,3-GlcNAc or ß1,3/4-Gal, and are usually terminally capped by the addition of SA or Fuc.
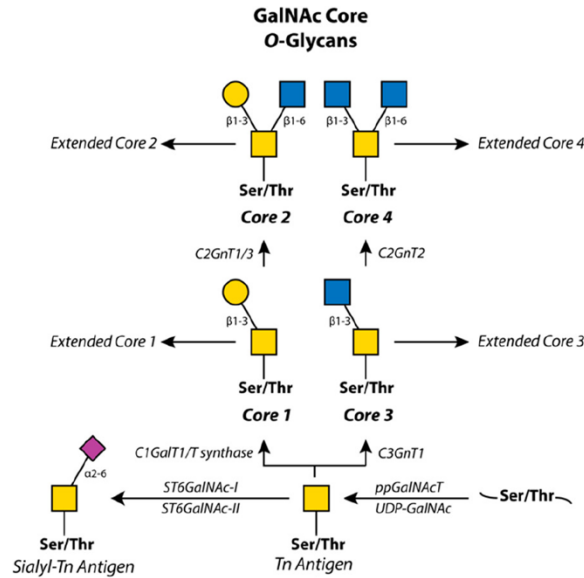
Figure 9. Common *O*-glycan cores. Synthesis pathways for the expression of the most common *O*-GalNAc glycan structures. Extracted and modified from Wilkinson *et al.*, 2020 [153].

O-glycosylation have many and varied functions. In proteins with many *O*-glycan structures, it is involved in hydration, structural support and protection from proteolysis. On the other hand, single *O*-glycosylation may regulate proprotein and ectodomains cleavage, tissue formation and differentiation, and cell-cell, cell-matrix and host-pathogen interactions [126].

As elongation and final maturation of *N*- and *O*-glycans occur simultaneously in the same location, and is mediated by a number of glycosyltransferases, both *N*- and *O*-glycans may share common terminal structures. The most external GlcNAcs found in cores are often galactosylated. Addition of Gal in ß1,4-linkage will generate type 2 chains (also called LacNAc). This can be further substituted with another ß1,3-GlcNAc, and again with ß1,4-Gal to form poly-LacNAc chains. On the other hand, addition of Gal in ß1,3-linkage will generate type 1 structures. Despite type 1 and type 2 chains can be found at both *N*- and *O*-glycans, in general

terms type 1 chains are common to *O*-glycans, while type 2 are usually found in *N*-glycans.

These elongated structures constitute a perfect backbone for the expression of terminal determinants such as the ABO blood group or Lewis family antigens (Figure 10). Addition of α1,2 Fuc to terminal Gal generates the H antigen. Further modification by A transferase or B transferase will result in the addition of an α1,3-GalNAc or α1,3-Gal, respectively, to produce the A and B blood antigens [126,154]. The genetic encoding of the glycosyltransferases responsible for the specific addition of these monosaccharides, and thus their presence in the red blood cells surface, will determine an individual blood group.
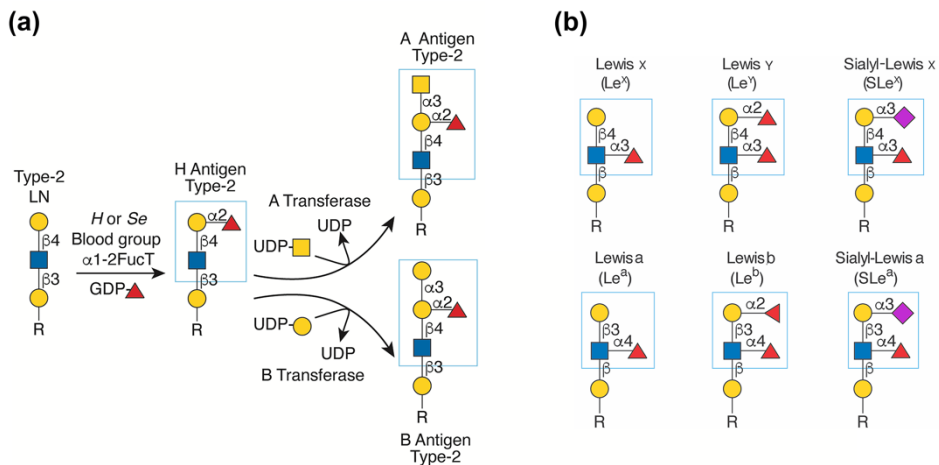
Figure 10. Fucosylated determinants expressed in human glycans. Representation and synthesis of **(a)** ABO blood group and **(b)** Lewis family antigens, glycan determinants that can be found in both *N*- and *O*-glycosylation. Extracted and modified from Varki *et al.*, 2022 [126].

The Lewis family comprises a group of fucosylated structures that carry α1,3/4Fuc residues [126]. The different linkage depends on the attachment of the Fuc residue to type 1 or type 2 chains. If a Galß1,4GlcNAc structure (type 2) is already formed, addition of Fuc to GlcNAc will necessarily be in α1,3-linkage to form the Lewis x (Lex) antigen. On the other hand, addition of Fuc to type 1 chains will be in α1,4-linkage, resulting in the Lewis a (Lea) antigen. Further terminal Gal fucosylation generates the Ley and Leb antigens, respectively. However, α2,3-sialylation of this Gal results in the formation of sialyl-Lewis x (sLex) and sLea. These determinants

are of special interest as have been involved in malignant processes such as metastasis and specific ligand recognition [155].

Altered glycosylation in cancer

Deregulation of the glycosylation process and altered glycan expression have been vastly described in cancer [135–137,145,156–159]. Its presence is so extensive that several cancer associated glycans have been detected in almost all cancer types (Figure 11) [135,157], with evidence supporting the implication of glycosylation in all steps of tumour progression. Actually, despite not being considered a hallmark of cancer itself, changes in glycosylation enable the acquisition of all the capabilities summarised in the hallmarks of cancer [160].
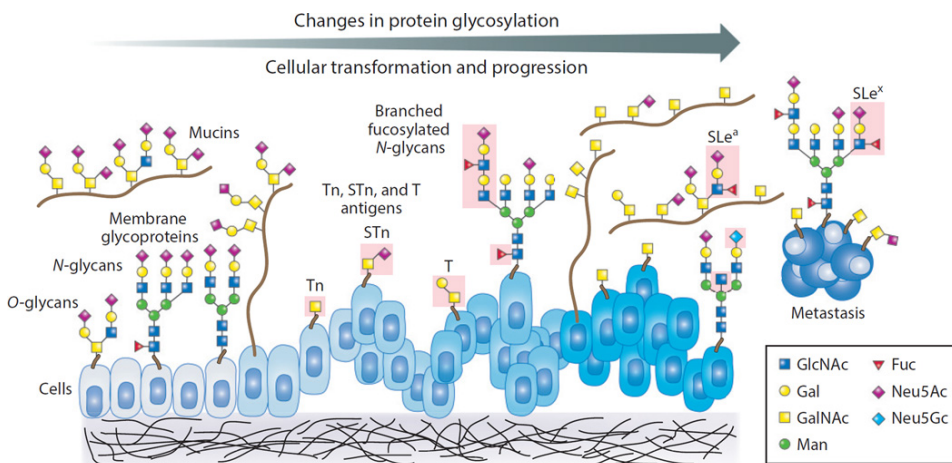
Figure 11. Altered glycosylation in cancer. Representation of the main changes in protein glycosylation along the tumorigenic process. Pink-boxed areas highlight the presence of truncated *O*-glycans (Tn, STn and T antigens) and the alteration of *N*-glycans branching and fucosylation, including the Lewis antigens (sLe[x] and sLe[a]). Extracted from Stowell *et al.*, 2015 [135].

There is an increasing interest to use altered glycosylation as biomarker and/or therapeutic target, while understanding its implication in cancer development is also crucial [158]. Cancer associated glycan determinants include short, truncated *O*-glycan structures such as Tn, STn and T antigens [161]. The presence of the Tn antigen has been detected in several cancers while it is rarely expressed in normal adult tissues. Its expression is correlated with metastatic potential and poor prognosis and may modulate immunosuppressive effects and extravasation [126].

Similar to Tn, the STn antigen is related to tumour growth and metastasis, promoting the invasiveness potential of tumour cells. It is implicated in the cell-matrix recognition, promoting cell disruption and migration. In addition, overexpression of STn in secreted glycoproteins as mucins has shown to help the tumour escape immunosurveillance [162]. In PaC, both Tn and STn promote EMT and stemness properties [163]. The T antigen also plays a crucial role in cell adhesion via interaction with galectins and may help in the establishment of metastatic tumours. It also may have a role in cell proliferation and angiogenesis [164].

Concerning *N*-glycosylation, a general increase in branching (specially the GlcNAcß1,6Man branch), corresponding to tri- and tetra-antennary structures, is detected in most malignancies [165]. *MGAT5* overexpression leading to the synthesis of ß1,6-branching is related to migration, wound healing and EMT increase [166]. It also regulates tumour growth, cell adhesion, invasion and metastasis [145]. In contrast, increased expression of the bisecting GlcNAc is associated with reduced malignancy, mainly through inhibition of the ß1,6-branch synthesis [166]. Branched *N*-glycans usually are further elongated by poly-LacNAc repetitions, which are involved in modulating cell adhesion and migration through interaction with galectins [167]. These structures also increase the metastatic potential by specific binding with galectin-3 [168] and may serve as scaffold for the expression of other tumour-associated carbohydrate antigens [169].

Increase in total sialylation and fucosylation has also been identified [135,158,166]. Fucosylation may occur at both external residues and at the inner *N*-glycan core, yielding the core fucose determinant. Actually, core fucose is known to play a role in cancer progression. For instance, induction of the EMT may involve core fucosylated E-cadherin [170] and TGF-ß receptor [166]. E-cadherin core fucosylation may also modulate cell-cell adhesion. Likewise, several proteins external fucosylation with α1,2-linked fucose promotes cancer stem cell potency and tumour metastasis [145].

Increased sialylation has been vastly addressed [171]. As a terminal determinant, SA plays a crucial role in signal interaction and recognition, and thus modification

in its expression alters malignant cells behaviour and capabilities. Overexpression of several STs has been associated with different cancers, both for the addition of SA in α2,3- or α2,6-linkage [172–174]. Deregulation of sialidases is also critical for differential sialylation [175,176]. Cell hypersialylation has been related to immune system evasion, apoptosis and cell death escaping, adhesion and invasion alteration, and circulation, extravasation and colonisation potential [177]. All these functions are regulated through SA action as a ligand for siglecs and selectins [178]. Recognition by siglecs, which are SA lectins preferentially expressed by immune cells, enhance cancer progression and immune evasion [179]. On the other hand, the interaction of sialylated structures with selectins (usually via the sLe^x tetrasaccharide) modulate different steps in cancer progression, mainly metastatic dissemination [180,181]. Sialylation of glycans may also alter the recognition by galectins and modulate these lectins functions. The identification of galactose by galectins is impaired by α2,6 sialylation, while α2,3-SA does not change their binding affinity [182]. In this regard, cancer sialylation may modulate galectin-1 role in immune surveillance [183] and angiogenesis [184].

As already hinted, the differential expression of the Lewis family determinants is also a characteristic feature of cancer cells, which is related with the overall upregulation in fucosylation and sialylation. Many tumours show the overexpression of some of these antigens [185–187]. The fucosylated antigens Le$^{x/y}$ are involved in adhesion activity, chemotherapy resistance, tumour progression and proliferation [188–190]. On the other hand, sialylated structures such as sLe^x and sLe^a are associated with tumour metastasis. It is well known that sLe^x act as a ligand for selectins, and aids metastatic cells in the extravasation process by mimicking leukocytes rolling [191]. SLe^a may also interact with E-selectin and modify cell-matrix recognition [192].

Specifically for PaC, glycosylation does not differ from general cancer glycan alterations, being the main observed features increased sLe^a and sLe^x, truncated *O*-glycans, sialylation, fucosylation and branched *N*-glycans [193–195]. For instance, modulation of migratory and invasive abilities have been observed in PaC cells

after upregulation [196,197], knockdown [198] and inhibition [199] of the STs involved in sLe^x formation. Also in PaC, expression of truncated *O*-glycans promote EMT and modulate AKT/mTOR or RAS/MAPK signalling pathways [163,200]. The analysis of PaC patients' serum *N*-glycome has shown higher levels of branching and antenna fucosylation, as well as a general increase in sialylation [201,202]. Interestingly, Vreeker and collaborators [201] observed raised α2,6 sialylation in all kinds of structures, while α2,3-SA was only elevated in fucosylated *N*-glycans (probably forming sialyl-Lewis structures).

Glycan analysis

Given the importance of glycosylation in human pathology, the study of glycans and glycoconjugates is of great interest, and several methodologies for such have been developed (Figure 12). It is interesting to note that many of these techniques can be performed for a global glycan profiling or for a more targeted glycoproteomic analysis.
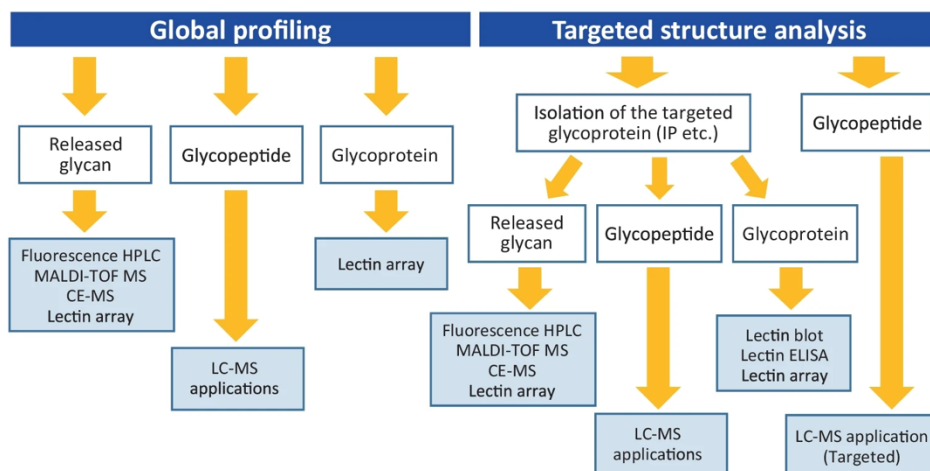
Figure 12. Current techniques for the study of protein glycosylation. Summary of available methodologies for the study of protein glycosylation under different strategies. Extracted and modified from Haga *et al.*, 2022 [158].

Since long ago, some particular glycan structures, such as red blood antigens, have been detected with biological probes such as antibodies. However, due to the low immunogenicity of most glycans, the use of lectins has been an alternative. Lectins

are carbohydrate binding proteins found in nature which serve various biological functions. Different lectins may recognise different carbohydrate structures (monosaccharides or small oligosaccharides) and might be selective for its linkage (Table 2). Unfortunately, lectins affinities are lower than those of antibodies. In addition, as most lectins only recognise terminal carbohydrates, combination with glycosidases is sometimes mandatory to reveal complex structures. Many lectin-based methodologies can be used, including lectin histochemistry, lectin microarrays, lectin blotting, a range of immunoassays (as enzyme-linked lectin assay (ELLA)) and lectin-affinity liquid chromatography (LC) [158,203–205].

Lectins are easily usable in histochemistry studies to unveil the general glycosylation pattern of a particular sample. As for antibodies, conjugation of lectins with fluorescent labels, enzyme labels or biotin let them be used as recognition probes on tissue slides. Thus, differential glycosylation in distinct tissues, development steps or physiological/pathological conditions may be mapped using this methodology [206]. A main limitation is the general staining of all glycoproteins bearing a specific carbohydrate motif, being the study of a single glycoprotein a challenge. However, dual lectin and antibody co-localisation may bring new insights to these applications [207]. Furthermore, methodologies such as proximity ligation assay arise as proper alternatives to identify protein post-translational modifications thanks to its ability to detect two epitopes in close proximity [208].

A different approximation like lectin microarrays is also useful for analysing the general glycomic profile of biological samples [209]. Lectin microarrays rely on the fluorescent dying of glycoconjugates prior to incubation in a glass plate with immobilised lectins, so that fluorescent spots will be observed for those lectins recognising a glycan [210]. Thus, several glycan structures can be detected in a single analysis. Despite not providing the complete structure of glycans, lectin arrays stand as a simple, high-sensitive and robust technique for comparative purposes [211].

Table 2. Commonly used lectins for the identification of glycan structures. Summary of the main lectins used as probes for the detection of glycans, with a schematic representation of the structure recognised following the SNFG guidelines. Data extracted from Varki *et al.*, 2022 [126] and Vector Laboratories catalogue.

Lectin	Common abbreviation	Preferred glycan	Detected structure
Aleuria aurantia	AAL	Fucose	
Concavalin A	ConA	Mannose	
Lens culinaris	LCA	*N*-glycans	
Maackia amurensis I	MAL-I	Galactose	
Maackia amurensis II	MAL-II	Sialic acid	
Peanut agglutinin	PNA	Galactose	
Phaseolus vulgaris erythtroagglutinin	PHA-E	GlcNAc-ß4	
Phaseolus vulgaris leucoagglutinin	PHA-L	GlcNAc-ß6	
Pholiota squarrosa	PhoSL	Core fucose	
Sambucus nigra	SNA	Sialic acid	
Ulex europaeus	UEA	Fucose	
Vicia villosa	VVL	GalNAc	
Wheat Germ	WGA	GlcNAc	
Wistera floribunda	WFA	GalNAc	

As an extension of western blotting (WB), lectin blotting relies on the use of lectins to detect glycoconjugates. This methodology, as other lectin-based protocols, requires the optimisation of lectins dilution to avoid false-positive detection. The main disadvantages include the need of previous protein purification for the study of single proteins' glycosylation pattern and the weak suitability to be used as a routine diagnostics tool [204].

Lectins can also be used as probes in a whole range of immunoassays. Development of an ELLA assay derive in a fast high-throughput methodology, which is quantitative and very cost effective. Like enzyme-linked immunosorbent assays (ELISAs), various ELLA approximations are feasible, such as direct detection of glycoconjugates in the well surface or sandwich ELLAs to detect particular carbohydrate structures. In addition, hybrid ELLA using an antibody for protein capture and a lectin for glycan recognition supports the quantification of specific proteins glycoforms, although cross-reactivity of lectins with the glycans present in the antibodies may represent a critical issue [212].

The ability of lectins to interact with carbohydrates can also be used for separation of glycoconjugates by affinity chromatography. Hence, lectins can be immobilised to a solid matrix to segregate glycoconjugates depending on their affinity to the lectin. This methodology presents high affinity and does not require a pre-purification step, but requires a large amount of sample, is time-consuming and only allows for individual samples analysis [213].

The use of lectins may present a problem as affinities are not high and specificities are not well defined for some of them. Fortunately, there are several strategies for the study of glycans that do not rely on lectins. LC separation prior to their detection is a commonly used strategy [214]. Release of glycans from the protein backbone is mandatory, and is usually accomplished by enzymatic digestion with Peptide:*N*-glycosydase F (PNGaseF) for *N*-glycans [215] and ß-elimination or oxidative release for *O*-glycans [153]. Once obtained, glycans require derivatisation with fluorescent labels that will be detected after separation. Typical labels include

2-aminobenzamide (2-AB), 2-aminobenzoic acid or procainamide [216–218]. Hydrophilic interaction liquid chromatography (HILIC) is the method of choice for glycan separation, which presents high selectivity. Briefly, glycans interact with a BEH amide column and are eluted with an acetonitrile (ACN) gradient, so glycans are separated by size and hydrophilicity. Other columns such as reverse phase or graphitised carbon may also be used, despite isomer separation and reproducibility may be compromised [218]. Ultra-performance liquid chromatography (UPLC) is the choice for such separation over high performance liquid chromatography (HPLC). UPLC allows the use of smaller beads size and higher pressures, which turns into reduced separation time and better resolution [214]. Combination of this separation with exoglycosidase arrays with linkage and monosaccharide specificity might allow accurate structural determination [219].

More detailed analyses can be performed with mass spectrometry (MS) techniques, which have greatly improved in recent years [134,153,158]. Furthermore, MS support the study of glycans and their corresponding glycoconjugates, in a global or targeted strategy. Thus, a comprehensive view on the alterations of glycosylation in a given sample can be assessed through glycomic and glycoproteomic approaches. The most used techniques include matrix assisted laser desorption/ionisation (MALDI) MS and electrospray ionisation (ESI) MS. MALDI MS (usually using a time-of-flight (Tof) detector) displays accurate mass determinations and is usually used to determine released glycans o small glycopeptides. Glycoprotein heterogeneity can also be detected through this methodology. Unfortunately, structural information as monosaccharide anomericity or site-specificity cannot be obtained.

Capillary electrophoresis coupled to MS may escape these limitations [220]. Electrophoretic separation, based on the charge and size of the analytes, can differentiate structural isoforms, even SA linkage [221]. In addition, capillary electrophoresis has demonstrated high sensitivity, high throughput, separation efficiency and reliable quantification [222]. Further, the use of MS/MS applications and fragmentation is the responsible for a detailed glycan structure characterisation [223].

The study of glycopeptides in a LC-MS tandem also lets the identification of different isoforms, which are separated in the chromatography, while site-specific glycosylation is observed in the peptide part. LC-ESI-MS/MS, often in triple quadrupole (TQ), quadrupole time-of-flight (QTof) or Orbitrap analysers, might also provide glycans structural information thanks to fragmentation, being collision induced dissociation the main method used to get this information. These strategies are extremely useful to detect already known glycoforms in complex matrices with targeted approximations such as operating in selected reaction monitoring (SRM) mode, in which selected glycopeptides are isolated in Q1 reducing interfering background signal, and allowing the identification and quantification of selected glycoforms as diagnostic tools [223].

Tumour markers

Biomarkers can be defined as functional/structural variants or quantitative/qualitative indexes of a biological process that predict or reflect the presence of, evolution of or predisposition to a disease, a clinical condition or a response to therapy [224]. Most common biomarkers are biological molecules found in blood, other body fluids or tissues that are a sign of a normal or abnormal process, or of a condition or disease. These substances are specifically named tumour markers when produced by cancer cells or other body cells in response to cancer, and thus give information about that malignancy.

The story of cancer biomarkers started long ago. Nevertheless, it was not since about two centuries that the use of tumour markers appeared in medicine, when a protein in urine from multiple myeloma patients was defined by Sir Bence Jones. More precisely, in 1847 the Bence-Jones protein was described as a light chain antibody of immunoglobulin G produced in tumours, excreted in urine and identified by heat denaturation [225]. Nevertheless, it was not until 1988 that an immunodiagnostic test was approved for its detection, two years after its discovery in blood samples [226,227]. Actually, tumour markers testing was transformed in the 1950s with the development of immunoassays and the use of polyclonal antibodies. Development of monoclonal antibodies in 1975 and sandwich immunoassays in 1982 expanded the field, which continued growing through recombinant antibody and other molecular biology techniques. At present, only very few tumour markers have been validated for its clinical use in assessing disease diagnosis or prognosis (Table 3), although in recent years many reports describing new potential biomarkers have been published [227].

Recently, the US National Institute of Health (NIH) and the Food and Drugs Administration (FDA) have listed up to seven categories of biomarkers: susceptibility/risk, diagnostic, monitoring, prognostic, predictive (outcomes and recurrence), response and safety [228]. Thus, a wide range of applications may be attributed to biomarkers, which can inform of a particular malignancy stage and be used in personalized medicine.

Table 3. Current commonly used tumour markers for the clinical management of cancer. Uses and organ specificity of the main cancer biomarkers used in clinical practice. Data extracted from [227,229,230].

Cancer biomarker	Cancer type	Applications
Prostate specific antigen (PSA)	Prostate	Screening, diagnosis, monitoring
Prostate cancer antigen 3 (PCA3)	Prostate	Prognosis
Carbohydrate antigen 125 (CA125)	Ovarian	Diagnosis, prognosis, detecting recurrence and monitoring therapy
Carcinoembryonic antigen (CEA)	Colorectal/hepatic	Monitoring therapy, prognosis, detecting recurrence
Carbohydrate antigen 15-3 (CA15-3)	Breast	Monitoring therapy
Estrogen, progesterone receptors (ER and PgR)	Breast	Patient stratification for therapy
HER2	Breast	Monitoring therapy
Carbohydrate antigen 27-29 (CA27-29)	Breast	Monitoring
Human chronic gonadotropin-ß (HCG-ß)	Testicular	Diagnosis, staging, detecting recurrence and monitoring therapy
Alfa-fetoprotein	Hepatocellular	Diagnosis, detecting recurrence and monitoring therapy
Calcitonin	Thyroid	Diagnosis and monitoring therapy
Thyroglobulin	Thyroid	Monitoring
CA19-9	Pancreatic	Monitoring therapy
Nuclear matrix protein 22 (NMP-22)	Bladder	Screening, monitoring and prognosis
ß2-microglobulin	Multiple myeloma	Diagnosis and prognosis

There are several considerations to bear in mind when developing a proper biomarker. Imperatively, the measured variable must be present in the patient, and must gather information about the irregular process. It needs to be evaluated with an optimal assay, both in terms of standardization, reproducibility, and accuracy. This means that tests performed in different hospitals/laboratories should always provide the same values as a result of well-defined procedures, assay results must be constant for a given sample, and small differences on the observed variant should be detected. In addition, the assay must be acceptable for the patient, minimizing invasive interventions, or holding a high profit to risk rate if invasiveness is unavoidable. Easily accessible fluids like serum or urine stand as the preferred source of biomarkers analysis. Finally, the observation or quantification of the biomarker should provide relevant clinical value, helping in the decision-making process of the medical personnel [227].

Biomarkers precision for discriminating between two conditions is assessed through several concepts, summarized in Figure 13 [224,227,231]. Individuals are usually considered "positive" when a condition (for instance being ill) is satisfied, while they are "negative" when it is not met (being healthy). Given a population where we know the actual condition of each individual and the predicted condition based on a biomarker value, each individual can be arranged in one of the following four groups: true positive, true negative, false positive and false negative.

Total population: P+N	Predicted condition			
	Positive	**Negative**		
Positive (P) (ill)	True positive (TP)	False negative (FN)	Sensitivity $\dfrac{TP}{TP + FN}$	Accuracy $\dfrac{TP + TN}{P + N}$
Negative (N) (non-ill)	False positive (FP)	True negative (TN)	Specificity $\dfrac{TN}{TN + FP}$	

Figure 13. Predictive analytics contingency matrix. Population groups according to actual and predicted conditions, and the corresponding analytical metrics for the evaluation of a biomarker potential.

True positives and true negatives gather people with concordant predicted and actual condition. Those individuals actually negative that are predicted positive integrate the false positive, while people actually positive and predicted as negative conform the false negative. Mathematically, sensitivity refers to the probability of correctly predict positive (ill/disease) conditions. Thus, high sensitivity tests would result in reduced undiagnosed patients. On the other hand, specificity refers to the probability of correctly predict negative outcomes (not ill). In this case, a higher value coincides with limited overdiagnosing [232]. Other aspects to consider include the positive predictive value, which reflects how many predicted positive tests are actually positive, and its opposite, the negative predictive value, which indicates the proportion of true negatives among all predicted negatives. Finally, the accuracy of a biomarker is defined by the number of correctly predicted cases (both true positive and true negative) among the overall population.

An ideal biomarker would distinguish the two conditions (ill and non-ill) with 100% sensitivity and specificity. However, this is never the real scenario, as perfect separation is often not possible. Thus, there is always a need to define a biomarker test cut-off value which maximizes both sensitivity and specificity (Figure 14.a). Depending on the situation, improved sensitivity might be preferred at the cost of decreased specificity, and vice versa. In these cases, a shift in the cut-off criteria might alter the distribution of false-positives and false-negatives, and thus provide new sensitivity and specificity values (Figure 14.b).
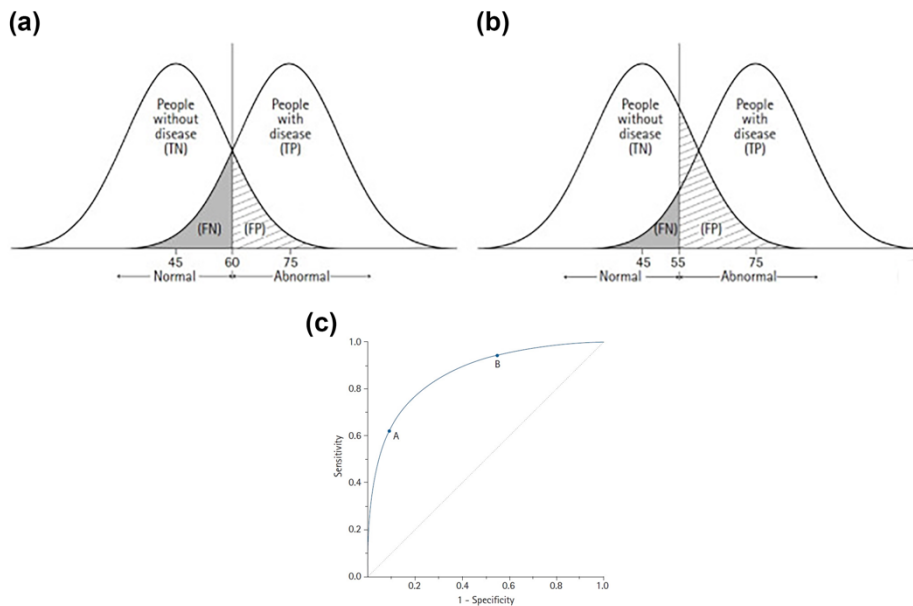
Figure 14. Optimisation of tumour markers accuracy. (a) Analytical grouping of a hypothetic population distribution for patients with or without a disease. **(b)** Variation of the cut-off value (vertical line) rearranges observed populations providing different sensitivity-specificity values. **(c)** Example of a generic ROC curve, which indicates paired sensitivities and specificities at different cut-offs. Extracted and modified from Nahm *et al.*, 2022 [231].

Receiver operating characteristic (ROC) curves are an excellent tool for assessing a biomarker performance [231]. In short, a ROC curve represents paired sensitivity and 1-specificity values for a test at different cut-offs (Figure 14.c). The closer the curve is to the upper left corner, the better accuracy of the test. Actually, an area under the curve (AUC) can be calculated to determine the test accuracy, being AUCs over 0.8 generally considered acceptable. An optimal cut-off value can

also be obtained from a ROC curve following diverse approximations. The Youden's J statistic defines the best cut-off at the point where the vertical distance from the ROC curve to the 45º reference line is maximal. On the other hand, the Euclidean distance establishes the cut-off at the point with minimal distance to the upper left corner. Less common approaches include the index of union, the maximum product of sensitivity and specificity or the maximum sum of sensitivity and specificity, among others [231].

Since long ago, the discovery of tumour markers has been focused on proteins expressed at higher amounts in malignant cells than normal ones [227,233–242]. However, other biomolecules were rapidly acquiring interest in the field of cancer biomarkers discovery, such as glycoproteins [243], glycans [244], metabolites [245,246], nucleic acids [247] (with special attention to non-coding RNAs [248], micro-RNAs (mi-RNAs) [249] and circulating tumour DNA (ctDNAs) [250,251]), or circulating tumour cells (CTCs) [252,253]. In addition, the evolution of genomic, metagenomic and transcriptomic profiling enlarge the spectrum of candidates to develop new biomarkers [254–257].

Despite the US national cancer institute lists almost a hundred commonly used tumour markers [258], not even fifteen are routinely and consistently used in the clinical practice [227,229], a number that has roughly increased in the last twenty years. These cancer biomarkers have different applications including diagnosis, prognosis and monitoring of patients. For instance, only five of them are used on the diagnostic step, which has been described as a critical one for the management of the disease [259]. Thus, the scientific and medical communities must gather resources and efforts to expand the sum of available tumour markers, which could improve cancer handling through earlier detection and improved decision-making procedures. Nonetheless, it is not easy to overcome all the steps to develop a cancer biomarker, which include discovery of the marker, development of an assessment method, preliminary clinical potential analysis, standardisation of the assay and final validation for clinical use [260].

Current situation of PaC biomarkers and proposed candidates

As mentioned before for the specific case of PaC, there are no tumour markers for detecting the disease in early stages, when the malignancy is asymptomatic [26,261,262]. Therefore, it is mandatory to find novel biomarkers to apply better management options. Only CA19-9 can be used in the clinical routine, helping in monitoring patient's response to therapy. Its potential use for PaC diagnosis has been extensively assessed, but it displays moderate accuracy values not acceptable for its clinical application. A meta-analysis containing 20.991 participants reported a pooled sensitivity of 72%, specificity of 86% and AUC of 0.847 [78]. Reasons hampering the use of CA19-9 as an early diagnosis tumour marker include reduced specificity and raised false negative outcomes. CA19-9 is also found in other cancers including gastrointestinal malignancies, ovarian mucinous carcinoma or lung cancer, decreasing its specificity and making it useless for PaC diagnosis [263]. It is also increased in benign pathologies like ChP or renal failure. In the clinical routine, choledocholithiasis and acute cholangitis are the main benign conditions with CA19-9 elevation, and pose a challenge in the differential diagnosis with malignant pathologies [264]. On the other hand, CA19-9 cannot be synthesized by genotypically deficient Le[a-b-] individuals, that account for 5-10% of general population, thus producing false negative outcomes [265]. This subpopulation presents a poorer outcome when affected by PaC, with higher metastatic rate than Lewis-positive patients [266]. Despite the impaired use for PaC diagnosis, CA19-9 correlates with the clinical response after pancreatectomy [77], making it a good biomarker for patient following-up [267]. It can also inform on tumour resectability, being 500 U/mL the consensus threshold for such application [268].

Diverse combinations of CA19-9 with other proteins have been proposed to enhance its performance in detecting PaC, but none of them has resulted in significant improved discrimination. Slight gain is obtained with CA19-9 combination with CEA, CA242, albumin or IGF-1 [269–271]. Combination with mHOXA1 and mSST improves diagnostic sensitivity [272], while CA19-9 combination with

MUC5AC shows improved general accuracy [273]. CA19-9 has also been combined with molecules other than proteins, and for instance combination with five miRNAs has shown excellent results [274].

However, in the need to overcome CA19-9 use, many candidates have been proposed for developing into novel PaC tumour markers. Actually, hundreds of molecules with differential expression between PaC and healthy conditions have already been described [275,276], most of which require further validation. For example, while the use of a single protein such as PIM-1 has exhibited 95% sensitivity and 100% specificity [277], proteins as CA125 require combination with CA19-9 to report adequate accuracy [278]. Other protein candidates include the trefoil factors family [279], regenerating protein family [280–283] or apolipoproteins family [284,285].

Proteins and glycoproteins are not the only molecules found in serum samples that can be used as tumour markers. The detection of non-coding RNAs have also shown promising results [286]. Among them, miRNA profiling have detected several species that could be useful for PaC diagnosis [287]. For instance, downregulation of miR-92, miR-132, miR-148a, miR-216a and miR-217 has been detected in PaC, contrarily to upregulation of miR-31, miR-143, miR-145, miR-146a, miR-150, miR-155, miR-194, miR-196a, miR-196b, miR-210, miR-222 and miR-223 [286]. Long non-coding RNAs like LINC01232 and HULC, among others, may also become tumour markers [288–290], as well as some circRNAs including circ_0007534 or circ-LDRAD3 [291,292].

Other approaches include serum liquid biopsies to detect differential expression of exosomes, CTCs and ctDNA [293], being exosomes the ones performing better discrimination. For, instance, Lai and collaborators have reported a miRNA signature found in exosomes with superior diagnostic potential than CA19-9 [294]. The detection of ctDNA is also promising, being *KRAS* the main identified fragment, together with *ADAMTS1* and *BNC1* [295,296]. The performance of CTCs to detect PaC still requires further validation, as their detection greatly varies depending on the used enrichment method [297].

Metabolomics is recently getting interest as a source of novel tumour markers. Mayerle and collaborators have described a metabolic signature including 9 metabolites and CA19-9 that can discriminate PaC patients from ChP with an AUC of 0.96 [298]. Another study has shown metabolite panels with 4 to 12 analytes that also yield promising results for PaC diagnosis [299]. Likewise, a seven-metabolite signature named CarboSign has been able to discriminate PaC from non-oncologic patients with an AUC of 0.97, and is already patented under ref. EP22382244.6 [300]. Novel methodologies are exploring the establishment of a panel of volatile organic compounds to detect PaC, which can be detected in the gas phase of urine samples [301] and even in the alveolar air [302]. These last examples remind us that blood serum is not the only sampling source for biomarkers detection, and other fluids such urine [303] or saliva [304,305] can be used, as well as easily accessible samples like stool [254]. In the specific case of PaC, minimally invasive obtention of pancreatic juice might also be considered to find novel tumour markers [280,306].

Finally, since the glycosylation process is altered in the tumour situation, glycomic approaches are also being explored to develop new biomarkers [201,307,308]. However, as exposed before, PaC altered glycosylation is similar to general cancer glycosylation, and despite it can discriminate between PaC patients and healthy individuals, glycomic analyses cannot distinguish different tumours. In this regard, glycoproteomic biomarkers (based on glycosylation changes on specific proteins) could be able to improve the potential of the proteins alone [230,309,310]. Several glycoforms on specific proteins have been found increased in PaC patients. These include core fucosylated RNase1 [311], fucosylated haptoglobin [312–314], fucosylated AGP [315], α2,6-sialylated alpha-1-ß glycoprotein [316,317], fucosylation and sialylation of hemopexin, kininogen-1 and antithrombin-III [318], or the presence of sLe[a] on MUC1, MUC5AC and MUC16 [319,320] and sLe[x] on ceruloplasmin [321], MUC5AC and MUC1 [322], and MFAP4 [323], among others.

New candidates as PaC biomarkers

As commented, the analysis of specific glycoforms in selected pancreas-derived glycoproteins could outperform the biomarker accuracy of the protein levels. Thus, this thesis has focused on analysing the glycosylation of two serum-secreted glycoproteins, mesothelin and regenerating islet-derived 1 protein, which are neo- or overexpressed in PaC, and evaluate the potential of their tumour-specific glycoforms as PaC tumour markers.

Mesothelin

Mesothelin (MSLN) is a glycoprotein *neo*-expressed in PaC that has been proposed as a diagnostic biomarker and as a therapeutic target [324]. It was first described by Chang & Pastan [325], who isolated a cDNA encoding a 40 kDa glycoprotein on the surface of mesothelial cells recognized by the K1 antibody. MSLN is synthesised as a 622-residue precursor of about 69 kDa, which is then cleaved to produce a C-terminal GPI-anchored 40 kDa mature form and a released 31 kDa fragment named megakaryocyte-potentiating factor (MPF) [326] (Figure 15.a). Eventually, the mature MSLN form can be released from the cell membrane and reach the bloodstream by the action of various proteases [327] or due to the formation of a soluble variant resulting from alternative splicing [328]. Actually, splicing events can provide several isoforms. A minor form with 24 extra base pairs in exon 13 has been detected in ovarian cancer transcripts, which results in a protein with eight extra aminoacidic residues [329]. The previously mentioned soluble isoform results from the insertion of an 82-base pair fragment, corresponding to intron 16, which produces a frameshift in the C-terminal region [330]. This change disrupts the signal sequence for the GPI-anchor, leading to the direct secretion of the protein.

At the structural level, MSLN and other members of its superfamily are predicted to have superhelical structures with Armadillo-type repeats (Figure 15.b), which are suggested to interact with carbohydrate moieties from other glycoproteins [331]. MSLN presents three putative *N*-glycosylation sites (N-X-S/T) according

to its aminoacidic sequence. Moreover, MSLN N-glycosylation has been confirmed by a molecular weight decrease in SDS-PAGE after the complete resection of its N-glycans [332]. Even so, its N-glycans structural characterization has not been performed up to date, and there is no detailed information about the carbohydrate determinants that are more frequently displayed on MSLN under different physiological conditions. Only Fujihira and collaborators, in a study with mesothelioma cell lines, have shown increase in bisecting GlcNAc [333].
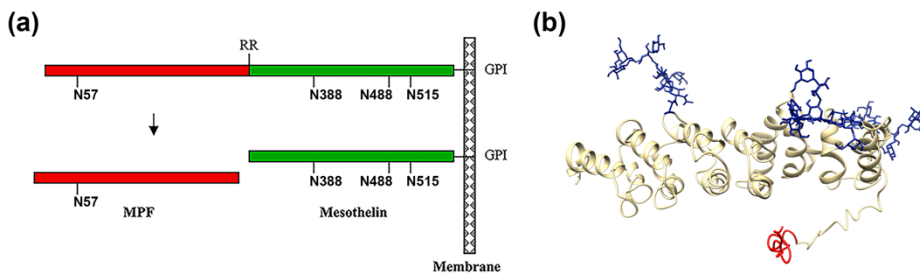
Figure 15. Synthesis, maturation and structure of MSLN. (a) MSLN is synthesised as a 622-aminoacid precursor, which is processed and attached to the cell membrane via a GPI anchor. It is further cleaved to give two mature proteins: while MSLN is kept at the extracellular part of the membrane, MPF is released. RR, precursor cleavage site; Nxxx, N-glycosylation site. Extracted and modified from *Hassan et al., 2004* [326]. **(b)** 3D predicted structure for MSLN based on its sequence (Uniprot Q13421) using I-TASSER [334]. An N-glycan core on the three predicted N-glycosylation sites is represented in blue, while the GPI anchor signal sequence is shown in red.

MSLN mediates cellular adhesion through binding to MUC16 [335], being the glycosylation of the latter essential for such interaction [336]. In cancer, MSLN is involved in several processes including resistance to cell death, cell proliferation, invasive and metastatic potential, angiogenesis, apoptosis regulation and EMT [337–341]. The expression of MSLN in healthy tissues is restricted to mesothelial cells of the pericardium, peritoneum and pleura, while it is found in various malignancies apart from PaC, including ovarian cancer or mesothelioma [325,342–344]. This limited expression in healthy tissues makes MSLN a useful target for therapeutic strategies. In this regard, various clinical trials against MSLN involving immunotoxins, monoclonal antibodies, antibody-drug conjugates, vaccines, or CAR-T cells have been conducted for PaC and other solid tumours [338,344,345].

The presence of MSLN in pancreatic tissues has been vastly examined by immunohistochemistry [346–350]. Abundant expression (>85% specimens) has been observed in PaC tissues, despite it is unrelated to cancer aggressiveness [351]. On the other hand, no presence has been found in normal pancreas or benign alterations as ChP [352]. However, it is not clear in which moment of cancer development MSLN expression begins: while PanINs does not show MSLN expression [346], MSLN is observed in intraductal papillary mucinous neoplasms [348], with higher frequency in invasive lesions [349,353]. When released from the cell membrane, MSLN can be detected in sera from different cancers patients [342,343,354]. Actually, its quantification with the Mesomark® assay is routinely used for the diagnosis of malignant pleural mesothelioma [355]. In addition, the analysis of circulating MSLN levels in blood was proposed for PaC diagnosis. However, despite MSLN was increased in PaC patients vs. healthy individuals (from 0.58 nmol/mL to 0.66 nmol/mL) [356], it showed no utility in the diagnosis of PaC due to low sensitivity and specificity values [354,357]. This slight increase arises from the capacity of the tumour microenvironment to retain shed MSLN and maintain its blood levels [358].

MSLN serum levels are not useful as a PaC biomarker. Here, we propose the study of MSLN glycosylation to search for specific glycoforms associated to PaC that could overcome the accuracy of current methodologies.

Regenerating islet-derived 1 proteins

Regenerating-islet derived (REG) proteins belong to a C-type lectin-like protein family involved in multiple functional roles involving pro-proliferative, anti-apoptotic and bactericidal activity. Since their discovery about forty years ago, REG proteins have been associated with diverse pathologies, but their exact roles, underlying mechanisms and expression regulation still remain elusive [359–361]. Based on their primary structure, REG proteins can be sorted into four groups: REG1, REG2, REG3 and REG4.

Several groups independently discovered REG1 and reported it under various names, including lithostathine, pancreatic stone protein (PSP), PTP or P19 [362].

Humans express two proteins from this first group, REG1A and REG1B, which are 89% homologous to each other [363]. REG2 is the most similar to REG1, but it is only expressed in mouse. REG3A and REG3G are those proteins from the third group produced in human cells. Finally, REG4 is the member that shows the lowest similarity with the rest of REG proteins, as it contains one extra exon and locates in a different chromosome [359]. Most human REG proteins have a trypsin cleavage site next to its *N*-terminus. Digestion of this terminal peptide results in an insoluble fragment that forms fibrils resistant to further protease digestion, in a similar way to those formed by ß-amyloid fibres in Alzheimer's disease.

REG proteins secretion is restricted to normal pancreatic cells, while they are also expressed in lower levels in brain, liver and the gastrointestinal tract. REG1A/B are secreted by pancreatic acinar cells, while no protein nor mRNA expression is observed in pancreatic ducts [364–369]. Increased expression has been reported during injury or inflammation, as well as in diabetes, ChP and PaC. In addition, some REG proteins are linked to pathological conditions of the digestive tract, the nervous and reproductive systems, and cancers including lung, breast, bladder or colorectal [359]. For all these reasons, REG proteins have been proposed as therapeutic targets and biomarkers. Especially, the use of REG1A appears to be successful in the detection and management of sepsis [370,371]. In this work, we will focus on the utility of REG1A and REG1B glycoforms as tumour markers for PaC.

REG1A is a 16 kDa glycoprotein mainly secreted by pancreatic acinar cells, and it is the most abundant non-enzymatic protein of pancreatic juice [372]. REG1A mRNA encodes a 166-residue protein with a 22-aminoacid signal peptide which is cleaved in the secreted form. The 144-residue mature form is stabilized by three disulfide bridges and presents a highly trypsin-sensitive arginine at residue 11 (Figure 16.a-b) [359]. Interestingly, this 11-residue N-terminal peptide starts with a pyroglutamic acid (5-oxiproline) and carries an *O*-linked glycan at Thr-5, which is responsible for the different molecular weight bands observed by SDS-PAGE corresponding to REG1A [373]. The glycan part was first characterized by NMR in

REG1A purified from the pancreatic juice of patients suffering diverse pancreatic diseases [374]. Up to eleven glycoforms were detected, seven of which could be fully sequenced. Most of them corresponded to an elongated core 2 structure (Glc-NAc(ß1,6)[Gal(ß1,3)]GalNAcα-), and ranged from 4 to 9 monosaccharide residues through elongation by *N*-acetyllactosamine units. Some structures carried a terminal H antigen with α1,2-linked Fuc, and all glycans except one presented α2,3-SA. A single disialylated form was observed, corresponding to a di-sialyl core 1. The complete X-ray structure of REG1A has also shown the presence of an *O*-glycan at the same position [375]. A 4-residue chain was detected and assigned to the di-sialyl core 1 after building three of its residues (NeuAc(α2,6)[Gal(ß1-3)]Gal-NAcα-). As stated before, REG1B has 89% homology with REG1A (Figure 16.c). REG1B is also a 166-aminoacid protein with 6 exons and 5 introns, with an *O*-glycan present in its *N*-terminal undecapeptide (Figure 16.d). It is predominantly expressed in human pancreas and gastrointestinal tract [367,369].

The exact roles of REG1A/B remain elusive. REG1 was first identified in pancreatic stones and assumed to inhibit its formation, but was later associated to its production, as the tryptic cleavage of its 11-residue *N*-terminal peptide results in an insoluble protein which causes protein aggregation, the first step of pancreatic calculi formation [372]. REG1 is also a promoter of islet ß-cells growth in response to inflammation and injury [372], and can even stimulate pancreatic ducts cell proliferation in cancer [376]. In colorectal carcinoma, REG1B silencing inhibits cancer cell proliferation and invasion [377]. The role of REG1 in cell differentiation and apoptosis prevention has also been suggested [359].

REG1 has been considered as a biomarker or therapeutic target due to its overexpression in diverse pathological conditions [359]. Immunohistochemistry analyses have demonstrated REG1 expression in PaC. REG1A is present in tumour-adjacent acinar areas and stains strongly positive in cancer glands. REG1 is already observed along the pancreatic carcinogenic process. REG1A expression increases alongside PanIN progression, while REG1B is elevated since the first stages of the process [378]. Studies involving other minor pancreatic tumours from Said *et al.*

[366] illustrated that REG1A shows diffuse strong cytoplasmatic staining in all normal acinar cells, while expression in PaC is of 65%. Similar results were previously reported by Satomura and collaborators, who observed REG1 in the 100% of normal pancreases and ChP, while it was found in 70% of PaC tissues [365]. The intensity pattern suggested a REG1 increase in mild pancreatic injuries and a decrease in more severe injuries.
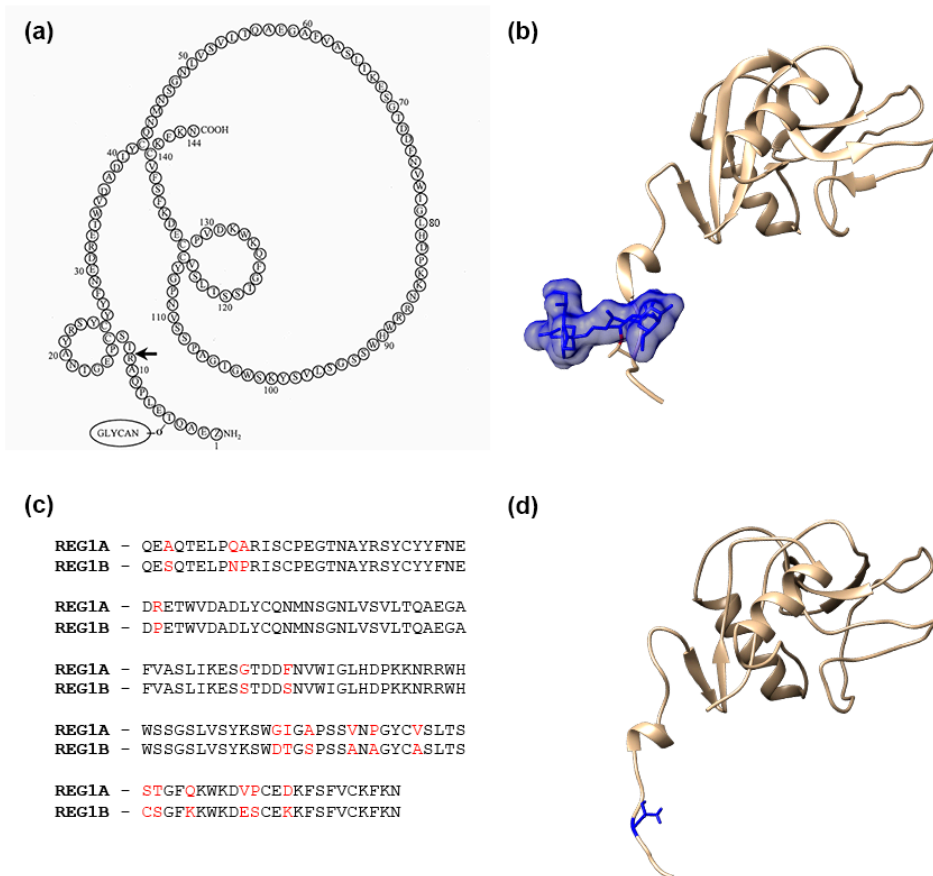
Figure 16. REG family proteins' structure. (**a**) Polypeptide representation of human REG1A, which contains three disulfide bonds and one glycan chain. An arrow indicates the trypsin-sensitive Arg-11. Extracted from *De Reggi, 2001* [372]. (**b**) REG1A X-ray 3D model at 1.3Å resolution (PDB: 1QDD) [375], with the three carbohydrate residues on its Thr-5 in blue. (**c**) Sequence alignment between REG1A and REG1B (89% homology), with their distinct residues highlighted in red. (**d**) 3D model for REG1B based on its sequence (Uniprot P48304) predicted with I-TASSER [334]. The potentially *O*-glycosylated Thr-5 is coloured in blue.

REG1 levels may also be increased due to pathologic conditions in body fluids such as blood or urine, and both REG1A and REG1B have been proposed to take part in several biomarker panels. A three-protein panel including REG1A in urine has shown promising results for the identification of PaC patients compared to healthy subjects and ChP [379]. Regarding REG1A performance in urine, basal levels of 113.8 ng/mL in controls are increased to 546.4 ng/mL in PaC, while remain at 127.8 ng/mL in ChP. Following studies suggested that substituting REG1A for REG1B in the proposed panel could enhance its performance [380]. REG1B levels in urine are of 41.3 ng/mL in control individuals, 64.2 ng/mL in benign pancreatic-related conditions (including 78.9 ng/mL for ChP) and 226.3 ng/mL in PaC patients. PancRISK, an algorithm that includes the urine biomarkers LYVE1, REG1B, TFF1 as well as age and creatinine, has been reported as a promising approach to early detect PaC [380]. Recently, a prospective study has demonstrated the use of this urinary panel in combination with plasma CA19.9 to predict PaC up to 2 years before its actual diagnosis with an AUC of 0.77 [381].

REG1A and B have also been detected in ductal fluid samples, and are increased in PaC compared to normal conditions [382]. In addition, their band pattern on the WB is different, suggesting that cancer samples present higher heterogeneity with a larger number of isoforms that could be attributed to a different glycosylation pattern of REG1 proteins. REG1A protein has also been reported in pancreatic juice from PaC patients [280]. In this study, the authors showed that, in blood serum, REG1 levels raised from 272.4 ng/mL in healthy individuals to 470.5 ng/mL and 481.7 ng/mL in ChP and PaC, respectively. Other authors have also shown elevated REG1A in serum, from 3271.4 ng/mL in healthy individuals to 3669.8 ng/mL in ChP and 4614.2 in PaC [378]. The same authors also revealed REG1B increased levels, from 134.1 ng/mL in normal conditions to 150.3 ng/mL and 159.7 ng/mL in ChP and PaC, respectively. Makawita and collaborators [281] also described increased serum REG1B in PaC patients, from 4582.0 ng/mL to 25380.0 ng/mL. They also reported elevated levels in benign pancreatic conditions (9344 ng/mL) and other cancers (16903 ng/mL), which included

colon, liver and stomach. A more general study on the presence of REG1A in serum from various cancers patients has also demonstrated its increase, from 32.9 ng/mL in non-cancer subjects to 60.6 ng/mL for those with cancer [383]. Diabetic people also present raised REG1A respect to control individuals, with levels going from 10.7 ng/mL to 13.5 ng/mL in control and diabetes, respectively [384]. It must be noted that REG1A/B values vary considerably between studies, making it challenging to find a reference level for each condition. Unfortunately, no consensus method is established for REG1 determination, and different studies use distinct ELISA methodologies, which can explain the variability in absolute serum values of REG1. In addition, results vary remarkably in different studies which use the same commercial assay, implying that lots of work are still needed in order to standardize REG1 quantification. However, all studies point out in the same direction, indicating an increase of REG1 in cancer conditions.

As stated, REG1 serum levels are altered in several diseases including both benign and malignant pancreatic disorders, and therefore their use as a biomarker of PaC is limited due to their low specificity. With the aim of overtaking this issue, we propose the analysis of REG1 *O*-glycosylation to identify those REG1 glycoforms that could be PaC-related and useful as PaC biomarkers.

Duran Sidera, A.

AIMS & SCOPES

Duran Sidera, A.

The need for a diagnostic test to detect PaC patients in the early stages of the disease is crucial to improve patients' survival. Protein glycosylation plays a fundamental role in the tumour process, being the aberrant expression of some glycan determinants a key target for developing new tumour markers. However, neither a single glycomic test nor any assay based on protein altered levels has yielded enough accuracy to discriminate PaC patients up to date. In this regard, we hypothesise that the combination of glycans and proteins detection, focusing on the identification of unique tumour-associated glycoforms on specific glycoproteins, could overcome the accuracy of current PaC biomarkers, mainly in terms of specificity. Thus, the main objective of this thesis is the assessment of tumour-associated glycoforms on MSLN and on REG1 as potential tumour markers for PaC. MSLN and REG1 are two glycoproteins *neo*-expressed and overexpressed in PaC, respectively, which can be detected in the bloodstream of PaC patients. To accomplish this aim, the following sub-objectives were proposed:

1. Determine the expression of both MSLN and REG1 in PaC cell lines and pancreatic tissues.

2. Characterise MSLN and REG1 glycosylation in PaC cell lines and pancreatic tissues through *N*-glycan sequencing, lectin blotting on the immunopurified proteins or MS-based methodologies.

3. Develop methodologies to immunpurify the selected glycoproteins from blood serum samples to further analyse their tumour-associated glycoforms by sandwich ELLAs or LC-MS.

4. Evaluate the serum levels of the selected MSLN and REG1 glycoforms in a cohort including PaC patients, ChP patients and healthy individuals to determine their usefulness as PaC tumour markers.

Duran Sidera, A.

MATERIALS & METHODS

Duran Sidera, A.

1.   Samples

Recombinant protein standards were acquired and used as controls. A recombinant human mesothelin standard (rMSLN) produced in a murine myeloma cell line (NS0) and a recombinant human REG1A (rREG1A) standard expressed on human embryonic kidney cells (HEK293) were obtained from R&D Systems, MN, USA. The recombinant human REG1B protein (rREG1B), also produced in HEK293 cells, was obtained from Abcam, Cambridge, UK. A human pooled serum (MP Biomedicals, Thermo Fisher Scientific, MA, USA) was obtained and used as a biological matrix control.

1.1.   Human cell lines

Seven PaC cell lines from the American Type Culture Collection (ATCC) and one ovarian cancer cell line (Ovcar-8) from the National Cancer Institute (NCI) were used in this study (Table 4). Capan-1, Capan-2, SW1990, BxPC-3 and AsPC-1 were maintained in Dulbecco's modified Eagle's Medium (DMEM) from Gibco (MA, USA), supplemented with 10% fetal bovine serum (FBS, Gibco) (20% FBS for Capan-1 cells), 1% penicillin/streptomycin and 2 mM L-glutamine (Gibco). HPAF-II cells were maintained in Eagle's Minimum Essential Medium (EMEM) supplemented with 10% FBS and 1% penicillin/streptomycin and Panc 10.05 cells in Roswell Park Memorial Institute Medium (RPMI) (Lonza, Basel, Switzerland) supplemented with 15% FBS, 1% penicillin/streptomycin and 10 U/mL insulin. Ovcar-8 cells were maintained in RPMI (Lonza) supplemented with 10% FBS and 1% penicillin/streptomycin. All cells were cultured at 37°C in a humidified atmosphere containing 5% CO_2. Cell growth and morphology were assessed daily using an inverted microscope, and the absence of mycoplasma was tested routinely with the Venor® GeM OneStep Mycoplasma Detection Kit (Minerva Biolabs, Berlin, Germany).

Table 4. Cell lines characteristics. Histology, source and gene mutations in the principal PaC affected genes. OvC: ovarian serous adenocarcinoma.

Cell line	Histology	Tumour source	Mutant genes	Reference
Capan-1	PaC	Metastasis: liver	*KRAS, SMAD4, TP53*	ATCC HTB-79
Capan-2	PaC	Primary	*KRAS*	ATCC HTB-80
SW1990	PaC	Metastasis: spleen	*KRAS, CDKN2A*	ATCC CRL-2172
BxPC-3	PaC	Primary	*CDKN2A, SMAD4, TP53*	ATCC CRL-1687
AsPC-1	PaC	Metastasis: ascites	*KRAS, CDKN2A, TP53*	ATCC CRL-1682
HPAF-II	PaC	Metastasis: ascites	*KRAS, CDKN2A, TP53*	ATCC CRL-1997
Panc 10.05	PaC	Primary	*KRAS, TP53*	ATCC CRL-2547
Ovcar-8	OvC	Primary	*KRAS, TP53*	NCI-DTP OVCAR-8

1.1.1. *Cell protein lysates*

Protein lysates were obtained from fresh cultures in a 100 mm plate. Following medium aspiration and PBS washing, 500 µL of ice-cold RIPA buffer (50 mM Tris pH 7.4, 150 mM NaCl, 0.1% NP-40, 0.5% Na-deoxycholate, 0.1% SDS, 2 mM EDTA, 50 mM NaF, 1 mM PMSF, 0.2 mM Na_3VO_4, and protease inhibitors Complete ULTRA Tablets™ (Roche Diagnostics, Mannheim, Germany)) were added. After 10 minutes incubation at 4 °C, plates were scratched and lysates recovered in 1.5mL tubes, for being next passed through a needle 25 times and incubated for 15 extra minutes at 4 °C. Finally, lysates were centrifuged for 10 minutes at 14000 g, and the supernatant recovered and stored at -20 °C until its use.

Total protein quantification was performed using the Bradford assay (Bio-Rad Laboratories, CA, USA) following manufacturer's instructions. Briefly, samples were diluted and loaded in triplicates in a 96-well plate. A calibration curve from 0 to 25 µg/mL of bovine serum albumin (BSA) was prepared in the same buffer (and dilution) than the samples to analyse and was also included in triplicates. After adding the Bradford dye reagent to each well, plates were incubated for about 10 minutes before determining the optical density in a plate reader (Synergy 4, Bio-Tek, VT, USA) at λ=595 nm with correction at λ=579 nm.

1.1.2. *Conditioned media collection*

For the study of cell secreted proteins, the conditioned media from the cell lines was collected. In short, cells were washed with PBS when reached 80% confluence. Next, cells were grown in free-FBS medium, which was collected 48 hours later and concentrated by centrifugation at 6500 rpm and 4 ºC in 10 kDa filters (Merck Millipore, MA, USA) previously passivated with 5% Brij-35. As for protein lysates, total protein concentration was measured by Bradford assay.

1.2. Human samples

Human tissue and serum samples were provided by the Hospital Dr. Josep Trueta (Girona, Spain) following the standard operation procedures of its ethics committee, in accordance with the current Declaration of Helsinki, the European Regulation (EU) 2016/679 and the Spanish Organic Law 3/2018 on data protection. The ethical approval for this study was obtained from the Comité de Ética de Investigación con Medicamentos Hospital Universitari Dr. Josep Trueta (Girona, Spain), reference number 2021.005. Informed written consent was obtained from all participants. Cancer tissue samples were confirmed by biopsy or image examination by the digestive and pathology units and were classified according to the Tumor Node Metastasis Classification of Malignant Tumors of the International Union Against Cancer (UICC) 8th edition.

1.2.1. *Pancreatic tissues*

A small piece of resected pancreatic tissues was immediately frozen in liquid nitrogen and kept at -80 ºC until its use for protein extraction. 55 frozen tissues were used: 38 PaC tissues of different stage, 5 tissues from other gastrointestinal malignancies and 12 control tissues (Table 5). Controls included 2 healthy pancreases from autopsy and 10 pancreatic non-tumor (NT) tissues adjacent to the cancer region (2 from cholangiocarcinomas, 2 from duodenum adenocarcinomas, 1 from ampullary carcinoma and 5 from PaC).

Duran Sidera, A.

Table 5. Frozen pancreatic tissue pieces used for protein lysate obtention. Patient's clinico-pathological characteristics.

Pathology	Cases	Male / Female	N	Age average	Age range
Control	Healthy (from autopsies)	M	1	50	-
		F	1	68	-
	NT cholangiocarcinoma	M	1	78	-
		F	1	71	-
	NT duodenum adenocarcinoma	M	0	-	-
		F	2	75	72-78
	NT ampullary carcinoma	M	1	67	-
		F	0	-	-
	NT PaC	M	4	70.3	64-77
		F	1	64	-
PaC	IB	M	1	76	-
		F	0	-	-
	IIA	M	4	68.8	65-70
		F	2	71	64-78
	IIB	M	11	69.7	57-79
		F	13	61.3	49-79
	III	M	4	63.3	52-73
		F	0	-	-
	IV	M	2	70.5	65-76
		F	1	59	-
Other gastrointestinal malignancies	Cholangiocarcinoma	M	1	78	-
		F	1	71	-
	Duodenum adenocarcinoma	M	0	-	-
		F	2	62	52-72
	Ampullary carcinoma	M	0	-	-
		F	1	77	-

Whole protein homogenates were obtained through lysis with lysing matrix beads D (MP Biomedicals, CA, USA) in a FastPrep-24 Instrument (MP Biomedicals, CA, USA) in a modified RIPA buffer (50 mM Tris pH7.4, 150 mM NaCl, 0.1% NP-40, 2 mM EDTA, 50 mM NaF, 1 mM PMSF, 0.2 mM Na_3VO_4, protease inhibitors Complete ULTRA Tablets™ (Roche Diagnostics, Mannheim, Germany)). Total protein quantification was determined by the Bradford assay (Bio-Rad Laboratories, CA, USA) following manufacturer's instructions.

1.2.2. Serum samples

Blood serum was collected from 49 individuals including 13 healthy donors, 13 ChP patients and 23 PaC patients of different stage (Table 6). Serum was collected following standard procedures, aliquoted and stored at -80 °C until its use. CA19-9 values of all samples were routinely determined by the Elecsys® CA19-9 assay in a Cobas E801 Module (Roche, Mannheim, Germany) at the clinic laboratory of the Hospital Universitari Josep Trueta. MSLN protein levels were quantified with Mesomark® (Fujirebio Diagnostics, PA, USA) by Reference Laboratory, Barcelona, Spain. Samples were immunopurified for specific proteins as detailed below (section 5).

Table 6. Serum samples for specific glycoforms quantification. Clinico-pathological characteristics of the population cohort.

Pathology	Cases	Male / Female	N	Age average	Age range
Control	Healthy	M	7	62.9	44-79
		F	6	63.2	51-75
	ChP	M	10	50.1	42-72
		F	3	53.3	46-60
PaC	IIA	M	2	66	61-71
		F	3	69.3	64-78
	IIB	M	2	72.5	64-81
		F	4	73.3	70-78
	III	M	4	63.3	52-73
		F	1	49	-
	IV	M	5	62.8	30-76
		F	2	65.5	59-72

2. Glycosidase digestion

Complete release of *N*-glycans from protein samples was achieved by digestion with *N*-glycosidase F (PNGaseF, New England Biolabs, MA, USA). Briefly, 1 µL PNGaseF was added to 10 µg of total protein in the manufacturer's supplied buffer conditions. The reaction took place at 37 °C overnight. For samples to be analysed by MS, previous reduction in 10 mM DTT for 15 minutes at 65 °C followed by alkylation in 50 mM IAA for 30 minutes in the dark was performed. 2 U PNGaseF

(Roche Diagnostics, Mannheim, Germany) per 10 µg protein were added and incubated overnight at 37 ℃.

SA was selectively digested with the use of different sialidases. For specific α2,3-linkage digestion, *Streptococcus pneumoniae* neuraminidase (NAN1, Agilent Technologies, CA, USA) was used, and 50 mU were added to 10 µg total protein. The reaction took place in 50 mM sodium phosphate pH 6 at 37 ℃ overnight. Digestion of total SA (cleavage of SA both α2,3- and α2,6-linked) was achieved through *Arthobacter ureafaciens* sialidase (ABS, Roche Diagnostics, Mannheim, Germany) digestion. 10 mU ABS per 10 µg total protein were used and incubated in 20 mM sodium phosphate buffer pH 7.2 at 37 ℃ overnight.

3. SDS-PAGE

Proteins were separated by discontinuous SDS-PAGE in order to assess their molecular weight. In brief, stacking gels were polymerized at 5% and separating gels at 10% acrylamide for MSLN analyses and 15% for REG1 evaluation. Samples were diluted in 4x Laemmli buffer (0.5 M Tris, 4% SDS, 40% glycerol, 0.1% bromophenol blue) supplemented with 5% ß-mercaptoethanol when reduced conditions were required and heated at 95 ℃ for 5 min. Electrophoresis was run at 200V until the electrophoretic front escaped the gel.

3.1. Coomassie staining

Acrylamide gels were Coomassie stained following standard protocols. In short, gels were incubated for 15 minutes in fixing solution (50% methanol, 7% acetic acid) followed by 15 minutes incubation in Coomassie blue dye (0.25% Coomassie blue (Merck Millipore, MA, USA), 50% methanol, 10% acetic acid). Gels were destained in 7% acetic acid at shaking overnight.

Gels undergoing band excision and further *N*-glycan sequencing were stained for 2 hours in the same dye solution, incubated in destaining solution 1 (50% methanol, 7% acetic acid) for 5 minutes and totally destained with destaining solution 2 (5% methanol, 7% acetic acid) until protein bands were clearly visible.

3.2. Silver staining

Proteins were fixed for 1 hour in 50% methanol, 12% acetic, 0.5 mL/L formaldehyde 37%. After three 20 minutes washes with 50% ethanol, gels were oxidized for 1 minute with 0.2 g/L $Na_2S_2O_3$, followed by 20 minutes impregnation with 2 g/L $AgNO_3$, 0.75 mL/L formaldehyde 37%. Samples were developed for about 10 minutes in 60 g/L Na_2CO_3, 4 mg/L $Na_2S_2O_3$, 0.5 mL/L formaldehyde 37% until protein bands were observed. Gels were washed twice in water and development was totally stopped with 50% methanol, 12% acetic for 10 minutes. Gels were finally washed in 50% methanol for 20 minutes.

4. Western blotting

WB was performed following standard procedures. Briefly, protein samples were loaded in Laemmli buffer and resolved by SDS-PAGE. Proteins were transferred to previously activated PVDF membranes for 4 hours at 100V in Towbin buffer (25 mM Tris, 192 mM glycine, 20% v/v methanol, pH 8.3) at 4ºC. Membranes were then blocked for 1 hour in 5% skimmed milk or 5% BSA in TBST (10 mM Tris, 100 mM NaCl, 0.1% v/v Tween-20, pH 7.4) with shaking. After TBST washing, membranes were incubated overnight at 4 ºC with the corresponding antibody: anti-MSLN, clone MN-1 (LifeSpan Biosciences, WA, USA) at 1 µg/mL in TBST with 3% skimmed milk; anti-GAPDH, clone 1E6D9 (Proteintech, IL, USA) at 0.1 µg/mL in TBST 5% skimmed milk.; anti-REG1A, ab47099 (Abcam, Cambridge, UK) at 1 µg/mL in TBST 3% BSA; anti-REG1B, ab233210 (Abcam, Cambridge, UK) at 0.5 µg/mL in TBST 3% BSA. Following short water and TBST washing, membranes were then incubated for 1 hour with an HRP-conjugated goat anti-mouse antibody (Merck Millipore, MA, USA) at 0.1 µg/mL in TBST with 0.5% skimmed milk or goat anti-rabbit antibody (Thermo Fisher Scientific, MA, USA) at 0.02 µg/mL in TBST with 0.5% BSA. Membranes were then washed with water and TBST and were finally developed with Immobilon™ Western Chemiluminescent HRP substrate (Merck Millipore, MA, USA). Chemiluminescence was visualized using the imaging system Fluorchem SP (Alpha Innotech, CA, USA) under non-saturating conditions.

For glycan determinants detection, membranes were blocked in TBST with 2% polyvinylpyrrolidone overnight with shaking at room temperature. They were washed in TBST, and then incubated for 2 hours with different lectins (Table 7) in lectin buffer (100 mM Tris, 150 mM NaCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, pH 7.6). After washing with water and TBST, depending on the label of the used lectin, membranes were incubated for 1 hour with HRP-conjugated streptavidin (GE Healthcare, IL, USA), anti-FITC HRP-conjugated antibodies or anti-DIG HRP-conjugated antibodies (Roche Diagnostics, Mannheim, Germany) diluted in TBST with 0.5% BSA. Following water and TBST washing, membranes were developed as for protein detection.

Table 7. Lectins used in WB. Conjugated label, recognised glycan and concentration used for each lectin in WB experiments.

Lectin	Label	Glycan determinant	Concentration	Reference
Sambucus nigra Agglutinin (SNA)	Biotin	Neu5Acα6Gal/GalNAc	2 µg/mL	B-1305 (Vector Laboratories)
Maackia amurensis Lectin II (MAL-II)	Biotin	Neu5Acα3Galß3GalNAc	2 µg/mL	B-1265 (Vector Laboratories)
Aleuria aurantia Lectin (AAL)	Biotin	Fucα3/6GlcNAc	2 µg/mL	B-1395 (Vector Laboratories)
Pholiota squarrosa Lectin (PhoSL)	Biotin	Core fucose (α1,6GlcNAc)	2 µg/mL	[385]
Peanut Agglutinin (PNA)	Biotin	Galß3GalNAc (T antigen)	5 µg/mL	B-1075 (Vector Laboratories)
	Digoxigenin		10 µg/mL	11 210 238 001 (Roche)
Phaseolus vulgaris Leucoagglutinin (PHA-L)	Biotin	Galß4GlcNAcß6(GlcNAcß2Manα3)Manα3	5 µg/mL	B-1075 (Vector Laboratories)
Phaseolus vulgaris Erythroagglutinin (PHA-E)	Fluorescein	Galß4GlcNAcß2Manα6(GlcNAcß4)(GlcNAcß4Manα3)Manß4	2 µg/mL	FL-1121 (Vector Laboratories)
Ulex europaeus Agglutinin (UEA)	Fluorescein	αFuc (preferably α1,2)	2 µg/mL	FL-1061 (Vector Laboratories)
Vicia villosa Lectin (VVL)	Fluorescein	GalNAc (preferably Tn antigen)	2 µg/mL	FL-1231 (Vector Laboratories)

4.1. Stripping

Membranes were reblotted using standard stripping protocols. These were washed with TBS (10 mM Tris, 100 mM NaCl pH 7.4) before incubation with Restore Western Blot Stripping buffer (Thermo Fisher Scientific, MA, USA) for 10

minutes at room temperature plus 10 minutes at 37 ºC. Next, membranes were washed with TBS and TBST and developed under the WB protocol starting from the blocking step.

5.   Protein immunopurification

Several immunopurification approaches were conducted throughout this thesis depending on the target protein, sample origin and subsequent analysis methodology. Hence, this section has been structured based on the protein to be isolated: MSLN or REG1.

5.1. MSLN

MSLN was immunopurified from cell lines conditioned media to determine its glycosylation by (a) N-glycan sequencing and (b) WB using lectins. Samples containing 2 µg of MSLN (for N-glycan sequencing analysis) or 100 µg total protein (for WB analysis) were diluted in 300 µL incubation buffer (50 mM Tris pH 7.4, 150 mM NaCl, 2 mM EDTA, 1 mM PMSF, 0.2 mM Na_3VO_4, protease inhibitors Complete ULTRA Tablets™ (Roche Diagnostics, Mannheim, Germany)) and added to 100 µL protein G-sepharose beads (GE Healthcare, IL, USA), previously washed with water and equilibrated with the same incubation buffer, in a spin-tube (Costar, Corning Inc., NY, USA). Samples were incubated for 2 hours at 4 ºC and were then centrifuged for 2 minutes at 5000 g to elute the precleared fraction. In parallel, anti-MSLN antibody clone MN-1 (LifeSpan Biosciences, WA, USA) was bound (a) to protein G-magnetic beads (SureBeads™, Bio-Rad Laboratories) in TBST, 0.001% BSA buffer for 45 minutes at shaking or (b) to protein G-sepharose beads (GE Healthcare, IL, USA) in TBST, 0.01% BSA for 2 hours at the roller mixer. Next, the previously precleared fractions were added to the antibody-beads complex and were incubated for 4 hours at shaking at room temperature. Samples were then eluted with (a) 40 µL 2x Laemmli buffer for 10 minutes at 70 ºC before beads magnetization or (b) 70 µL 3x Laemmli buffer for 5 minutes at 95 ºC before centrifugation at 2000 g for 5 minutes.

For MSLN analysis by ELISA/ELLA, MSLN was immunopurified from both tissue protein lysates and blood serum samples. First, a mouse anti-MSLN antibody (clone MN-1, LifeSpan Biosciences, WA, USA) was covalently bound to magnetic beads (Dynabeads™ M-270 Epoxy, Thermo Fisher Scientific, MA, USA) following the manufacturer's procedures. Briefly, beads were washed with the supplemented buffer before addition of 7 µg antibody per mg of beads, which were incubated at shaking overnight at 37 °C. Beads were washed, blocked and finally resuspended at 10 mg/mL using the manufacturer's buffers. For each sample, 50 µL of the antibody-beads complex were equilibrated twice with PBST (Phosphate buffered saline, 0.05% Tween) with 0.1% BSA. Next, 50-80 µg of total protein from tissue lysates (corresponding to 6 ng of MSLN) in 1 mL modified RIPA buffer, or 1 mL blood serum were incubated for 1 hour at room temperature with shaking. Tubes were magnetized, the supernatant removed, and beads washed with PBST. MSLN was eluted in 44 µL citrate buffer 0.1 M, pH 3.1. After 2 min incubation, tubes were magnetized, samples collected and neutralized in 16 µL Tris 1 M pH 9. In the case of serum samples, this eluted fraction was diluted with 900 µL modified RIPA buffer and incubated again for 1 hour with the same antibody-beads complex previously equilibrated. Beads were washed and MSLN finally eluted following the same steps described before.

5.2. REG1

REG1 was immunopurified from tissue samples for the analysis of its glycans by WB with lectins. 100 µg of total protein were diluted in 300 µL incubation buffer and precleared for 2 hours at 4 °C in 100 µL protein A-sepharose beads (GE Healthcare, IL, USA) previously washed with water and equilibrated in RIPA buffer. In parallel, 4 µg anti-REG1A antibody (ab47099, Abcam, Cambridge, UK) or anti-REG1B antibody (ab233210, Abcam, Cambridge, UK) were coupled to 1 mg protein A-magnetic beads (Bio-Rad Laboratories, CA, USA) for 45 minutes at room temperature. Precleared samples were transferred to the antibody-magnetic beads complexes and incubated overnight at 4 °C. After magnetization and washing in TBST, REG1 was eluted by incubation with 40 µL 2x

Laemmli buffer (+2.5% ß-mercaptoetanol) for 10 minutes at 70 °C before beads magnetization.

Immunopurification of REG1A from serum samples was performed with two objectives: the quantification of specific glycoforms by ELISA/ELLA and the detection of particular glycopeptides by MS. A rabbit anti-REG1A antibody (ab47099, Abcam, Cambridge, UK) was covalently bound to magnetic beads (Dynabeads™ M-270 Epoxy, Thermo Fisher Scientific, MA, USA) as briefly described above. Then, 1 mL blood serum was incubated with 50 µL of antibody-beads complex for 1 hour at room temperature. After beads magnetisation, the supernatant was removed and beads were washed three times with PBST.

For the subsequent analysis by ELISA/ELLA, elution of REG1A was achieved by incubation with 44 µL citrate buffer 0.1 M, pH 3.1 for 2 minutes. Samples were neutralised in 16 µL Tris 1M pH 9 after beads magnetisation. The eluted fraction was diluted in 900 µL RIPA and incubated for an extra hour with the same antibody-magnetic beads conjugate. Then, beads were magnetised, supernatant removed and protein-antibody complexes washed with PBST. Final elution of REG1A was performed again with 2 minutes incubation in 44 µL citrate buffer which were neutralised with 16 µL Tris. On the other hand, for the analysis by MS, REG1A glycopeptides were eluted from the magnetic beads by digestion with trypsin (Promega, WI, USA) at a 1:2 enzyme:substrate ratio in a total volume of 35 µL of 100 mM $NaHCO_3$ buffer for 15 minutes at 37 °C. Samples were kept at 4 °C until their analysis.

6. Mass spectrometry (MS) analysis

6.1. Ultra-performance liquid chromatography – Electrospray ionisation – Quadrupole Time-of-flight (UPLC-ESI-QTof)

6.1.1. *Intact protein analysis*

MS analysis of intact MSLN was performed on rMSLN. Samples were prepared in a total volume of 10 µL and analysed with a Waters Acquity UPLC system

coupled to a Waters XEvo G2 QTof under the control of MassLynx 4.1 software (Waters, MA, USA). The chromatography separation was performed using an Acquity UPLC Protein BEH C4 column, 300 Å, 1.7 μm, 2.1x50 mm, set at 80 °C with a flow rate of 0.2 mL/min. Mobile phases were [A] water and [B] acetonitrile (ACN), both containing 0.1% formic acid (FA). A 25-minute linear gradient from 30 to 45% solvent B was used for protein separation. The mass spectrometer operated in the positive ionization mode. The capillary voltage was set at 1.5 kV, the cone voltage at 40 V and the extraction cone at 4.0 V. Source and desolvation temperatures were set at 120 °C and 600 °C, respectively. Gas flows were 20 L/hour for the cone gas and 500 L/hour for the desolvation gas. The scan range (500-5000 m/z) was calibrated with NaI. A lock mass correction with leucine-enkephalin was used throughout the experiments. Extracted ion chromatograms were peak-detected and noise-reduced in both the LC and MS domains, and then deconvoluted using the mMass software.

6.1.2. Peptide analysis

MSLN peptide analyses were performed from both (a) solution and (b) gel digested samples. For (a) solution digested samples, these were dried, resuspended in NH_4HCO_3, DTT reduced and IAA alkylated before digestion with trypsin (Promega, WI, USA) at a trypsin:protein ratio of 1:10 in 50 mM NH_4HCO_3 for 16 hours at 37 °C. For (b) gel pieces, which were already reduced and alkylated, gel bands were washed with 50 mM NH_4HCO_3, dried with ACN, and rehydrated with trypsin 1:50 under the same conditions described in (a). Peptides were extracted from the gel pieces with three washes by sonication in 200 μL $ACN:H_2O:FA$ 50:50:1. Samples were dried and resuspended in water containing 0.1% FA to the desired concentration.

Peptides were separated and detected in the same Acquity UPLC system used for intact protein experiments, but using an Acqyuity UPLC Peptide BEH C18 column, 130 Å, 1.7 μm, 2.1x150 mm. Mobil phases [A] and [B] were water 0.1% FA and ACN 0.1% FA, respectively. Flow rate was set at 0.2 mL/min, and column temperature at 40 °C. Peptides were separated in a 40-minute linear gradient from

10 to 28% solvent B. The mass spectrometer operated in the positive ionization mode. The capillary voltage was set at 3.0 kV, the cone voltage at 25 V and the extraction cone at 4.0 V. Source and desolvation temperatures were set at 120 ºC and 350 ºC, respectively. Gas flows were 2 L/hour for the cone gas and 800 L/hour for the desolvation gas. The scan range (200-2000 m/z) was calibrated with leucine-enkephalin, which was also used for the lock mass correction. Extracted ion chromatograms were peak-detected and noise-reduced in both the LC and MS domains.

6.2. Matrix-assisted laser desorption/ionisation − Time-of-flight (MALDI-Tof)

The analysis of rREG1A glycoforms was performed by MALDI-Tof both at the whole protein and glycopeptide level. Samples were loaded on a MTP 384 target plate ground steel BC and analysed in an Autoflex max MALDI-Tof-MS (smartbeam™ II Nd: YAG solid state laser, 2000 Hz) under the command of FlexControl software (Bruker, MA, USA).

The evaluation of total rREG1A was performed mixing the commercial glycoprotein with a sinapinic acid matrix (10 g/L in ACN, 0.1% TFA) in a 1:1 ratio before loading 1 µL onto the plate and letting it dry at room temperature. Data acquisition was achieved in the positive lineal mode, accumulating mass scans in a range of 10000-20000 m/z calibrated with Protein Standard I (Bruker). Laser intensity was set at 75% and a minimum of 10 shots manually obtained over the spot were acquired for each sample.

For the analysis of its glycopeptides, rREG1A was digested with trypsin (Promega, WI, USA) in an enzyme:substrate ratio of 1:10. The reaction was performed in 100 mM $NaHCO_3$ pH 7.6 for 15 minutes at 37 ºC, generating the cleavage of only the 11-residue N-terminal peptide. The obtained glycopeptides were purified in a 96-well plate format hydrophilic interaction chromatography (Waters, MA, USA). Briefly, the plate was conditioned with water and equilibrated with 90% ACN before adding the samples also in 90% ACN. After 15 minutes, wells were

washed with ACN and glycopeptides were finally eluted in 100mM NH$_4$AcO. These were totally dried in the vacuum centrifuge and finally resuspended with 10 µL of a solution containing 40% ACN and 0.1% FA. An α-cyano-4-hydroxycinnamic acid matrix (10 g/L in 30% ACN, 0.1% TFA) was used for the study of purified glycopeptides. Data acquisition was performed as for the total protein, but in a mass scan range of 1000-3000 m/z calibrated with Peptide Standard II (Bruker). All the obtained data was processed and curated with the FlexAnalysis 3.4 software (Bruker, MA, USA).

6.3. Ultra-performance liquid chromatography – Electrospray ionisation – Triple quadrupole (UPLC-ESI-TQ)

REG1A glycopeptides were additionally analysed by UPLC-TQ-MS. This methodology presents a low background signalling due to the separation of compounds in the liquid chromatography and the possibility to select specific mass-to-charge transitions to be analysed at the quadrupole, reason why it was chosen to analyse REG1A glycopeptides from serum samples. These glycopeptides were obtained by trypsin digestion on rREG1A or serum immunopurified REG1A as stated before in section 5.2, and 2 µL were injected to a Waters Acquity UPLC system coupled to a Waters XEvo TQ-S mass spectrometer under the control of Mass-Lynx software (Waters, MA, USA). An Acquity Premier Glycoprotein BEH amide column, 300 Å, 1.7 µm, 2.1x150 mm was used, with a 0.3 mL/min flow at 55 °C. Solvent A was 10mM ammonium formate with 0.04% FA and solvent B was 10 mM ammonium formate in 90% ACN. A 20-minute gradient from 80 to 55% solvent B was used. The mass spectrometer operated in the positive ionization mode, using nitrogen as a nebulising gas and argon as a collision gas. The capillary voltage was set at 3.0 kV. Source and desolvation temperatures were set at 150 °C and 600 °C, respectively, with their corresponding gas flows at 50 L/h in the cone and 1200 L/h for desolvation. Data was acquired in a selected reaction monitoring (SRM) acquisition mode, and processed using the TargetLynx XS application manager (Waters, MA, USA). Selected transitions in SRM for each glycopeptide were: non-glycosylated, 627.4 m/z; 3 NAcHex + Hex, 675.8 m/z; 3 NAcHex +

Hex + Fuc, 724.5 m/z; NAcHex, 729.1 m/z; 3 NAcHex + Hex + Neu5Ac, 772.9 m/z; 3 NAcHex + Hex + Fuc + Neu5Ac, 821.6 m/z; NAcHex + Neu5Ac, 874.4 m/z; NAcHex + Hex + Neu5Ac, 955.7 m/z; NAcHex + Hex + 2 Neu5Ac, 1101.3 m/z.

7.   *N*-glycan sequencing

MSLN *N*-glycans were sequenced using the methodology described by Royle et al. [219] with slight modifications. Briefly, immunopurified MSLN was separated on SDS-PAGE and stained with Coomassie blue. Bands corresponding to MSLN were excised, washed with 20 mM NaHCO$_3$ and ACN before reduction with 100 µL 50 mM DTT for 10 minutes at 65 ℃ and alkylation with 100 µL 20 mM IAA for 30 minutes in the dark. Then, gels were washed several times with ACN and 20 mM NaHCO$_3$. After totally drying gel pieces in the vacuum centrifuge, 100 µL PNGaseF (New England Biolabs, MA, USA) diluted 1/400 in 20 mM NaHCO$_3$ were added, plus 100 µL extra 15 minutes later. Gels were topped up with 20 mM NaHCO$_3$ buffer and incubated at 37 ℃ overnight. Glycans were totally eluted with water and ACN sonication. They were then filtered and completely dried in the vacuum centrifuge overnight. *N*-glycans were then labelled with (2-AB) as described in [386].

Derivatized 2-AB-*N*-glycans were digested for 16 hours at 37 ℃ using several exoglycosidases from Prozyme unless stated otherwise, either alone or in combination: 0.5 U/mL ABS (*Arthobacter ureafaciens* sialidase, digests α2,3/6/8/9-SA), 5 U/mL NAN1 (Sialidase S, digests α2,3-linked *N*-acetylneuraminic acid), 1 U/ml BTG (Bovine testes ß-galactosidase, digests ß1,3/4 galactose), 400 U/mL AMF (Almond meal α-fucosidase, digests α1,3/4 fucose; from New England Biolabs, MA, USA), 1 U/mL BKF (Bovine kidney α-fucosidase, digests α1,2/3/4/6 fucose, also core fucose), 8 U/mL GUH (Streptococcus pneumoniae hexosaminidase, digests ß-N-acetylglucosamine) and 25 U/mL CBG (Coffee bean α-galactosidase, digests α1,3/4 galactose). Enzymes were added for a total volume of 10 µL in

50 mM sodium acetate pH 5.5 buffer. After digestion, exoglycosidases were inactivated at 65 ºC for 15 minutes and removed through filtration in 10K microcentrifuge filtration devices (Pall Corporation, NY, USA).

2-AB labelled *N*-glycans were analysed by UPLC with fluorescence detection on a Waters Acquity UPLC H-Class system consisting of a quaternary solvent manager, sample manager and fluorescence detector under the control of Empower3 software (Waters, MA, USA). *N*-glycans were separated through HILIC in an Acquity UPLC Glycan BEH amide column, 130 Å, 1.7 µM, 2.1x150 mm. The temperature of the column was set at 40 ºC. Solvent A was 50 mM ammonium formate pH 4.4 and solvent B was ACN. A 30-minute method with a linear gradient 70-53% solvent B at 0.56 mL/min was used. Samples were injected in 20 µL 70% ACN. The fluorescence excitation/emission wavelengths were λ_{ex}=330 nm and λ_{em}=420 nm, respectively. Retention times were converted into glucose units (GU) by time-based standardization against a dextran ladder.

8. Enzyme-linked immunosorbent assay (ELISA)

Sandwich ELISAs were developed to quantify MSLN and REG1A concentration. ELISA 96-well plates (Thermo Fisher Scientific, MA, USA) were coated with 300 ng anti-MSLN rabbit polyclonal antibody (LifeSpan Biosciences, WA, USA) or anti-REG1A rabbit polyclonal antibody (Abcam, Cambridge, UK) in 100 µL Na_2CO_3-$NaHCO_3$ pH 9.6 buffer at 4 ºC overnight. After washing with saline solution (0.9% NaCl, 0.05% Tween-20), wells were blocked with 400 µL 1% BSA PBST (Phosphate buffered saline, 0.05% Tween) for 1 hour. Then, samples diluted in PBST were incubated in triplicates for 2 hours. For MSLN quantification, a calibration curve with rMSLN ranging from 0.25 ng/mL to 10 ng/mL in PBS-T was included. For REG1A, the calibration curve with rREG1A ranged from 1 ng/mL to 40 ng/mL. Following washes, 100 µL anti-MSLN antibody clone MN-1 (LifeSpan Biosciences, WA, USA) or anti-REG1A antibody (MA5-29515, Thermo Fisher Scientific, MA, USA) at 1 µg/mL in 1% BSA PBST were added to each well and incubated for 2 hours. Next, HRP-conjugated goat anti-mouse

antibody (Merck Millipore, MA, USA) at 0.33 µg/mL 1% BSA PBST was incubated for 1 hour. All incubations were undergone in a humid chamber. Plates were developed with 100 µL/well BM blue HRP substrate (Roche Diagnostics, Mannheim, Germany), and the reaction was stopped after 10 minutes for MSLN or 2 minutes for REG1A with 100 µL 1M H_2SO_4. Optical density was determined in a plate reader (Synergy 4, BioTek, VT, USA) at λ=450 nm with correction at λ=690 nm. Negative controls were wells without sample or capture antibody. Intra-assay variation was calculated from the three technical replicates in all performed experiments. Inter-assay variation was calculated for each sample with data of at least two independent assays. Data for each sample is expressed as the mean of at least two independent assays.

9.   Enzyme-linked lectin assay (ELLA)

9.1.   Cf-MSLN quantification

MSLN core fucosylated glycoforms were quantified using an enzyme-linked lectin assay (ELLA). Protein G-coated ELISA plates (Thermo Fisher Scientific, MA, USA) were washed with saline solution (0.9% NaCl, 0.05% Tween-20) and incubated with 300 ng anti-MSLN antibody clone MN-1 (LifeSpan Biosciences, WA, USA) in 100 µL PBST for 1 hour. After washing, plates were blocked with 400 µL 2% polyvinylpyrrolidone in PBS for 1 hour, followed by the incubation of samples in triplicates diluted in PBST for 2 hours. A calibration curve with rMSLN ranging from 1 ng/mL to 40 ng/mL was used. Next, 100 µL of biotinylated PhoSL [385] at 1 µg/mL in lectin buffer (100 mM Tris pH 7.6, 150 mM NaCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$) were added for 2 hours. Streptavidin-HRP conjugated from Vectastain® ABC-HRP kit (Vector Laboratories, CA, USA) was used for signal amplification, following manufacturers' procedures. All incubations were undergone in a humid chamber. Wells were developed with 100 µL BM blue HRP substrate (Roche Diagnostics, Mannheim, Germany) for about 2 minutes before stopping the reaction with 100 µL 1 M H_2SO_4. Optical density was determined in a plate reader (Synergy 4, BioTek, VT, USA) at λ=450 nm with correction at λ=690 nm. Negative controls were wells without sample. Intra-assay variation was calculated

from samples' technical replicates in all performed experiments. Inter-assay variation was calculated for each sample with data of at least two independent assays. Data for each sample is expressed as the mean of at least two independent assays.

9.2. REG1A-glycoforms quantification

Several REG1A glycoforms were attempted to be quantified by ELLAs. Transparent (for colorimetric development) or white (for luminescence assays) ELISA 96-well plates (Thermo Fisher Scientific, MA, USA) were coated with 300 ng anti-REG1A antibody (Abcam, Cambridge, UK) for in 100 µL Na_2CO_3-$NaHCO_3$ pH 9.6 buffer at 4 ºC overnight. After washing, wells were blocked with 400 µL 2% polyvinylpyrrolidone in PBS for 1 hour. Next, samples diluted in PBST were incubated in triplicates for 2 hours, as well as different calibration curves with rREG1A. Following washes, 100 µL detection lectins were added in lectin buffer and incubated for 2 hours. The used lectins included: biotinylated MAL-II, biotinylated, digoxigenin labelled and fluorescein labelled PNA, and fluorescein labelled VVL. Then, 100µL streptavidin-HRP conjugated from Vectastain® ABC-HRP kit (Vector Laboratories, CA, USA) following manufacturers' procedures, anti-digoxigenin-HRP antibody (Roche Diagnostics, Mannheim, Germany) at 16.7 mU/mL in 1% BSA PBST or anti-fluorescein-HRP antibody (Roche Diagnostics, Mannheim, Germany) at 15 mU/mL in 1% BSA PBST were added for 1 hour. All incubations were performed in a humid chamber. Two development strategies were applied: colorimetric development was achieved with 100 µL BM blue HRP substrate (Roche Diagnostics, Mannheim, Germany) for about 8 minutes before stopping the reaction with 100 µL 1 M H_2SO_4. Optical density was determined in a plate reader (Synergy 4, BioTek, VT, USA) at λ=450 nm with correction at λ=690 nm. Luminescence acquisition was performed by adding 100 µL Immobilon™ Western Chemiluminescent HRP substrate (Merck Millipore, MA, USA) just before luminescence acquisition in the plate reader (1 second integration time per well). Negative controls were wells without sample or capture antibody.

10. Statistics

Statistical analyses of the obtained results were performed using IBM SPSS Statistics, version 28 (IBM, NY, USA) and Prism 9 (GraphPad Software, CA, USA). On WB analyses, differences in protein expression between groups were assessed through the Pearson Chi-square test. Data from ELISA/ELLA involving biomarkers quantification was analysed for normality with the Shapiro-Wilk test. After meeting this criteria, one-way ANOVA with Tukey's post-hoc comparison was used when comparing three groups, while unpaired t test was performed for two-groups mean comparison. For non-normally distributed data, Kruskal-Wallis and Mann-Whitney tests were performed for the same comparisons. Correlation of independent variables was determined by Pearson correlation when values were normally distributed, and by Spearman correlation when this criteria was not met. Combination of biomarkers was achieved through a multiple logistic regression model in which the response variable corresponded to the probability of being diagnosed with PaC (variable taking value 1) or being in the control group including healthy and ChP individuals (variable taking value 0). Receiver operating characteristic (ROC) curves were analysed for each biomarker to distinguish between PaC and control groups. Comparison of the area under the curve (AUC) among different ROC curves was performed with the DeLong's method [387] using an R statistical package with the pROC library [388]. For all the analyses, $p < 0.05$ was considered statistically significant. Graphs were generated with Prism 9 (GraphPad Software, CA, USA).

Duran Sidera, A.

RESULTS

Duran Sidera, A.

CHAPTER 1. Characterisation of mesothelin glycosylation in pancreatic cancer: decreased core fucosylated glycoforms in pancreatic cancer patients' sera

Mesothelin expression in cell lines

As exposed above, the determination and quantification of MSLN glycoforms could provide new biomarkers for PaC. In order to identify the best source for MSLN glycans analysis, we first performed WB to evaluate its expression in a panel of seven PaC cell lines of diverse differentiation degree and genetic background. MSLN was detected, at different intensities, in the protein lysates from all the analysed PaC cell lines: SW1990, BxPC-3, Capan-1, Capan-2, AsPC-1, Panc 10.05 and HPAF-II (Figure 17.a). Recombinant mesothelin (rMSLN) and the ovarian cancer cell line Ovcar-8, which has been found to express MSLN, were used as MSLN expression controls. In some cell lines, MSLN precursor (a band of around 70 kDa) was also detected in addition to the mature form of the protein. It should be emphasised that mature MSLN showed up as a broad band (between 40 and 55 kDa), with modest molecular weight variations between cells, indicating a potential range of isoforms.
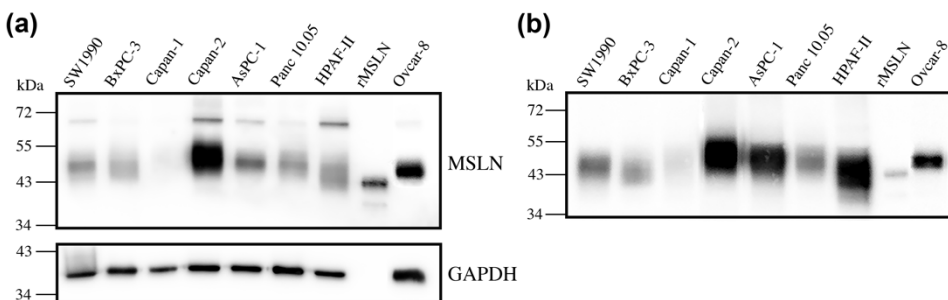

Figure 17. MSLN expression in PaC and ovarian cancer cell lines. (a) WB against MSLN in cell lines lysates (top), stripping and reblotting with GAPDH as a loading control (bottom): 20 µg of total protein loaded per lane. **(b)** WB against MSLN from cells conditioned media: 25 µg of total protein loaded per lane. In both analyses, 25 ng of rMSLN were used as a positive control.

Despite being a GPI-anchored glycoprotein, MSLN can be shed from cells yielding the secreted mature form of the protein of around 40-55 kDa. In addition, this secreted form is the most interesting from the perspective of the biomarker field, as it should be more similar to the one found in serum. For this reason, we also analysed the presence of secreted MSLN in the cell lines conditioned media (Figure 17.b). MSLN expression from cells' conditioned media showed a similar pattern than the one obtained for the whole protein lysates, but as expected, MSLN precursor form was not observed in this case. MSLN abundance in the culture supernatant was higher than that in the cell lysates when loading similar amounts of total protein. Therefore, for further MSLN-specific glycosylation studies, we selected the conditioned media of the cell lines with higher MSLN expression (Capan-2, AsPC-1 and HPAF-II PaC cells and Ovcar-8). The commercial rMSLN used as a positive control was also added to the glycosylation analyses because it is produced in a murine myeloma cell line (NS0-derived), which is likewise capable of glycosylating its secreted proteins.

Characterisation of MSLN *N*-glycosylation

MSLN immunopurification from biological samples

The study of a protein specific glycosylation relies upon the ability of obtaining pure proteins so that the detected glycans are those carried by the glycoprotein of interest. Thus, MSLN purification was a crucial step to properly assess its glycosylation. Several approaches to immunopurify MSLN from different sample sources were used alongside this study, which depended on the amount of protein needed and the analytical method to be subsequently performed.

Our first aim was to immunopurify MSLN from cell lines conditioned media using specific antibodies against MSLN bound to protein G-agarose beads. Several steps were considered in order to optimise the purification protocol. Two different antibodies were proposed for the purification. While a commercial agarose-bound antibody (clone K1, SantaCruz) was not able to capture MSLN, the use of the clone MN-1 bound to protein G-sepharose beads displayed promising results with

the presence of the protein in the purified fraction (Figure 18.a-b). However, the addition of ß-mercaptoethanol as a reducing agent in the elution buffer represented a problem due to the release of the capture antibody heavy chains, which could provide potential signal interferences with the purified MSLN. Hence, samples were finally eluted with Laemmli buffer under non reducing conditions. This provided good protein recovery and specificity, as observed by WB (Figure 18.c). This protocol was used to purify MSLN from cell lines conditioned media before the analysis of its N-glycan determinants by WB with lectins.

Next, we required to purify higher amounts of MSLN to perform N-glycan sequencing, which needs a clear and defined Coomassie band of minimum 1-2 µg in order to perform the N-glycans extraction. The first approximation continued with the use of coupled antibodies to protein G-agarose beads, scaling the quantities of sample and antibody used. WB on the obtained purified and unbound fractions showed poor MSLN recovery with antibody:antigen mass ratios of 5:1, 4:1 and 3:1 (Figure 18.d). Even a 10:1 ratio was not successful in capturing all the loaded sample, providing diffuse faint Coomassie bands and loss of protein in the unbound fraction (Figure 18.d-e). Next, we modified the methodology replacing agarose beads with protein G-magnetic beads . Despite the new approach continued leaving too much MSLN in the unbound fraction, the possibility to elute the purified sample in a minor volume displayed a clearer and more intense MSLN Coomassie band (Figure 18.f). Thus, we decided to perform the immunopurification of MSLN from the cell lines conditioned media using this last approximation. In order to optimise MSLN recovery yield, the flowthrough of the first immunopurification was subjected to a second purification step using the same beads and protocol. Coomassie staining reported high specificity as only a solid single band corresponding to MSLN, in addition to the used capture antibody, was observed on the immunopurified fractions (Figure 18.f). Interestingly, while Capan-2 and Ovcar-8 required this extra step, all MSLN was captured from AsPC-1 and HPAF-II in a single immunopurification step.
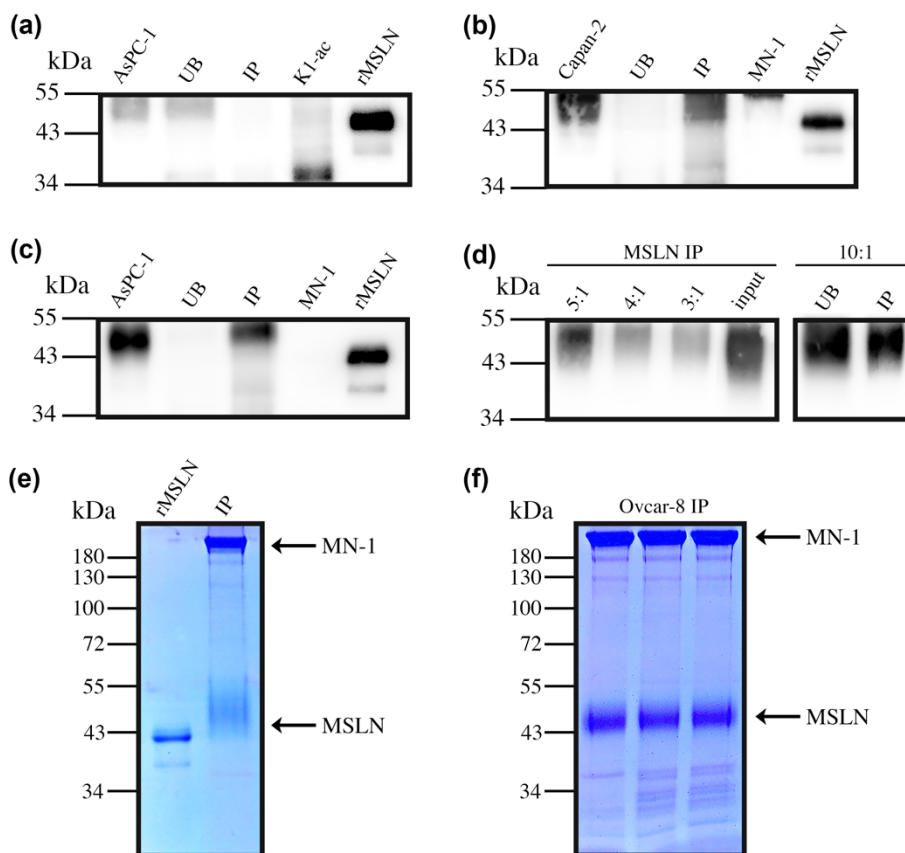
Figure 18. Optimisation of MSLN immunopurification from cell lines conditioned media. (**a**) Anti-MSLN WB of the obtained fractions during MSLN immunopurification from AsPC-1 using the K1 antibody bound to agarose beads; UB: unbound fraction, IP: immunopurified fraction. (**b**) Anti-MSLN WB of the obtained fractions during MSLN immunopurification from Capan-2 using the MN-1 antibody bound to agarose beads. (**c**) Anti-MSLN WB of the obtained fractions during MSLN immunopurification from AsPC-1 using the MN-1 antibody bound to agarose beads and eluting the samples under non-reducing conditions. (**d**) Anti-MSLN WB of the obtained fractions during high amounts MSLN immunopurification from Capan-2 with protein G-sepharose beads using different antibody:antigen mass ratios. (**e**) Coomassie staining of immunopurified MSLN from Capan-2 conditioned media using protein G-sepharose beads and a 10:1 antibody:antigen ratio. (**f**) Representative Coomassie staining of immunopurified MSLN from cell lines conditioned media (Ovcar-8) with protein G-magnetic beads.

MSLN immunopurification from biological samples presented a new challenge, as tissue lysates and serum samples comprise a rather complex biological matrix. To quantify specific MSLN glycoforms with an ELISA-like methodology, we

needed to pre-purify the samples in order to lower the background. In addition, we required to elute the sample in its native conformation, and to avoid the co-elution of the sample and the capture antibody together. This was achieved by covalently binding the antibody against MSLN to magnetic beads. After optimisation of the beads:antibody:sample ratios, incubation times, incubation buffers and elution method (acidic elution with citrate buffer 0.1 M, pH 3.1), the recovery and purity of the immunopurification process were checked by WB and silver staining over rMSLN spiked in a control tissue lysate not expressing MSLN (50 ng rMSLN in 943 µg total protein (1 mL tissue lysate) (Figure 19.a). Most MSLN was observed in the immunopurified fraction (IP), while very little was lost in the unbound fraction (UB). In addition, a main protein band at 43 kDa was observed by silver staining in the IP fraction, confirming high MSLN purity.
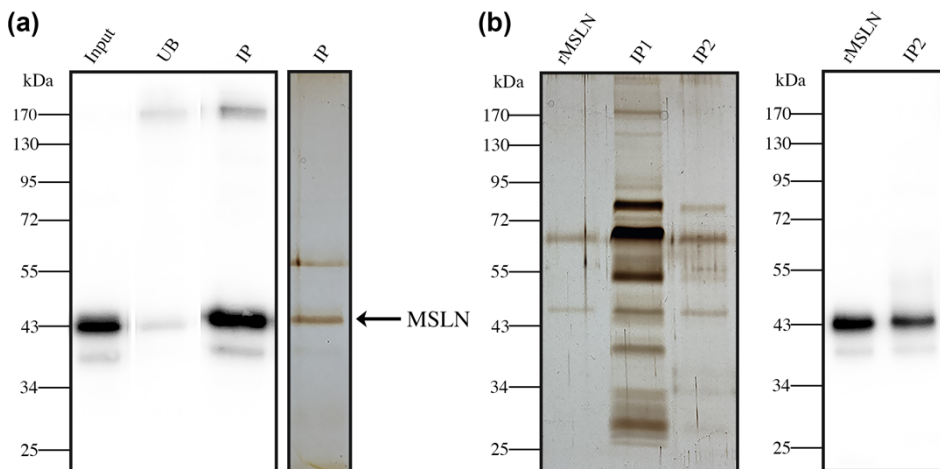

Figure 19. Patients' samples pre-purification before Cf-MSLN quantification by ELISA/ELLA. (a) MSLN immunopurification from PaC tissue lysates. WB with anti-MSLN antibody of rMSLN spiked in a tissue lysate (input), unbound (UB) and immunopurified (IP) fractions (left); and silver staining of the immunopurified (IP) fraction (right). **(b)** MSLN immunopurification from blood serum. Silver staining of rMSLN spiked in serum after one (IP1) or two (IP2) successive immunopurification steps (left); and WB with anti-MSLN antibody of the double immunopurified (IP2) fraction (right).

Protein concentration and matrix complexity is much higher in blood serum than in tissue lysates. For this reason, the pre-purification step proposed for tissue lysates was unable to properly purify MSLN and avoid background in serum samples

(Figure 19.b, IP1). Thus, a double pre-purification process was needed on these samples, which consisted in a second MSLN immunopurification using the same antibody-magnetic beads conjugates. As a result, the second immunopurification reduced impurities and background without losing much protein in the process (Figure 19.b, IP2).

Mesothelin N-glycosylation: site occupancy

N-glycan digestion with PNGaseF was first used to assess MSLN *N*-glycosylation. After PNGaseF treatment, a molecular weight decrease of about 20 kDa was observed by WB against MSLN, indicating that *N*-glycans were removed in the four selected cell lines protein lysates and in rMSLN (Figure 20.a). In addition, *N*-glycans release from rMSLN was confirmed by WB with AAL (specific lectin for Fucα1,3/6GlcNAc detection), which binds to fucosylated glycans of rMSLN as shown below. A single band at the molecular weight of glycosylated rMSLN was detected in the untreated sample, while no signal was observed after PNGaseF digestion, confirming the removal of *N*-glycans in the digested sample (Figure 20.b).
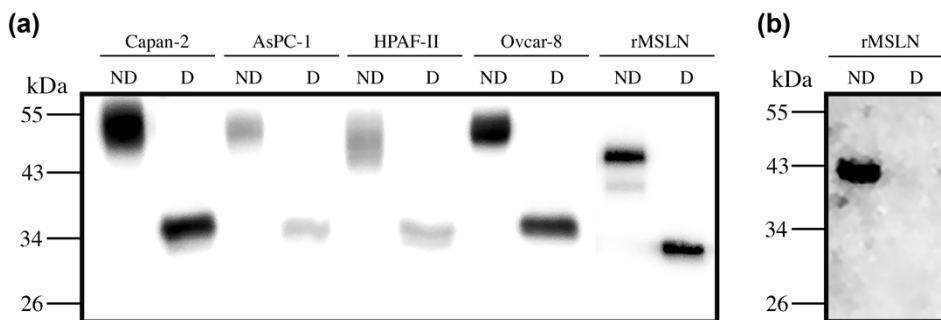

Figure 20. Digestion of MSLN *N*-glycans with PNGaseF. (a) WB of 10 μg protein lysates or 25 ng rMSLN digested (D) or not-digested (ND) with PNGaseF and detected with an anti-MSLN antibody (clone MN-1) under reducing conditions. **(b)** WB of 50 ng rMSLN detected with *Aleuria aurantia* lectin (Fucose detection) under non-reducing conditions.

In cell lines, glycosylated MSLN was detected as a single broad band with a mass between 40 to 55 kDa, indicating the presence of several isoforms. A single band at around 35 kDa was detected after PNGaseF digestion, which corresponds with

the computed mass for the human MSLN polypeptide sequence (Uniprot reference Q13421, 296-606aa). Regarding rMSLN, a major band of 43 kDa and a faint band of 39 kDa were shifted to 33 kDa after PNGaseF digestion, which is the theoretical mass of non-glycosylated rMSLN. The molecular weight change following PNGaseF digestion for both the cell lines and rMSLN confirmed the presence of *N*-glycans on MSLN and suggested a relatively high glycosylation rate because of the remarkable mobility shift.

The exact mass change on the commercial rMSLN before and after PNGaseF digestion was determined through intact protein analysis by MS (ESI-QTof). The glycoprotein mass shifted from 43.9 kDa to 33.3 kDa due to *N*-glycan release upon PNGaseF digestion, confirming the results obtained by WB (Figure 20). According to this 10.6 kDa reduction, rMSLN would probably have the three potential *N*-glycan sites occupied with highly branched and elongated structures, as the masses for complex core fucosylated and sialylated bi-/tri-/tetra-antennary glycan chains are 2368.84 kDa, 3025.07 kDa and 3681.29 kDa, respectively.

We next examined the occupancy percentage of the three theoretical MSLN *N*-glycosylation sites by peptide mapping (Table 8). For that, immunopurified MSLN SDS-PAGE bands from all samples (Figure 18.f) were excised and in-gel digested with PNGaseF to remove *N*-glycans, and then peptides were generated by trypsin digestion. The removal of *N*-glycans by PNGaseF causes the deamidation of the glycan-linked asparagine (N) to aspartic acid (D) in the resulting peptides, which induces a +0.98 Da shift in their molecular weight compared to the not glycosylated peptide. Whether the asparagine was glycosylated or not is determined by the detection (or lack thereof) of this mass shift in a peptide containing a potential *N*-glycosylation site.

The three potentially *N*-glycosylated peptides (peptide 92-101, peptide 187-200 and peptide 215-230) were detected in rMSLN by UPLC-ESI-QTof MS. All three were only identified with a mass corresponding to the D-X-S/T sequence, indicating that all the *N*-glycan sites were fully occupied (Table 8). Concerning MSLN secreted in the cell lines conditioned media, peptide 92-101 was detected in all

samples with full occupancy. Peptide 187-200 was also found in Capan-2 and was again completely occupied, but could not be detected in the rest of samples. Peptide 215-230 could not be identified in any form. Low MSLN quantities and weaker ionisation of peptides 187-200 and 215-230 compared to peptide 92-101 could explain the difficulty to detect them.

Table 8. MSLN tryptic peptides containing an *N*-glycosylation motif (N-X-S/T). Expected masses for non-glycosylated peptides (N) or de-N-glycosylated peptides (D) after PNGaseF digestion. Crossed boxes correspond to the identified peptides on Capan-2 (C), AsPC-1 (A), HPAF-II (H), Ovcar-8 (O) and rMSLN (r) after PNGaseF digestion of the corresponding MSLN bands.

Peptide ID	Sequence	N mass (Da)	D mass (Da)	Identification N					Identification D				
				C	A	H	O	r	C	A	H	O	r
92-101	W**NVT**SLETLK	1189.64	1190.62						X	X	X	X	X
187-200	LAFQNM**NGS**EYFVK	1646.78	1647.76						X				X
215-230	ALSQQ**NVS**MDLATFMK	1782.87	1783.85										X

To confirm that the gel digested bands corresponded only to MSLN, we searched for all its peptides. Fourteen peptides in rMSLN (52.3% coverage) and nine peptides in cancer cell lines (ten in the case of Capan-2) were identified, with a sequence coverage of 30.7% and 35.3% respectively (Table 9), corroborating MSLN presence.

N-glycan sequencing by HILIC-UPLC

N-glycan sequencing was used to characterise MSLN exact *N*-glycan structures from the different cancer cell lines and rMSLN. Glycan sequencing consisted in the release, purification and fluorescent labelling of *N*-glycans, followed by their separation in a hydrophilic column. The more polar *N*-glycans are, the more retained in the chromatographic column. Before the chromatographic analysis, purified *N*-glycans were digested using linkage- and saccharide-specific exoglycosidases. The decrease in the retention time of a peak (measured in glucose units (GU) obtained from standardisation with a dextran ladder) after an exoglycosidase digestion indicated the number of the specific monosaccharides present in that glycan structure. Subsequent treatments with the corresponding specific exoglycosidases could digest the structure until the basic *N*-glycan core. Finally, stacking

of all the obtained chromatograms and analysis of GU shifts let the reconstruction of glycan structures, being the areas under each peak the relative abundance of that particular *N*-glycan. Furthermore, to confirm the ascribed structures, the acquired GUs were further confirmed with our *N*-glycan GU database, Glycostore (https://glycostore.org) [389,390].

Table 9. MSLN tryptic peptides detected by UPLC-ESI-QTof. Crossed boxes indicate peptides identified in Capan-2 (C), AsPC-1 (A), HPAF-II (H), Ovcar-8 (H), rMSLN (rM) and deglycosylated rMSLN (dM).

Peptide ID	Sequence	Retention time (min)	Theoretical mass (Da)	Experimental mass (Da)	Mass deviation	C	A	H	O	rM	dM
84-90	MSPEDIR	7.06	846.3905	846.4038 (424.2097 + 2H$^+$)	+0.0133					X	X
102-108	ALLEVNK	8.94	785.4647	785.4818 (393.7487 + 2H$^+$)	+0.0171	X	X	X	X	X	X
247-256	LLGPHVEGLK	12.11	1061.6233	1061.6517 (354.8917 + 3H$^+$)	+0.0284	X	X	X	X	X	X
177-184	QLDVLYPK	15.16	974.5437	974.5490 (488.2823 + 2H$^+$)	+0.0053	X	X	X	X	X	X
109-122	GHEMSPQVATLIDR	17.05	1552.7667	1552.7793 (518.6009 + 3H$^+$)	+0.0126					X	X
266-270	DWILR	17.57	701.3860	701.4058 (351.7107 + 2H$^+$)	+0.0198	X	X	X	X	X	X
201-214	IQSFLGGAPTEDLK	22.15	1474.7667	1474.7676 (738.3916 + 2H$^+$)	+0.0009	X	X	X	X	X	X
233-246	TDAVLPLTVAEVQK	24.67	1482.8293	1482.8290 (742.4223 + 2H$^+$)	-0.0003	X	X	X	X	X	X
187-200	LAFQNM**NGS**EYFVK	26.14	1646.7763	1647.7828 (824.8922 + 2H$^+$)	+1.0065	X					X
92-101	W**NVT**SLETLK	27.01	1189.6343	1190.6230 (596.3193 + 2H$^+$)	+0.9887	X	X	X	X		X
15-24	EIDESLIFYK	27.12	1255.6336	1255.6336 (628.8246 + 2H$^+$)	±0.0000	X	X	X	X		
13-24*	APEIDESLIFYK	28.40	1423.7235	1423.7188 (712.8672 + 2H$^+$)	-0.0047					X	X
215-230	ALSQQ**NVS**MDLATFMK	34.27	1782.8644	1783.8636 (595.6290 + 3H$^+$)	+0.9992						X
44-58	VNAIPFTYEQLDVLK	35.96	1748.9348	1748.9320 (875.4738 + 2H$^+$)	-0.0028	X	X	X	X	X	X
26-43	WELEACVDAALLATQMDR	42.51	2090.9765	2090.9950 (698.0062 + 3H$^+$)	+0.0185					X	X
61-83	LDELYPQGYPESVIQHLGYLFLK	46.30	2721.4053	2721.3921 (908.1385 + 3H$^+$)	-0.0132				X		

*rMSLN contains a residue swap in position 14 (arginine to proline) which removes a cleavage site for trypsin, thus giving rise a different peptide that the one observed in cell lines.

A key point for such analyses was the obtention of pure MSLN, which was accomplished by immunopurification with anti-MSLN antibodies coupled to protein G-magnetic beads, as stated before (Figure 18.f). As expected for their distinct origin, *N*-glycan sequencing clearly showed different structures between MSLN purified from human cancer cell lines and rMSLN, produced in murine cells. Throughout this results' section, structures will be denoted as follows: *F* at the start of the abbreviation indicates a core Fuc α1,6-linked to the inner GlcNAc; *Ax*, the number (x) of antenna (GlcNAc) on the trimannosyl core: A1, mono-antennary; A2, biantennary with both GlcNAcs ß1,2-linked; A3, tri-antennary with two GlcNAc linked ß1,2 to both mannoses and the third GlcNAc linked ß1,4 to the α1,3-linked mannose; A4, tetra-antennary with GlcNAcs linked as in A3, and with an additional GlcNAc ß1,6-linked to α1,6-mannose; *B* indicates bisecting GlcNAc linked ß1,4 to the ß-mannose of the *N*-glycan core; *Fx*, number (x) of Fuc linked α1,3/4 to antenna GlcNAc; *Gx*, number (x) of ß1,3/4-linked Gal on antenna; *Lacx*, number (x) of *N*-Acetyllactosamine (LacNAc) repeats; *Gax*, number (x) of terminal α1,3-Gal; *Sx*, number (x) of SA linked to Gal and (3),(6) indicates the SA linkage, α2,3 or α2,6 respectively. Structures' relative abundance, calculated by the area under each peak, is represented by their percentage among all assigned structures. A schematic representation for all the assigned structures, their retention time and relative abundance in every profile is shown in Table 10, Table 11 and Table 12, at the end of this chapter.

The complete identification of rMSLN *N*-glycans through exoglycosidase digestions showed several complex core fucosylated (α1,6-linked) tri-antennary structures with terminal α1,3-Gal moieties, differing in the number and position of these terminal α1,3-Gal and the presence or absence of external Fuc and SA (Figure 21).

Figure 21. HILIC-UPLC profiling of rMSLN *N*-glycans. From top to bottom, chromatograms for the undigested profile (UND) and after digestion with the specified exoglycosidases. Retention times were standardised against a dextran hydrolysate with glucose units (GU). Most abundant *N*-glycan structures are shown and represented following symbol nomenclature for glycans (SNFG) guidelines. Dotted arrows represent main peaks mobility after subsequent exoglycosidases digestions. All glycans with their corresponding GUs and abundance are depicted in Table 10. ▶
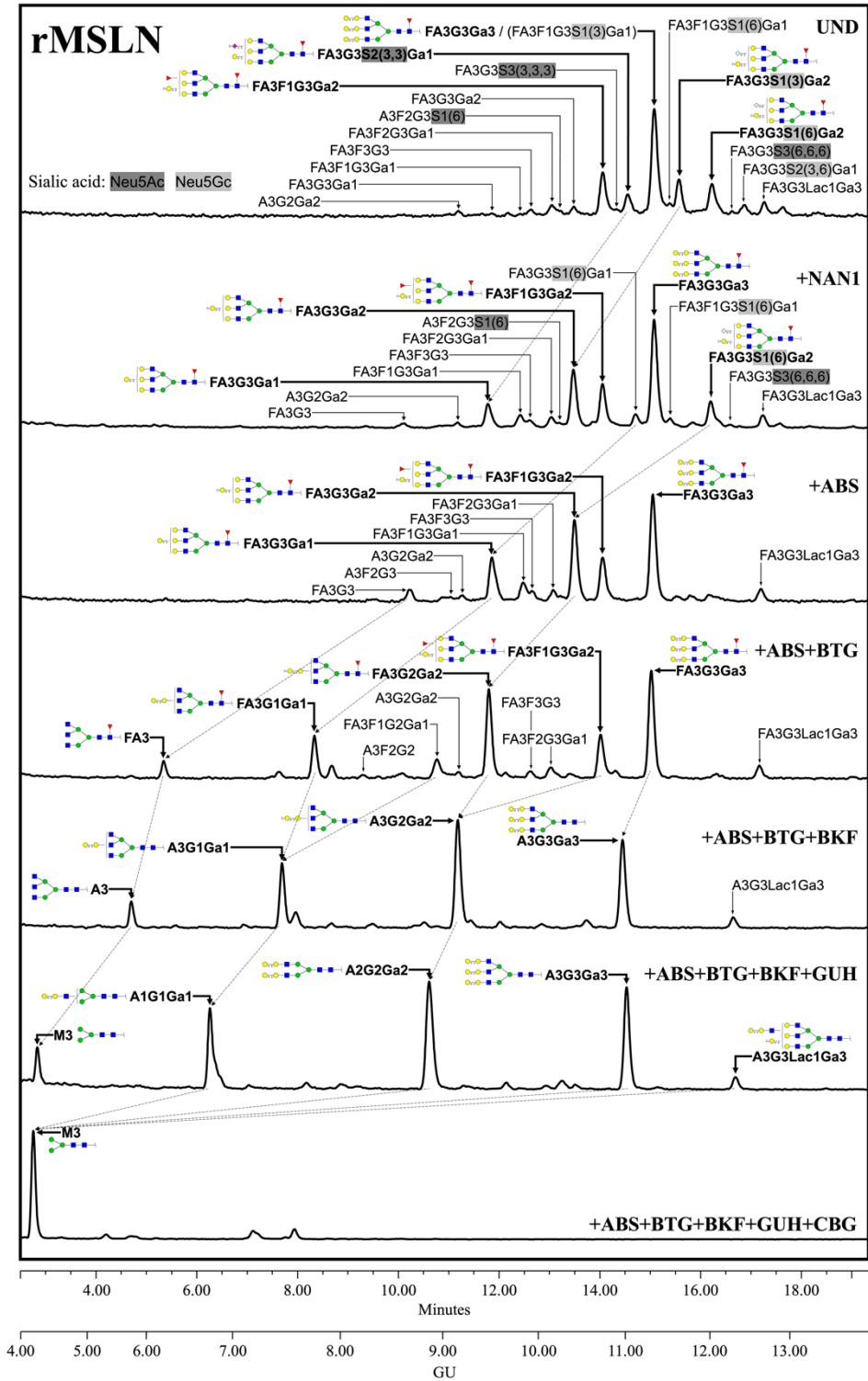
SA was found on 42.2% of structures, with a predominance of Neu5Gc (31.2%) over Neu5Ac (11.0%). Neu5Gc presents an extra carbonyl group than Neu5Ac, making it more polar. Accordingly, the digestion of Neu5Gc produces a higher GU decrease, which enables its differential assignment. Regarding the SA linkage, α2,3-SA was overexpressed against α2,6-linked SA (23.3% vs. 15.3%). SA α2,3- or α2,6-linked were assigned by differential shift after NAN1 or ABS digestion, which remove α2,3-SA or both α2,3-SA and α2,3-SA, respectively. A single disialylated structure carried both α2,3- and α2,6-SA (3.6%). We also observed several *N*-glycans with external fucosylation on Gal (27.1%), as well as a structure carrying a LacNAc repeat (4.0%). Identification of *N*-glycan isomers consisting of tetra-antennary vs. tri-antennary structures containing LacNAc repeats were performed by analysing their GU shifts after BTG digestion, as those structures with the LacNAc motif have one less terminal Gal to be digested.

Regarding human MSLN from cells' conditioned media, most structures corresponded to complex core fucosylated *N*-glycans with terminal sialylation. The main glycans were assigned from the identification obtained after specific exoglycosidases digestion. Conditioned media from Ovcar-8 cells yielded higher amount of MSLN *N*-glycans than the one obtained from PaC cells, and this allowed to perform a complete panel of exoglycosidase digestions to characterise Ovcar-8 glycans (Figure 22). The undigested profile showed high *N*-glycan heterogeneity, mainly due to differences in SA as the sialidase treatment gave rise to a more reduced number of complex *N*-glycans. Actually, sialylation confered such a massive heterogeneity that *N*-glycan structures could poorly be assigned in the undigested profile. SA was mainly found in α2,3 linkage, as the digested profile for NAN1 (α2,3-SA digestion) only differed in a single peak (GU=11.16, 18.3%) from ABS digestion (both α2,3- and α2,6-SA digestion). The main peak obtained after ABS digestion was at GU=10.11, which contained a core fucosylated tetra-antennary structure (23.1%) together with its bisected form (8.0%). The relative quantification of these two glycans was performed in further digestions (+ABS+BTG+BKF) that yielded two separate peaks corresponding to these two structures. In addition, core fucosylated tetra-antennary *N*-glycans with a LacNAc repeat were identified
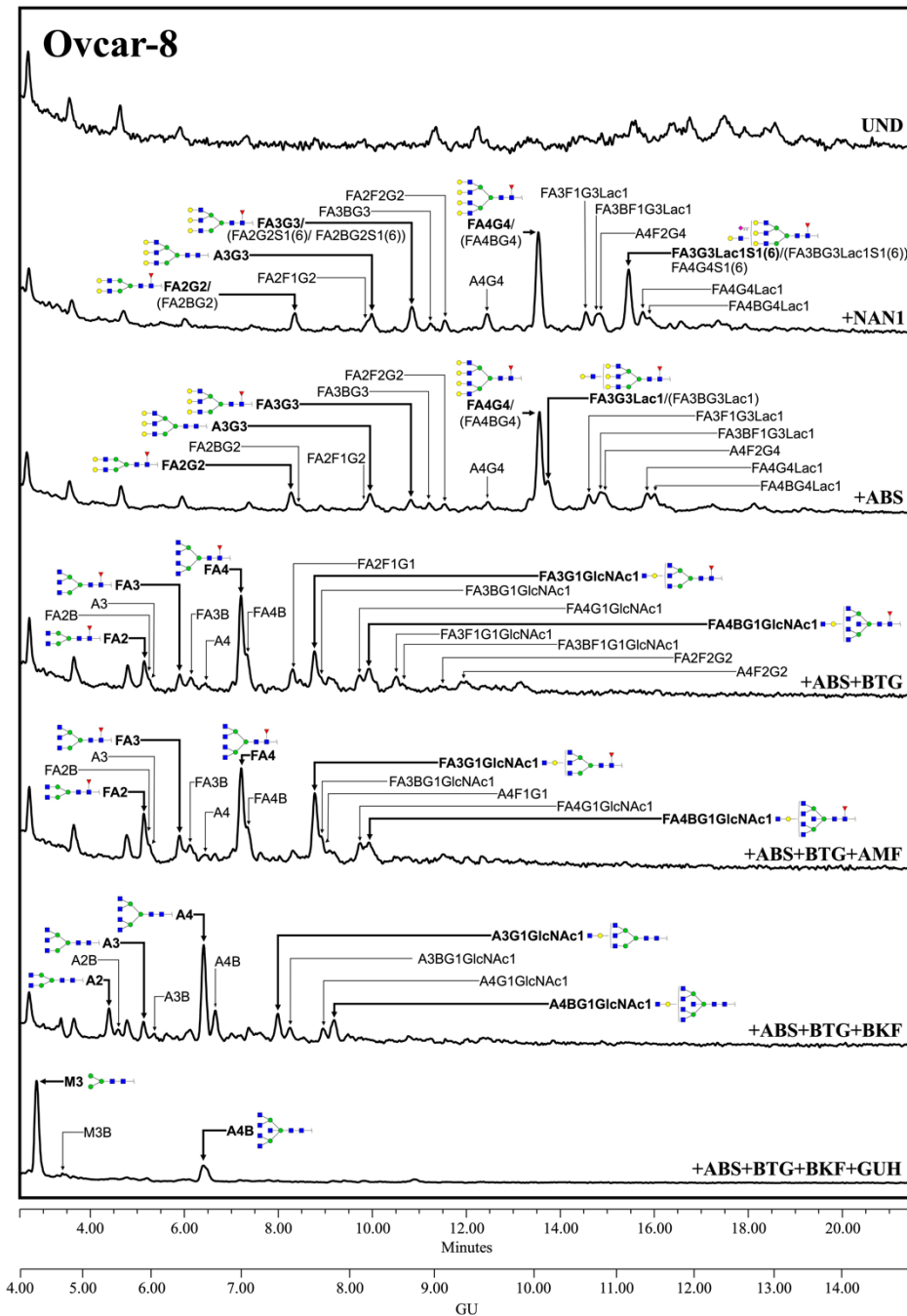
Figure 22. HILIC-UPLC profiling of MSLN *N*-glycans from Ovcar-8 conditioned media. From top to bottom, chromatograms for the undigested profile (UND) and after digestion with the specified exoglycosidases. Retention times were standardised against a dextran hydrolysate with glucose units (GU). Most abundant *N*-glycan structures are shown and represented following SNFG guidelines. All glycans with their corresponding GUs are depicted in Table 11.

(15.1%), as well as other non-core fucosylated tetra-antennary structures (11.6%). In total, tetra-antennary *N*-glycans accounted for about 57.8% of all glycans. Complex bi- and tri-antennary structures were less prevalent, accounting for about 13.5% and 28.7%, respectively. Approximately two thirds of these tri-antennary glycans carried a LacNAc repeat. The bisecting GlcNAc determinant was also present in tri- and bi-antennary glycans, and in total accounted for about the 28% of all *N*-glycans. Core fucosylation was found in most glycans (about 86%), which digested to the unfucosylated ones after BKF digestion. Some glycans (19.5%) also showed outer arm fucosylation, as shown by the digestion with AMF that do not digests core-fucosylated ones.

MSLN *N*-glycans from PaC cells were analysed after ABS digestion (Figure 23) because terminal sialylation caused such heterogeneity in the undigested profile that most structures could not be distinguished from background in that profile (data not shown), which points out the vast extension of sialylated structures. The major neutral structures obtained after ABS digestion were assigned from their GUs in Glycostore and through comparison with the structures obtained in the ABS digestion of Ovcar-8 cells. We observed several complex fucosylated *N*-glycan structures assigned to bi-, tri- and tetra-antennary *N*-glycans, but none of them were as prevalent as the core fucosylated tetra-antennary structure found in Ovcar-8 cells after ABS digestion. Capan-2 and AsPC-1 MSLN shared a similar *N*-glycan profile, which highly differed from that of HPAF-II. The latter only displayed three main structures, which could not be assigned, and so results' interpretation is focused on the formers. In Capan-2 and AsPC-1, only complex core fucosylated structures were identified: approximately 8.5% bi-antennary (10.5/6.7), 21.5% tri-antennary (21.8/21.1) and 70% tetra-antennary (67.7/72.1) *N*-glycans. Among these tetra-antennary structures, 23% (15.0/31.0) carried a LacNAc repeat and 21% (27.1/14.5) carriedtwo repetitions, while the remaining 26% (25.7/26.7) were not elongated by these repetitions. The low levels of *N*-glycans precluded further exoglycosidase digestions and therefore the presence of outer arm fucosylation could not be assessed on PaC samples.
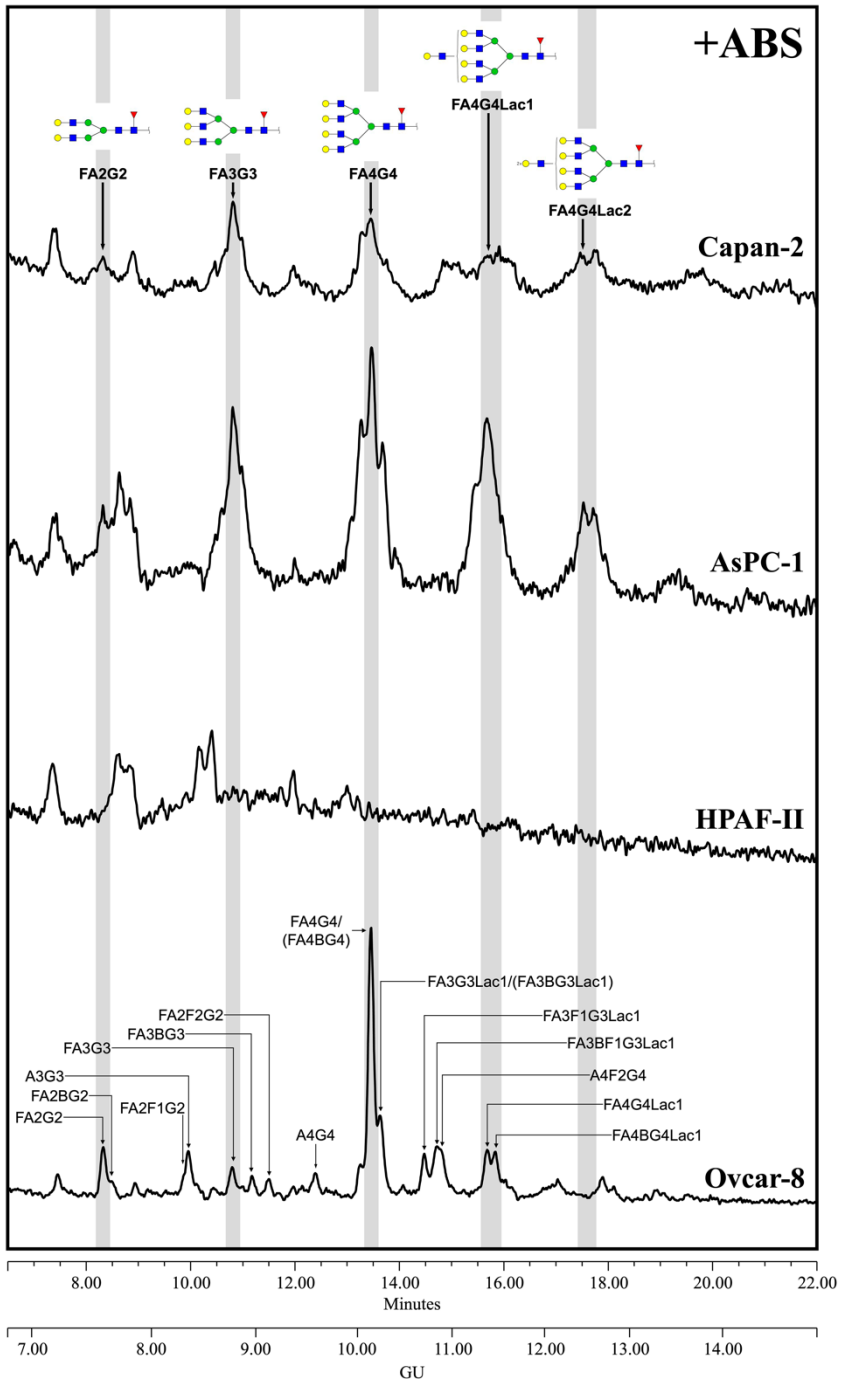
Figure 23. HILIC-UPLC profiling of MSLN *N*-glycans from the different cancer cell lines after ABS digestion. Retention times were standardised against a dextran hydrolysate with glucose units (GU). Most abundant *N*-glycan structures are shown and represented following SNFG guidelines. All glycans with their corresponding GUs are depicted in Table 12.

Glycan determinants' expression by WB with lectins and sialidase digestion

Although the characterisation of MSLN N-glycans showed the major structural glucidic/carbohydrate components on MSLN, it is not an option for a regular clinical PaC diagnosis assay based on specific MSLN glycoforms due to the inability to obtain sufficient amounts of purified MSLN from human serum or tissue. Thus, we focused on using specific lectins to analyse those MSLN glycan determinants previously described by N-glycan sequencing, in particular sialylation, fucosylation, antenna branching and bisecting GlcNAc. The selected lectins were SNA, which binds to α2,6-SA; MAL-II, which recognises most α2,3-SA glycans; AAL, which binds to Fuc determinants; PhoSL, which is specific for core fucosylation; PHA-L, which recognises ß1,6-antenna on tri- and tetra-antennary structures; and PHA-E, which binds to bisecting GlcNAc structures. We also used PNA, which recognises Galß3GalNAc, including the T antigen.

All these glycosylation determinants were analysed on immunopurified MSLN, which was obtained using protein G-agarose beads coupled with anti-MSLN antibodies. Immunopurified proteins were then separated by SDS-PAGE, blotted, and detected with the specified lectins. All membranes were stripped and reblotted for MSLN, confirming the co-localisation of both glycans and protein. To assess how much carbohydrate determinant was present on each sample, the intensity of lectin-positive bands was qualitatively relativised against the corresponding MSLN protein band. As a result, we could evaluate the level of the specific carbohydrate structures' expression on the cell lines' MSLN, as well as on rMSLN. MSLN sialylation was analysed using SNA and MAL-II lectins (Figure 24.a-b). To detect α2,6-SA using SNA, samples had to be previously reduced. Under this condition, α2,6-SA could be detected in rMSLN, but not in the immunoprecipitated MSLN from the cell lines' conditioned media. The quantity of MSLN detected after reblotting the membranes with anti-MSLN antibodies was higher in the rMSLN lane than in those of the cell lines, which could indicate that higher amounts of the immunopurified MSLN were required to detect α2,6-SA MSLN

glycans with SNA. Regarding MAL-II, none of the MSLN samples showed binding to this lectin. As *N*-glycan sequencing had shown SA on MSLN *N*-glycans in all samples, we attempted to confirm SA presence through specific sialidase digestions with NAN1 and ABS followed by WB with anti-MSLN antibodies (Figure 24.c). The results showed that, in all cell lines except for HPAF-II, digestion of α2,3-SA by NAN1 decreased MSLN molecular weight as much as total release of SA with ABS digestion. This indicated that MSLN from all samples, excluding HPAF-II, was sialylated, primarily presenting α2,3-linked SA as was described for Ovcar-8 in the *N*-glycan sequencing. Only neutral structures were found in HPAF-II since sialidase digestions failed to reveal any molecular weight shift.
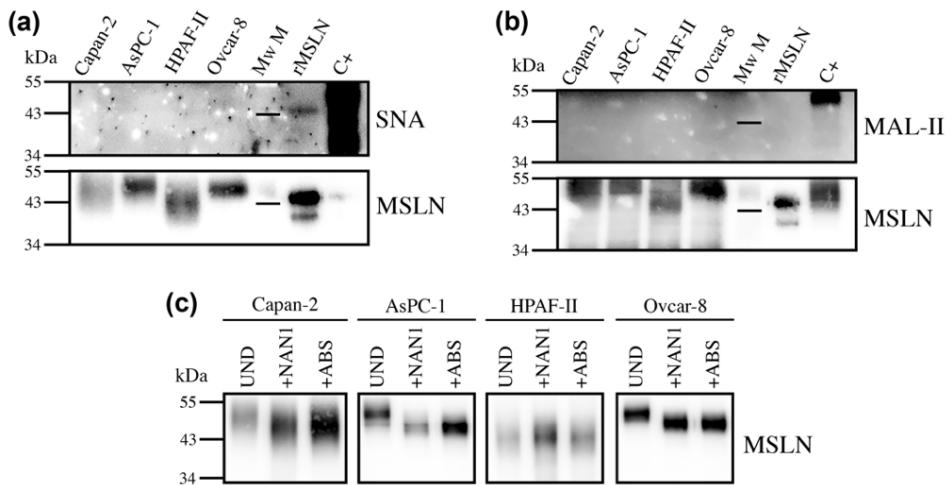
Figure 24. Analysis of sialylation on immunopurified MSLN from cell lines conditioned media by WB with lectins and sialidases digestion. Samples loaded under reducing conditions except for panel (c). For each lectin, top panel shows the specific glycan determinant recognition while bottom panel shows the result after membrane stripping and MSLN detection. Mw M: molecular weight marker -with the 43 kDa line represented-, C+: protein lysate used as positive control for glycan-lectin recognition. **(a)** SNA for α2,6-SA. **(b)** MAL-II for α2,3-SA. **(c)** WB of MSLN from conditioned media treated with sialidases. UND: undigested, +NAN1: α2,3-SA digestion, +ABS: all SA digestion.

To go further in the cancer cell lines MSLN glycophenotyping, we also assessed the levels of fucosylation, branching and bisecting GlcNAc using lectins (Figure 25). All cell lines were found to have core fucosylation, detected with PhoSL (spe-

cific for core fucosylation). Overall fucosylation (detected by AAL) was also observed in all cells except for AsPC-1. Antenna branching was assessed through the detection with PHA-L, which detected the ß1-6 ramification in branched *N*-glycans (Galß4GlcNAcß6(GlcNAcß2Manɑ3)Manɑ3). This feature was shown in all cell lines, being especially remarkable in AsPC-1. Bisecting GlcNAc, evaluated with PHA-E, was also found on MSLN from all cell lines at different degrees. Finally, the presence of terminal Gal with the PNA lectin, which preferentially binds to Galß3GalNAc, could not be detected (Figure 25.e). Altogether, PhoSL, PHA-L and PHA-E stood as the best options to develop a specific assay to detect core fucose, antenna branching and bisecting GlcNAc on MSLN, respectively.
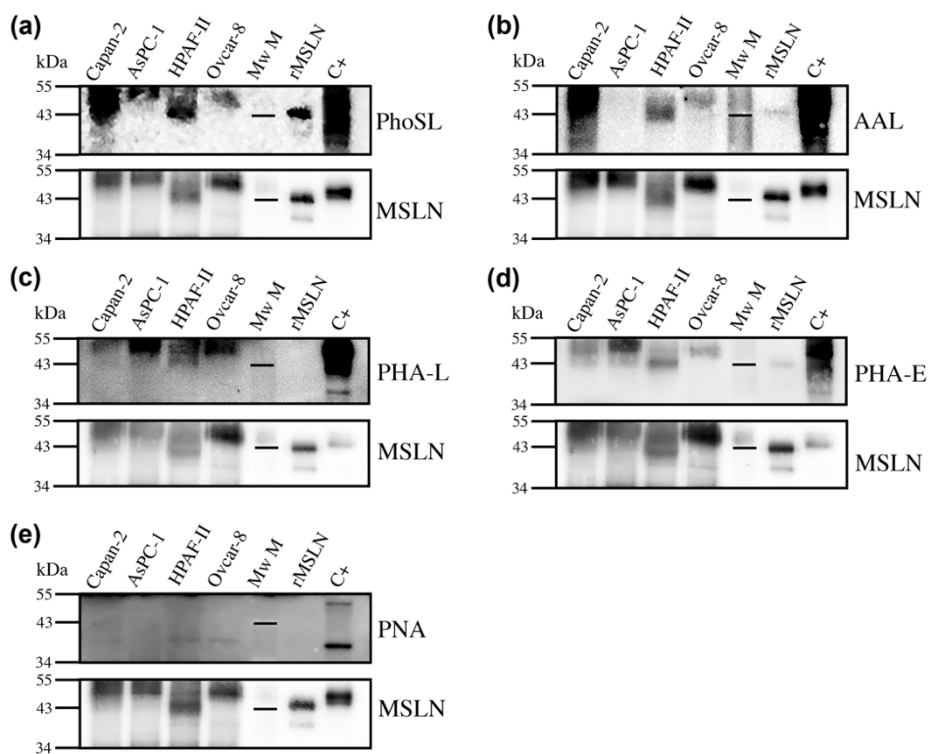
Figure 25. Analysis of glycan determinants on immunopurified MSLN from cell lines conditioned media by WB with lectins. Samples loaded under non-reducing conditions. For each lectin, top panel shows the specific glycan determinant recognition, while bottom panel shows the result after membrane stripping and MSLN detection. Mw M: molecular weight marker -with the 43kDa line represented-, C+: protein lysate used as positive control for glycan-lectin recognition. (**a**) PhoSL for core Fuc. (**b**) AAL for Fuc. (**c**) PHA-L for ß1,6-antenna. (**d**) PHA-E for bisecting GlcNAc. (**e**) PNA for terminal Gal.

Mesothelin expression in pancreatic tissues

After the characterisation of MSLN N-glycans from PaC cell lines, we assessed the presence of specific MSLN glycoforms in PaC tissues. First, MSLN expression was evaluated by WB on tissue lysates corresponding to different staging PaC patients (n=31) and control individuals (n=10) (Figure 26). MSLN was found in 77.4% (24 out of 31) of the PaC tissues and only in one control tissue, which was a non-tumour tissue adjacent to a pancreatic tumour. In two tissues from other gastrointestinal cancers, MSLN was also slightly expressed. MSLN bands in PaC samples were observed in 5/6 (83.3%), 16/21 (76.2%), 2/2 (100%) and 1/2 (50%) of stages IIA, IIB, III and IV patients, respectively. Thereby, MSLN was significatively expressed in PaC tissues, despite no differences were detected between staging groups.
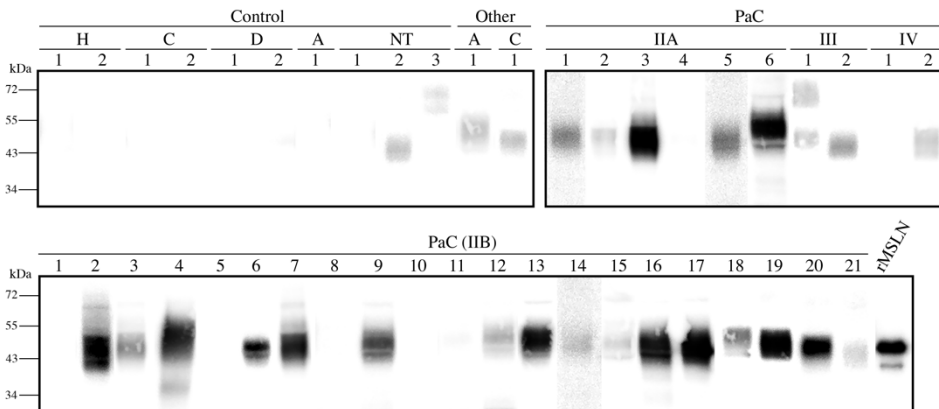
Figure 26. MSLN expression in pancreatic tissue lysates. WB under non-reducing conditions on 20 µg total protein from pancreatic tissue lysates. Each number represents a different patient in each condition. 25 ng rMSLN were used as a positive control. All membranes were equally exposed to chemiluminescence for lane comparison. Control samples included two healthy pancreas (H); pancreatic non-tumour tissues adjacent to the cancer region from: two cholangiocarcinomas (C), two duodenum adenocarcinomas (D), one ampuloma (A) and three PaC (NT). Cancer tissues included thirty-one PaC of different stage (IIA, IIB, III and IV) and other tumours: one ampuloma (A) and one cholangiocarcinoma (C)).

Core fucosylated mesothelin abundance in pancreatic tissues

After examining MSLN expression in pancreatic tissues, we attempted to determine its specific glycoforms in those tissue samples with higher MSLN expression using lectins. As WB is a semiquantitative methodology, we aimed to develop a combined ELISA/ELLA assay to quantify the specific MSLN glycoforms. A sandwich ELISA was developed to measure MSLN concentration and an ELLA to determine particular glycan levels on MSLN. The abundance of a particular glycan on MSLN would correspond to the ratio of the MSLN glycan to MSLN protein concentration.

In the light of previous results with cell lines using lectins, fucosylation, branching and bisecting GlcNAc expression using the lectins AAL/PhoSL, PHA-L and PHA-E, respectively, stood as the best options to develop the methodology. Preliminary assays using PHA-E on rMSLN did not provide enough sensitivity to detect MSLN glycoforms with bisecting GlcNAc. A calibration curve with a modest slope (Figure 27.a), combined with the low signal obtained in WB bands (Figure 25.d), refrained us from further explore this detection. Similar problems were met when using PHA-L to detect ß1,6-branched structures, where the slope of the calibration curve was even lower (Figure 27.b). In order to enhance the detection, protein G-coated ELISA plates were used, which facilitate the proper orientation of the capture antibody. However, no improvement was observed. We next tried to quantify MSLN sialylated glycoforms, despite no signal had been previously obtained by WB on the MSLN from the cell lines. Detection of α2,3-SA with MAL-II was again unsuccessful in providing a calibration curve with a satisfactory slope (Figure 27.c). Moreover, plate development faced rapid saturation at low absorbance levels. On the other hand, the use of SNA to quantify α2,6-sialylated MSLN glycoforms yielded promising results, with sufficient slope on the calibration curve (Figure 27.d). Unfortunately, the analysis of such glycoforms in conditioned media from the four studied cell lines was not possible, as the limit of detection (LOD) was too high and not enough amounts of MSLN could be loaded for its analysis.

Finally, we assessed MSLN fucosylation, and specific core fucosylation, with AAL and PhoSL. The former displayed a modest slope in the calibration curve (Figure 127.e), while the use of PhoSL reported almost a 6-fold slope increase (Figure 27.f). Thus, we optimised the calibration curve, antibody/lectin dilution and incubation times for an assay using PhoSL to determine core fucosylated MSLN (Cf-MSLN).
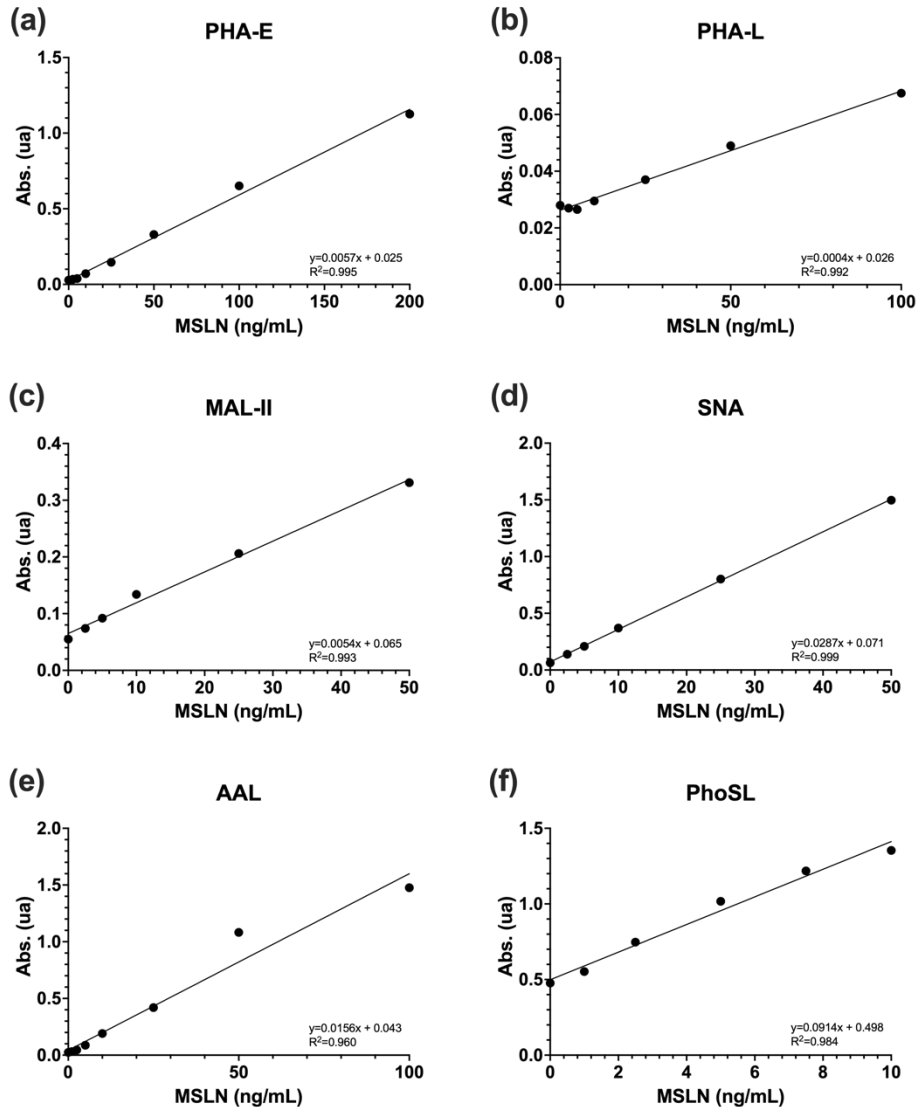

Figure 27. Calibration curves for ELLAs with different lectins. Representative calibration curves obtained during the development of ELLA assays to quantify specific MSLN glycoforms with a panel of lectins including (**a**) PHA-E, (**b**) PHA-L, (**c**) MAL-II, (**d**) SNA, (**e**) AAL and (**f**) PhoSL. For each curve, the equation of its linear range and the coefficient of determination (R^2) are indicated.

An 8-point calibration curve (0-10 ng/mL) was used for the ELISA to quantify MSLN. The LOD was 0.11 ng/mL, and linearity was demonstrated across 0.25-5 ng/mL. Intra- and inter-assay variation was of <5% and <17%, respectively. A different 8-point calibration curve was used for the detection of Cf-MSLN with PhoSL (0-40 ng/mL). The LOD was 0.41 ng/mL and the lineal range 1-10 ng/mL. There was <4% and <19% intra- and inter-assay variation, respectively.

This ELISA/ELLA methodology was used to quantify Cf-MSLN in 8 tissue samples (2 IIA (patients 3, 6), 6 IIB (patients 2, 4, 7, 16, 17, 20)), which had previously shown high MSLN expression by WB (Figure 26). Before the analysis of the samples, a pre-purification step to lower the background, especially needed for the detection with the lectin, was accomplished by immunoprecipitation with anti-MSLN antibodies covalently coupled to magnetic beads. Cf-MSLN from PaC tissue lysates was quantified and relativised to MSLN protein levels. The Cf-MSLN/MSLN ratio (0.66±0.26 and 0.41±0.11 for IIA and IIB patients, respectively) was lower than the obtained for rMSLN, which was used as standard (Figure 28). Comparison with a healthy group was not possible as healthy pancreatic tissues do not express MSLN. Our results suggested that Cf-MSLN abundance in PaC tissues could decrease along the disease progression, as levels in IIB patients were inferior to those in IIA individuals, despite no conclusions could be obtained due to the low number of samples. However, the determination of Cf-MSLN levels on immunopurified MSLN from human PaC tissues using ELISA/ELLA provided the methodological basis to analyse them in serum samples.
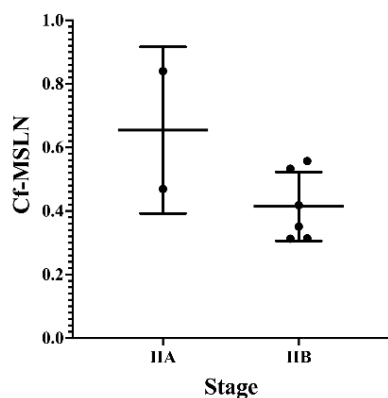
Figure 28. Analysis of Cf-MSLN in PaC tissues. Results obtained for the quantification of Cf-MSLN in eight PaC samples with the developed ELISA/ELLA methodology, represented as the mean ± SD for each PaC stage.

Serum core fucosylated mesothelin as a PaC biomarker

As MSLN is shed from the cell membrane, we next analysed Cf-MSLN abundance on serum samples from 15 PaC patients and compared it with the levels of 5 healthy donors and 11 ChP patients. As in the case of tissue samples, a pre-purification step was mandatory to reduce background. Due to its higher protein concentration, serum samples required a double immunopurification process with antibodies covalently coupled to magnetic beads.

The analysis of the ratio Cf-MSLN/MSLN on immunopurified MSLN from serum samples showed that PaC patients' serum Cf-MSLN ratio was significantly decreased respect the control groups (from 0.93±0.13 to 0.74±0.20) (Figure 29.a). When dividing the control group in healthy individuals and ChP patients, statistical differences were observed between the latter and PaC patients, while not among healthy and PaC patients, which could be attributed to the lack of statistic power caused by the low numbers of healthy individuals (p=0.053). No differences were observed among PaC staging. A ROC curve was calculated to determine the diagnostic potential of Cf-MSLN, which could discriminate PaC patients from the rest of control individuals with an AUC of 0.806. Reported sensitivity and specificity with a cut-off value of 0.84 (determined by the Euclidean distance) was of 73.3% and 75.0%, respectively (Figure 29.d).

Cf-MSLN ratio performance was compared to that of MSLN protein levels in the studied cohort. These were quantified with the commercial ELISA kit Mesomark®, which recognises soluble MSLN related peptides. In this case, no significant differences were observed between healthy, ChP and PaC individuals (Figure 29.b,e) and, as described, MSLN levels were not useful for PaC diagnosis. It must be noted that MSLN levels in serum correlated with MSLN protein levels quantified after immunopurification (Figure 29.g), indicating that no bias was induced during samples' processing. Altogether, while Cf-MSLN had the potential to discriminate PaC from control patients, MSLN protein levels did not have this ability.
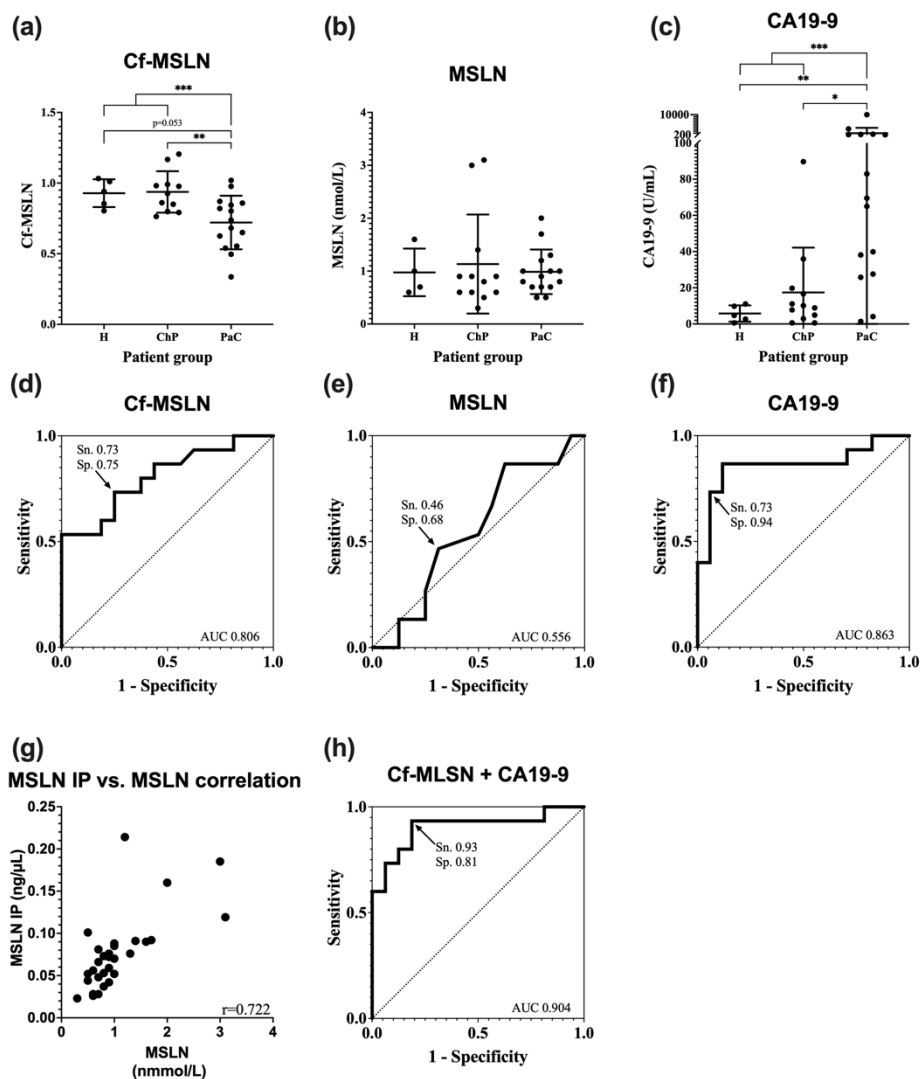
Figure 29. Biomarkers analysis in serum samples from PaC patients and controls.
(**a**) Cf-MSLN levels for healthy, ChP and PaC patients, represented as the mean ± SD for each group. **p<0.01, ***p<0.001. (**b**) MSLN protein levels for healthy, ChP and PaC patients, represented as the mean ± SD for each group. No significant differences were detected among groups. (**c**) CA19-9 levels for healthy, ChP and PaC patients, represented as the mean ± SD for each group. *p<0.05, **p<0.01, ***p<0.001. (**d**) ROC curve for discriminating between PaC and control group using Cf-MSLN ratios. (**e**) ROC curve for discriminating between PaC and control group using MSLN protein levels. (**f**) ROC curve for discriminating between PaC and control group using CA19-9. (**g**) Correlation analysis of serum MSLN levels with levels quantified after MSLN immuno-purification. (**h**) ROC curve for discriminating between PaC and control group using a combinatory biomarker comprising Cf-MSLN ratios and CA19-9.

Next, Cf-MSLN diagnostic accuracy was confronted to the current PaC bi-omarker, CA19-9. As previously reported, elevated CA19-9 was observed in PaC patients versus control groups (Figure 29.c). Unusually, no significant differences were seen between healthy and ChP individuals. PaC patients were differentiated from controls with a sensitivity of 73.3% and a specificity of 94.1% using the stand-ardised 37 U/mL cut-off value (Figure 29.f). In our study cohort, a cut-off value of 22 U/mL (established by the Euclidean distance) yielded a sensitivity and spec-ificity of 86.7% and 88.2%, respectively. Therefore, Cf-MSLN could not outper-form CA19-9 accuracy, which presented higher specificity at equal sensitivity.
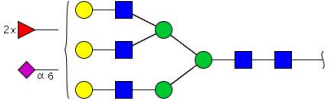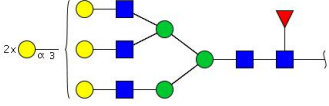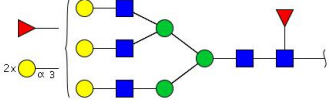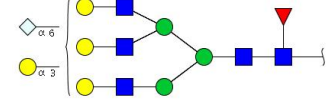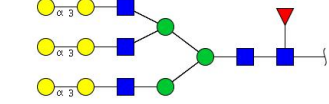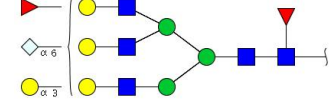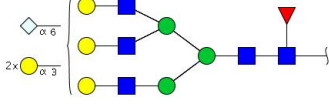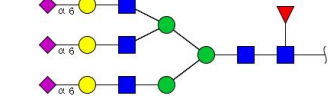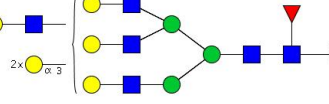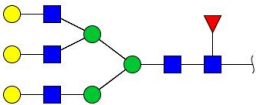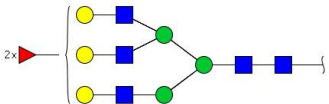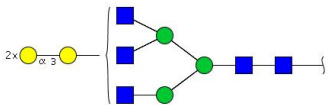
Finally, we investigated whether the combination of both biomarkers could en-hance their clinical value and diagnostic potential. A multiple logistic regression incorporating CA19-9 and Cf-MSLN was performed to assess the risk of suffering PaC. This model yielded 93.3% and 81.2% sensitivity and specificity, respectively, with an AUC of 0.904 (Figure 29.h). Interestingly, the analysis of paired values for individual samples demonstrated that the use of Cf-MSLN could correct the mis-placement of several individuals. The only ChP patient with elevated CA19-9 (89.65 U/mL) would not be assigned as PaC due to high Cf-MSLN (0.97). On the other hand, three PaC patients with low CA19-9 (<37 U/mL) would be correctly assigned regarding their Cf-MSLN values (<0.84). Only one patient, correspond-ing to a IIA PaC, would be misdiagnosed for both biomarkers, as it presented low CA19-9 and high Cf-MSLN.
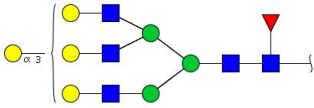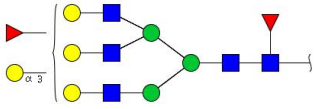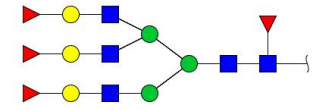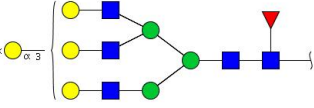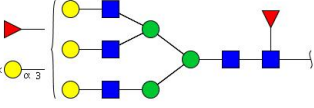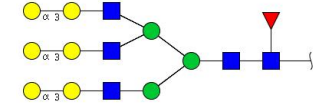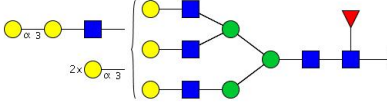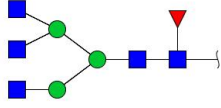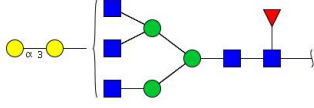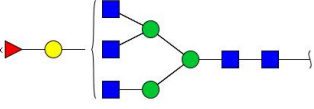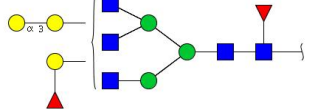
In summary, this study showed the potential of Cf-MSLN as a biomarker for PaC. However, it should be further validated in studies involving a larger cohort of pa-tients, which include sera from other benign pancreatic disorders, individuals from risk groups and other types of cancers in order to assess its specificity.

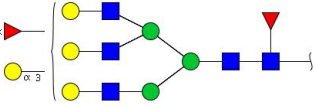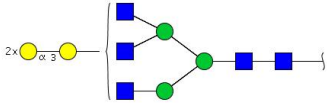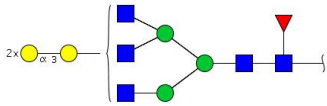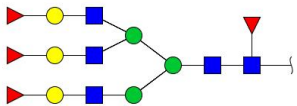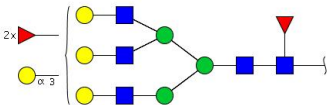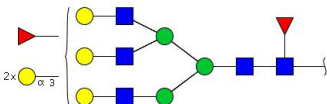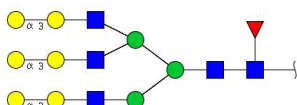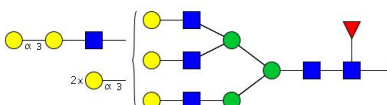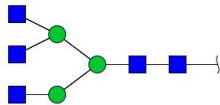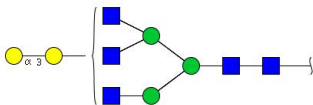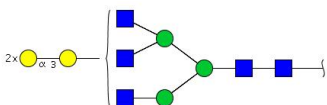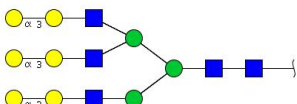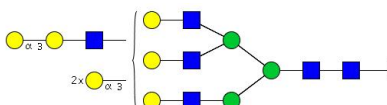**Table 10. *N*-glycans identified in rMSLN for each exoglycosidase digestion.** Structures arranged by retention time (GU), representation following symbol nomenclature for glycans (SNFG) guidelines [149] and abundance expressed as the percentage among all assigned *N*-glycans in that profile. Sialic acid form: Neu5Ac, Neu5Gc.

Retention time (GU)	Structure	Representation	Abundance (%)
		UNDIGESTED profile	
9.25	A3G2Ga2		1.14
9.58	FA3G3Ga1		0.88
9.92	FA3F1G3Ga1		0.48
10.02	FA3F3G3		2.18
10.25	FA3F2G3Ga1		3.26
10.30	A3F2G3S1(6)		1.52
10.48	FA3G3Ga2		2.85
10.80	FA3F1G3Ga2		13.56
10.93	FA3G3S3(3,3,3)		1.38
11.06	FA3G3S2(3,3)Ga1		6.80
11.39	FA3G3Ga3		29.44

11.39	FA3F1G3S1(3)Ga1		3.49
11.59	FA3F1G3S1(6)Ga1		2.61
11.67	FA3G3S1(3)Ga2		11.64
12.08	FA3G3S1(6)Ga2		9.89
12.28	FA3G3S3(6,6,6)		1.26
12.38	FA3G3S2(3,6)Ga1		3.58
12.67	FA3G3Lac1Ga3		4.05

+NAN1 digestion profile

8.76	FA3G3		1.53
9.25	A3G2Ga2		1.46
9.58	FA3G3Ga1		7.20
9.92	FA3F1G3Ga1		3.95
10.02	FA3F3G3		2.12
10.25	FA3F2G3Ga1		2.97

10.30	A3F2G3S1(6)		1.46
10.48	FA3G3Ga2		16.35
10.80	FA3F1G3Ga2		12.50
11.32	FA3G3S1(6)Ga1		3.86
11.39	FA3G3Ga3		31.20
11.59	FA3F1G3S1(6)Ga1		2.28
12.08	FA3G3S1(6)Ga2		8.59
12.28	FA3G3S3(6,6,6)		1.25
12.67	FA3G3Lac1Ga3		3.27
+ABS digestion profile			
8.76	FA3G3		3.37
9.13	A3F2G3		1.50
9.25	A3G2Ga2		1.47

9.58	FA3G3Ga1		13.32
9.92	FA3F1G3Ga1		6.66
10.02	FA3F3G3		2.34
10.25	FA3F2G3Ga1		3.41
10.48	FA3G3Ga2		24.10
10.80	FA3F1G3Ga2		12.55
11.39	FA3G3Ga3		28.16
12.67	FA3G3Lac1Ga3		3.13
+ABS+BTG digestion profile			
6.20	FA3		3.82
7.78	FA3G1Ga1		10.39
8.24	A3F2G2		1.12
8.97	FA3F1G2Ga1		7.06

9.25	A3G2Ga2		1.60
9.58	FA3G2Ga2		23.63
10.02	FA3F3G3		2.37
10.25	FA3F2G3Ga1		3.31
10.80	FA3F1G3Ga2		13.09
11.39	FA3G3Ga3		29.79
12.67	FA3G3Lac1Ga3		3.83
+ABS+BTG+BKF digestion profile			
5.79	A3		6.94
7.46	A3G1Ga1		20.54
9.25	A3G2Ga2		37.48
11.04	A3G3Ga3		31.08
12.35	A3G3Lac1Ga3		3.97
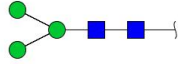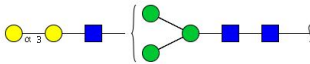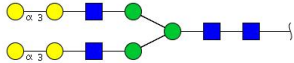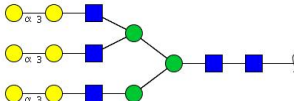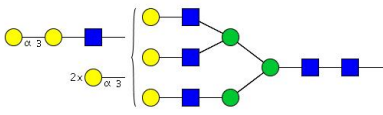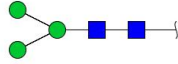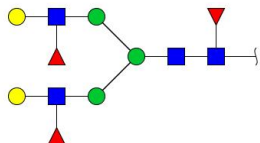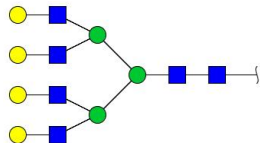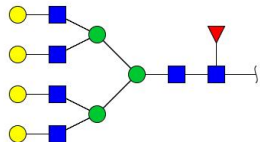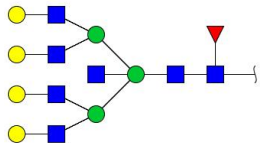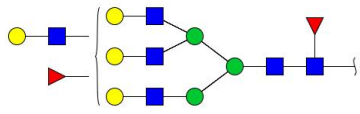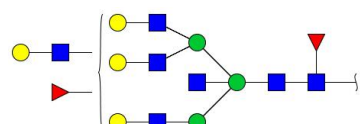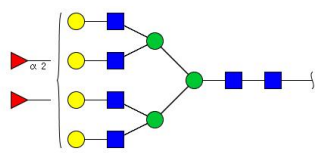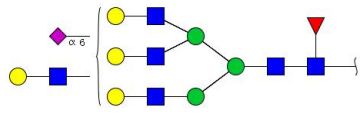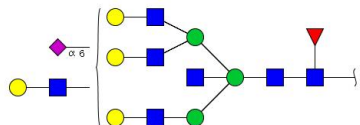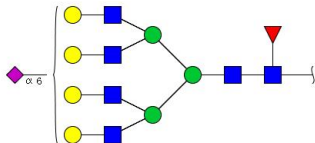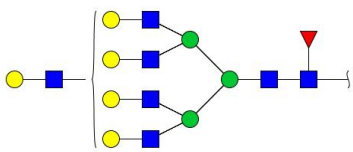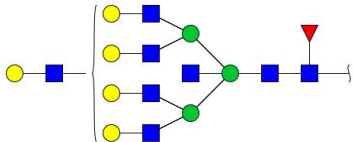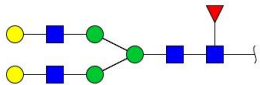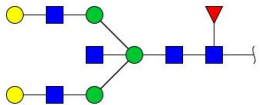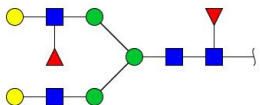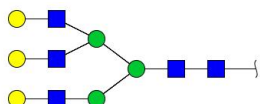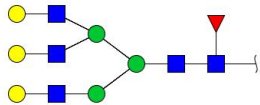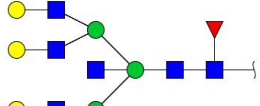
	+ABS+BTG+BKF+GUH digestion profile		
4.30	M3		7.12
6.66	A1G1Ga1		20.88
8.88	A2G2Ga2		38.34
11.04	A3G3Ga3		30.22
12.35	A3G3Lac1Ga3		3.44
	+ABS+BTG+BKF+GUH+CBG digestion profile		
4.30	M3		100.00
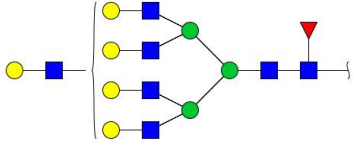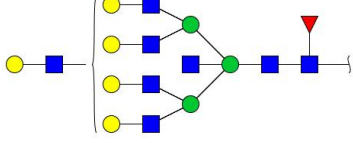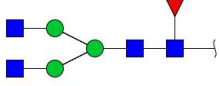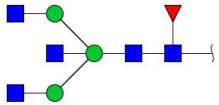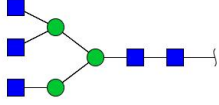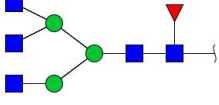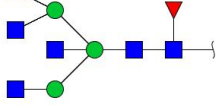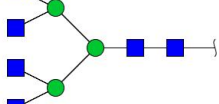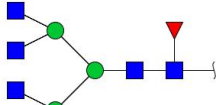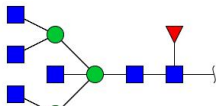
Table 11. *N*-glycans identified in MSLN from Ovcar-8 conditioned media for each exoglycosidase digestion. Structures ordered by retention time (GU), representation following SNFG guidelines [149] and abundance expressed as the percentage among all assigned *N*-glycans in that profile.
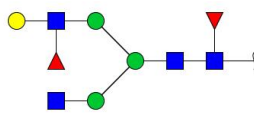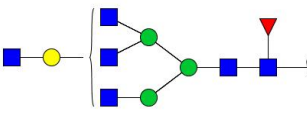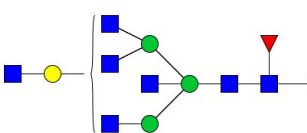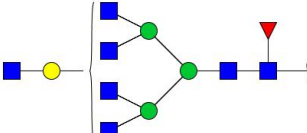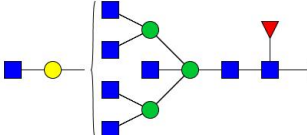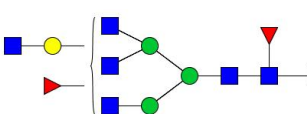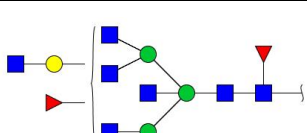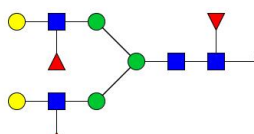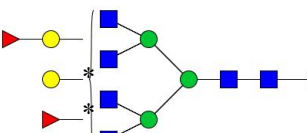
Retention time (GU)	Structure	Representation	Abundance (%)
		+NAN1 digestion profile	
7.55	FA2G2		2.12
7.59	FA2BG2		2.12
8.32	FA2F1G2	*Fuc linked to antenna GlcNAc	2.95
8.32	A3G3		3.50
8.72	FA3G3		4.41
8.72	FA2G2S1(6)		2.41
8.72	FA2BG2S1(6)		1.13
8.92	FA3BG3		2.94

9.00	FA2F2G2		2.89
9.53	A4G4		5.55
10.11	FA4G4		18.73
10.11	FA4BG4		8.67
10.67	FA3F1G3Lac1	*Fuc linked to antenna GlcNAc	5.38
10.83	FA3BF1G3Lac1	*Fuc linked to antenna GlcNAc	2.06
10.83	A4F2G4	*α2 Fuc linked to Gal *Fuc linked to antenna GlcNAc	6.01
11.16	FA3G3Lac1S1(6)		8.05
11.16	FA3BG3Lac1S1(6)		2.94
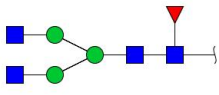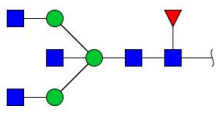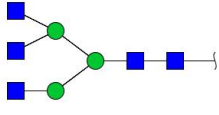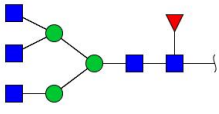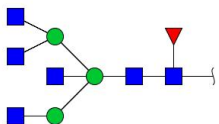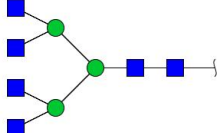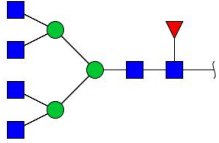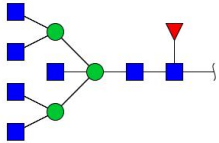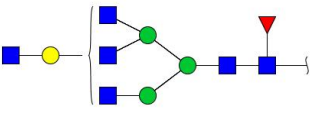
11.16	FA4G4S1(6)		7.35
11.37	FA4G4Lac1		5.16
11.45	FA4BG4Lac1		5.64

+ABS digestion profile

7.55	FA2G2		4.49
7.59	FA2BG2		3.17
8.32	FA2F1G2		3.20
8.32	A3G3		3.50
8.72	FA3G3		4.44
8.92	FA3BG3		2.73
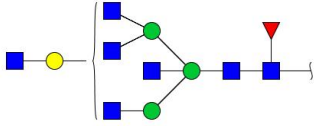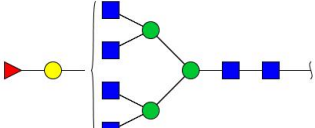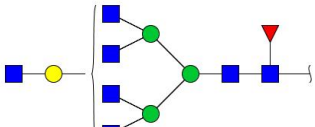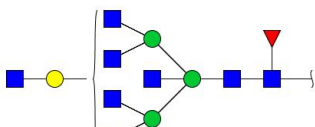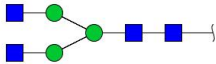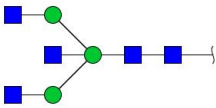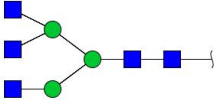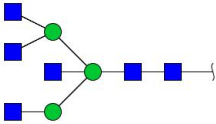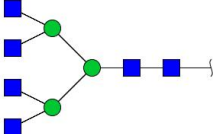
9.00	FA2F2G2		2.71
9.53	A4G4		3.69
10.11	FA4G4		23.12
10.11	FA4BG4		8.03
10.23	FA3G3Lac1		8.19
10.23	FA3BG3Lac1		2.60
10.67	FA3F1G3Lac1	 *Fuc linked to antenna GlcNAc	5.28
10.83	FA3BF1G3Lac1	 *Fuc linked to antenna GlcNAc	2.01
10.83	A4F2G4	 *α2 Fuc linked to Gal *Fuc linked to antenna GlcNAc	7.76

11.37	FA4G4Lac1		5.71
11.45	FA4BG4Lac1		9.37

+ABS+BTG digestion profile

5.82	FA2		4.40
5.82	FA2B		3.16
5.82	A3		1.76
6.25	FA3		4.98
6.38	FA3B		2.45
6.54	A4		4.03
6.93	FA4		23.30
6.98	FA4B		8.88

7.44	FA2F1G1		3.23
7.65	FA3G1GlcNAc1		8.87
7.68	FA3BG1GlcNAc1		4.54
8.08	FA4G1GlcNAc1		5.14
8.17	FA4BG1GlcNAc1		8.75
8.43	FA3F1G1GlcNAc1		4.86
8.49	FA3BF1G1GlcNAc1		2.00
9.00	FA2F2G2		3.10
9.08	A4F2G2		6.55

*Gal and *Fuc linked to same GlcNAc

+ABS+BTG+AMF digestion profile			
5.82	FA2		9.75
5.82	FA2B		3.24
5.82	A3		1.60
6.25	FA3		4.60
6.38	FA3B		2.52
6.54	A4		3.23
6.93	FA4		23.00
6.98	FA4B		10.84
7.65	FA3G1GlcNAc1		11.08

7.68	FA3BG1GlcNAc1		6.64
7.68	A4F1G1		7.77
8.08	FA4G1GlcNAc1		5.62
8.17	FA4BG1GlcNAc1		10.11

+ABS+BTG+BKF digestion profile

5.36	A2		10.13
5.48	A2B		3.37
5.82	A3		5.68
5.86	A3B		2.61
6.54	A4		33.86
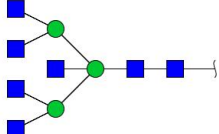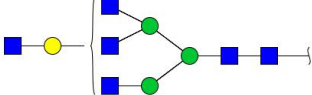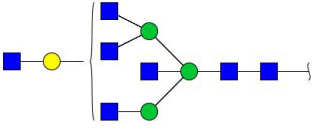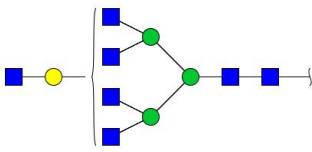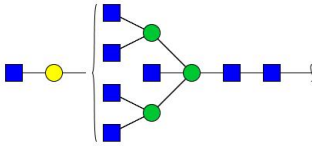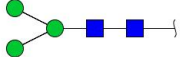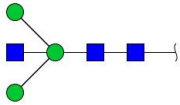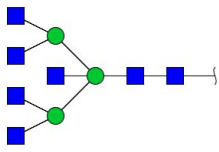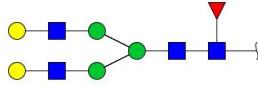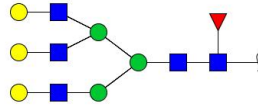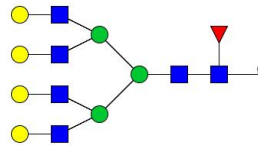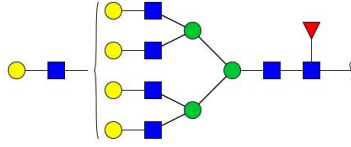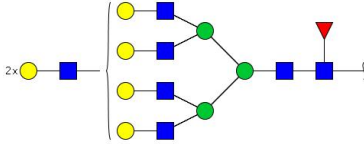
6.62	A4B		10.91
7.29	A3G1GlcNAc1		11.16
7.41	A3BG1GlcNAc1		6.83
7.73	A4G1GlcNAc1		5.48
7.84	A4BG1GlcNAc1		9.97
	+ABS+BTG+BKF+GUH digestion profile		
4.25	M3		67.03
4.65	M3B		12.46
6.59	A4B		20.51

Table 12. *N*-glycans identified in MSLN from PaC cell lines conditioned media after ABS digestion. Structures ordered by retention time (GU), representation following SNFG guidelines [149] and abundance expressed as the percentage among all assigned *N*-glycans.

Retention time (GU)	Structure	Representation	Abundance (%)	
			Capan-2	AsPC-1
7.55	FA2G2		10.49	6.74
8.72	FA3G3		21.80	21.12
10.11	FA4G4		25.66	26.69
11.37	FA4G4Lac1		14.98	30.94
12.61	FA4G4Lac2		27.07	14.51

Duran Sidera, A.

CHAPTER 2. Analysis of Regenerating islet-derived protein 1 (REG1) altered glycosylation in pancreatic cancer

REG1 expression in pancreatic cells

REG1 is overexpressed in a large number of conditions, which reduces its specificity when proposed as a biomarker for PaC. In this regard, we have hypothesized that REG1 glycosylation could be more specific and could provide new biomarkers based on tumour specific REG1 glycoforms.

First, the expression of REG1A was assessed by WB in cell lysates from seven ductal PaC cell lines and a control pancreas lysate obtained from a healthy donor (Figure 30.a).
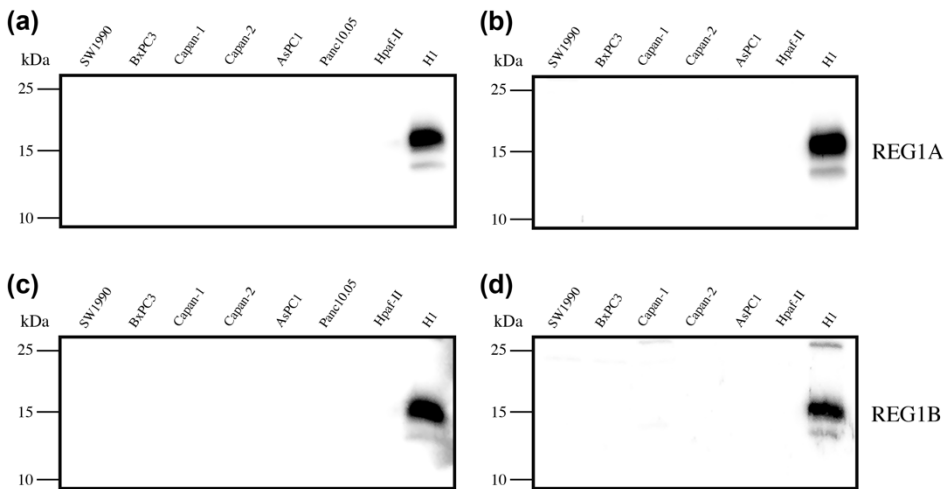
Figure 30. REG1 expression in PaC cell lines. WB analysis for REG1A and REG1B detection on PaC cell lines (20 µg of total protein). Analysis of (**a**) REG1A and (**c**) REG1B on PaC cell lines lysates and analysis of (**b**) REG1A and (**d**) REG1B on PaC cell lines conditioned media. All samples were loaded in Laemmli buffer under reducing conditions. H1: healthy pancreas tissue lysate.

None of the PaC cell lines expressed the glycoprotein while the pancreas lysate showed high REG1A expression. As REG1A is secreted, we also analysed its presence in the conditioned media of these PaC cells. Again, no bands were observed by WB (Figure 30.b), indicating that malignant ductal cells did not express it. The

same methodology was performed to analyse REG1B expression. Similar results were obtained, with no observed bands in protein lysates (Figure 30.c) or conditioned media (Figure 30.d), but there was a high expression in the healthy pancreas lysate. Despite no signal was observed in cell lines, REG1 expression was assessed in pancreatic tissue lysates (Figure 31).
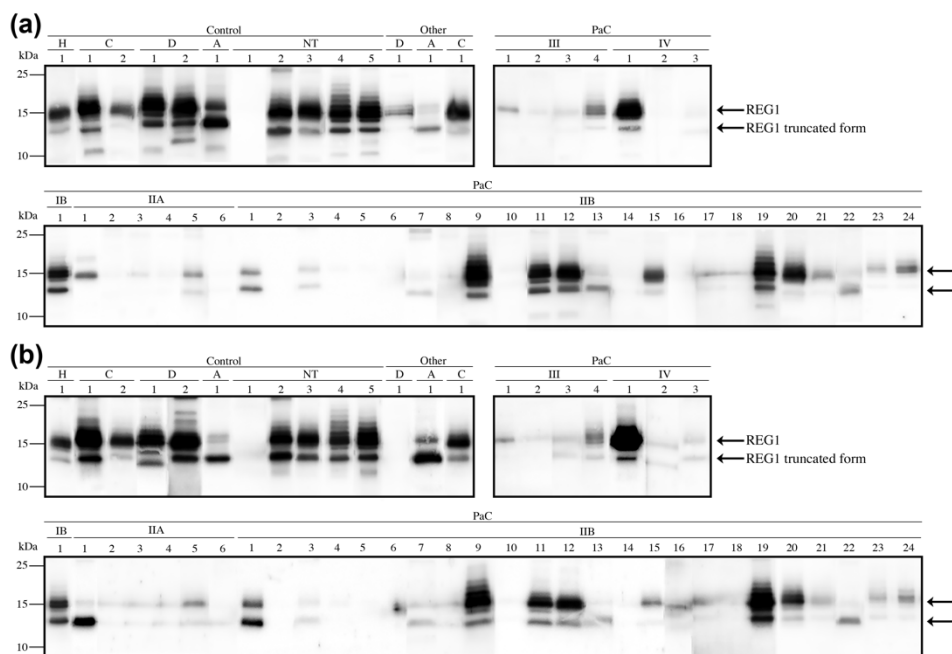
Figure 31. REG1 expression on healthy and cancerous pancreatic tissues. Detection of **(a)** REG1A and **(b)** REG1B by WB on 10 µg total protein. Samples loaded in Laemmli buffer under reducing conditions. Each number represents a different patient in each condition. Control samples included one healthy pancreas (H) and pancreatic non-tumour tissues adjacent to the cancer region from cholangiocarcinomas (C), from duodenum adenocarcinomas (D), from ampuloma (A) and from PaC (NT). Cancer tissues included forty-one PaC of different stage and other tumours: one cholangiocarcinoma (C), one duodenum adenocarcinoma (D) and one ampuloma (A). 5 µg of the healthy pancreas (H1) tissue lysate were loaded at all membranes as a control, which were equally exposed to chemiluminescence for lane comparison.

We observed REG1A expression in 60.5% (n=38) of PaC tissues (Figure 31.a). No differences were detected between different cancer stages, but there was a trend depending on the differentiation degree. While well differentiated tissues only presented 33.3% REG1A positivity, REG1A expression increased to 62.1% in moderately differentiated and to 80.0% in poorly differentiated samples. The full

length REG1A band, over 15 kDa, was predominant in most samples, being quite heterogenous in many of them. Three samples (PaC IIB 7, PaC IIB 13 and PaC IIB 22) only showed the truncated form (a band around 13-14 kDa). Regarding the healthy tissues, 90.9% (n=11) showed the presence of the glycoprotein. These results sustained that REG1A was produced and secreted by acinar cells, vastly represented in healthy tissues, while the predominance of ductal cells in the tumour tissue decreased REG1A presence in these samples.

Regarding REG1B, similar results were obtained (Figure 31.b). Different expression (presence/absence) was only observed in samples D1, PaC IIA 3, IIB 16, IV 2 and IV 3. This was translated into a slight increase of REG1B expression compared to REG1A in PaC samples (65.8% vs. 60.5%). However, overall expression trends did not vary.

The similar expression pattern obtained by WB for REG1A and REG1B moved us to analyse if there could be cross-reactivity between the used anti-REG1A and anti-REG1B antibodies. For that, we tested both antibodies with recombinant REG1A (rREG1A) and recombinant REG1B (rREG1B) (Figure 32). Both antibodies recognized both rREG1A and rREG1B protein bands, contrary to manufacturers' specifications.
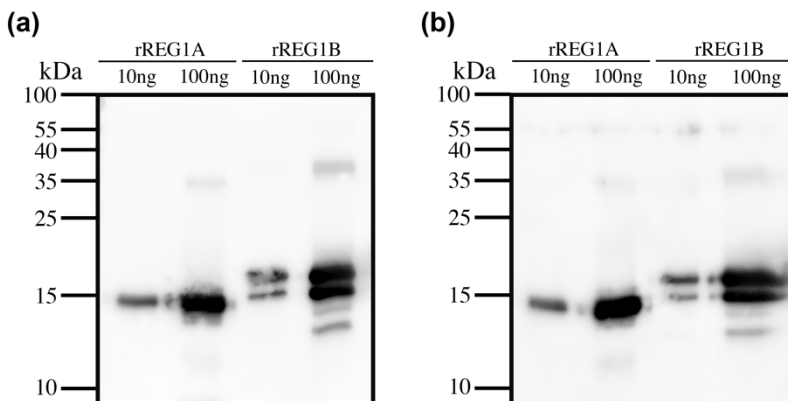
Figure 32. Analysis of REG1 antibodies cross-reactivity. Evaluation of REG1A (ab47099) and REG1B (ab233210) antibodies reactivity by WB. **(a)** WB with anti-REG1A antibody on rREG1A and rREG1B. **(b)** WB with anti-REG1B antibody on rREG1A and rREG1B.

Regardless of their similarity, rREG1A and rREG1B showed a different band pattern. A single band at 15 kDa was observed for rREG1A, which should correspond to residues 23-166 according to the product datasheet. On the other hand, rREG1B presented two bands. The lower one, slightly above rREG1A, could correspond to the protein with the His tag, while the upper one could be the protein maintaining its signal peptide (aa 1-22). However, the manufacturer indicated differential glycosylation as the source of this double band.

Considering these results, we could not assure if we were detecting REG1A, REG1B or both when using these antibodies. Therefore, from here on, we will refer to general REG1 expression without specifying whether it is REG1A or REG1B.

REG1 glycan analysis from pancreatic tissues

REG1 is a glycoprotein with sialylated glycans

Once stated the expression of REG1 in PaC tissues, we focused on analysing its glycosylation using exoglycosidases and lectins. First, we performed specific sialidase digestions using NAN1 (only digests terminal α2,3-SA) and ABS (digests all terminal SA) to check the presence of terminal SA on the *O*-glycan chain. For that, we determined whether REG1 electrophoretic mobility was increased after sialidase digestions, since the removal of the negatively charged terminal SA contributes to increased mobility in SDS-PAGE (Figure 33).

The removal of SA caused a decrease in REG1 molecular weight, confirming the presence of sialylated glycoforms. The selective digestion of α2,3-SA showed a slight increase in the mobility of upper REG1 band compared to the undigested samples. Total SA digestion (ABS) did not yield significant change in band mobility compared to NAN1 digestion, which indicated that most (if not all) SA was α2,3-linked. Interestingly, SA digestion was unable to reduce the glycoform heterogeneity observed in the samples, thus suggesting that complex structures differing in monosaccharides other than SA were present in REG1. It is also important to note that the low REG1 band at 14 kDa did not vary after SA removal, confirming

that the *O*-glycan was found at the N-terminal peptide (Thr-5) (which is absent in this truncated form).
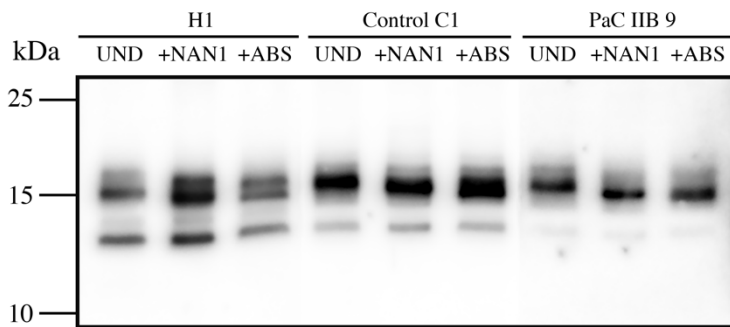
Figure 33. SA analysis of REG1 from pancreatic tissue lysates. WB using anti-REG1A antibody on pancreatic tissue lysates (3 µg total protein) after digestion of SA with specific sialidases. UND: undigested, +NAN1: α2,3-SA digestion, +ABS: all SA digestion. Samples were loaded in Laemmli buffer under reducing conditions. Samples numeration corresponds to samples in Figure 31; H1: healthy pancreas; Control C1: non-tumour tissue adjacent to a cholangiocarcinoma; PaC IIB 9: PaC IIB tumour.

REG1 glycan determinants analysis on pancreatic tissues

To obtain more information on REG1 glycosylation, we next analysed its glycan determinants through WB using specific lectins. Thus, we first set-up an immu-noaffinity method to purify REG1. For that, we compared three protocols using anti-REG1A or anti-REG1B antibodies (Abcam, Cambridge, UK), either coupled (1) to protein A-sepharose beads, (2) to protein A-magnetic beads, or (3) covalently bounded to magnetic beads. REG1-reactive tissue lysates were incubated with the antibody-beads complexes, and the glycoprotein was then eluted in Laemmli buffer under reducing conditions. Obtained fractions were evaluated by WB and silver staining to assess the protein recovery and purity for each immunoaffinity method (Figure 34). WB results of immunoaffinity purification using anti-REG1A antibodies showed similar recoveries for the three approaches, with a positive band around 15 kDa corresponding to REG1 (plus two bands at 55 and 25 kDa, from the heavy and light chains of the capture antibody, respectively). However, when using the antibody covalently bounded to magnetic beads, some REG1 was still found in the unbound fraction and this protocol was thus discarded. Silver

staining showed a distinct band at 15 kDa corresponding to REG1 as well as the bands of the capture antibody. Although the use of protein A-sepharose beads and protein A magnetic beads yielded quite similar results, the later was more straightforward to perform and was selected for REG1 immunopurification from tissues.
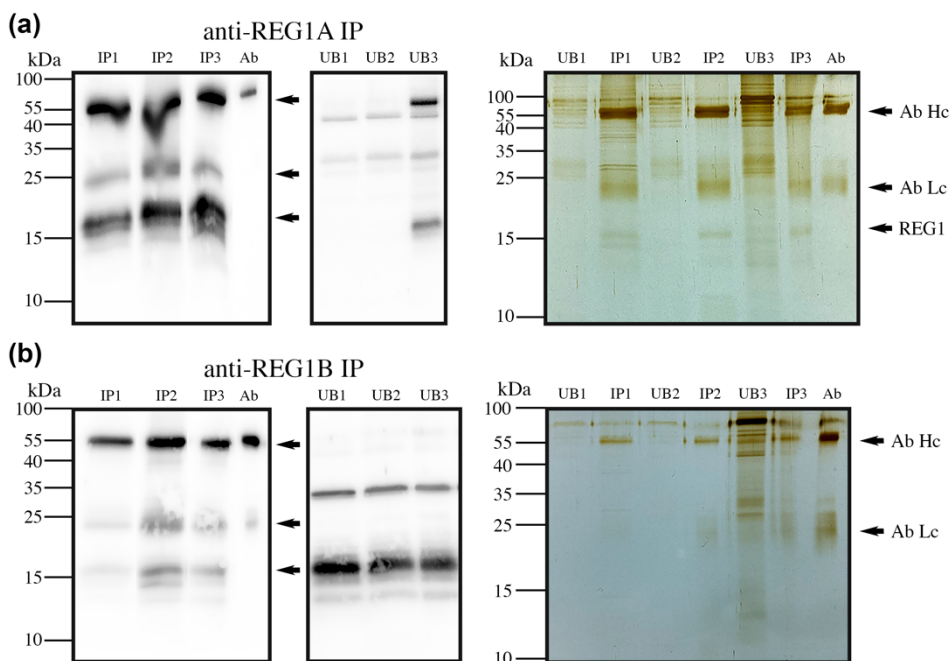
Figure 34. REG1 immunopurification from pancreatic tissues. REG1 was immunopurified from 100 µg total protein of the H1 tissue lysate (used as control along the study) following different methodologies. The obtained fractions for immunopurification with **(a)** anti-REG1A antibody or **(b)** anti-REG1B antibody were analysed by WB with anti-REG1A and anti-REG1B antibodies respectively (left panels), and by silver staining (right panels). IP1: immunopurification with protein A-sepharose beads, IP2: immunopurification with protein A-magnetic beads, IP3: immunopurification with antibodies covalently coupled to magnetic beads, UB1/UB2/UB3: corresponding unbound fractions, Ab: capture antibody (anti-REG1A *ab47099*; anti-REG1B *ab233210*), Ab Hc: capture antibody heavy chain, Ab Lc: capture antibody light chain.

On the other hand, immunopurification using the anti-REG1B antibody only yielded low amounts of REG1 in the immunopurified fractions while the unbound fractions still showed substantial REG1 protein amounts by WB (Figure 34.b). These results indicated that the antibody was not able to capture all the protein

present in the lysates. In addition, silver staining was also unable to detect the purified protein.

Next, using the anti-REG1A antibody, REG1 was immunopurified from seven pancreatic tissues which were selected based on their high expression levels of REG1 (Figure 31). These included one non-tumour pancreatic tissue (NT 3), one healthy tissue adjacent to a cholangiocarcinoma (C1) and five pancreatic tumours of different stage (one IB, three IIB (patients 9, 11 and 12) and one IV (patient 1)). Immunopurified REG1 was blotted with several lectins to determine its specific glycan determinants. Those lectins were selected based on previous results that showed the presence of SA on REG1 (Figure 33) and previous studies underlying REG1A glycosylation [374,375]. They were MAL-II which recognises α2,3-SA, SNA for α2,6-SA, VVL for αGalNAc (Tn antigen), PNA that binds terminal Galß1,3GalNAc and UEA for α1,2-fucose. Then, membranes underwent a stripping protocol and were reblotted for REG1, thus confirming their colocalization and protein load.

Regarding sialylated *O*-glycans, we detected α2,3-SA in all samples (Figure 35.a), corroborating the results obtained by WB after sialidases digestions (Figure 33). Different expression levels were observed among samples, which did not correlate with the presence of tumour or PaC staging. On the other hand, no α2,6-SA expression was detected in healthy tissues, while three out of five PaC tissues contained this structure on REG1 (Figure 35.b). Terminal GalNAc was only observed in some PaC samples, although its expression was very weak (Figure 35.c). Regarding terminal Gal, an opposite trend was detected, being it exclusively present in non-PaC tissues (Figure 35.d). Finally, *O*-glycans with terminal α1,2-fucose detected with UEA were observed in all samples except one tumour (PaC IIB 11) (Figure 35.e), being the amount of this residue much higher in PaC IV 1 over the rest of samples. In summary, α2,6-sialylated glycoforms, as well as those presenting terminal Gal or GalNAc, stood out as the best options to discriminate PaC from control specimens. Interestingly, the heterogeneous pattern of REG1 upper band

at 15 kDa when detecting REG1 was much more simplified when recognised with lectins (especially remarkable in sample PaC IIB 9).
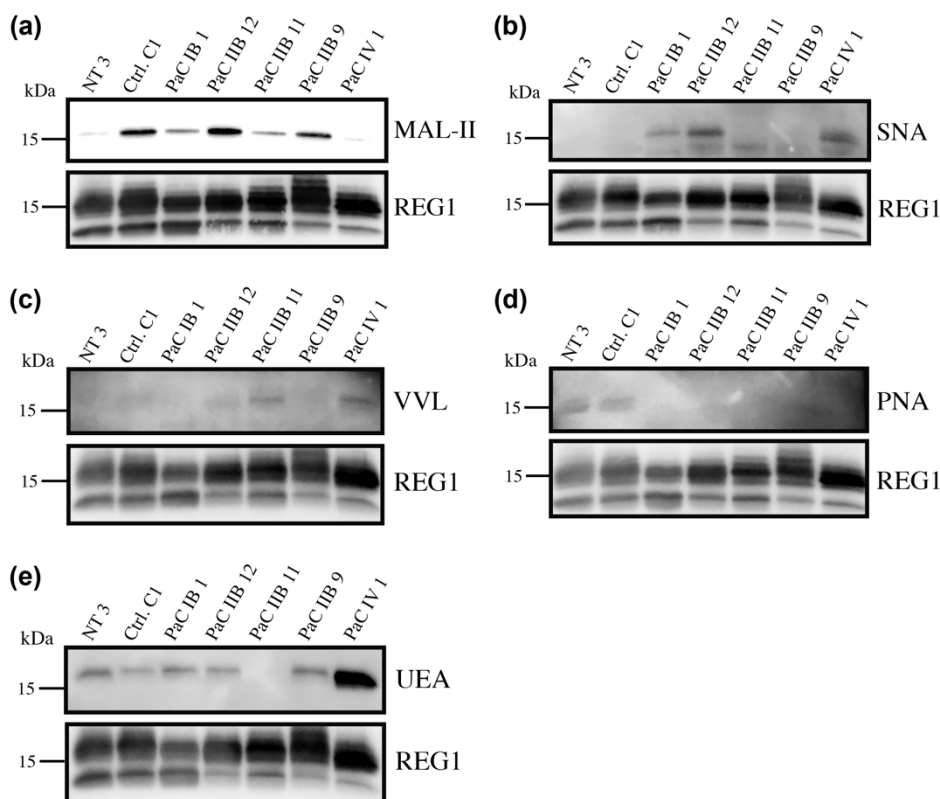
Figure 35. Analysis of glycan determinants by WB with lectins on REG1 immunopurified from tissue samples. For each lectin, top panel shows the specific glycan determinant recognition, while bottom panel shows the result after membrane stripping with anti-REG1A antibodies. (**a**) MAL-II for α2,3-SA. (**b**) SNA for α2,6-SA. (**c**) VVL for αGalNAc. (**d**) PNA for Galß1,3GalNAc. (**e**) UEA for α1,2-fucose. Samples numeration correspond to samples in Figure 31; NT3: non-tumour tissue adjacent to PaC; Control C1: non-tumour tissue adjacent to a cholangiocarcinoma; PaC: different stage PaC tissues.

Evaluation of blood serum REG1 glycoforms

Purification of REG1 from blood serum samples

REG1 is a secreted glycoprotein. Thus, we next focused on developing a methodology to properly measure the protein and its specific glycoforms levels in blood serum samples, which reduces the invasiveness of sample collection and gets closer

to the clinical translation. As WB is just a semiquantitative methodology that requires relatively high amounts of protein, we chose to develop an ELISA-type immunoassay using the previously assayed lectins that better discriminated between PaC and non-tumour pancreatic tissues.
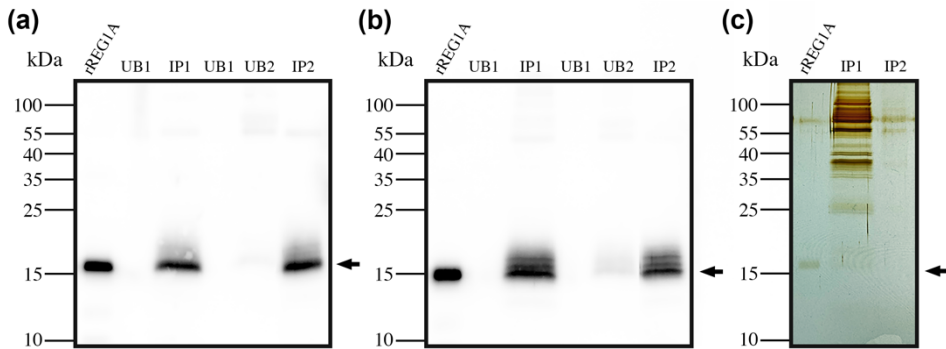
Figure 36. REG1 immunopurification from serum samples. Evaluation by WB of the obtained fractions after REG1 immunopurification from (**a**) a control serum spiked with 50 ng of rREG1A and (**b**) a serum form a PaC patient (stage IIB). (**c**) Silver staining analysis of the obtained fractions for REG1 purification from the PaC serum. rREG1A: 25 ng recombinant REG1A, UB1: unbound fraction, IP1: immunopurified fraction, UB2: unbound fraction after the second incubation, IP2: immunopurified fraction after two consecutive incubations. All samples were loaded in Laemmli buffer under reducing conditions. Arrows indicate REG1 bands.

Due to the complexity and high glycoprotein concentration found in blood serum samples, a pre-purification step was mandatory to avoid the interference of other glycoproteins, especially when using lectins to quantify glycans. For that, the immunopurification of REG1 was performed using the anti-REG1A antibodies covalently bound to magnetic beads to avoid the elution of antibodies that could interfere in the subsequent REG1 glycan detection using lectins. A double incubation with the antibody-beads conjugate was needed to properly deplete major serum proteins and to obtain a fraction containing highly enriched REG1. WB of the obtained fractions after one (IP1) or two (IP2) incubation steps showed high REG1 recovery, while very few REG1 amounts were lost in the unbound fraction (Figure 36). Regarding the specificity of the purification protocol, silver staining of the purified fractions still revealed the presence of several proteins after the first

immunopurification step (IP1), which were mostly removed after the second incubation (IP2). Altogether, these results endorse the use of this methodology as a previous step to determine serum REG1 glycans. Thus, REG1 was immunopurified using this methodology on 41 serum samples from healthy individuals, ChP and PaC patients for its subsequent glycan quantification.

Development of sandwich ELISA/ELLA assays for REG1 and its glycoforms quantification

Measurement of REG1 concentration was performed after the establishment of an appropriate ELISA protocol using two commercial antibodies against the protein: a rabbit polyclonal anti-REG1A antibody for the capture and a mouse monoclonal anti-REG1A antibody as detection. An 8-point calibration curve ranging from 0-40 ng/mL rREG1A was used, with linearity between 2.5-20 ng/mL. The LOQ was 0.52 ng/mL, while intra-assay variation was of <6%. This methodology was used to quantify REG1 levels in blood serum samples in order to determine the amount loaded for the detection with lectins, and to normalise the results obtained by ELLA.

We next developed several sandwich ELLAs using a capture antibody against REG1A and a specific labelled lectin to determine the glycan determinants on REG1. As mentioned before, the determination of α2,6-sialylated REG1 glycoforms stood as the best choice to discriminate PaC patients, as well as the determination of terminal Gal or GalNAc residues. Thus, we selected the lectins SNA, PNA and VVL, respectively, which recognize these glycan moieties, to be used as detection probes in the ELLA assays. To explore these lectins applicability in the ELLA assay we first determined by WB whether the potential antibodies used to capture REG1A (rabbit polyclonal (*ab47099, Abcam*) or mouse monoclonal (*MA5-29515, Thermo Fisher Scientific*)) expressed the glycan moiety recognized by the lectin. In that case, this lectin could not be used due to its binding to the capture antibodies, what would cause a background noise. In addition, we determined the binding of those lectins to the rREG1A that would be used for the assay calibration curve (Figure 37). They should react with rREG1A to be used in the ELLA.

As shown in Figure 37.a, SNA was not suitable for the ELLA because it was reactive with both REG1A antibodies and it could not detect rREG1A. Despite presumably presenting a lower discriminative potential, the use of MAL-II to detect α2,3-SA was also investigated. The analysis by WB showed that this lectin could detect rREG1A and was not recognising the antibodies (Figure 37.b). However, its use in an ELLA assay resulted in a calibration curve with elevated background signalling, reason why it was also discarded.
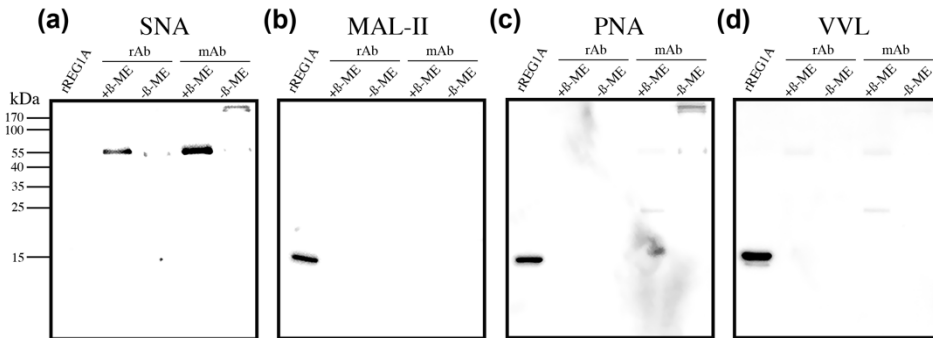
Figure 37. Glycosylation patterns of rREG1A and anti-REG1A antibodies. Analysis by WB of the expression of several glycan determinants on rREG1A (500 ng) and two anti-REG1A antibodies (300 ng) that could be used as capture antibodies in an ELLA assay under reducing (+ß-ME) or non-reducing (-ß-ME) conditions. **(a)** SNA for α2,6-SA. **(b)** MAL-II for α2,3-SA. **(c)** PNA for Galß3GalNAc. **(d)** VVL for αGalNAc. rAb: rabbit anti-REG1A antibody (ab47099, Abcam), mAb: mouse anti-REG1A antibody (MA5-29515, Thermo).

PNA, which detects terminal Gal (T antigen), presented high reactivity with rREG1A and was not reactive with the rabbit capture antibody (Figure 37.c). Therefore, a sandwich ELLA was developed using a biotinylated PNA lectin. However, no signal was observed either using normal plates or protein G-coated ones (designed to bind the Fc region of the antibody ensuring maximum exposure of the antigen binding site). An ELLA with biotinylated PNA but without capture antibody could neither detect rREG1A. To improve the sensitivity of the assay, several streptavidin-peroxidase conjugates were tested, but no reactivity with rREG1A was detected. A distinct PNA, labelled with digoxigenin, was reactive with rREG1A in ELLA, although the sensitivity of the assay was poor due to the low slope of the calibration curve. To increase it, we got back to the sandwich

ELLA method and we changed the peroxidase substrate from a colorimetric to a chemiluminescent one. Under the assayed conditions, this change resulted in high signalling but also increased background signal, leaving a too short luminescence units' range for sample testing. Dilution of capture antibody, lectin or peroxidase anti-digoxigenin secondary antibody did not show any improvement, always producing the loss of the standard rREG1A signalling but not blank noise reduction. The use of different development substrates neither improved the results. A final attempt using fluorescein isothiocyanate (FITC)-labelled PNA was performed, as other lectins with this labelling had worked properly in our lab. However, this change did not alter previously observed outcomes.

Since the determination of REG1 terminal Gal with PNA was not successful, we finally tested the use of VVL to detect terminal GalNAc residues, which is reactive with truncated *O*-glycans such as the Tn antigen. WB analysis of rREG1A and anti-REG1A antibodies showed that the standard could be detected with VVL, while faint bands were detected on the tested antibodies (Figure 37.d). As for PNA, an ELLA to detect rREG1A with FITC-labelled VVL showed poor sensitivity. Progression to a sandwich ELLA, on this occasion, did show more promising results when using the rabbit antibody (less reactive to VVL) and a chemiluminescent substrate. A calibration curve from 0-100 ng/mL rREG1A was established, with a lineal range from 20 to 100 ng/mL. The LOQ was 9.87 ng/mL, with intra-assay variations of <7%. However, serum samples could not be quantified. Values similar to the blank were obtained in all cases (even loading a sample at 200 ng/mL of immunopurified REG1), suggesting that human REG1 glycoforms with terminal GalNAc were not predominant and that the developed methodology was not sensitive enough.

As REG1 glycoforms quantification was not possible with an ELLA approximation, we explored the feasibility of performing a semi-quantitative detection through WB with lectins on the immunopurified serum samples. However, quantification of REG1 protein levels from the tissue samples used for the glycan determinants characterization above (Figure 35) showed that immunopurified

REG1 amounts from the serum were not enough to perform WB analyses. While we obtained between 15 and 75 ng of REG1 from 1 ml blood serum, between 70 and over 200 ng were required for the lectin recognition by WB.

Whatever the case, the use of lectins presented serious limitations for the detection of REG1 glycoforms. A different approach was considered in order to evaluate the glycan moieties carried on this interesting glycoprotein, as distinct determinants had been observed in pancreatic tissues. In this regard, it must be noted that a collection of about forty immunopurified REG1 samples was quantified and is ready-to-use as soon as a suitable methodology allows its glycoforms analysis.

REG1A glycoforms analysis by MS

Determination of rREG1A glycoforms by MALDI-Tof

Due to the limitations to develop a proper immunoassay to quantify REG1 glycoforms, a different approach to detect them was considered by using mass spectrometry-based methodologies. Therefore, we could get back to the analysis of a specific REG1 protein, REG1A in this case, as the polypeptide sequence variations could be detected in the mass spectrometer.

First, we focused on detecting the different glycoforms by MALDI-Tof. The direct analysis of rREG1A allowed the identification of 14 peaks, indicating that several glycoforms could be detected (Figure 38). In addition, selective digestion of REG1A sialylation with NAN1 and ABS before its analysis yielded a new peak (peak 8) and changes in the proportion of most peaks, which helped in the structures' assignments and also in the identification of the SA linkages.

According to the mass shift respect to the non-glycosylated protein (16.247 kDa, Peak 1), the glycan moiety of these diverse glycoforms could be assigned. Although the MS data obtained did not allow the exact identification of glycan isomers that could differ in their type of monosaccharide (as for instance N-acetylglucosamine and N-acetylgalactosamine have the same molecular mass), their position or linkage, various structures were proposed (Table 13) based on previous rREG1A glycoprotemic analyses [374,375] and on common O-glycan structures. On one side,

we found the truncated Tn-antigen (Peak 2) and core 1 glycoforms corresponding to the T-antigen (Peak 3), sialyl core 1 (Peak 7), di-sialyl T-antigen (Peak 12) and a minor elongated core 1 (Peak 10), being the sialylated forms the most abundant. On the other hand, glycans corresponding to core 2 (Peaks 13 and 17) and core 4 (Peaks 9, 11, 14, 15 and 16) structures, with prevalence of the latter, were found.
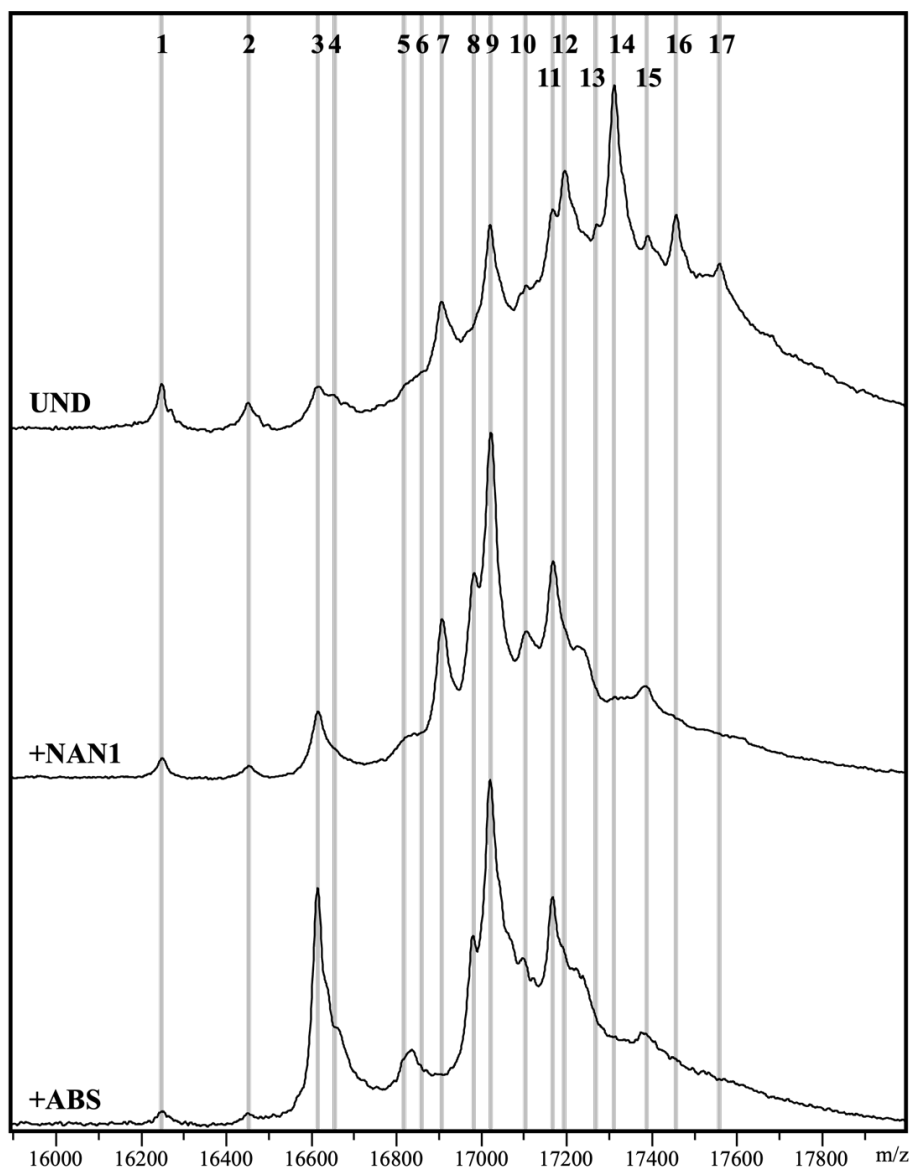
Figure 38. Analysis of rREG1A heterogeneity by MS. Aligned MALDI-Tof mass spectras of undigested rREG1A (UND) and after sialidase digestion with NAN1 (α2,3-SA) and ABS (total SA). Assigned peaks mass and proposed structures can be found at Table 13.

Table 13. rREG1A glycoforms detected by mass spectrometry. Monosaccharide composition of the assigned rREG1A glycoforms detected by MALDI-ToF based on the experimental mass. Glycans represented following SNFG guidelines. NAcHex: *N*-acetylhexosamine, Hex: hexosamine, Neu5Ac: *N*-acetylneu-raminic acid, Fuc: fucose, n.d.: not detected.

Peak n°	Glycoprotein experimental mass (Da)	Glycoprotein theoretical mass (Da)	Glycopeptide experimental mass (Da)	Glycopeptide theoretical mass (Da)	Assigned structure	Proposed structure
1	16247.094	16247.653	1253.346	1252.605	Non-glycosylated peptide	ZEAQTELPQAR
2	16451.450	16450.732	1456.020	1455.684	NAcHex	
3	16615.012	16612.785	1618.217	1617.737	NAcHex + Hex	
4	16651.967	16653.812	1659.210	1658.764	2 NAcHex	
5	n.d.	16815.864	1821.305	1820.816	2 NAcHex + Hex	
6	n.d.	16856.891	1862.220	1861.843	3 NAcHex	
7	16905.531	16903.880	1909.117	1908.832	NAcHex + Hex + Neu5Ac	
8	16979.305 *	16977.917	1983.632	1982.869	2 NAcHex + 2 Hex	
9	17019.286	17018.944	2024.599	2023.896	3 NAcHex + Hex	

No.	Structure	Composition			
10		2 NAcHex + Hex + Neu5Ac	2111.912	n.d.	17106.960	17103.933
11		3 NAcHex + Hex + Fuc	2169.954	2170.812	17165.002	17166.296
12		NAcHex + Hex + 2 Neu5Ac	2199.928	2199.046	17194.976	17194.844
13		2 NAcHex+ 2 Hex + Neu5Ac	2273.965	2274.378	17269.013	17268.579
14		3 NAcHex + Hex + Neu5Ac	2314.991	2315.516	17310.039	17310.302
15		4 NAcHex + 2 Hex	2389.028	n.d.	17384.076	17389.470
16		3 NAcHex + Hex + Neu5Ac + Fuc	2461.049	2461.242	17456.097	17456.013
17		2 NAcHex + 2 Hex + 2 Neu5Ac	2565.060	2565.926	17560.108	17558.598

* Structure only detected after sialidase digestion.

Next, in order to improve the resolution of rREG1A glycoforms to confirm those previous assignments, we performed the analysis of the glycopeptides obtained by trypsin digestion-sensitive cleavage site at the aminoacidic residue R11 of the protein. A short 15-minute digestion with trypsin showed the complete release of the *N*-terminal rREG1A glycopeptide, while the rest of the protein remained undigested and detected in a main peak at 15015.79 Da (theoretical mass 15013.05 Da) (Figure 39.a). The presence of this peak corroborated that the previously observed rREG1A heterogeneity mainly corresponded to different glycoforms at Thr-5, now contained in the digested glycopeptide. MALDI-Tof-MS analysis of the released rREG1A glycopeptide displayed several peaks, which corresponded to the previous detected glycoforms on the whole glycoprotein (Figure 39.b), with similar intensity ratios. Just two peaks were not found (Peaks 10 and 15), while three new ones were detected (Peaks 5, 6 and 8) (Table 13).
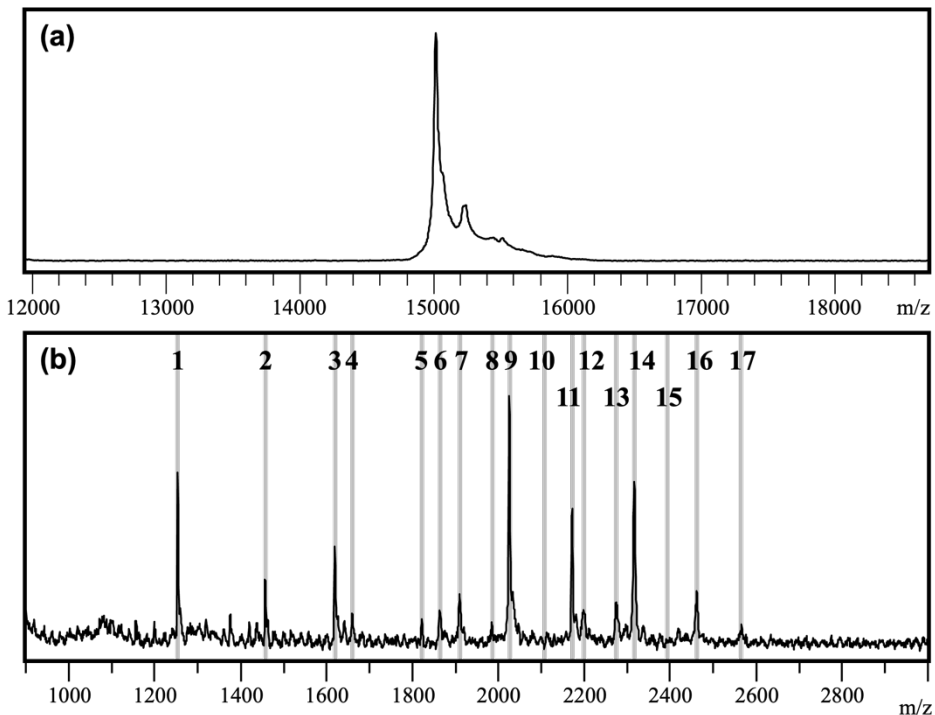
Figure 39. Analysis of trypsin-digested rREG1A. (a) Determination of rREG1A by MALDI-Tof after digestion of its *N*-terminal glycopeptide with trypsin. **(b)** Detection of rREG1A glycopeptides obtained after trypsin digestion and HILIC purification. Assigned peaks correspond to the same glycoforms observed in Figure 38 and detailed in Table 13.

Analysis of REG1A glycopeptides from biological samples by LC-MS

Glycopeptide analysis by MALDI-Tof was a good approach to detect different glycoforms, which served for the assignment of many glycan structures on rREG1A. However, it presented some limitations, being the most critical its application for the study of complex biological samples such as serum because it is necessary to have the target protein purified in order to avoid signals from other proteins. To overcome this limitation, we next developed a methodology to detect such rREG1A glycoforms in complex biological samples (blood serum) by combining glycopeptides separation in a UPLC system (equipped with an HILIC column) with their detection in a triple quadrupole after electrospray ionisation (UPLC-ESI-TQ), a methodology that should confer enhanced sensitivity and reduced background.

First, we checked whether the rREG1A glycoforms on the *N*-terminal glycopeptide identified in the MALDI-Tof instrument could also be detected by this UPLC-ESI-TQ methodology. In a sample of trypsin-digested pure rREG1A, we could identify all the main glycoforms (Peaks 1, 2, 7, 9, 11, 12, 14 and 16) except two of them (Peaks 3 and 5) using a selected reaction monitoring (SRM) data acquisition mode (Figure 40). Some of these glycoforms masses contained more than one isomer, as shown by two or more different retention time peaks in the chromatogram. This is the case of peak 2, which contained two isomers with the most abundant probably corresponding to the Tn antigen. Peak 7 also contained two isomers, the most abundant assigned to sialyl core 1 and the other to sialyl T-antigen, as α2,3-SA presents a lower retention time than α2,6-SA. Furthermore, these results agreed with those observed by MALDI-Tof, in which the sialyl core 1 (which was majoritarian and presented α2,6-SA) was identified. Peak 9 contained two isomers with diverse NAcHex and Hex arrangements. Peak 14 also contained 2 isomers. Considering their separation and relative abundance, the addition of a SA did not add any heterogeneity, indicating that it was always included in the same linkage (α2,3-SA according to MALDI-Tof).
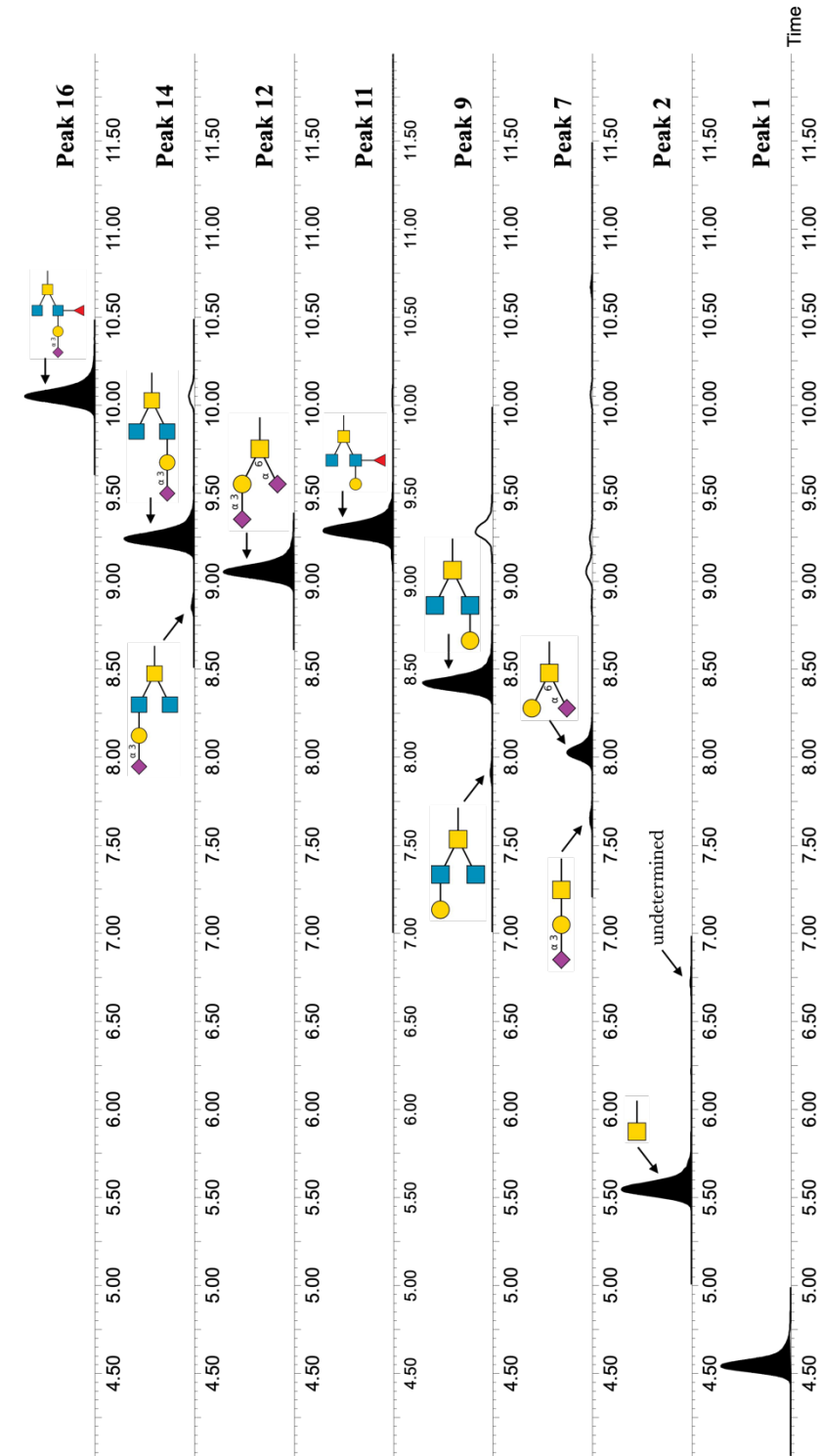
Figure 40. rREG1A glycopeptides analysis by LC-MS. Obtained chromatograms for the transitions detected in rREG1A by SRM in the UPLC-ESI-TQ. It must be noted that some peaks really corresponded to the selected mass (in black), while others came from fragmentation of larger structures in the ionisation source (in white).

Interestingly, some transitions displayed more peaks than the expected possible isoforms for a selected mass. This phenomenon was caused by the fragmentation of larger structures, that were detected as smaller ones after being partially reduced in the ionisation source. These were identified through the observation of the larger glycoform in the same retention time for a different transition and were therefore not considered (peaks in white in Figure 40). It is also important to note that all analysed glycopeptides could generate a 1253.3 m/z fragment, corresponding to the non-glycosylated peptide, thus we could confirm that they were specific rREG1A glycopeptides. Finally, optimization of the solvation temperature, HILIC chromatographic gradient and mobile phase led to the detection of all these rREG1A glycoforms from just 18 ng rREG1A.

Next, we focused on the analysis of the glycoprotein in a complex matrix. The first step was the obtention of the REG1A glycopeptide from serum, which was achieved through REG1 immunopurification using magnetic beads and subsequent trypsin digestion. We evaluated two different protocols to isolate the glycopeptides: trypsin digestion on the immunopurified REG1, previously eluted from the magnetic beads using acidic citrate buffer (pH 3.1), neutralised with Tris 1M pH 9 and desalted with centrifugal filters; or trypsin digestion on the magnetic beads bound REG1. This way, we could recover the generated glycopeptide in solution without additional steps. When comparing both methodologies, the former yielded a poor glycopeptide recovery, which was 4-fold lower than the latter (Table 14). Then, we tested three different trypsin:substrate ratios, starting from 1:10 (used previously to generate the REG1 glycopeptide for MALDI-Tof analyses) and increasing the ratio to 1:2 and 1:1 to ensure enough enzyme activity when working with 100 ng of rREG1A spiked in 1 mL of pooled serum (as a matrix). The enzyme:substrate ratio 1:2 increased approximately 2-fold the MS signal of the different glycosylated peptides compared to 1:10 ratio, while the ratio 1:1 did not show any significant improvement compared with 1:2. When analysing 100 ng rREG1A in 1 mL serum following the previously described methodology, we were able to detect all glycopeptides, but only the isomer 2 of the glycopeptide contained in peak 7 (Table 14).

Table 14. Intensity of the detected glycopeptides by LC-MS. Detailed quantification of the glycopeptides detected by UPLC-ESI-TQ with the SRM data acquisition for various experiments. For each transition, the retention time and predicted *O*-glycan structure of detected isomers is indicated. n.d.: not determined; -: not included in the SRM method.

Peak n°	Retention time (min)	Proposed structure	rREG1A	1 µg rREG1A IP citrate elution	1 µg rREG1A IP trypsin elution	100 ng rREG1A IP trypsin 1:10	100 ng rREG1A IP trypsin 1:2	50 ng rREG1A IP trypsin 1:2	50 ng rREG1A IP lyophilisation	50 ng rREG1A IP speedvac	2 mL serum control	2 mL serum PaC IIB
1	4.54	non-glycosylated	68160	244	763	131	214	17	27	36	62	624
2 .1	5.55	(structure)	153615	261	1176	388	563	7	20	41	18	72
2 .2	6.71	undetermined	2836	n.d.	n.d.	n.d.	11	n.d.	n.d.	n.d.	n.d.	n.d.
STn	7.04	(structure)	n.d.	-	-	-	-	n.d.	n.d.	n.d.	n.d.	1091
7 .1	7.64	(structure)	15693	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	716
7 .2	8.03	(structure)	161912	296	938	n.d.	575	n.d.	n.d.	n.d.	n.d.	1654
9* .1	7.91	(structure)	8410	n.d.	42	n.d.	55	n.d.	n.d.	n.d.	n.d.	n.d.
9* .2	8.42	(structure)	292763	896	3506	859	1565	n.d.	n.d.	n.d.	n.d.	n.d.
11	9.28	(structure)	193132	-	-	-	-	4	n.d.	n.d.	31	151
12	9.05	(structure)	519557	597	3177	673	1726	23	85	98	360	5911
14* .1	8.86	(structure)	56557	381	1610	278	415	n.d.	n.d.	n.d.	n.d.	n.d.
14* .2	9.24	(structure)	1661805	6141	22128	5574	10203	226	549	647	40	n.d.
16	10.05	(structure)	519557	1215	4807	1381	2533	8	118	181	3	157

* Despite their representation, we cannot assure which isoform correspond to each chromatographic peak.

To assess whether the pooled serum contained endogenous REG1A that could interfere with the analysed rREG1A glycopeptides, REG1 levels of this pooled serum were quantified by ELISA and showed a concentration of 58.5 ng/mL. The analysis of 1 mL of this serum by LC-MS following the described methodology for serum samples showed only one glycopeptide, corresponding to peak 12, with very low intensity. These results indicated that the endogenous REG1 levels of this serum should not affect rREG1A glycopeptides quantification.

The next step was to determine these glycopeptides in real samples from patients. We first addressed the sensitivity of the methodology. While the addition of 100 ng rREG1A to 1 mL of healthy pooled serum (matrix) yielded quantifiable intensities for all selected glycopeptides, 50 ng rREG1A spiked in serum just allowed the detection of peaks 1, 2.1, 11, 12, 14.2 and 16 (Table 14).

Overall, these results suggested that, at this stage, we required about 150 ng/mL REG1A to obtain enough sensitivity to quantify the different REG1 glycopeptides by LC-MS. Thus, we needed to improve the sensitivity before analysing real samples, which could have lower REG1 levels (as the commercial pooled serum). To increase the REG1 concentration, we focused on reducing the final volume of the sample after REG1A immunopurification. In a sample of 50 ng of rREG1A in 1 mL serum, we changed the trypsin elution buffer from $NaHCO_3$ to NH_4HCO_3, so that samples could be completely dried (either by lyophilisation or in a vacuum centrifuge) and further resuspended in just 15 µL NH_4HCO_3 instead of 35 µL. We observed that both methodologies yielded higher peak intensities than the previous protocol, being the vacuum centrifuge the best option with at least an approximately 4-fold increase (Table 14).

A pilot test with one control sample (commercial pooled serum, 58.5 ng/mL REG1) and one PaC IIB sample (217.2 ng/mL REG1) was performed to analyse REG1A glycopeptides in patients' samples. To guarantee the detection of glycopeptides, REG1 was initially immunopurified from 2 mL serum. The control serum underwent the concentration processing by lyophilisation, while the PaC sample was not concentrated. Several glycopeptides could be detected (Table 14),

corresponding to the non-glycosylated (peak 1), the peak 2.1, peak 11, peak 12 and peak 16. Peak 14.2 was only observed in the control serum, while both isoforms of peak 7 were exclusively observed in the tumour sample. The SRM method included a new transition, corresponding to the STn structure, not previously observed in rREG1A (either by MALDI-Tof or LC-MS) nor the pooled serum, but with biologic significance and possible overexpression in cancer samples. Actually, this structure was only observed in the PaC serum. Another interesting feature of this PaC sample was observed in peak 7, which presented an altered isoform ratio respect the rREG1A, translated in the increase of the sialyl T-antigen (isomer 7.1). Thus, serum REG1A glycopeptides in a PaC IIB patient were characterised by the increase in STn and sialyl-T antigens, two structures widely related to cancer progression.

These results demonstrated that this methodology was successful in analysing patients' serum and paved the way for the analysis of REG1A glycoforms in a panel of samples. Unfortunately, the optimisation of the analytical method with a recombinant standard directed the SRM acquisition to the structures more abundantly found in this, overlooking other glycans that could be found in natural human produced REG1A. In this regard, assuring that the SRM method can detect all relevant REG1A glycoforms (those expressed in rREG1A and those expressed by humans in healthy and malignant conditions) is the last step before the analysis of a cohort of patients.

Duran Sidera, A.

DISCUSSION

Duran Sidera, A.

Despite overall advances in cancer diagnosis and treatment in the last decades, which have reduced the mortality of most malignancies, PaC still pose a worldwide healthcare issue, currently being the fourth cause of cancer death in western countries and the tumour with the worst prognosis [17]. In addition, it is projected to be the second cause of death in the US by 2030 [391]. The lack of tumour markers for its early diagnosis [26,262] and adequate therapeutic strategies for late stages of the disease [112] are responsible for the poorest survival among all cancers.

As previously stated, despite a high number of molecules and assays have been proposed for the diagnosis of PaC [275,276], none of them have been validated by the FDA, which still presents CA19-9 as the only option for PaC management [392]. Actually, the development of new non-invasive biomarkers has proven to be extremely challenging in recent years, and the number of clinically relevant tumour markers has barely raised [229,393]. It must be noted that it is not easy for a biomarker to progress from the laboratory to the clinical practice. Researchers do not just need to find a molecule differently expressed among healthy and pathologic individuals, but also a proper analytical method with enough sensitivity to discriminate both populations. Once accomplished, years of validation studies (including protocol optimisation and standardisation, high-throughput adaptation, retrospective clinical studies, prospective screening studies and randomised control trials) will eventually confirm the use of such biomarker [260,394].

It is difficult to imagine that a single biomarker can provide optimal sensitivity and specificity for the detection of PaC. While organ-specific molecules may be altered in other pancreatic disorders, those more cancer-specific might be found in several cancer types. Additionally, population genetic variation may block the synthesis of tumour-associated variants, like the inhibition of CA19-9 expression in 5-10% population [154]. In this regard, the use of biomarker combinations, panels or clinical nomograms is acquiring interest to compensate the limitations given by a single analyte and yield better accuracy. An alternative methodology relies on the detection of biomolecules that can bear double information. Glycoproteins may fulfil this requisite, as the protein part may be organ-specific while alteration in

glycosylation could be cancer-associated [230]. Actually, the study of altered glycosylation on glycoproteins, which had already been validated by the FDA as cancer biomarkers, was proposed long ago to enhance the accuracy of such tumour markers [395]. Other authors have also reviewed the potential of altered glycosylation on specific glycoproteins as biomarkers, and even the need to combine several glycoproteomic markers to develop a clinically relevant diagnostic tool [230,243,396,397].

Previous studies have already been successful in developing a cancer biomarker based on the altered glycosylation of a specific protein. This is the case of alpha-fetoprotein (AFP), which is a glycoprotein overexpressed in hepatocellular carcinoma, but also in other cancers and benign liver disorders [398]. Nowadays, it still remains as the most used method for hepatocellular carcinoma diagnosis despite a sensitivity and specificity of just 41-65% and 80-94%, respectively [399]. A specific core fucosylated AFP glycoform, reactive with the lectin *Lens culinaris* agglutinin (LCA) and named AFP-L3, has demonstrated enhanced accuracy for assessing the risk of developing hepatocellular carcinoma [400]. Moreover, AFP-L3 has been combined with other parameters to develop the GALAD score, which has been approved by the FDA as an early diagnosis tool for hepatocarcinoma [401].

Similarly, altered glycoforms are being studied in other cancers, such as altered PSA glycoforms in prostate cancer including altered α2,3-sialylation, LacdiNAc and core fucosylation [402–410]. Another example is found over specific CA125 glycoforms bearing the STn antigen, which also hold a promising potential to improve the diagnosis of ovarian cancer [411].

In this thesis, we have focused on the development of new methodologies to identify novel biomarkers for PaC based on the altered glycosylation of two glycoprotein candidates, MSLN and REG1. For such, we have assessed their cell and tissue expression, characterised their glycan moieties and developed new methodologies to detect and quantify, in blood serum, those PaC-related glycoforms which could be useful as tumour markers.

The analysis of single proteins glycosylation posed several challenges compared to the evaluation of the overall glycophenotype on a given situation, being the purification of the glycoprotein of interest a crucial step for that purpose. Almost all methods for the structural glycan analysis of a target protein, including lectin, chromatographic or MS-based methodologies, require the isolation of such glycoprotein [158]. If not properly isolated, obtained results would be masked by the glycan component of other glycoconjugates. Only glycopeptide targeted MS methods may skip this purification step, despite it is recommended for reducing background signalling. Several strategies are available for the separation of glycoproteins, being chromatographic techniques those featuring higher resolution [412]. For subsequential glycan analysis, which requires really accurate isolation, the use of immunoaffinity chromatography is the standard methodology of use [413]. In this thesis we have performed protein immunoprecipitation, which maintain the advantages of the immunoaffinity chromatography regarding specificity while being a less time-consuming and cost-effective method.

Main considerations for designing an immunoaffinity protocol include the immobilisation of the antibody and the conditions for the elution of the protein, which will depend on the following analytical technique. Coupling of antibodies to protein A/G beads is the most classical approach [414,415]. However, co-elution of the antibody might limit further analysis. Covalent binding of the antibody to the beads is a choice to overcome this issue [415]. Common elution conditions include pH variation, chaotropic agents (usually high salt concentrations), denaturing effects and competitive ligands [413]. Again, the preferential method relies on the sample post-processing.

Unfortunately, purification methodologies not always display high yields, so glycoprotein recovery is usually poor and subsequent analyses might be compromised. The obtention of enough purified protein quantity is crucial to reach posterior analytical sensitivity, a not so frequent problem in the analysis of global glycosylation profiles.

MSLN glycoforms as PaC tumour markers

MSLN is a glycoprotein *de novo* expressed in a number of cancers including PaC [344], which makes it attractive as a diagnostic and therapeutic target [337]. Despite its role in the malignant transformation is not totally understood, several authors have proved its influence in the invasive and metastatic potential, growth and apoptosis regulation and EMT in PaC [338].

Due to its limited expression in healthy tissues, MSLN has grown into a therapeutic target for different types of cancers and therapeutic strategies, like vaccines [416], immunotoxins [417–420], monoclonal antibodies [421,422], antibody-drug conjugates [122,423] or CAR-T cells [424]. On the other hand, MSLN has also been proposed as a biomarker for various malignancies. Actually, there is an FDA-approved test for the diagnosis of malignant pleural mesothelioma based on soluble MSLN related peptides serum levels [342,355,425]. Nonetheless, it has not been useful to diagnose other cancers, including PaC [354,426], as we also showed in the cohort of patients analysed in this work.

The expression of MSLN and the analysis of its glycosylation from PaC cells and tissues was addressed in this study to determine the main MSLN glycan determinants in malignant conditions (Figure 41). The acquired knowledge might not only be useful as a discriminatory tool between patient conditions, but it may also aid in the development of new MSLN-based targeted therapies [324,344]. It is known that differential glycosylation, apart from a target itself, may vary the affinity of an antibody for a given glycoprotein, as has been previously observed [427,428]. Thus, when designing antibodies for a directed therapy, it is crucial to understand how the target glycosylation can affect the therapeutic agent. For instance, in the case of MSLN, the recognition by the antibody MORAb-009 is independent of MSLN glycosylation [429]. MORAb-009 identifies a non-linear epitope in the *N*-terminal region [430], before the first *N*-glycosylation sequon (Asn388). The monoclonal antibody used in this work, MN-1, recognises an epitope in the same region with much higher affinity [431], so MSLN glycosylation should not interfere

in its recognition. However, other antibodies binding different epitopes might be affected.

In this work, we showed that PaC cell lines express MSLN, which is in agreement with other authors [421,432]. In addition, MSLN expression was also analysed in PaC tissue lysates, which were mainly positive by WB (77%). This percentage is in accordance with other studies that report high expression levels in PaC (60-95%), while the presence in normal pancreatic tissues or other controls was scarce [349–351,433–437]. Unfortunately, the need to biopsy these tissues make this differentiation ineffective for a routinely clinical test.

PNGaseF digestion on MSLN from PaC cells confirmed the existence of N-glycans as previously described [438]. The molecular weight difference from the glycosylated to the PNGaseF digested MSLN, observed by SDS-PAGE and MS, suggested a full occupancy of the three N-glycosylation sites and was confirmed by peptide mapping after PNGaseF digestion. However, these N-glycans had not been characterised up to date. Only Fujihira and collaborators [333], in a study regarding mesothelioma, had assessed MSLN glycosylation before. A lectin microarray showed increased expression of bisecting GlcNAc in MSLN from mesothelioma cell lines versus normal mesothelial cells. Glycopeptide analysis on H226 mesothelioma cell line showed the presence of N-glycans in the three N-glycosylation motifs, which were complex, sialylated and core fucosylated. Bi-, tri- and tetra-antennary structures were observed, being the N-glycan at position Asn388 the one bearing the higher number of structures including the bisecting GlcNAc.

The N-glycans from PaC cell lines immunopurified MSLN, which were characterised by N-glycan sequencing and by WB with several lectins, also showed to be complex-type and sialylated, most of which highly branched and with core fucosylation (Figure 41). α2,3-SA was predominant on MSLN glycans for all cell lines except for HPAF-II, which only showed neutral N-glycan structures. These glycan determinants coincide with the overall pattern found in the PaC glycome, which has been extensively reviewed [193,307,439,440]. Interestingly, MSLN N-glycosylation was highly conserved in AsPC-1 and Capan-2, despite these two cell lines

present quite different properties. While AsPC-1 cells are poorly differentiated and derive from the ascites of a female with metastasis, Capan-2 cells are well differentiated and come from the primary tumour of a male without metastasis [441]. Thus, it seems that MSLN glycosylation could be independent of the differentiation degree and metastatic potential of the tumour.

The *N*-glycans detected in the commercial recombinant MSLN standard, produced in NS0-derived cells, were totally different from those in human cells MSLN, being the presence of terminal α-Gal and *N*-glycoylneuraminic acid (Neu5Gc) the most remarkable features. Actually, these epitopes are not naturally found in human, and are responsible for inflammation when identified by the immune system [442]. The absence of Neu5Gc in human glycans is due to the loss of function of CMAH, the enzyme responsible for the conversion of Neu5Ac to Neu5Gc [443]. However, by means of unknown mechanisms, dietary Neu5Gc can be incorporated into human glycans, despite its abundance is 10000 times lower than Neu5Ac [444]. Similarly, inactivation of the *GGTA1* gene in primates precludes the synthesis of α-Gal in humans [445,446]. Additionally, this is a highly immunogenic epitope, being anti-Gal the most abundant antibody in humans and accounting for about 1% of total immunoglobulins [447]. However, murine myeloma cell lines (such as NS0), which are often used as a production platform for recombinant glycoproteins, yield considerably high levels of terminal α-Gal and Neu5Gc, with sialylation predominantly in α2,6-linkage [448]. Recombinant MSLN glycosylation mostly agreed with all these statements, except for the SA linkage, which was mainly α2,3-linked.

In order to evaluate MSLN glycoforms from PaC tissues and serum, a quantitative methodology based on an antibody-lectin immunoassay was developed aiming to analyse those glycoforms previously identified in PaC cells (Figure 41). The main challenge was overcoming the low-binding affinity of lectins, found at the low millimolar range (while antibody-antigen interactions are usually around the nanomolar range) [449]. Despite optimisation of several conditions including capture

antibody and lectins dilution, buffers compositions, incubation times and temperatures, this issue could not be solved for several lectins. However, in the case of PhoSL, a core fucose-specific lectin with a dissociation constant of 3 μM [385], proper analytical sensitivity was achieved. Hence, the analysis of core fucosylated MSLN glycoforms was assessed.

Cf-MSLN levels were lower in PaC tissues than the full core fucosylated rMSLN. When analysing the ratio of Cf-MSLN/MSLN in serum samples from PaC patients, other pancreatic disorders and healthy volunteers, a significant decrease in Cf-MSLN glycoforms was observed in malignant stages, which agreed with the low Cf-MSLN levels described in PaC tissues.

Figure 41. Flowchart of the main findings of MSLN glycosylation in PaC. Schematic representation of the methodology followed and main findings regarding the analysis of MSLN glycosylation as a PaC biomarker, which include the analysis of MSLN expression, characterisation of its *N*-glycans, development of a methodology for the quantification of Cf-MSLN and validation of its discriminatory potential in a cohort of blood serum samples.

Other authors have also described changes in the core fucosylation degree of serum proteins in PaC patients. This is the case for haptoglobin, a glycoprotein mainly produced by the liver with increased fucosylation in PaC patients compared to healthy individuals [312,313,450]. Specifically, core fucosylated haptoglobin, assessed through a sandwich ELLA using PhoSL (as we have performed Cf-MSLN detection), showed increased levels in PaC patients, and even higher in

ChP [451]. Improvement of the methodology, adding high urea concentration to denaturalise the protein and facilitate lectin detection, displayed higher levels of core fucosylated haptoglobin in PaC [452]. However, elevated levels in ChP, cholangiocarcinoma and hepatocellular carcinoma impaired its usefulness as a PaC biomarker.

Elevated haptoglobin core fucosylation coincides with the general increase in core fucosylation observed in the PaC serum glycome [201], mainly in tri- and tetra-antennary structures, while Cf-MSLN levels follow a contrary trend. These controversial results are suitable to illustrate the complexity of the glycosylation process and, as detailed in this thesis introduction, understand that all the agents involved in protein glycosylation can distinctly glycosylate two proteins in similar conditions. Hence, it is important to emphasise that a single glycoprotein does not necessarily reflect the overall glycome features from PaC glycoproteins.

We demonstrated that Cf-MSLN had the potential to discriminate PaC from control patients, and 93.3% sensitivity and 81.2% specificity was obtained with the combination with CA19-9. On the other hand, MSLN protein levels did not show this ability in the same patients cohort, as reported previously by other authors [354,357]. Recent studies suggest that low MSLN serum levels in PaC patients result from the glycoprotein retention in the tumour microenvironment via collagen density and MUC16 interaction [358]. This condition, which is impairing MSLN usefulness as a PaC biomarker, enhances its potential as a target for directed therapeutic strategies.

Considering the prevalence of PaC in Europe, which is of 0.0138% [25], the positive predictive value (PPV) and negative predictive value (NPV) of the combination of Cf-MSLN and CA19.9 were 6.5% and 99.6%, respectively. Despite proper accuracy for the combination of Cf-MLSN + CA19-9, a low PPV of just 6.5% impairs its use in a general population screening approach. However, the faecal immunochemical test, widely used to screen for colorectal cancer, presents a similar PPV of 7% in low-risk population [453]. This shows the need to narrow the population group to screen. Considering the increased prevalence of PaC in the

population at risk, as those patients over 50 years with diabetes [454], patients with ChP [455] or first-degree relatives of PaC patients [456], higher PPVs would be obtained. Thus, a diagnostic test for PaC could really be useful to screen population at risk of suffering PaC and help in the selection of patients to undergo further confirmation of the disease with imaging techniques.

It would be interesting to compare the accuracy of this newly proposed biomarker combination with that of other biomarker panels which are in the way of validation by international consortiums, such as the UroPanc test, based on a three urinary-biomarkers panel plus plasma CA19-9 [380], or the IMMray PanCan-d test, based on a 9-plex biomarker signature [457]. However, it is difficult to compare biomarkers accuracy assayed in different cohorts, because they could provide different sensitivity and specificity values depending on the type of cohorts analysed. For instance, UroPanc separately compares healthy and benign conditions versus PaC. Regarding the IMMray PanCan-d test, the presented results do not include benign pancreatic conditions. Ideally, cohorts comprising the groups at PaC risk and patients with benign pancreatic disorders should be included and tested together to compare biomarkers test accuracy.

Altogether, the obtained results emphasise the need for further studies evaluating serum Cf-MSLN in a larger cohort of patients which include sera from other group of patients with high PaC risk as well as other types of cancers, thus aiding in the evaluation of Cf-MSLN specificity for PaC. Moreover, further optimisation of the current protocol to increase its methodological sensitivity and high-throughputness would be needed for its clinical translation. In this regard, with the arrival of new glycoproteomic strategies [458,459], we consider that the development of an MS-based glycoproteomic pipeline could provide this improvement.

REG1 glycoforms as PaC tumour markers

REG1 glycoproteins, secreted by pancreatic acinar cells, are the principal non-enzymatic component of pancreatic juice [372]. Increased expression is observed in several pathological conditions of the pancreas, digestive tract, nervous system and a number of cancers [359], reason why they have been studied as potential

biomarkers. The exact function of REG1 is not clearly understood. In normal situation, it is related with pancreatic calcium calculi formation and regeneration of brain and pancreas tissues [372]. It might also play a role in differentiation induction and have mitogenic and antiapoptotic activity [460].

In this work, the expression of REG1A and REG1B was assessed in PaC cell lines and pancreatic tissues (Figure 42). Unfortunately, due to the high similarity between both glycoproteins (89% of homology [363]), differential assessment was not possible. Despite denying any cross-reactivity, the commercial antibodies tested in this work could recognise both glycoproteins, which led us to the general analysis of REG1 glycoproteins. We did not detect REG1 in either the protein lysates or secreted media of seven different PaC cell lines, in contrast with previous studies reporting REG1B expression in Capan-1 [363] or REG1A expression in BxPC-3 and SW1990 [376]. However, it must be noted that these last studies relied on an immunised rabbit antisera dilution and a currently discontinued antibody, respectively, so results should be carefully interpreted.

Regarding REG1 expression in pancreatic tissues, it was observed in 60.5-65.8% of PaC tissue lysates and 90.9% of healthy tissues. Previous immunohistochemistry studies reported similar results: 100% staining in control tissues and 65-70% in PaC [365,366]. Interestingly, the study of Satomura and collaborators observed a correlation between samples staining grade and tissue differentiation degree [365]. Staining was strong in well-differentiated tissues and mild in poorly differentiated cells, contrarily to the trend observed in this work. However, we must be careful when comparing immunohistochemistry and WB determinations. While the former provide detailed information of protein expression on different observed tissue/cell structures within a slide, the latter relied on the lysis of tissue pieces with unknown abundancies of pancreatic tumour cells, tumour stroma, new blood vessels or adjacent healthy cells.

The study of REG1A glycosylation is of utmost interest given the crucial role it might play in its physiological activity. Digestion of the REG1 N-terminal peptide derives in an insoluble form that can induce fibril-like structures formation in brain

[461]. Lebart and collaborators have recently described that REG1A can be digested by calpain-2, a protease with a cleavage site between Gln4 and Thr5 [462]. Interestingly, this cleavage depends on REG1A glycosylation status, and only the non-glycosylated protein is cleaved. However, no fibril formation has been observed *in vitro* after this digestion. Nevertheless, this study illustrates how REG1A glycosylation can modulate its activity and points the importance of its study.

Due to the limited sample availability, REG1 *O*-glycosylation could not be assessed through glycan sequencing in a UPLC system and it was first evaluated by lectin detection on immunopurified REG1 from pancreatic tissues and glycosidase digestions (Figure 42). SA digestion with NAN1 and ABS revealed the presence of SA, mainly in α2,3-linkage. WB with lectins also showed the presence of α2,3-SA, in varying amounts depending on the sample, while α2,6-SA was only observed in some tumour samples. Accordingly, those samples that did not show a molecular mass decrease due to α2,6-SA digestion could not be detected with SNA. These results agreed with previous published works underlying REG1 glycosylation [374], which also showed a predominance of α2,3-SA. Similarly, expression of terminal α1,2-Fuc was also observed with the lectin UEA.

Our results showed the presence of terminal α2,6-SA and GalNAc only in PaC samples, while terminal Gal was exclusively observed in controls. We hypothesised that the observed GalNAc in PaC REG1 could correspond to the Tn antigen, which is widely expressed in tumour conditions [161]. The expression of α2,6-SA could also be related to the hypersialylation normally observed in malignant situations [202,463], while the absence of this SA would reveal the subterminal galactose detected in control samples.

The quantitative analysis of these REG1 glycoforms required the development of ELISA and ELLA immunoassays to properly measure the protein and its specific glycoforms, respectively. Our optimised ELISA protocol reported a LOD of 0.16 ng/mL, slightly higher than other developed or commercial available kits, which showed LODs ranging from 0.01 to 0.09 ng/mL (USCN, China; MyBioSource, CA, USA; Novus Biological, MN, USA; [383,464]). As no consensus nor validated

ELISA kit is predominant in the published literature, our developed protocol seemed good enough to quantify REG1.
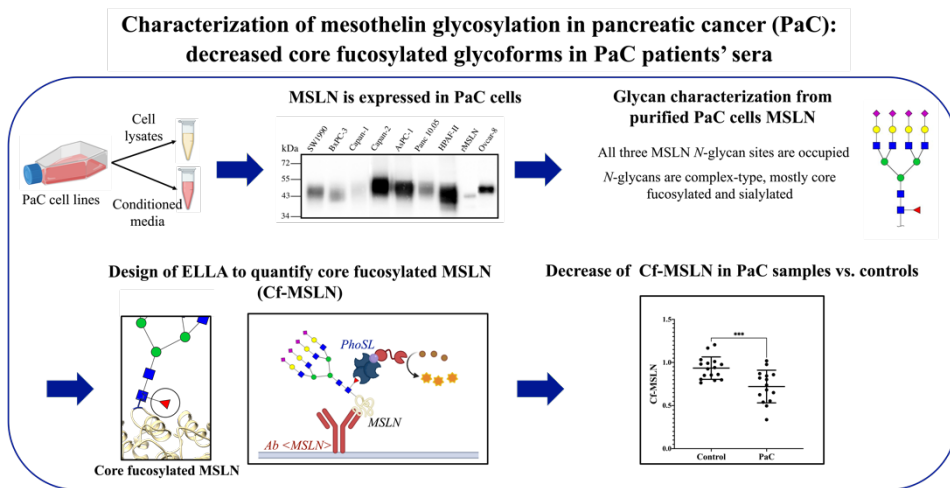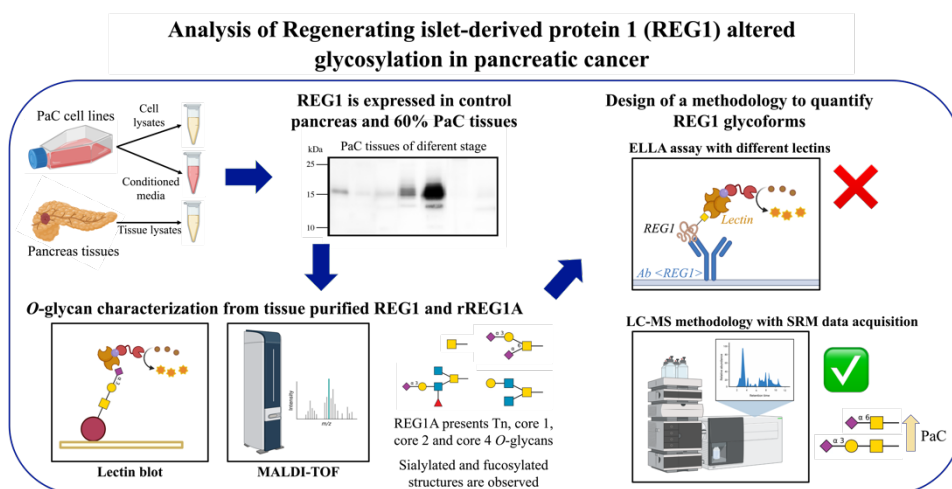

Figure 42. Flowchart of the main findings of REG1 glycosylation in PaC. Schematic representation of the used methodology and main findings regarding the analysis of REG1 glycosylation as a PaC biomarker, which include the analysis of REG1 expression, characterisation of its *O*-glycans and development of a method for the quantification of REG1 tumour-associated glycoforms.

On the other hand, we were not successful in developing an ELLA assay for detecting REG1 specific glycoforms. As mentioned before for the case of MSLN, lectin low-binding affinities may pose a challenge for its use in the detection of glycoproteins, especially in proteins with short glycan structures and a single glycosylation site as REG1. Only the detection of rREG1 glycoforms with terminal GalNAc using the lectin VVL yielded an accurate calibration curve, despite all the assayed serum samples did not reach the LOQ. Regarding the other tested lectins, increased background or poor analytical sensitivity (by means of low calibration curve slope) was observed.

To overcome lectins limitations, we performed the analysis of REG1 glycans by MS, starting from the characterisation of rREG1A in a MALDI-Tof application (Figure 42). This MS methodology also presented some limitations, being the differentiation of isoforms with the same molecular mass the main issue. However, based on general described *O*-glycan structures and results obtained after specific sialidase digestions, several structures could be proposed. We observed truncated

core 1 structures, as well as core 2 and core 4 *O*-glycans, with prevalence of the latter. These structures did not agree with previous studies regarding REG1A glycosylation, which mainly showed core 2 *O*-glycans [374]. Only the di-sialyl T was detected in both studies. However, it was not rare to find these differences in glycosylation. We must bear in mind that rREG1A is expressed in kidney cells (HEK293), which may present a glycosylation machinery regulation different from human pancreatic samples. HEK293 cells *O*-glycosylation is mainly characterised by the expression of mono- and di-sialylated T, followed by the T antigen and other core 2 structures [465,466]. We also detected these structures, in the addition to core 4 *O*-glycans. Nevertheless, comparison of the general *O*-glycan pattern with that of a single protein is controversial, even more considering that rREG1A is being expressed in a cell type that does not express it normally.

A LC-MS method was developed to reduce background signalling and provide a sensitive and specific methodology for the analysis of REG1 glycopeptides from serum samples. The study of rREG1A glycopeptides in a LC (HILIC)-TQ operating in SRM mode with the transitions selected from the MALDI-Tof results could detect all the main glycoforms except the T antigen. In addition, different isoforms were observed in several masses, that could be attributed to different SA linkage or Gal/GlcNAc branching arrangement.

Interestingly, some selected masses showed a quite large number of isoforms, even more than the expected for the number and possible linkage between monomers. We observed that, in most cases, additional peaks were observed due to fragmentation of terminal SA. Large structures, with relatively high retention times, could loss a SA in the ionisation source and provide signals for lower masses at unexpected retention times. Optimisation of the source temperature could not totally prevent this event, as maximal signal of the ionised glycopeptides was preferred to yield enough analytical sensitivity. Reduction of SA loss also resulted in a decrease in the glycopeptides ionisation and corresponding signal intensity.

Actually, SA analysis by MS is quite challenging [467]. The negative charge and unique structure of SA provide properties such as instability, ionisation discrimination and mixed adducts, which alter MS results, especially in MALDI ionisation. These problematics could explain why we observed the T antigen in MALDI-Tof and only its sialylated variants in the LC-ESI-TQ analyses. Common strategies to overcome these issues include sialoglycoforms enrichment (not applicable if the detection of neutral glycoforms is also desired) or SA derivatisation. Several methodologies are currently available, and are used to avoid the loss of SA, enhance the ionisation efficiency and facilitate isomers profiling. For example, derivatisation with dimethylamine in the presence of EDC and HOBt results in lactonization and dimethylamidation of $\alpha2,3$- and $\alpha2,6$-SA, respectively, yielding different glycopeptide masses and enhancing the stability [468]. Other derivatisation methodologies include ethyl esterification and amidation of sialoglycans [469], SA alkylamidation [470] or permethylation [471]. However, our attempts to analyse dimethylamine derivatised REG1A glycopeptides were unsuccessful.

The analysis of sialylated glycopeptides by MALDI-Tof is recommended in the negative-ion mode using 3-Aminoquinoline/p-Coumaric acid as a matrix [472]. However, we obtained no signal in the negative-ion mode with the use of α-cyano-4-hydroxycinnamic acid, another commonly used matrix for the detection of glycopeptides [473], reason why we needed to operate in the positive-ion mode. Other common used methodologies for the separation and differential analysis of $\alpha2,3$- and $\alpha2,6$-SA include capillary electrophoresis coupled to MS [221] or LC-MS approaches using chromatographic columns with the ability to separate sialylated glycopeptide isoforms, for instance with a HILIC column [474,475], as we used in this study.

We optimised a workflow for human serum samples analysis consisting of the immunopurification of REG1A from 2 mL serum followed by elution with trypsin, concentration and LC-MS analysis. The study of a control and a PaC IIB sample showed differences in the abundance of some glycoforms, being the most relevant the increase of sialyl-T and STn antigens in the tumour sample. Thus, REG1A

O-glycosylation in the PaC patient serum corresponded with the tumour-associated structures found in *O*-glycans [476,477]. Analyses of more PaC and control samples are required to confirm the presence of these tumour-associated glycoforms in most PaC samples and their differential expression from control ones.

Concluding remarks and future directions

The study of specific glycoforms on glycoproteins with altered expression in PaC could yield novel tumour markers for its diagnosis. In this work, we in-depth analysed the glycan structures carried by two glycoproteins (MSLN and REG1) in diverse physiological conditions. Moreover, we designed new assays for the determination of those glycoforms able to discriminate between benign and malignant conditions.

Cf-MSLN showed promising results in diagnosing PaC when combined with CA19-9 levels. The usefulness of this biomarker should be further validated in a larger cohort of patients including other cancer types as well as other benign situations. Development of other strategies for the assessment of Cf-MSLN, minimising sample handling, would also be appreciated for its translation into the clinical routine. Many other glycan determinants were detected in MSLN apart from its core fucosylation, such as highly branched and sialylated structures. In this regard, the development of other assays to evaluate these glycoforms could also yield encouraging results.

A MS strategy for determining several REG1A glycoforms in biological samples was developed and could be used to analyse a large number of serum samples in a high-throughput and automated mode. Glycopeptide analysis of REG1A by MS provided the identification of various particular glycan structures in contraposition to the use of biological probes such as lectins, which have intrinsic limited affinity and specificity. The increase in REG1A glycopeptides containing sialyl-T and STn in a PaC sample pointed the potential of these glycoforms as tumour markers. Hence, the analysis of REG1A glycoforms in a cohort of patients and control individuals to determine their accuracy in detecting PaC should be regarded as the next step in this research line.

To summarise, this thesis provides novel knowledge on the structure of MSLN and REG1 glycosylation in PaC and describes useful methodologies for their analysis in complex biological samples such as blood serum. At the same time, it opens the way to their future validation as PaC diagnostic tests based on these glycoproteins altered glycosylation.

Duran Sidera, A.

CONCLUSIONS

Duran Sidera, A.

In regard to **the analysis of MSLN glycoforms as PaC tumour markers**:

1. MSLN is vastly expressed in PaC cell lines, both in protein lysates and their secreted media, being the concentration in the latter much higher. MSLN is also expressed in 77.4% of PaC tissues and only in 10% of healthy pancreas lysates.

2. At the macroheterogeneity level, MSLN N-glycosylation is found at its three potential N-glycosylation sites. Regarding its microheterogeneity, MSLN N-glycans from PaC cell lines are mainly composed of sialylated complex-type structures, most of which with core fucosylation and highly branched.

3. A sandwich ELISA for the quantification of MSLN and a hybrid ELLA immunoassay to specifically quantify core fucosylated MSLN glycoforms (Cf-MSLN) in complex biological samples have been established. The former relies in the use of two different species anti-MSLN antibodies, while the latter requires an anti-MSLN capture antibody and the PhoSL lectin as the detection probe.

4. The analysis of serum Cf-MSLN in a cohort of healthy individuals, ChP and PaC patients has shown a decrease of Cf-MSLN glycoforms in PaC patients. This novel test can discriminate PaC patients from ChP patients and healthy individuals with a sensitivity and specificity of 73.3% and 75%, respectively. Combination of Cf-MSLN with CA19-9 increase the accuracy of the biomarker up to 93.3% and 81.2% sensitivity and specificity.

Related to **the analysis of REG1 glycoforms as PaC tumour markers**:

5. REG1 is not expressed in the analysed PaC cell lines, either in protein lysates or secretion media. In tissues, MSLN is expressed in 60.5-65.8% of PaC samples and 90.9% of control tissues.

6. Pancreatic tissues REG1 *O*-glycans analysis with lectins has shown them to be sialylated and fucosylated. The expression of terminal Gal is only observed in control tissues, while the expression of GalNAc and α2,6-SA is only detected in PaC tissues.

7. The development of an ELLA assay to quantify those REG1 glycoforms differentially expressed in benign and malignant situations has not been possible due to the lack of affinity and sensitivity of commercially available lectins and secondary reagents. Nonetheless, a sandwich ELISA to quantify REG1 has been established.

8. The analysis of a recombinant REG1A standard *O*-glycosylation by MS-based glycoproteomic approaches has showenup to 16 glycan structures, including truncated, core 2 and core 4 *O*-glycans.

9. A targeted LC-MS methodology consisting in a HILIC separation, ESI ionisation and TQ detection (with SRM data acquisition) to analyse REG1A glycopeptides from complex biological samples has detected up to eight REG1A glycoforms.

10. REG1A glycoforms that contain tumour-associated glycans (STn and sialyl-T) are *neo*- and over-expressed in a PaC serum sample compared to a pooled control serum. Therefore, these REG1A glycoforms stand as candidates for PaC biomarkers.

Duran Sidera, A.

REFERENCES

Duran Sidera, A.

1.   Fares, J.; Fares, M.Y.; Khachfe, H.H.; Salhab, H.A.; Fares, Y. Molecular Principles of Metastasis: A Hallmark of Cancer Revisited. *Sig Transduct Target Ther* **2020**, *5*, 28, doi:10.1038/s41392-020-0134-x.

2.   Nenclares, P.; Harrington, K.J. The Biology of Cancer. *Medicine* **2020**, *48*, 67–72, doi:10.1016/j.mpmed.2019.11.001.

3.   Suhail, Y.; Cain, M.P.; Vanaja, K.; Kurywchak, P.A.; Levchenko, A.; Kalluri, R.; Kshitiz Systems Biology of Cancer Metastasis. *Cell Systems* **2019**, *9*, 109–127, doi:10.1016/j.cels.2019.07.003.

4.   Massagué, J.; Obenauf, A.C. Metastatic Colonization by Circulating Tumour Cells. *Nature* **2016**, *529*, 298–306, doi:10.1038/nature17038.

5.   Matthews, H.K.; Bertoli, C.; De Bruin, R.A.M. Cell Cycle Control in Cancer. *Nat Rev Mol Cell Biol* **2022**, *23*, 74–88, doi:10.1038/s41580-021-00404-3.

6.   Li, T.; Kang, G.; Wang, T.; Huang, H. Tumor Angiogenesis and Anti-Angiogenic Gene Therapy for Cancer. *Oncol Lett* **2018**, *16*, 687–702, doi:10.3892/ol.2018.8733.

7.   Pastushenko, I.; Blanpain, C. EMT Transition States during Tumor Progression and Metastasis. *Trends in Cell Biology* **2019**, *29*, 212–226, doi:10.1016/j.tcb.2018.12.001.

8.   Jolly, M.K.; Ware, K.E.; Gilja, S.; Somarelli, J.A.; Levine, H. EMT and MET: Necessary or Permissive for Metastasis? *Mol Oncol* **2017**, *11*, 755–769, doi:10.1002/1878-0261.12083.

9.   Cheng, X.; Cheng, K. Visualizing Cancer Extravasation: From Mechanistic Studies to Drug Development. *Cancer Metastasis Rev* **2021**, *40*, 71–88, doi:10.1007/s10555-020-09942-2.

10.   Ganesh, K.; Massagué, J. Targeting Metastatic Cancer. *Nat Med* **2021**, *27*, 34–44, doi:10.1038/s41591-020-01195-4.

11.   Hanahan, D.; Weinberg, R.A. The Hallmarks of Cancer. *Cell* **2000**, *100*, 57–70, doi:10.1016/S0092-8674(00)81683-9.

12.   Hanahan, D. Hallmarks of Cancer: New Dimensions. *Cancer Discovery* **2022**, *12*, 31–46, doi:10.1158/2159-8290.CD-21-1059.

13.   Murphy, S.L. *Mortality in the United States, 2020*; National center for health statistics, 2021;

14.   Causes of Death Statistics Available online: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Causes_of_death_statistics (accessed on 6 February 2023).

15.   Statistics | Eurostat Available online: https://ec.europa.eu/eurostat/databrowser/view/HLTH_CD_ARO__custom_4843203/default/table?lang=en (accessed on 6 February 2023).

16. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A Cancer J Clin* **2021**, *71*, 209–249, doi:10.3322/caac.21660.

17. Siegel, R.L.; Miller, K.D.; Wagle, N.S.; Jemal, A. Cancer Statistics, 2023. *CA A Cancer J Clinicians* **2023**, *73*, 17–48, doi:10.3322/caac.21763.

18. Dyba, T.; Randi, G.; Bray, F.; Martos, C.; Giusti, F.; Nicholson, N.; Gavin, A.; Flego, M.; Neamtiu, L.; Dimitrova, N.; et al. The European Cancer Burden in 2020: Incidence and Mortality Estimates for 40 Countries and 25 Major Cancers. *Eur J Cancer* **2021**, *157*, 308–347, doi:10.1016/j.ejca.2021.07.039.

19. Kleeff, J.; Korc, M.; Apte, M.; La Vecchia, C.; Johnson, C.D.; Biankin, A.V.; Neale, R.E.; Tempero, M.; Tuveson, D.A.; Hruban, R.H.; et al. Pancreatic Cancer. *Nat Rev Dis Primers* **2016**, *2*, 16022, doi:10.1038/nrdp.2016.22.

20. Wolfgang, C.L.; Herman, J.M.; Laheru, D.A.; Klein, A.P.; Erdek, M.A.; Fishman, E.K.; Hruban, R.H. Recent Progress in Pancreatic Cancer. *CA: a cancer journal for clinicians* **2013**, *63*, 318–348, doi:10.3322/caac.21190.

21. Singh, D.; Upadhyay, G.; Srivastava, R.K.; Shankar, S. Recent Advances in Pancreatic Cancer: Biology, Treatment, and Prevention. *Biochimica et biophysica acta* **2015**, *1856*, 13–27, doi:10.1016/j.bbcan.2015.04.003.

22. Scott, A.T.; Howe, J.R. Evaluation and Management of Neuroendocrine Tumors of the Pancreas. *Surg Clin North Am* **2019**, *99*, 793–814, doi:10.1016/j.suc.2019.04.014.

23. Fang, J.M.; Shi, J. A Clinicopathologic and Molecular Update of Pancreatic Neuroendocrine Neoplasms With a Focus on the New World Health Organization Classification. *Arch Pathol Lab Med* **2019**, *143*, 1317–1326, doi:10.5858/arpa.2019-0338-RA.

24. Konukiewitz, B.; Jesinghaus, M.; Kasajima, A.; Klöppel, G. Neuroendocrine Neoplasms of the Pancreas: Diagnosis and Pitfalls. *Virchows Arch* **2022**, *480*, 247–257, doi:10.1007/s00428-021-03211-5.

25. International Agency for Research on Cancer Cancer Today Available online: http://gco.iarc.fr/today/home (accessed on 14 September 2023).

26. Yang, J.; Xu, R.; Wang, C.; Qiu, J.; Ren, B.; You, L. Early Screening and Diagnosis Strategies of Pancreatic Cancer: A Comprehensive Review. *Cancer Communications* **2021**, *41*, 1257–1274, doi:10.1002/cac2.12204.

27. Torphy, R.J.; Fujiwara, Y.; Schulick, R.D. Pancreatic Cancer Treatment: Better, but a Long Way to Go. *Surg Today* **2020**, *50*, 1117–1125, doi:10.1007/s00595-020-02028-0.

28. Leung, P.S. Overview of the Pancreas. In *The Renin-Angiotensin System: Current Research Progress in The Pancreas*; Advances in Experimental Medicine and Biology; Springer Netherlands: Dordrecht, 2010; Vol. 690, pp. 3–12 ISBN 978-90-481-9059-1.

29.   Capurso, G.; Traini, M.; Piciucchi, M.; Signoretti, M.; Arcidiacono, P.G. Exocrine Pancreatic Insufficiency: Prevalence, Diagnosis, and Management. *Clin Exp Gastroenterol* **2019**, *12*, 129–139, doi:10.2147/CEG.S168266.

30.   Leung, P.S. Physiology of the Pancreas. In *The Renin-Angiotensin System: Current Research Progress in The Pancreas*; Advances in Experimental Medicine and Biology; Springer Netherlands: Dordrecht, 2010; Vol. 690, pp. 13–27 ISBN 978-90-481-9059-1.

31.   Chen, N.; Unnikrishnan I, R.; Anjana, R.M.; Mohan, V.; Pitchumoni, C.S. The Complex Exocrine–Endocrine Relationship and Secondary Diabetes in Exocrine Pancreatic Disorders. *Journal of Clinical Gastroenterology* **2011**, *45*, 850–861, doi:10.1097/MCG.0b013e31822a2ae5.

32.   Yatchenko, Y.; Horwitz, A.; Birk, R. Endocrine and Exocrine Pancreas Pathologies Crosstalk: Insulin Regulates the Unfolded Protein Response in Pancreatic Exocrine Acinar Cells. *Experimental Cell Research* **2019**, *375*, 28–35, doi:10.1016/j.yexcr.2019.01.004.

33.   Rickels, M.R.; Norris, A.W.; Hull, R.L. A Tale of Two Pancreases: Exocrine Pathology and Endocrine Dysfunction. *Diabetologia* **2020**, *63*, 2030–2039, doi:10.1007/s00125-020-05210-8.

34.   Li, L.; Pandol, S.J. Editorial: Interaction between Endocrine and Exocrine Pancreas. *Front. Endocrinol.* **2022**, *13*, 967066, doi:10.3389/fendo.2022.967066.

35.   Li, S.; Xie, K. Ductal Metaplasia in Pancreas. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **2022**, *1877*, 188698, doi:10.1016/j.bbcan.2022.188698.

36.   Connor, A.A.; Gallinger, S. Pancreatic Cancer Evolution and Heterogeneity: Integrating Omics and Clinical Data. *Nat Rev Cancer* **2022**, *22*, 131–142, doi:10.1038/s41568-021-00418-1.

37.   Orth, M.; Metzger, P.; Gerum, S.; Mayerle, J.; Schneider, G.; Belka, C.; Schnurr, M.; Lauber, K. Pancreatic Ductal Adenocarcinoma: Biological Hallmarks, Current Status, and Future Perspectives of Combined Modality Treatment Approaches. *Radiation Oncology* **2019**, *14*, 141, doi:10.1186/s13014-019-1345-6.

38.   Nilsson, L.N.; Keane, M.G.; Shamali, A.; Millastre Bocos, J.; Marijinissen Van Zanten, M.; Antila, A.; Verdejo Gil, C.; Del Chiaro, M.; Laukkarinen, J. Nature and Management of Pancreatic Mucinous Cystic Neoplasm (MCN): A Systematic Review of the Literature. *Pancreatology* **2016**, *16*, 1028–1036, doi:10.1016/j.pan.2016.09.011.

39.   Distler, M.; Aust, D.; Weitz, J.; Pilarsky, C.; Grützmann, R. Precursor Lesions for Sporadic Pancreatic Cancer: PanIN, IPMN, and MCN. *Biomed Res Int* **2014**, *2014*, 474905, doi:10.1155/2014/474905.

40.   Sipos, B.; Frank, S.; Gress, T.; Hahn, S.; Klöppel, G. Pancreatic Intraepithelial Neoplasia Revisited and Updated. *Pancreatology* **2009**, *9*, 45–54, doi:10.1159/000178874.

41.   Gnoni, A.; Licchetta, A.; Scarpa, A.; Azzariti, A.; Brunetti, A.; Simone, G.; Nardulli, P.; Santini, D.; Aieta, M.; Delcuratolo, S.; et al. Carcinogenesis of Pancreatic Adenocarcinoma: Precursor Lesions. *International Journal of Molecular Sciences* **2013**, *14*, 19731–19762, doi:10.3390/ijms141019731.

42.   Wang, S.S.; Xu, J.; Ji, K.Y.; Hwang, C.-I. Epigenetic Alterations in Pancreatic Cancer Metastasis. *Biomolecules* **2021**, *11*, 1082, doi:10.3390/biom11081082.

43.   Waters, A.M.; Der, C.J. KRAS: The Critical Driver and Therapeutic Target for Pancreatic Cancer. *Cold Spring Harb Perspect Med* **2018**, *8*, a031435, doi:10.1101/cshperspect.a031435.

44.   Buscail, L.; Bournet, B.; Cordelier, P. Role of Oncogenic KRAS in the Diagnosis, Prognosis and Treatment of Pancreatic Cancer. *Nat Rev Gastroenterol Hepatol* **2020**, *17*, 153–168, doi:10.1038/s41575-019-0245-4.

45.   Luo, J. KRAS Mutation in Pancreatic Cancer. *Semin Oncol* **2021**, *48*, 10–18, doi:10.1053/j.seminoncol.2021.02.003.

46.   Gisselsson, D.; Jonson, T.; Petersén, Å.; Strömbeck, B.; Dal Cin, P.; Höglund, M.; Mitelman, F.; Mertens, F.; Mandahl, N. Telomere Dysfunction Triggers Extensive DNA Fragmentation and Evolution of Complex Chromosome Abnormalities in Human Malignant Tumors. *Proc Natl Acad Sci U S A* **2001**, *98*, 12683–12688, doi:10.1073/pnas.211357798.

47.   Hruban, R.H.; Maitra, A.; Goggins, M. Update on Pancreatic Intraepithelial Neoplasia. *Int J Clin Exp Pathol* **2008**, *1*, 306–316.

48.   Hosoda, W.; Chianchiano, P.; Griffin, J.F.; Pittman, M.E.; Brosens, L.A.; Noë, M.; Yu, J.; Shindo, K.; Suenaga, M.; Rezaee, N.; et al. Genetic Analyses of Isolated High-Grade Pancreatic Intraepithelial Neoplasia (HG-PanIN) Reveal Paucity of Alterations in TP53 and SMAD4. *J Pathol* **2017**, *242*, 16–23, doi:10.1002/path.4884.

49.   Jones, S.; Zhang, X.; Parsons, D.W.; Lin, J.C.-H.; Leary, R.J.; Angenendt, P.; Mankoo, P.; Carter, H.; Kamiyama, H.; Jimeno, A.; et al. Core Signaling Pathways in Human Pancreatic Cancers Revealed by Global Genomic Analyses. *Science* **2008**, *321*, 1801–1806, doi:10.1126/science.1164368.

50.   Saiki, Y.; Horii, A. Molecular Pathology of Pancreatic Cancer. *Pathology International* **2014**, *64*, 10–19, doi:10.1111/pin.12114.

51.   Waddell, N.; Pajic, M.; Patch, A.-M.; Chang, D.K.; Kassahn, K.S.; Bailey, P.; Johns, A.L.; Miller, D.; Nones, K.; Quek, K.; et al. Whole Genomes Redefine the Mutational Landscape of Pancreatic Cancer. *Nature* **2015**, *518*, 495–501, doi:10.1038/nature14169.

52.   Flowers, B.M.; Xu, H.; Mulligan, A.S.; Hanson, K.J.; Seoane, J.A.; Vogel, H.; Curtis, C.; Wood, L.D.; Attardi, L.D. Cell-of-Origin Influences Pancreatic Cancer Subtype. *Cancer Discov* **2021**, *11*, 660–677, doi:10.1158/2159-8290.CD-20-0633.

53.   Bailey, J.M.; Hendley, A.M.; Lafaro, K.J.; Pruski, M.A.; Jones, N.C.; Alsina, J.; Younes, M.; Maitra, A.; McAllister, F.; Iacobuzio-Donahue, C.A.; et al. P53 Mutations Cooperate with Oncogenic Kras to Promote Adenocarcinoma from Pancreatic Ductal Cells. *Oncogene* **2016**, *35*, 4282–4288, doi:10.1038/onc.2015.441.

54.   Lee, A.Y.L.; Dubois, C.L.; Sarai, K.; Zarei, S.; Schaeffer, D.F.; Sander, M.; Kopp, J.L. Cell of Origin Affects Tumour Development and Phenotype in Pancreatic Ductal Adenocarcinoma. *Gut* **2019**, *68*, 487–498, doi:10.1136/gutjnl-2017-314426.

55.   Ferreira, R.M.M.; Sancho, R.; Messal, H.A.; Nye, E.; Spencer-Dene, B.; Stone, R.K.; Stamp, G.; Rosewell, I.; Quaglia, A.; Behrens, A. Duct- and Acinar-Derived Pancreatic Ductal Adenocarcinomas Show Distinct Tumor Progression and Marker Expression. *Cell Rep* **2017**, *21*, 966–978, doi:10.1016/j.celrep.2017.09.093.

56.   Cros, J.; Raffenne, J.; Couvelard, A.; Poté, N. Tumor Heterogeneity in Pancreatic Adenocarcinoma. *Pathobiology* **2018**, *85*, 64–71, doi:10.1159/000477773.

57.   Peng, J.; Sun, B.-F.; Chen, C.-Y.; Zhou, J.-Y.; Chen, Y.-S.; Chen, H.; Liu, L.; Huang, D.; Jiang, J.; Cui, G.-S.; et al. Single-Cell RNA-Seq Highlights Intra-Tumoral Heterogeneity and Malignant Progression in Pancreatic Ductal Adenocarcinoma. *Cell Res* **2019**, *29*, 725–738, doi:10.1038/s41422-019-0195-y.

58.   Yachida, S.; Jones, S.; Bozic, I.; Antal, T.; Leary, R.; Fu, B.; Kamiyama, M.; Hruban, R.H.; Eshleman, J.R.; Nowak, M.A.; et al. Distant Metastasis Occurs Late during the Genetic Evolution of Pancreatic Cancer. *Nature* **2010**, *467*, 1114–1117, doi:10.1038/nature09515.

59.   Halbrook, C.J.; Lyssiotis, C.A.; Pasca di Magliano, M.; Maitra, A. Pancreatic Cancer: Advances and Challenges. *Cell* **2023**, *186*, 1729–1754, doi:10.1016/j.cell.2023.02.014.

60.   Collisson, E.A.; Bailey, P.; Chang, D.K.; Biankin, A.V. Molecular Subtypes of Pancreatic Cancer. *Nat Rev Gastroenterol Hepatol* **2019**, *16*, 207–220, doi:10.1038/s41575-019-0109-y.

61.   Han, J.; DePinho, R.A.; Maitra, A. Single-Cell RNA Sequencing in Pancreatic Cancer. *Nat Rev Gastroenterol Hepatol* **2021**, *18*, 451–452, doi:10.1038/s41575-021-00471-z.

62.   Aung, K.L.; Fischer, S.E.; Denroche, R.E.; Jang, G.-H.; Dodd, A.; Creighton, S.; Southwood, B.; Liang, S.-B.; Chadwick, D.; Zhang, A.; et al. Genomics-Driven Precision Medicine for Advanced Pancreatic Cancer: Early Results from the COMPASS Trial. *Clinical Cancer Research* **2018**, *24*, 1344–1354, doi:10.1158/1078-0432.CCR-17-2994.

63.   Zhan, H.; Zhou, B.; Cheng, Y.; Xu, J.; Wang, L.; Zhang, G.; Hu, S. Crosstalk between Stromal Cells and Cancer Cells in Pancreatic Cancer: New Insights into Stromal Biology. *Cancer Letters* **2017**, *392*, 83–93, doi:10.1016/j.canlet.2017.01.041.

64.   Velasco, R.M.; García, A.G.; Sánchez, P.J.; Sellart, I.M.; Sánchez-Arévalo Lobo, V.J. Tumour Microenvironment and Heterotypic Interactions in Pancreatic Cancer. *J Physiol Biochem* **2023**, *79*, 179–192, doi:10.1007/s13105-022-00875-8.

65.   Klein, A.P. Pancreatic Cancer Epidemiology: Understanding the Role of Lifestyle and Inherited Risk Factors. *Nat Rev Gastroenterol Hepatol* **2021**, *18*, 493–502, doi:10.1038/s41575-021-00457-x.

66.   Zhao, Z.; Liu, W. Pancreatic Cancer: A Review of Risk Factors, Diagnosis, and Treatment. *Technol Cancer Res Treat* **2020**, *19*, 1533033820962117, doi:10.1177/1533033820962117.

67.   Wada, K.; Takaori, K.; Traverso, L.W. Screening for Pancreatic Cancer. *Surgical Clinics of North America* **2015**, *95*, 1041–1052, doi:10.1016/j.suc.2015.05.010.

68.   Barone, E.; Corrado, A.; Gemignani, F.; Landi, S. Environmental Risk Factors for Pancreatic Cancer: An Update. *Arch Toxicol* **2016**, *90*, 2617–2642, doi:10.1007/s00204-016-1821-9.

69.   Schmidt-Hansen, M.; Berendse, S.; Hamilton, W. Symptoms of Pancreatic Cancer in Primary Care: A Systematic Review. *Pancreas* **2016**, *45*, 814–818, doi:10.1097/MPA.0000000000000527.

70.   Chang, V.T.; Sandifer, C.; Zhong, F. GI Symptoms in Pancreatic Cancer. *Clinical Colorectal Cancer* **2023**, *22*, 24–33, doi:10.1016/j.clcc.2022.12.002.

71.   Roche, S.P.; Kobos, R. Jaundice in the Adult Patient. **2004**, *69*.

72.   Hanada, K.; Okazaki, A.; Hirano, N.; Izumi, Y.; Teraoka, Y.; Ikemoto, J.; Kanemitsu, K.; Hino, F.; Fukuda, T.; Yonehara, S. Diagnostic Strategies for Early Pancreatic Cancer. *J Gastroenterol* **2015**, *50*, 147–154, doi:10.1007/s00535-014-1026-z.

73.   Kitano, M.; Yoshida, T.; Itonaga, M.; Tamura, T.; Hatamaru, K.; Yamashita, Y. Impact of Endoscopic Ultrasonography on Diagnosis of Pancreatic Cancer. *J Gastroenterol* **2019**, *54*, 19–32, doi:10.1007/s00535-018-1519-2.

74.   Lee, E.S.; Lee, J.M. Imaging Diagnosis of Pancreatic Cancer: A State-of-the-Art Review. *World J Gastroenterol* **2014**, *20*, 7864–7877, doi:10.3748/wjg.v20.i24.7864.

75.   Banafea, O.; Mghanga, F.P.; Zhao, J.; Zhao, R.; Zhu, L. Endoscopic Ultrasonography with Fine-Needle Aspiration for Histological Diagnosis of Solid Pancreatic Masses: A Meta-Analysis of Diagnostic Accuracy Studies. *BMC Gastroenterol* **2016**, *16*, 108, doi:10.1186/s12876-016-0519-z.

76.   Maemura, K.; Takao, S.; Shinchi, H.; Noma, H.; Mataki, Y.; Kurahara, H.; Jinnouchi, S.; Aikou, T. Role of Positron Emission Tomography in Decisions on Treatment Strategies for Pancreatic Cancer. *Journal of Hepato-Biliary-Pancreatic Surgery* **2006**, *13*, 435–441, doi:10.1007/s00534-006-1102-8.

77.   Kirwan, A.; Utratna, M.; O'Dwyer, M.E.; Joshi, L.; Kilcoyne, M. Glycosylation-Based Serum Biomarkers for Cancer Diagnostics and Prognostics. *BioMed research international* **2015**, *2015*, 490531, doi:10.1155/2015/490531.

78.   Zhao, B.; Zhao, B.; Chen, F. Diagnostic Value of Serum Carbohydrate Antigen 19-9 in Pancreatic Cancer: A Systematic Review and Meta-Analysis. *European Journal of Gastroenterology & Hepatology* **2022**, *34*, 891, doi:10.1097/MEG.0000000000002415.

79.   Planz, V.; Galgano, S.J. Percutaneous Biopsy and Drainage of the Pancreas. *Abdom Radiol (NY)* **2022**, *47*, 2584–2603, doi:10.1007/s00261-021-03244-z.

80.   Brierley, J.D.; Gospodarowicz, M.K.; Wittekind, C. *TNM Classification of Malignant Tumours*; 8th ed.; Wiley, 2016; ISBN 978-1-119-26357-9.

81.   Mahul B. Amin; Stephen B. Edge; Frederick L. Greene; David R. Byrd; Robert K. Brookland; Mary Kay Washington; Jeffrey E. Gershenwald; Carolyn C. Compton; Kenneth R. Hess; Daniel C. Sullivan; et al. *AJCC Cancer Staging Manual*; 8th ed.; Springer Cham, 2017; ISBN 978-3-319-40617-6.

82.   Shin, D.W.; Kim, J. The American Joint Committee on Cancer 8th Edition Staging System for the Pancreatic Ductal Adenocarcinoma: Is It Better than the 7th Edition? *Hepatobiliary Surg Nutr* **2020**, *9*, 98–100, doi:10.21037/hbsn.2019.08.06.

83.   *Digestive System Tumours*; Organisation mondiale de la santé, Centre international de recherche sur le cancer, Eds.; World health organization classification of tumours; 5th ed.; International agency for research on cancer: Lyon, 2019; ISBN 978-92-832-4499-8.

84.   Loveday, B.P.; Lipton, L.; Thomson, B.N. Pancreatic Cancer: An Update on Diagnosis and Management. *Aust J Gen Pract* **2019**, *48*, 826–831, doi:10.31128/AJGP-06-19-4957.

85.   Tempero, M.A.; Malafa, M.P.; Al-Hawary, M.; Behrman, S.W.; Benson, A.B.; Cardin, D.B.; Chiorean, E.G.; Chung, V.; Czito, B.; Chiaro, M.D.; et al. Pancreatic Adenocarcinoma, Version 2.2021, NCCN Clinical Practice Guidelines in Oncology. *Journal of the National Comprehensive Cancer Network* **2021**, *19*, 439–457, doi:10.6004/jnccn.2021.0017.

86.   McGuigan, A.; Kelly, P.; Turkington, R.C.; Jones, C.; Coleman, H.G.; McCain, R.S. Pancreatic Cancer: A Review of Clinical Diagnosis, Epidemiology, Treatment and Outcomes. *World J Gastroenterol* **2018**, *24*, 4846–4861, doi:10.3748/wjg.v24.i43.4846.

87.   Jiang, Y.-L.; Zhang, R.-C.; Zhou, Y.-C. Comparison of Overall Survival and Perioperative Outcomes of Laparoscopic Pancreaticoduodenectomy and Open Pancreaticoduodenectomy for Pancreatic Ductal Adenocarcinoma: A Systematic Review and Meta-Analysis. *BMC Cancer* **2019**, *19*, 781, doi:10.1186/s12885-019-6001-x.

88.   Simon, R. Complications After Pancreaticoduodenectomy. *Surgical Clinics of North America* **2021**, *101*, 865–874, doi:10.1016/j.suc.2021.06.011.

89.   Del Chiaro, M.; Rangelova, E.; Segersvärd, R.; Arnelo, U. Are There Still Indications for Total Pancreatectomy? *Updates Surg* **2016**, *68*, 257–263, doi:10.1007/s13304-016-0388-6.

90. Riviere, D.; Gurusamy, K.S.; Kooby, D.A.; Vollmer, C.M.; Besselink, M.G.; Davidson, B.R.; van Laarhoven, C.J. Laparoscopic versus Open Distal Pancreatectomy for Pancreatic Cancer. *Cochrane Database Syst Rev* **2016**, *2016*, CD011391, doi:10.1002/14651858.CD011391.pub2.

91. Lof, S.; Van Der Heijde, N.; Abuawwad, M.; Al-Sarireh, B.; Boggi, U.; Butturini, G.; Capretti, G.; Coratti, A.; Casadei, R.; D'Hondt, M.; et al. Robotic *versus* Laparoscopic Distal Pancreatectomy: Multicentre Analysis. *British Journal of Surgery* **2021**, *108*, 188–195, doi:10.1093/bjs/znaa039.

92. Schneider, M.; Hackert, T.; Strobel, O.; Büchler, M.W. Technical Advances in Surgery for Pancreatic Cancer. *British Journal of Surgery* **2021**, *108*, 777–785, doi:10.1093/bjs/znab133.

93. Zhang, B.; Zhou, F.; Hong, J.; Ng, D.M.; Yang, T.; Zhou, X.; Jin, J.; Zhou, F.; Chen, P.; Xu, Y. The Role of FOLFIRINOX in Metastatic Pancreatic Cancer: A Meta-Analysis. *World J Surg Oncol* **2021**, *19*, 182, doi:10.1186/s12957-021-02291-6.

94. Burris, H.A.; Moore, M.J.; Andersen, J.; Green, M.R.; Rothenberg, M.L.; Modiano, M.R.; Cripps, M.C.; Portenoy, R.K.; Storniolo, A.M.; Tarassoff, P.; et al. Improvements in Survival and Clinical Benefit with Gemcitabine as First-Line Therapy for Patients with Advanced Pancreas Cancer: A Randomized Trial. *J Clin Oncol* **1997**, *15*, 2403–2413, doi:10.1200/JCO.1997.15.6.2403.

95. Rocha Lima, C.M.; Green, M.R.; Rotche, R.; Miller, W.H.; Jeffrey, G.M.; Cisar, L.A.; Morganti, A.; Orlando, N.; Gruia, G.; Miller, L.L. Irinotecan Plus Gemcitabine Results in No Survival Advantage Compared With Gemcitabine Monotherapy in Patients With Locally Advanced or Metastatic Pancreatic Cancer Despite Increased Tumor Response Rate. *JCO* **2004**, *22*, 3776–3783, doi:10.1200/JCO.2004.12.082.

96. Louvet, C.; Labianca, R.; Hammel, P.; Lledo, G.; Zampino, M.G.; André, T.; Zaniboni, A.; Ducreux, M.; Aitini, E.; Taïeb, J.; et al. Gemcitabine in Combination With Oxaliplatin Compared With Gemcitabine Alone in Locally Advanced or Metastatic Pancreatic Cancer: Results of a GERCOR and GISCAD Phase III Trial. *JCO* **2005**, *23*, 3509–3516, doi:10.1200/JCO.2005.06.023.

97. Heinemann, V.; Quietzsch, D.; Gieseler, F.; Gonnermann, M.; Schönekäs, H.; Rost, A.; Neuhaus, H.; Haag, C.; Clemens, M.; Heinrich, B.; et al. Randomized Phase III Trial of Gemcitabine Plus Cisplatin Compared With Gemcitabine Alone in Advanced Pancreatic Cancer. *JCO* **2006**, *24*, 3946–3952, doi:10.1200/JCO.2005.05.1490.

98. Ciliberto, D.; Botta, C.; Correale, P.; Rossi, M.; Caraglia, M.; Tassone, P.; Tagliaferri, P. Role of Gemcitabine-Based Combination Therapy in the Management of Advanced Pancreatic Cancer: A Meta-Analysis of Randomised Trials. *European Journal of Cancer* **2013**, *49*, 593–603, doi:10.1016/j.ejca.2012.08.019.

99.   Von Hoff, D.D.; Ervin, T.; Arena, F.P.; Chiorean, E.G.; Infante, J.; Moore, M.; Seay, T.; Tjulandin, S.A.; Ma, W.W.; Saleh, M.N.; et al. Increased Survival in Pancreatic Cancer with Nab-Paclitaxel plus Gemcitabine. *N Engl J Med* **2013**, *369*, 1691–1703, doi:10.1056/NEJMoa1304369.

100.   Conroy, T.; Desseigne, F.; Ychou, M.; Bouché, O.; Guimbaud, R.; Bécouarn, Y.; Adenis, A.; Raoul, J.-L.; Gourgou-Bourgade, S.; de la Fouchardière, C.; et al. FOLFIRINOX versus Gemcitabine for Metastatic Pancreatic Cancer. *N Engl J Med* **2011**, *364*, 1817–1825, doi:10.1056/NEJMoa1011923.

101.   Klein-Brill, A.; Amar-Farkash, S.; Lawrence, G.; Collisson, E.A.; Aran, D. Comparison of FOLFIRINOX vs Gemcitabine Plus Nab-Paclitaxel as First-Line Chemotherapy for Metastatic Pancreatic Ductal Adenocarcinoma. *JAMA Netw Open* **2022**, *5*, e2216199, doi:10.1001/jamanetworkopen.2022.16199.

102.   Mahaseth, H.; Brutcher, E.; Kauh, J.; Hawk, N.; Kim, S.; Chen, Z.; Kooby, D.A.; Maithel, S.K.; Landry, J.; El-Rayes, B.F. Modified FOLFIRINOX Regimen With Improved Safety and Maintained Efficacy in Pancreatic Adenocarcinoma. *Pancreas* **2013**, *42*, 1311–1315, doi:10.1097/MPA.0b013e31829e2006.

103.   Ozaka, M.; Ishii, H.; Sato, T.; Ueno, M.; Ikeda, M.; Uesugi, K.; Sata, N.; Miyashita, K.; Mizuno, N.; Tsuji, K.; et al. A Phase II Study of Modified FOLFIRINOX for Chemotherapy-Naïve Patients with Metastatic Pancreatic Cancer. *Cancer Chemother Pharmacol* **2018**, *81*, 1017–1023, doi:10.1007/s00280-018-3577-9.

104.   Conroy, T.; Hammel, P.; Hebbar, M.; Ben Abdelghani, M.; Wei, A.C.; Raoul, J.-L.; Choné, L.; Francois, E.; Artru, P.; Biagi, J.J.; et al. FOLFIRINOX or Gemcitabine as Adjuvant Therapy for Pancreatic Cancer. *N Engl J Med* **2018**, *379*, 2395–2406, doi:10.1056/NEJMoa1809775.

105.   Yang, F.; Jin, C.; Fu, D.-L.; Warshaw, A.L. Modified FOLFIRINOX for Resected Pancreatic Cancer: Opportunities and Challenges. *World J Gastroenterol* **2019**, *25*, 2839–2845, doi:10.3748/wjg.v25.i23.2839.

106.   Tempero, M.A.; Pelzer, U.; O'Reilly, E.M.; Winter, J.; Oh, D.-Y.; Li, C.-P.; Tortora, G.; Chang, H.-M.; Lopez, C.D.; Bekaii-Saab, T.; et al. Adjuvant Nab-Paclitaxel + Gemcitabine in Resected Pancreatic Ductal Adenocarcinoma: Results From a Randomized, Open-Label, Phase III Trial. *J Clin Oncol* **2023**, *41*, 2007–2019, doi:10.1200/JCO.22.01134.

107.   Springfeld, C.; Ferrone, C.R.; Katz, M.H.G.; Philip, P.A.; Hong, T.S.; Hackert, T.; Büchler, M.W.; Neoptolemos, J. Neoadjuvant Therapy for Pancreatic Cancer. *Nat Rev Clin Oncol* **2023**, *20*, 318–337, doi:10.1038/s41571-023-00746-1.

108.   Borazanci, E.; Sckolnik, S.; Amini, A. Neo-Adjuvant Therapy for Pancreatic Cancer: Hope for the Future. *Expert Review of Gastroenterology & Hepatology* **2019**, *13*, 579–589, doi:10.1080/17474124.2019.1607294.

109. Temraz, S.; Nassar, F.; Hammoud, M.S.; Mukherji, D.; O'Reilly, E.M.; Dbouk, H.; Farhat, F.; Charafeddine, M.; Faraj, W.; Khalifeh, M.J.; et al. Neo-Adjuvant FOLFIRINOX in Borderline Resectable and Locally Advanced Pancreatic Adenocarcinoma. *Asia-Pacific Journal of Clinical Oncology* **2022**, *18*, 735–742, doi:10.1111/ajco.13775.

110. Seufferlein, T.; Uhl, W.; Kornmann, M.; Algül, H.; Friess, H.; König, A.; Ghadimi, M.; Gallmeier, E.; Bartsch, D.K.; Lutz, M.P.; et al. Perioperative or Only Adjuvant Gemcitabine plus Nab-Paclitaxel for Resectable Pancreatic Cancer (NEONAX)—a Randomized Phase II Trial of the AIO Pancreatic Cancer Group. *Annals of Oncology* **2023**, *34*, 91–100, doi:10.1016/j.annonc.2022.09.161.

111. Qian, Y.; Gong, Y.; Fan, Z.; Luo, G.; Huang, Q.; Deng, S.; Cheng, H.; Jin, K.; Ni, Q.; Yu, X.; et al. Molecular Alterations and Targeted Therapy in Pancreatic Ductal Adenocarcinoma. *J Hematol Oncol* **2020**, *13*, 130, doi:10.1186/s13045-020-00958-3.

112. Kolbeinsson, H.M.; Chandana, S.; Wright, G.P.; Chung, M. Pancreatic Cancer: A Review of Current Treatment and Novel Therapies. *Journal of Investigative Surgery* **2023**, *36*, 2129884, doi:10.1080/08941939.2022.2129884.

113. Adamska, A.; Domenichini, A.; Falasca, M. Pancreatic Ductal Adenocarcinoma: Current and Evolving Therapies. *Int J Mol Sci* **2017**, *18*, 1338, doi:10.3390/ijms18071338.

114. Liu, X.; Li, Z.; Wang, Y. Advances in Targeted Therapy and Immunotherapy for Pancreatic Cancer. *Advanced Biology* **2021**, *5*, 1900236, doi:10.1002/adbi.201900236.

115. Bear, A.S.; Vonderheide, R.H.; O'Hara, M.H. Challenges and Opportunities for Pancreatic Cancer Immunotherapy. *Cancer Cell* **2020**, *38*, 788–802, doi:10.1016/j.ccell.2020.08.004.

116. Andrikou, K.; Peterle, C.; Pipitone, S.; Salati, M.; Cascinu, S. Emerging Antibodies for the Treatment of Pancreatic Cancer. *Expert Opinion on Emerging Drugs* **2017**, *22*, 39–51, doi:10.1080/14728214.2017.1293649.

117. O'Reilly, E.M.; Oh, D.-Y.; Dhani, N.; Renouf, D.J.; Lee, M.A.; Sun, W.; Fisher, G.; Hezel, A.; Chang, S.-C.; Vlahovic, G.; et al. Durvalumab With or Without Tremelimumab for Patients With Metastatic Pancreatic Ductal Adenocarcinoma. *JAMA Oncol* **2019**, *5*, 1431–1438, doi:10.1001/jamaoncol.2019.1588.

118. Padrón, L.J.; Maurer, D.M.; O'Hara, M.H.; O'Reilly, E.M.; Wolff, R.A.; Wainberg, Z.A.; Ko, A.H.; Fisher, G.; Rahma, O.; Lyman, J.P.; et al. Sotigalimab and/or Nivolumab with Chemotherapy in First-Line Metastatic Pancreatic Cancer: Clinical and Immunologic Analyses from the Randomized Phase 2 PRINCE Trial. *Nat Med* **2022**, *28*, 1167–1177, doi:10.1038/s41591-022-01829-9.

119. Cardillo, T.M.; Govindan, S.V.; Sharkey, R.M.; Trisal, P.; Arrojo, R.; Liu, D.; Rossi, E.A.; Chang, C.-H.; Goldenberg, D.M. Sacituzumab Govitecan (IMMU-132), an Anti-Trop-2/SN-38 Antibody–Drug Conjugate: Characterization and Efficacy in Pancreatic, Gastric, and Other Cancers. *Bioconjugate Chem.* **2015**, *26*, 919–931, doi:10.1021/acs.bioconjchem.5b00223.

120. Challita-Eid, P.M.; Satpayev, D.; Yang, P.; An, Z.; Morrison, K.; Shostak, Y.; Raitano, A.; Nadell, R.; Liu, W.; Lortie, D.R.; et al. Enfortumab Vedotin Antibody–Drug Conjugate Targeting Nectin-4 Is a Highly Potent Therapeutic Agent in Multiple Preclinical Cancer Models. *Cancer Research* **2016**, *76*, 3003–3013, doi:10.1158/0008-5472.CAN-15-1313.

121. Jin, Y.; Zhang, Z.; Zou, S.; Li, F.; Chen, H.; Peng, C.; Deng, X.; Wen, C.; Shen, B.; Zhan, Q. A Novel C-MET-Targeting Antibody-Drug Conjugate for Pancreatic Cancer. *Front Oncol* **2021**, *11*, 634881, doi:10.3389/fonc.2021.634881.

122. Golfier, S.; Kopitz, C.; Kahnert, A.; Heisler, I.; Schatz, C.A.; Stelte-Ludwig, B.; Mayer-Bartschmid, A.; Unterschemmann, K.; Bruder, S.; Linden, L.; et al. Anetumab Ravtansine: A Novel Mesothelin-Targeting Antibody–Drug Conjugate Cures Tumors with Heterogeneous Target Expression Favored by Bystander Effect. *Mol Cancer Ther* **2014**, *13*, 1537–1548, doi:10.1158/1535-7163.MCT-13-0926.

123. Guerrero-Ochoa, P.; Ibáñez-Pérez, R.; Berbegal-Pinilla, G.; Aguilar, D.; Marzo, I.; Corzana, F.; Minjárez-Sáenz, M.; Macías-León, J.; Conde, B.; Raso, J.; et al. Preclinical Studies of Granulysin-Based Anti-MUC1-Tn Immunotoxins as a New Antitumoral Treatment. *Biomedicines* **2022**, *10*, 1223, doi:10.3390/biomedicines10061223.

124. D\elm, C.; Tano, Z.E.; Varghese, A.M.; Adusumilli, P.S. CAR T-Cell Therapy for Pancreatic Cancer. *J Surg Oncol* **2017**, *116*, 63–74, doi:10.1002/jso.24627.

125. Rojas, L.A.; Sethna, Z.; Soares, K.C.; Olcese, C.; Pang, N.; Patterson, E.; Lihm, J.; Ceglia, N.; Guasp, P.; Chu, A.; et al. Personalized RNA Neoantigen Vaccines Stimulate T Cells in Pancreatic Cancer. *Nature* **2023**, *618*, 144–150, doi:10.1038/s41586-023-06063-y.

126. *Essentials of Glycobiology*; Varki, A., Cummings, R.D., Esko, J.D., Stanley, P., Hart, G.W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N.H., Prestegard, J.H., Schnaar, R.L., Seeberger, P.H., Eds.; 4th ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2022; ISBN 978-1-62182-421-3.

127. Roach, P.J.; Depaoli-Roach, A.A.; Hurley, T.D.; Tagliabracci, V.S. Glycogen and Its Metabolism: Some New Developments and Old Themes. *Biochem J* **2012**, *441*, 763–787, doi:10.1042/BJ20111416.

128. Schjoldager, K.T.; Narimatsu, Y.; Joshi, H.J.; Clausen, H. Global View of Human Protein Glycosylation Pathways and Functions. *Nat Rev Mol Cell Biol* **2020**, *21*, 729–749, doi:10.1038/s41580-020-00294-x.

129. Datta, A. Comparative Sequence Analysis in the Sialyltransferase Protein Family: Analysis of Motifs. *CDT* **2009**, *10*, 483–498, doi:10.2174/138945009788488422.

130. Mehboob, M.Z.; Lang, M. Structure, Function, and Pathology of Protein O-Glucosyltransferases. *Cell Death Dis* **2021**, *12*, 1–13, doi:10.1038/s41419-020-03314-y.

131. Chandel, N.S. Carbohydrate Metabolism. *Cold Spring Harb Perspect Biol* **2021**, *13*, a040568, doi:10.1101/cshperspect.a040568.

132.   Varki, A. Biological Roles of Glycans. *Glycobiology* **2017**, *27*, 3–49, doi:10.1093/glycob/cww086.

133.   Hart, G.W.; Copeland, R.J. Glycomics Hits the Big Time. *Cell* **2010**, *143*, 672–676, doi:10.1016/j.cell.2010.11.008.

134.   Illiano, A.; Pinto, G.; Melchiorre, C.; Carpentieri, A.; Faraco, V.; Amoresano, A. Protein Glycosylation Investigated by Mass Spectrometry: An Overview. *Cells* **2020**, *9*, 1986, doi:10.3390/cells9091986.

135.   Stowell, S.R.; Ju, T.; Cummings, R.D. Protein Glycosylation in Cancer. *Annu Rev Pathol* **2015**, *10*, 473–510, doi:10.1146/annurev-pathol-012414-040438.

136.   Pinho, S.S.; Reis, C.A. Glycosylation in Cancer: Mechanisms and Clinical Implications. *Nature Reviews Cancer* **2015**, *15*, 540–555, doi:10.1038/nrc3982.

137.   Oliveira-Ferrer, L.; Legler, K.; Milde-Langosch, K. Role of Protein Glycosylation in Cancer Metastasis. *Seminars in Cancer Biology* **2017**, *44*, 141–152, doi:10.1016/j.semcancer.2017.03.002.

138.   Schedin-Weiss, S.; Winblad, B.; Tjernberg, L.O. The Role of Protein Glycosylation in Alzheimer Disease. *The FEBS Journal* **2014**, *281*, 46–62, doi:10.1111/febs.12590.

139.   Haukedal, H.; Freude, K.K. Implications of Glycosylation in Alzheimer's Disease. *Front Neurosci* **2021**, *14*, 625348, doi:10.3389/fnins.2020.625348.

140.   Mueller, A.-L.; Payandeh, Z.; Mohammadkhani, N.; Mubarak, S.M.H.; Zakeri, A.; Alagheband Bahrami, A.; Brockmueller, A.; Shakibaei, M. Recent Advances in Understanding the Pathogenesis of Rheumatoid Arthritis: New Treatment Strategies. *Cells* **2021**, *10*, 3017, doi:10.3390/cells10113017.

141.   Zhou, X.; Motta, F.; Selmi, C.; Ridgway, W.M.; Gershwin, M.E.; Zhang, W. Antibody Glycosylation in Autoimmune Diseases. *Autoimmun Rev* **2021**, *20*, 102804, doi:10.1016/j.autrev.2021.102804.

142.   Wang, G.; Yuan, J.; Luo, J.; Ocansey, D.K.W.; Zhang, X.; Qian, H.; Xu, W.; Mao, F. Emerging Role of Protein Modification in Inflammatory Bowel Disease. *J Zhejiang Univ Sci B* **2022**, *23*, 173–188, doi:10.1631/jzus.B2100114.

143.   Bhat, A.H.; Maity, S.; Giri, K.; Ambatipudi, K. Protein Glycosylation: Sweet or Bitter for Bacterial Pathogens? *Critical Reviews in Microbiology* **2019**, *45*, 82–102, doi:10.1080/1040841X.2018.1547681.

144.   Feng, T.; Zhang, J.; Chen, Z.; Pan, W.; Chen, Z.; Yan, Y.; Dai, J. Glycosylation of Viral Proteins: Implication in Virus–Host Interaction and Virulence. *Virulence* **2022**, *13*, 670–683, doi:10.1080/21505594.2022.2060464.

145.   Thomas, D.; Rathinavel, A.K.; Radhakrishnan, P. Altered Glycosylation in Cancer: A Promising Target for Biomarkers and Therapeutics. *Biochim Biophys Acta Rev Cancer* **2021**, *1875*, 188464, doi:10.1016/j.bbcan.2020.188464.

146. Ramazi, S.; Zahiri, J. Post-Translational Modifications in Proteins: Resources, Tools and Prediction Methods. *Database* **2021**, *2021*, baab012, doi:10.1093/database/baab012.

147. Esmail, S.; Manolson, M.F. Advances in Understanding N-Glycosylation Structure, Function, and Regulation in Health and Disease. *European Journal of Cell Biology* **2021**, *100*, 151186, doi:10.1016/j.ejcb.2021.151186.

148. Petrescu, A.-J.; Milac, A.-L.; Petrescu, S.M.; Dwek, R.A.; Wormald, M.R. Statistical Analysis of the Protein Environment of N-Glycosylation Sites: Implications for Occupancy, Structure, and Folding. *Glycobiology* **2004**, *14*, 103–114, doi:10.1093/glycob/cwh008.

149. Varki, A.; Cummings, R.D.; Aebi, M.; Packer, N.H.; Seeberger, P.H.; Esko, J.D.; Stanley, P.; Hart, G.; Darvill, A.; Kinoshita, T.; et al. Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* **2015**, *25*, 1323–1324, doi:10.1093/glycob/cwv091.

150. de las Rivas, M.; Lira-Navarrete, E.; Gerken, T.A.; Hurtado-Guerrero, R. Polypeptide GalNAc-Ts: From Redundancy to Specificity. *Curr Opin Struct Biol* **2019**, *56*, 87–96, doi:10.1016/j.sbi.2018.12.007.

151. Bennett, E.P.; Mandel, U.; Clausen, H.; Gerken, T.A.; Fritz, T.A.; Tabak, L.A. Control of Mucin-Type O-Glycosylation: A Classification of the Polypeptide GalNAc-Transferase Gene Family. *Glycobiology* **2012**, *22*, 736–756, doi:10.1093/glycob/cwr182.

152. Daniel, E.J.P.; Las Rivas, M.; Lira-Navarrete, E.; García-García, A.; Hurtado-Guerrero, R.; Clausen, H.; Gerken, T.A. Ser and Thr Acceptor Preferences of the GalNAc-Ts Vary among Isoenzymes to Modulate Mucin-Type O-Glycosylation. *Glycobiology* **2020**, *30*, 910–922, doi:10.1093/glycob/cwaa036.

153. Wilkinson, H.; Saldova, R. Current Methods for the Characterization of *O*-Glycans. *J. Proteome Res.* **2020**, *19*, 3890–3905, doi:10.1021/acs.jproteome.0c00435.

154. Dotz, V.; Wuhrer, M. Histo-Blood Group Glycans in the Context of Personalized Medicine. *Biochim Biophys Acta Gen Subj* **2016**, *1860*, 1596–1607, doi:10.1016/j.bbagen.2015.12.026.

155. Natoni, A.; Macauley, M.S.; O'Dwyer, M.E. Targeting Selectins and Their Ligands in Cancer. *Front Oncol* **2016**, *6*, 93, doi:10.3389/fonc.2016.00093.

156. Peracaula, R.; Barrabés, S.; Sarrats, A.; Rudd, P.M.; Llorens, R.D. Altered Glycosylation in Tumours Focused to Cancer Diagnosis. *Disease Markers* **2008**, *25*, 207–218, doi:10.1155/2008/797629.

157. Häuselmann, I.; Borsig, L. Altered Tumor-Cell Glycosylation Promotes Metastasis. *Front Oncol* **2014**, *4*, 28, doi:10.3389/fonc.2014.00028.

158. Haga, Y.; Ueda, K. Glycosylation in Cancer: Its Application as a Biomarker and Recent Advances of Analytical Techniques. *Glycoconj J* **2022**, *39*, 303–313, doi:10.1007/s10719-022-10043-1.

159.   Freire-de-lima, L. Sweet and Sour : The Impact of Differential Glycosylation in Cancer Cells Undergoing Epithelial – Mesenchymal Transition. *Frontiers in oncology* **2014**, *4*, 1–10, doi:10.3389/fonc.2014.00059.

160.   Munkley, J.; Elliott, D.J. Hallmarks of Glycosylation in Cancer. *Oncotarget* **2016**, *7*, 35478–35489, doi:10.18632/oncotarget.8155.

161.   Fu, C.; Zhao, H.; Wang, Y.; Cai, H.; Xiao, Y.; Zeng, Y.; Chen, H. Tumor-Associated Antigens: Tn Antigen, sTn Antigen, and T Antigen. *HLA* **2016**, *88*, 275–286, doi:10.1111/tan.12900.

162.   Ju, T.; Otto, V.I.; Cummings, R.D. The Tn Antigen—Structural Simplicity and Biological Complexity. *Angew Chem Int Ed Engl* **2011**, *50*, 1770–1791, doi:10.1002/anie.201002313.

163.   Thomas, D.; Sagar, S.; Caffrey, T.; Grandgenett, P.M.; Radhakrishnan, P. Truncated O-glycans Promote Epithelial-to-mesenchymal Transition and Stemness Properties of Pancreatic Cancer Cells. *J Cell Mol Med* **2019**, *23*, 6885–6896, doi:10.1111/jcmm.14572.

164.   Yu, L.-G. The Oncofetal Thomsen–Friedenreich Carbohydrate Antigen in Cancer Progression. *Glycoconj J* **2007**, *24*, 411–420, doi:10.1007/s10719-007-9034-3.

165.   de-Souza-Ferreira, M.; Ferreira, É.E.; de-Freitas-Junior, J.C.M. Aberrant N-Glycosylation in Cancer: MGAT5 and B1,6-GlcNAc Branched N-Glycans as Critical Regulators of Tumor Development and Progression. *Cell Oncol.* **2023**, *46*, 481–501, doi:10.1007/s13402-023-00770-4.

166.   Pucci, M.; Malagolini, N.; Dall'Olio, F. Glycobiology of the Epithelial to Mesenchymal Transition. *Biomedicines* **2021**, *9*, 770, doi:10.3390/biomedicines9070770.

167.   Thiemann, S.; Baum, L.G. Galectins and Immune Responses—Just How Do They Do Those Things They Do? *Annu. Rev. Immunol.* **2016**, *34*, 243–264, doi:10.1146/annurev-immunol-041015-055402.

168.   Srinivasan, N.; Bane, S.M.; Ahire, S.D.; Ingle, A.D.; Kalraiya, R.D. Poly N-Acetyllactosamine Substitutions on N- and Not O-Oligosaccharides or Thomsen–Friedenreich Antigen Facilitate Lung Specific Metastasis of Melanoma Cells via Galectin-3. *Glycoconj J* **2009**, *26*, 445–456, doi:10.1007/s10719-008-9194-9.

169.   Zhou, D. Why Are Glycoproteins Modified by Poly-N-Acetyllactosamine Glycoconjugates? *CPPS* **2003**, *4*, 1–9, doi:10.2174/1389203033380304.

170.   Shao, K.; Chen, Z.Y.; Gautam, S.; Deng, N.H.; Zhou, Y.; Wu, X.Z. Posttranslational Modification of E-Cadherin by Core Fucosylation Regulates Src Activation and Induces Epithelial–Mesenchymal Transition-like Process in Lung Cancer Cells. *Glycobiology* **2016**, *26*, 142–154, doi:10.1093/glycob/cwv089.

171.   Zhang, Z.; Wuhrer, M.; Holst, S. Serum Sialylation Changes in Cancer. *Glycoconj J* **2018**, *35*, 139–160, doi:10.1007/s10719-018-9820-0.

172.   Wu, X.; Zhao, J.; Ruan, Y.; Sun, L.; Xu, C.; Jiang, H. Sialyltransferase ST3GAL1 Promotes Cell Migration, Invasion, and TGF-B1-Induced EMT and Confers Paclitaxel Resistance in Ovarian Cancer. *Cell Death Dis* **2018**, *9*, 1102, doi:10.1038/s41419-018-1101-0.

173.   Dorsett, K.A.; Marciel, M.P.; Hwang, J.; Ankenbauer, K.E.; Bhalerao, N.; Bellis, S.L. Regulation of ST6GAL1 Sialyltransferase Expression in Cancer Cells. *Glycobiology* **2021**, *31*, 530–539, doi:10.1093/glycob/cwaa110.

174.   Venturi, G.; Gomes Ferreira, I.; Pucci, M.; Ferracin, M.; Malagolini, N.; Chiricolo, M.; Dall'Olio, F. Impact of Sialyltransferase ST6GAL1 Overexpression on Different Colon Cancer Cell Types. *Glycobiology* **2019**, *29*, 684–695, doi:10.1093/glycob/cwz053.

175.   Miyagi, T.; Takahashi, K.; Hata, K.; Shiozaki, K.; Yamaguchi, K. Sialidase Significance for Cancer Progression. *Glycoconj J* **2012**, *29*, 567–577, doi:10.1007/s10719-012-9394-1.

176.   Glanz, V.Yu.; Myasoedova, V.A.; Grechko, A.V.; Orekhov, A.N. Sialidase Activity in Human Pathologies. *European Journal of Pharmacology* **2019**, *842*, 345–350, doi:10.1016/j.ejphar.2018.11.014.

177.   Dobie, C.; Skropeta, D. Insights into the Role of Sialylation in Cancer Progression and Metastasis. *Br J Cancer* **2021**, *124*, 76–90, doi:10.1038/s41416-020-01126-7.

178.   Pearce, O.M.T.; Läubli, H. Sialic Acids in Cancer Biology and Immunity. *Glycobiology* **2016**, *26*, 111–128, doi:10.1093/glycob/cwv097.

179.   van Houtum, E.J.H.; Büll, C.; Cornelissen, L.A.M.; Adema, G.J. Siglec Signaling in the Tumor Microenvironment. *Front Immunol* **2021**, *12*, 790317, doi:10.3389/fimmu.2021.790317.

180.   Läubli, H.; Borsig, L. Altered Cell Adhesion and Glycosylation Promote Cancer Immune Suppression and Metastasis. *Front Immunol* **2019**, *10*, 2120, doi:10.3389/fimmu.2019.02120.

181.   Julien, S.; Ivetic, A.; Grigoriadis, A.; QiZe, D.; Burford, B.; Sproviero, D.; Picco, G.; Gillett, C.; Papp, S.L.; Schaffer, L.; et al. Selectin Ligand Sialyl-Lewis x Antigen Drives Metastasis of Hormone-Dependent Breast Cancers. *Cancer Res* **2011**, *71*, 7683–7693, doi:10.1158/0008-5472.CAN-11-1139.

182.   Zhuo, Y.; Bellis, S.L. Emerging Role of A2,6-Sialic Acid as a Negative Regulator of Galectin Binding and Function. *J Biol Chem* **2011**, *286*, 5935–5941, doi:10.1074/jbc.R110.191429.

183.   Huang, Y.; Wang, H.-C.; Zhao, J.; Wu, M.-H.; Shih, T.-C. Immunosuppressive Roles of Galectin-1 in the Tumor Microenvironment. *Biomolecules* **2021**, *11*, 1398, doi:10.3390/biom11101398.

184.   Croci, D.O.; Cerliani, J.P.; Dalotto-Moreno, T.; Méndez-Huergo, S.P.; Mascan-froni, I.D.; Dergan-Dylon, S.; Toscano, M.A.; Caramelo, J.J.; García-Vallejo, J.J.; Ouyang, J.; et al. Glycosylation-Dependent Lectin-Receptor Interactions Preserve Angiogenesis in Anti-VEGF Refractory Tumors. *Cell* **2014**, *156*, 744–758, doi:10.1016/j.cell.2014.01.043.

185.   Nakagoe, T.; Fukushima, K.; Nanashima, A.; Sawai, T.; Tsuji, T.; Jibiki, M.; Yamaguchi, H.; Yasutake, T.; Ayabe, H.; Matuo, T.; et al. Sialyl Lewisx Antigens as Prognostic Factors in Patients with Colorectal Cancer. **2000**, *14*.

186.   Cohen, E.N.; Fouad, T.M.; Lee, B.-N.; Arun, B.K.; Liu, D.; Tin, S.; Gutierrez Barrera, A.M.; Miura, T.; Kiyokawa, I.; Yamashita, J.; et al. Elevated Serum Levels of Sialyl Lewis X (sLeX) and Inflammatory Mediators in Patients with Breast Cancer. *Breast Cancer Res Treat* **2019**, *176*, 545–556, doi:10.1007/s10549-019-05258-0.

187.   Kolben, T.; Müller, L.; Meister, S.; Keilmann, L.; Buschmann, C.; Trillsch, F.; Burges, A.; Czogalla, B.; Mitter, S.; Schmoeckel, E.; et al. Blood Group Antigens SLeX, SLeA, and LeY as Prognostic Markers in Endometrial Cancer. *J Cancer Res Clin Oncol* **2022**, *148*, 3323–3335, doi:10.1007/s00432-022-04098-8.

188.   Elola, M.T.; Capurro, M.I.; Barrio, M.M.; Coombs, P.J.; Taylor, M.E.; Drickamer, K.; Mordoh, J. Lewis x Antigen Mediates Adhesion of Human Breast Carcinoma Cells to Activated Endothelium. Possible Involvement of the Endothelial Scavenger Receptor C-Type Lectin. *Breast Cancer Res Treat* **2007**, *101*, 161–174, doi:10.1007/s10549-006-9286-9.

189.   Liu, J.; Zheng, M.; Qi, Y.; Wang, H.; Liu, M.; Liu, Q.; Lin, B. Lewis(y) Antigen-Mediated Positive Feedback Loop Induces and Promotes Chemotherapeutic Resistance in Ovarian Cancer. *Int J Oncol* **2018**, *53*, 1774–1786, doi:10.3892/ijo.2018.4496.

190.   Hou, R.; Jiang, L.; Liu, D.; Lin, B.; Hu, Z.; Gao, J.; Zhang, D.; Zhang, S.; Iwamori, M. Lewis(y) Antigen Promotes the Progression of Epithelial Ovarian Cancer by Stimulating MUC1 Expression. *Int J Mol Med* **2017**, *40*, 293–302, doi:10.3892/ijmm.2017.3009.

191.   Ley, K.; Laudanna, C.; Cybulsky, M.I.; Nourshargh, S. Getting to the Site of Inflammation: The Leukocyte Adhesion Cascade Updated. *Nat Rev Immunol* **2007**, *7*, 678–689, doi:10.1038/nri2156.

192.   Scarà, S.; Bottoni, P.; Scatena, R. CA 19-9: Biochemical and Clinical Aspects. In *Advances in Cancer Biomarkers*; Scatena, R., Ed.; Advances in Experimental Medicine and Biology; Springer Netherlands: Dordrecht, 2015; Vol. 867, pp. 247–260 ISBN 978-94-017-7214-3.

193.   Munkley, J. The Glycosylation Landscape of Pancreatic Cancer (Review). *Oncology Letters* **2019**, *17*, 2569–2575, doi:10.3892/ol.2019.9885.

194.   Gupta, R.; Leon, F.; Thompson, C.M.; Nimmakayala, R.; Karmakar, S.; Nalla-samy, P.; Chugh, S.; Prajapati, D.R.; Rachagani, S.; Kumar, S.; et al. Global Analysis of Human Glycosyltransferases Reveals Novel Targets for Pancreatic Cancer Pathogenesis. *Br J Cancer* **2020**, *122*, 1661–1672, doi:10.1038/s41416-020-0772-3.

195.   Lumibao, J.C.; Tremblay, J.R.; Hsu, J.; Engle, D.D. Altered Glycosylation in Pancreatic Cancer and Beyond. *J Exp Med* **2022**, *219*, e20211505, doi:10.1084/jem.20211505.

196.   Pérez-Garay, M.; Arteta, B.; Pagés, L.; Llorens, R. de; Bolós, C. de; Vidal-Vana-clocha, F.; Peracaula, R. A2,3-Sialyltransferase ST3Gal III Modulates Pancreatic Cancer Cell Motility and Adhesion in Vitro and Enhances Its Metastatic Potential in Vivo. *PLoS ONE* **2010**, *5*, e12524, doi:10.1371/journal.pone.0012524.

197.   Pérez-Garay, M.; Arteta, B.; Llop, E.; Cobler, L.; Pagès, L.; Ortiz, R.; Ferri, M.J.; Bolós, C.D.; Figueras, J.; Llorens, R.D.; et al. A2,3-Sialyltransferase ST3Gal IV Promotes Migration and Metastasis in Pancreatic Adenocarcinoma Cells and Tends to Be Highly Expressed in Pancreatic Adenocarcinoma Tissues. *International Journal of Biochemistry and Cell Biology* **2013**, *45*, 1748–1757, doi:10.1016/j.biocel.2013.05.015.

198.   Guerrero, P.E.; Miró, L.; Wong, B.S.; Massaguer, A.; Martínez-Bosch, N.; Llo-rens, R. de; Navarro, P.; Konstantopoulos, K.; Llop, E.; Peracaula, R. Knockdown of A2,3-Sialyltransferases Impairs Pancreatic Cancer Cell Migration, Invasion and E-Selec-tin-Dependent Adhesion. *IJMS* **2020**, *21*, 6239, doi:10.3390/ijms21176239.

199.   Miró, L.; López, J.; Guerrero, P.E.; Martínez-Bosch, N.; Manero-Rupérez, N.; Moreno, M.; Ortiz, M.R.; Llop, E.; Navarro, P.; Peracaula, R. Sialyltransferase Inhibitor Ac53FaxNeu5Ac Reverts the Malignant Phenotype of Pancreatic Cancer Cells, and Re-duces Tumor Volume and Favors T-Cell Infiltrates in Mice. *Cancers (Basel)* **2022**, *14*, 6133, doi:10.3390/cancers14246133.

200.   Hofmann, B.T.; Picksak, A.-S.; Kwiatkowski, M.; Grupp, K.; Jücker, M.; Bach-mann, K.; Mercanoglu, B.; Izbicki, J.R.; Kahlert, C.; Bockhorn, M.; et al. Truncated O-GalNAc Glycans Impact on Fundamental Signaling Pathways in Pancreatic Cancer. *Gly-cobiology* **2021**, cwab088, doi:10.1093/glycob/cwab088.

201.   Vreeker, G.C.M.; Hanna-Sawires, R.G.; Mohammed, Y.; Bladergroen, M.R.; Nicolardi, S.; Dotz, V.; Nouta, J.; Bonsing, B.A.; Mesker, W.E.; Burgt, Y.E.M.; et al. Serum $\mathcal{N}$-Glycome Analysis Reveals Pancreatic Cancer Disease Signatures. *Cancer Med* **2020**, *9*, 8519–8529, doi:10.1002/cam4.3439.

202.   Rodrigues, E.; Macauley, M.S. Hypersialylation in Cancer: Modulation of Infla-mmation and Therapeutic Opportunities. *Cancers (Basel)* **2018**, *10*, 207, doi:10.3390/can-cers10060207.

203.   Dan, X.; Liu, W.; Ng, T.B. Development and Applications of Lectins as Biologi-cal Tools in Biomedical Research. *Medicinal Research Reviews* **2016**, *36*, 221–247, doi:10.1002/med.21363.

204.   Hashim, O.H.; Jayapalan, J.J.; Lee, C.-S. Lectins: An Effective Tool for Screening of Potential Cancer Biomarkers. *PeerJ* **2017**, *5*, e3784, doi:10.7717/peerj.3784.

205.   Hendrickson, O.D.; Zherdev, A.V. Analytical Application of Lectins. *Critical Reviews in Analytical Chemistry* **2018**, *48*, 279–292, doi:10.1080/10408347.2017.1422965.

206.   Brooks, S.A. Lectin Histochemistry: Historical Perspectives, State of the Art, and the Future. In *Histochemistry of Single Molecules*; Pellicciari, C., Biggiogera, M., Eds.; Methods in Molecular Biology; Springer New York: New York, NY, 2017; Vol. 1560, pp. 93–107 ISBN 978-1-4939-6787-2.

207.   McMahon, S.S.; Kilcoyne, M. Lectin Histochemistry for Tissues and Cells, and Dual Lectin and Antibody Co-Localization. In *Glycosylation: Methods and Protocols*; Davey, G.P., Ed.; Methods in Molecular Biology; Springer US: New York, NY, 2022; pp. 281–299 ISBN 978-1-07-161685-7.

208.   Souza de Oliveira, F.M.; Mereiter, S.; Lönn, P.; Siart, B.; Shen, Q.; Heldin, J.; Raykova, D.; Karlsson, N.G.; Polom, K.; Roviello, F.; et al. Detection of Post-Translational Modifications Using Solid-Phase Proximity Ligation Assay. *New Biotechnology* **2018**, *45*, 51–59, doi:10.1016/j.nbt.2017.10.005.

209.   Yu, H.; Shu, J.; Li, Z. Lectin Microarrays for Glycoproteomics: An Overview of Their Use and Potential. *Expert Review of Proteomics* **2020**, *17*, 27–39, doi:10.1080/14789450.2020.1720512.

210.   Hirabayashi, J.; Yamada, M.; Kuno, A.; Tateno, H. Lectin Microarrays: Concept, Principle and Applications. *Chem. Soc. Rev.* **2013**, *42*, 4443–4458, doi:10.1039/C3CS35419A.

211.   Dang, K.; Zhang, W.; Jiang, S.; Lin, X.; Qian, A. Application of Lectin Microarrays for Biomarker Discovery. *ChemistryOpen* **2020**, *9*, 285–300, doi:10.1002/open.201900326.

212.   Thompson, R.; Creavin, A.; O'Connell, M.; O'Connor, B.; Clarke, P. Optimization of the Enzyme-Linked Lectin Assay for Enhanced Glycoprotein and Glycoconjugate Analysis. *Analytical Biochemistry* **2011**, *413*, 114–122, doi:10.1016/j.ab.2011.02.013.

213.   O'Connor, B.F.; Monaghan, D.; Cawley, J. Lectin Affinity Chromatography (LAC). In *Protein Chromatography*; Walls, D., Loughran, S.T., Eds.; Methods in Molecular Biology; Springer New York: New York, NY, 2017; Vol. 1485, pp. 411–420 ISBN 978-1-4939-6410-9.

214.   Rowe, L.; Burkhart, G. Analyzing Protein Glycosylation Using UHPLC: A Review. *Bioanalysis* **2018**, *10*, 1691–1703, doi:10.4155/bio-2018-0156.

215.   Royle, L.; Radcliffe, C.M.; Dwek, R.A.; Rudd, P.M. Detailed Structural Analysis of *N*-Glycans Released From Glycoproteins in SDS-PAGE Gel Bands Using HPLC Combined With Exoglycosidase Array Digestions. In *Glycobiology Protocols*; Humana Press: New Jersey, 2006; Vol. 347, pp. 125–144 ISBN 978-1-59745-167-3.

216.   Bui, D.T.; Jung, J.; Kitova, E.N.; Li, Z.; Willows, S.D.; Boddington, M.E.; Kitov, P.I.; Mason, A.L.; Capicciotti, C.J.; Mahal, L.K.; et al. Mass Spectrometry-Based Shotgun Glycomics Using Labeled Glycan Libraries. *Anal. Chem.* **2022**, *94*, 4997–5005, doi:10.1021/acs.analchem.1c04779.

217.   Keser, T.; Pavić, T.; Lauc, G.; Gornik, O. Comparison of 2-Aminobenzamide, Procainamide and RapiFluor-MS as Derivatizing Agents for High-Throughput HILIC-UPLC-FLR-MS N-Glycan Analysis. *Front Chem* **2018**, *6*, 324, doi:10.3389/fchem.2018.00324.

218.   Ruhaak, L.R.; Zauner, G.; Huhn, C.; Bruggink, C.; Deelder, A.M.; Wuhrer, M. Glycan Labeling Strategies and Their Use in Identification and Quantification. *Anal Bioanal Chem* **2010**, *397*, 3457–3481, doi:10.1007/s00216-010-3532-z.

219.   Royle, L.; Campbell, M.P.; Radcliffe, C.M.; White, D.M.; Harvey, D.J.; Abrahams, J.L.; Kim, Y.-G.; Henry, G.W.; Shadick, N.A.; Weinblatt, M.E.; et al. HPLC-Based Analysis of Serum N-Glycans on a 96-Well Plate Platform with Dedicated Database Software. *Analytical Biochemistry* **2008**, *376*, 1–12, doi:10.1016/j.ab.2007.12.012.

220.   Stolz, A.; Jooß, K.; Höcker, O.; Römer, J.; Schlecht, J.; Neusüß, C. Recent Advances in Capillary Electrophoresis-Mass Spectrometry: Instrumentation, Methodology and Applications. *ELECTROPHORESIS* **2019**, *40*, 79–112, doi:10.1002/elps.201800331.

221.   Kammeijer, G.S.M.; Jansen, B.C.; Kohler, I.; Heemskerk, A.A.M.; Mayboroda, O.A.; Hensbergen, P.J.; Schappler, J.; Wuhrer, M. Sialic Acid Linkage Differentiation of Glycopeptides Using Capillary Electrophoresis – Electrospray Ionization – Mass Spectrometry. *Sci Rep* **2017**, *7*, 3733, doi:10.1038/s41598-017-03838-y.

222.   Makrydaki, E.; Kotidis, P.; Polizzi, K.M.; Kontoravdi, C. Hitting the Sweet Spot with Capillary Electrophoresis: Advances in N-Glycomics and Glycoproteomics. *Current Opinion in Biotechnology* **2021**, *71*, 182–190, doi:10.1016/j.copbio.2021.07.013.

223.   Peng, W.; Gutierrez Reyes, C.D.; Gautam, S.; Yu, A.; Cho, B.G.; Goli, M.; Donohoo, K.; Mondello, S.; Kobeissy, F.; Mechref, Y. MS-Based Glycomics and Glycoproteomics Methods Enabling Isomeric Characterization. *Mass Spectrometry Reviews* **2023**, *42*, 577–616, doi:10.1002/mas.21713.

224.   Aronson, J.K.; Ferner, R.E. Biomarkers—A General Review. *Current Protocols in Pharmacology* **2017**, *76*, 9.23.1-9.23.17, doi:10.1002/cpph.19.

225.   Sewpersad, S.; Pillay, T.S. Historical Perspectives in Clinical Pathology: Bence Jones Protein—Early Urine Chemistry and the Impact on Modern Day Diagnostics. *J Clin Pathol* **2021**, *74*, 212–215, doi:10.1136/jclinpath-2020-206675.

226.   Polanski, M.; Anderson, N.L. A List of Candidate Cancer Biomarkers for Targeted Proteomics. *Biomark Insights* **2006**, *1*, 117727190600100, doi:10.1177/117727190600100001.

227.   Kamel, H.F.M.; Al-Amodi, H.S.B. Cancer Biomarkers. In *Role of Biomarkers in Medicine*; Wang, M., Witzmann, F.A., Eds.; InTech, 2016 ISBN 978-953-51-2505-1.

228.   Califf, R.M. Biomarker Definitions and Their Applications. *Exp Biol Med (Maywood)* **2018**, *243*, 213–221, doi:10.1177/1535370217750088.

229.   Dasgupta, A.; Wahed, A. Tumor Markers. In *Clinical Chemistry, Immunology and Laboratory Quality Control*; Elsevier, 2021; pp. 269–293 ISBN 978-0-12-815960-6.

230.   Llop, E.; Guerrero, P.E.; Duran, A.; Barrabés, S.; Massaguer, A.; Ferri, M.J.; Albiol-Quer, M.; Llorens, R.D.; Peracaula, R. Glycoprotein Biomarkers for the Detection of Pancreatic Ductal Adenocarcinoma. *World Journal of Gastroenterology* **2018**, *24*, 2537–2554, doi:10.3748/wjg.v24.i24.2537.

231.   Nahm, F.S. Receiver Operating Characteristic Curve: Overview and Practical Use for Clinicians. *Korean J Anesthesiol* **2022**, *75*, 25–36, doi:10.4097/kja.21209.

232.   Monaghan, T.F.; Rahman, S.N.; Agudelo, C.W.; Wein, A.J.; Lazar, J.M.; Everaert, K.; Dmochowski, R.R. Foundational Statistical Principles in Medical Research: Sensitivity, Specificity, Positive Predictive Value, and Negative Predictive Value. *Medicina (Kaunas)* **2021**, *57*, 503, doi:10.3390/medicina57050503.

233.   Wu, L.; Qu, X. Cancer Biomarker Detection: Recent Achievements and Challenges. *Chem. Soc. Rev.* **2015**, *44*, 2963–2997, doi:10.1039/C4CS00370E.

234.   Hristova, V.A.; Chan, D.W. Cancer Biomarker Discovery and Translation: Proteomics and Beyond. *Expert Review of Proteomics* **2019**, *16*, 93–103, doi:10.1080/14789450.2019.1559062.

235.   Hu, B.; Niu, X.; Cheng, L.; Yang, L.-N.; Li, Q.; Wang, Y.; Tao, S.-C.; Zhou, S.-M. Discovering Cancer Biomarkers from Clinical Samples by Protein Microarrays. *Prot. Clin. Appl.* **2015**, *9*, 98–110, doi:10.1002/prca.201400094.

236.   Singhi, A.D.; Koay, E.J.; Chari, S.T.; Maitra, A. Early Detection of Pancreatic Cancer: Opportunities and Challenges. *Gastroenterology* **2019**, *156*, 2024–2040, doi:10.1053/j.gastro.2019.01.259.

237.   Barzaman, K.; Karami, J.; Zarei, Z.; Hosseinzadeh, A.; Kazemi, M.H.; Moradi-Kalbolandi, S.; Safari, E.; Farahmand, L. Breast Cancer: Biology, Biomarkers, and Treatments. *International Immunopharmacology* **2020**, *84*, 106535, doi:10.1016/j.intimp.2020.106535.

238.   Matsuoka, T.; Yashiro, M. Biomarkers of Gastric Cancer: Current Topics and Future Perspective. *WJG* **2018**, *24*, 2818–2832, doi:10.3748/wjg.v24.i26.2818.

239.   Arend, R.; Martinez, A.; Szul, T.; Birrer, M.J. Biomarkers in Ovarian Cancer: To Be or Not to Be. *Cancer* **2019**, *125*, 4563–4572, doi:10.1002/cncr.32595.

240.   Villalobos, P.; Wistuba, I.I. Lung Cancer Biomarkers. *Hematology/Oncology Clinics of North America* **2017**, *31*, 13–29, doi:10.1016/j.hoc.2016.08.006.

241.   Uhr, A.; Glick, L.; Gomella, L.G. An Overview of Biomarkers in the Diagnosis and Management of Prostate Cancer. **2020**, 4.

242. Landegren, U.; Hammond, M. Cancer Diagnostics Based on Plasma Protein Biomarkers: Hard Times but Great Expectations. *Mol Oncol* **2021**, *15*, 1715–1726, doi:10.1002/1878-0261.12809.

243. Song, E.; Mechref, Y. Defining Glycoprotein Cancer Biomarkers by MS in Conjunction with Glycoprotein Enrichment. *Biomark Med* **2015**, *9*, 835–844, doi:10.2217/bmm.15.55.

244. Silsirivanit, A. Glycosylation Markers in Cancer. In *Advances in Clinical Chemistry*; Elsevier, 2019; Vol. 89, pp. 189–213 ISBN 978-0-12-817145-5.

245. Kumar, A.; Misra, B.B. Challenges and Opportunities in Cancer Metabolomics. *PROTEOMICS* **2019**, *19*, 1900042, doi:10.1002/pmic.201900042.

246. Dinges, S.S.; Hohm, A.; Vandergrift, L.A.; Nowak, J.; Habbel, P.; Kaltashov, I.A.; Cheng, L.L. Cancer Metabolomic Markers in Urine: Evidence, Techniques and Recommendations. *Nat Rev Urol* **2019**, *16*, 339–362, doi:10.1038/s41585-019-0185-3.

247. Singhi, A.D.; Wood, L.D. Early Detection of Pancreatic Cancer Using DNA-Based Molecular Approaches. *Nat Rev Gastroenterol Hepatol* **2021**, *18*, 457–468, doi:10.1038/s41575-021-00470-0.

248. Slack, F.J.; Chinnaiyan, A.M. The Role of Non-Coding RNAs in Oncology. *Cell* **2019**, *179*, 1033–1055, doi:10.1016/j.cell.2019.10.017.

249. He, B.; Zhao, Z.; Cai, Q.; Zhang, Y.; Zhang, P.; Shi, S.; Xie, H.; Peng, X.; Yin, W.; Tao, Y.; et al. miRNA-Based Biomarkers, Therapies, and Resistance in Cancer. *Int. J. Biol. Sci.* **2020**, *16*, 2628–2647, doi:10.7150/ijbs.47203.

250. Pessoa, L.S.; Heringer, M.; Ferrer, V.P. ctDNA as a Cancer Biomarker: A Broad Overview. *Critical Reviews in Oncology/Hematology* **2020**, *155*, 103109, doi:10.1016/j.critrevonc.2020.103109.

251. Stadler, J.-C.; Belloum, Y.; Deitert, B.; Sementsov, M.; Heidrich, I.; Gebhardt, C.; Keller, L.; Pantel, K. Current and Future Clinical Applications of ctDNA in Immuno-Oncology. *Cancer Res* **2022**, *82*, 349–358, doi:10.1158/0008-5472.CAN-21-1718.

252. Lin, D.; Shen, L.; Luo, M.; Zhang, K.; Li, J.; Yang, Q.; Zhu, F.; Zhou, D.; Zheng, S.; Chen, Y.; et al. Circulating Tumor Cells: Biology and Clinical Significance. *Signal Transduct Target Ther* **2021**, *6*, 404, doi:10.1038/s41392-021-00817-8.

253. Vasseur, A.; Kiavue, N.; Bidard, F.; Pierga, J.; Cabel, L. Clinical Utility of Circulating Tumor Cells: An Update. *Mol Oncol* **2021**, *15*, 1647–1666, doi:10.1002/1878-0261.12869.

254. Kartal, E.; Schmidt, T.S.B.; Molina-Montes, E.; Rodríguez-Perales, S.; Wirbel, J.; Maistrenko, O.M.; Akanni, W.A.; Alashkar Alhamwe, B.; Alves, R.J.; Carrato, A.; et al. A Faecal Microbiota Signature with High Specificity for Pancreatic Cancer. *Gut* **2022**, *71*, 1359–1372, doi:10.1136/gutjnl-2021-324755.

255. Wirbel, J.; Pyl, P.T.; Kartal, E.; Zych, K.; Kashani, A.; Milanese, A.; Fleck, J.S.; Voigt, A.Y.; Palleja, A.; Ponnudurai, R.P.; et al. Meta-Analysis of Fecal Metagenomes Reveals Global Microbial Signatures That Are Specific for Colorectal Cancer. *Nat Med* **2019**, *25*, 679–689, doi:10.1038/s41591-019-0406-6.

256. Hackshaw, A.; Clarke, C.A.; Hartman, A.-R. New Genomic Technologies for Multi-Cancer Early Detection: Rethinking the Scope of Cancer Screening. *Cancer Cell* **2022**, *40*, 109–113, doi:10.1016/j.ccell.2022.01.012.

257. Hong, M.; Tao, S.; Zhang, L.; Diao, L.-T.; Huang, X.; Huang, S.; Xie, S.-J.; Xiao, Z.-D.; Zhang, H. RNA Sequencing: New Technologies and Applications in Cancer Research. *J Hematol Oncol* **2020**, *13*, 166, doi:10.1186/s13045-020-01005-x.

258. National cancer institute Tumor Markers in Common Use - NCI Available online: https://www.cancer.gov/about-cancer/diagnosis-staging/diagnosis/tumor-markers-list (accessed on 28 June 2023).

259. Crosby, D.; Bhatia, S.; Brindle, K.M.; Coussens, L.M.; Dive, C.; Emberton, M.; Esener, S.; Fitzgerald, R.C.; Gambhir, S.S.; Kuhn, P.; et al. Early Detection of Cancer. *Science* **2022**, *375*, eaay9040, doi:10.1126/science.aay9040.

260. Hammond, M.E.H.; Taube, S.E. Issues and Barriers to Development of Clinically Useful Tumor Markers: A Development Pathway Proposal. *Seminars in Oncology* **2002**, *29*, 213–221, doi:10.1053/sonc.2002.32896.

261. Wood, L.D.; Canto, M.I.; Jaffee, E.M.; Simeone, D.M. Pancreatic Cancer: Pathogenesis, Screening, Diagnosis, and Treatment. *Gastroenterology* **2022**, *163*, 386-402.e1, doi:10.1053/j.gastro.2022.03.056.

262. Zhang, L.; Sanagapalli, S.; Stoita, A. Challenges in Diagnosis of Pancreatic Cancer. **2018**, *24*, 2047–2060, doi:10.3748/wjg.v24.i19.2047.

263. Molina, V.; Visa, L.; Conill, C.; Navarro, S.; Escudero, J.M.; Auge, J.M.; Filella, X.; Lopez-Boado, M.A.; Ferrer, J.; Fernandez-Cruz, L.; et al. CA 19-9 in Pancreatic Cancer: Retrospective Evaluation of Patients with Suspicion of Pancreatic Cancer. *Tumor Biology* **2012**, *33*, 799–807, doi:10.1007/s13277-011-0297-8.

264. Tsen, A.; Barbara, M.; Rosenkranz, L. Dilemma of Elevated CA 19-9 in Biliary Pathology. **2018**, *18*, 862–867, doi:10.1016/j.pan.2018.09.004.

265. Ballehaninna, U.K.; Chamberlain, R.S. The Clinical Utility of Serum CA 19-9 in the Diagnosis, Prognosis and Management of Pancreatic Adenocarcinoma: An Evidence Based Appraisal. *Journal of Gastrointestinal Oncology* **2012**, *3*, 15, doi:10.3978/j.issn.2078-6891.2011.021.

266. Liu, C.; Deng, S.; Jin, K.; Gong, Y.; Cheng, H.; Fan, Z.; Qian, Y.; Huang, Q.; Ni, Q.; Luo, G.; et al. Lewis Antigen-Negative Pancreatic Cancer: An Aggressive Subgroup. *Int J Oncol* **2020**, *56*, 900–908, doi:10.3892/ijo.2020.4989.

267. Chan, A.; Diamandis, E.P.; Blasutig, I.M. Strategies for Discovering Novel Pancreatic Cancer Biomarkers. *Journal of Proteomics* **2013**, *81*, 126–134, doi:10.1016/j.jprot.2012.09.025.

268. Isaji, S.; Mizuno, S.; Windsor, J.A.; Bassi, C.; Fernández-del Castillo, C.; Hackert, T.; Hayasaki, A.; Katz, M.H.G.; Kim, S.-W.; Kishiwada, M.; et al. International Consensus on Definition and Criteria of Borderline Resectable Pancreatic Ductal Adenocarcinoma 2017. *Pancreatology* **2018**, *18*, 2–11, doi:10.1016/j.pan.2017.11.011.

269. Reitz, D.; Gerger, A.; Seidel, J.; Kornprat, P.; Samonigg, H.; Stotz, M.; Szkandera, J.; Pichler, M. Combination of Tumour Markers CEA and CA19-9 Improves the Prognostic Prediction in Patients with Pancreatic Cancer. *J Clin Pathol* **2015**, *68*, 427–433, doi:10.1136/jclinpath-2014-202451.

270. Zhang, Y.; Yang, J.; Li, H.; Wu, Y.; Zhang, H.; Chen, W. Tumor Markers CA19-9 , CA242 and CEA in the Diagnosis of Pancreatic Cancer : A Meta-Analysis. *Iternational journal of clinical and experimental medicine* **2015**, *8*, 11683–11691.

271. Ferri, M.J.; Saez, M.; Figueras, J.; Fort, E.; Sabat, M.; López-Ben, S.; de Llorens, R.; Aleixandre, R.N.; Peracaula, R. Improved Pancreatic Adenocarcinoma Diagnosis in Jaundiced and Non-Jaundiced Pancreatic Adenocarcinoma Patients through the Combination of Routine Clinical Markers Associated to Pancreatic Adenocarcinoma Pathophysiology. *PLoS ONE* **2016**, *11*, e0147214, doi:10.1371/journal.pone.0147214.

272. Suehiro, Y.; Suenaga, S.; Kunimune, Y.; Yada, S.; Hamamoto, K.; Tsuyama, T.; Amano, S.; Matsui, H.; Higaki, S.; Fujii, I.; et al. CA19-9 in Combination with Methylated HOXA1 and SST Is Useful to Diagnose Stage I Pancreatic Cancer. *Oncology* **2022**, *100*, 674–684, doi:10.1159/000527342.

273. Zhang, J.; Wang, Y.; Zhao, T.; Li, Y.; Tian, L.; Zhao, J.; Zhang, J. Evaluation of Serum MUC5AC in Combination with CA19-9 for the Diagnosis of Pancreatic Cancer. *World J Surg Oncol* **2020**, *18*, 31, doi:10.1186/s12957-020-1809-z.

274. Kim, M.W.; Koh, H.; Kim, J.Y.; Lee, S.; Lee, H.; Kim, Y.; Hwang, H.K.; Kim, S.I. Tumor-Specific miRNA Signatures in Combination with CA19-9 for Liquid Biopsy-Based Detection of PDAC. *Int J Mol Sci* **2021**, *22*, 13621, doi:10.3390/ijms222413621.

275. O'Neill, R.S.; Stoita, A. Biomarkers in the Diagnosis of Pancreatic Cancer: Are We Closer to Finding the Golden Ticket? *WJG* **2021**, *27*, 4045–4087, doi:10.3748/wjg.v27.i26.4045.

276. Al-shaheri, F.N.; Alhamdani, M.S.S.; Bauer, A.S.; Giese, N. Blood Biomarkers for Differential Diagnosis and Early Detection of Pancreatic Cancer. *Cancer treatment reviews* **2021**, *96*, 102193, doi:10.1016/j.ctrv.2021.102193.

277. Xu, J.; Xiong, G.; Cao, Z.; Huang, H.; Wang, T.; You, L.; Zhou, L.; Zheng, L.; Hu, Y.; Zhang, T.; et al. PIM-1 Contributes to the Malignancy of Pancreatic Cancer and Displays Diagnostic and Prognostic Value. *J Exp Clin Cancer Res* **2016**, *35*, 133, doi:10.1186/s13046-016-0406-z.

278.   Meng, Q.; Shi, S.; Liang, C.; Xiang, J.; Liang, D.; Zhang, B.; Qin, Y.; Ji, S.; Xu, W.; Xu, J.; et al. Diagnostic Accuracy of a CA125-Based Biomarker Panel in Patients with Pancreatic Cancer: A Systematic Review and Meta-Analysis. *J Cancer* **2017**, *8*, 3615–3622, doi:10.7150/jca.18901.

279.   Jahan, R.; Ganguly, K.; Smith, L.M.; Atri, P.; Carmicheal, J.; Sheinin, Y.; Rachagani, S.; Natarajan, G.; Brand, R.E.; Macha, M.A.; et al. Trefoil Factor(s) and CA19.9: A Promising Panel for Early Detection of Pancreatic Cancer. *EBioMedicine* **2019**, *42*, 375–385, doi:10.1016/j.ebiom.2019.03.056.

280.   Park, J.Y.; Kim, S.-A.; Chung, J.W.; Bang, S.; Park, S.W.; Paik, Y.-K.; Song, S.Y. Proteomic Analysis of Pancreatic Juice for the Identification of Biomarkers of Pancreatic Cancer. *J Cancer Res Clin Oncol* **2011**, *137*, 1229–1238, doi:10.1007/s00432-011-0992-2.

281.   Makawita, S.; Dimitromanolakis, A.; Soosaipillai, A.; Soleas, I.; Chan, A.; Gallinger, S.; Haun, R.S.; Blasutig, I.M.; Diamandis, E.P. Validation of Four Candidate Pancreatic Cancer Serological Biomarkers That Improve the Performance of CA19.9. *BMC Cancer* **2013**, *13*, 404, doi:10.1186/1471-2407-13-404.

282.   Takayama, R.; Nakagawa, H.; Sawaki, A.; Mizuno, N.; Kawai, H.; Tajika, M.; Yatabe, Y.; Matsuo, K.; Uehara, R.; Ono, K.; et al. Serum Tumor Antigen REG4 as a Diagnostic Biomarker in Pancreatic Ductal Adenocarcinoma. *J Gastroenterol* **2010**, *45*, 52–59, doi:10.1007/s00535-009-0114-y.

283.   Saukkonen, K.; Hagström, J.; Mustonen, H.; Lehtinen, L.; Carpen, O.; Andersson, L.C.; Seppänen, H.; Haglund, C. Prognostic and Diagnostic Value of REG4 Serum and Tissue Expression in Pancreatic Ductal Adenocarcinoma. *Tumour Biol.* **2018**, *40*, 101042831876149, doi:10.1177/1010428318761494.

284.   Honda, K.; Kobayashi, M.; Okusaka, T.; Rinaudo, J.A.; Huang, Y.; Marsh, T.; Sanada, M.; Sasajima, Y.; Nakamori, S.; Shimahara, M.; et al. Plasma Biomarker for Detection of Early Stage Pancreatic Cancer and Risk Factors for Pancreatic Malignancy Using Antibodies for Apolipoprotein-AII Isoforms. *Sci Rep* **2015**, *5*, 15921, doi:10.1038/srep15921.

285.   Liu, X.; Zheng, W.; Wang, W.; Shen, H.; Liu, L.; Lou, W.; Wang, X.; Yang, P. A New Panel of Pancreatic Cancer Biomarkers Discovered Using a Mass Spectrometry-Based Pipeline. *Br J Cancer* **2017**, *117*, 1846–1854, doi:10.1038/bjc.2017.365.

286.   Sharma, G.G.; Okada, Y.; Von Hoff, D.; Goel, A. Non-Coding RNA Biomarkers in Pancreatic Ductal Adenocarcinoma. *Semin Cancer Biol* **2021**, *75*, 153–168, doi:10.1016/j.semcancer.2020.10.001.

287.   Daoud, A.Z.; Mulholland, E.J.; Cole, G.; McCarthy, H.O. MicroRNAs in Pancreatic Cancer: Biomarkers, Prognostic, and Therapeutic Modulators. *BMC Cancer* **2019**, *19*, 1130, doi:10.1186/s12885-019-6284-y.

288.   Ou, Z.-L.; Luo, Z.; Lu, Y.-B. Long Non-Coding RNA HULC as a Diagnostic and Prognostic Marker of Pancreatic Cancer. *World Journal of Gastroenterology* **2019**, *25*, 6728–6742, doi:10.3748/wjg.v25.i46.6728.

289.   Du, W.; Lei, C.; Wang, Y.; Ding, Y.; Tian, P. LINC01232 Sponges Multiple miRNAs and Its Clinical Significance in Pancreatic Adenocarcinoma Diagnosis and Prognosis. *Technol Cancer Res Treat* **2021**, *20*, 153303382098852, doi:10.1177/1533033820988525.

290.   Ghafouri-Fard, S.; Fathi, M.; Zhai, T.; Taheri, M.; Dong, P. LncRNAs: Novel Biomarkers for Pancreatic Cancer. *Biomolecules* **2021**, *11*, 1665, doi:10.3390/biom11111665.

291.   Yang, F.; Liu, D.-Y.; Guo, J.-T.; Ge, N.; Zhu, P.; Liu, X.; Wang, S.; Wang, G.-X.; Sun, S.-Y. Circular RNA Circ-LDLRAD3 as a Biomarker in Diagnosis of Pancreatic Cancer. *World Journal of Gastroenterology* **2017**, *23*, 8345, doi:10.3748/wjg.v23.i47.8345.

292.   Hao, L.; Rong, W.; Bai, L.; Cui, H.; Zhang, S.; Li, Y.; Chen, D.; Meng, X. Upregulated Circular RNA Circ_0007534 Indicates an Unfavorable Prognosis in Pancreatic Ductal Adenocarcinoma and Regulates Cell Proliferation, Apoptosis, and Invasion by Sponging miR-625 and miR-892b. *Journal of Cellular Biochemistry* **2019**, *120*, 3780–3789, doi:10.1002/jcb.27658.

293.   Zhu, Y.; Zhang, H.; Chen, N.; Hao, J.; Jin, H.; Ma, X. Diagnostic Value of Various Liquid Biopsy Methods for Pancreatic Cancer. *Medicine (Baltimore)* **2020**, *99*, e18581, doi:10.1097/MD.0000000000018581.

294.   Lai, X.; Wang, M.; McElyea, S.D.; Sherman, S.; House, M.; Korc, M. A microRNA Signature in Circulating Exosomes Is Superior to Exosomal Glypican-1 Levels for Diagnosing Pancreatic Cancer. *Cancer Lett* **2017**, *393*, 86–93, doi:10.1016/j.canlet.2017.02.019.

295.   Eissa, M.A.L.; Lerner, L.; Abdelfatah, E.; Shankar, N.; Canner, J.K.; Hasan, N.M.; Yaghoobi, V.; Huang, B.; Kerner, Z.; Takaesu, F.; et al. Promoter Methylation of ADAMTS1 and BNC1 as Potential Biomarkers for Early Detection of Pancreatic Cancer in Blood. *Clin Epigenetics* **2019**, *11*, 59, doi:10.1186/s13148-019-0650-0.

296.   Maire, F.; Micard, S.; Hammel, P.; Voitot, H.; Lévy, P.; Cugnenc, P.-H.; Ruszniewski, P.; Puig, P.L. Differential Diagnosis between Chronic Pancreatitis and Pancreatic Cancer: Value of the Detection of KRAS2 Mutations in Circulating DNA. *Br J Cancer* **2002**, *87*, 551–554, doi:10.1038/sj.bjc.6600475.

297.   Martini, V.; Timme-Bronsert, S.; Fichtner-Feigl, S.; Hoeppner, J.; Kulemann, B. Circulating Tumor Cells in Pancreatic Cancer: Current Perspectives. *Cancers (Basel)* **2019**, *11*, 1659, doi:10.3390/cancers11111659.

298.   Mayerle, J.; Kalthoff, H.; Reszka, R.; Kamlage, B.; Peter, E.; Schniewind, B.; González Maldonado, S.; Pilarsky, C.; Heidecke, C.-D.; Schatz, P.; et al. Metabolic Biomarker Signature to Differentiate Pancreatic Ductal Adenocarcinoma from Chronic Pancreatitis. *Gut* **2018**, *67*, 128–137, doi:10.1136/gutjnl-2016-312432.

299.   Mahajan, U.M.; Oehrle, B.; Sirtl, S.; Alnatsha, A.; Goni, E.; Regel, I.; Beyer, G.; Vornhülz, M.; Vielhauer, J.; Chromik, A.; et al. Independent Validation and Assay Standardization of Improved Metabolic Biomarker Signature to Differentiate Pancreatic Ductal Adenocarcinoma From Chronic Pancreatitis. *Gastroenterology* **2022**, *163*, 1407–1422, doi:10.1053/j.gastro.2022.07.047.

300.   Pozo, Ó.; Gomez, À.; Gil, G.; Pera, M.; Visa, L.; Arriola, E.; Peracaula, R.; Llop, E.; Castro, E.; Garcia, A.; et al. Method for Diagnosing Cancer 2022. Patent: EP22382244.6.

301.   Daulton, E.; Wicaksono, A.N.; Tiele, A.; Kocher, H.M.; Debernardi, S.; Crnogorac-Jurcevic, T.; Covington, J.A. Volatile Organic Compounds (VOCs) for the Non-Invasive Detection of Pancreatic Cancer from Urine. *Talanta* **2021**, *221*, 121604, doi:10.1016/j.talanta.2020.121604.

302.   Princivalle, A.; Monasta, L.; Butturini, G.; Bassi, C.; Perbellini, L. Pancreatic Ductal Adenocarcinoma Can Be Detected by Analysis of Volatile Organic Compounds (VOCs) in Alveolar Air. *BMC Cancer* **2018**, *18*, 529, doi:10.1186/s12885-018-4452-0.

303.   Weeks, M.E.; Hariharan, D.; Petronijevic, L.; Radon, T.P.; Whiteman, H.J.; Kocher, H.M.; Timms, J.F.; Lemoine, N.R.; Crnogorac-Jurcevic, T. Analysis of the Urine Proteome in Patients with Pancreatic Ductal Adenocarcinoma. *PROTEOMICS – Clinical Applications* **2008**, *2*, 1047–1057, doi:10.1002/prca.200780164.

304.   Zhang, L.; Farrell, J.J.; Zhou, H.; Elashoff, D.; Akin, D.; Park, N.-H.; Chia, D.; Wong, D.T. Salivary Transcriptomic Biomarkers for Detection of Resectable Pancreatic Cancer. *Gastroenterology* **2010**, *138*, 949-957.e7, doi:10.1053/j.gastro.2009.11.010.

305.   Humeau, M.; Vignolle-Vidoni, A.; Sicard, F.; Martins, F.; Bournet, B.; Buscail, L.; Torrisani, J.; Cordelier, P. Salivary MicroRNA in Pancreatic Cancer Patients. *PLoS One* **2015**, *10*, e0130996, doi:10.1371/journal.pone.0130996.

306.   Nakamura, S.; Sadakari, Y.; Ohtsuka, T.; Okayama, T.; Nakashima, Y.; Gotoh, Y.; Saeki, K.; Mori, Y.; Nakata, K.; Miyasaka, Y.; et al. Pancreatic Juice Exosomal MicroRNAs as Biomarkers for Detection of Pancreatic Ductal Adenocarcinoma. *Ann Surg Oncol* **2019**, *26*, 2104–2111, doi:10.1245/s10434-019-07269-z.

307.   Park, H.-M.; Hwang, M.P.; Kim, Y.-W.; Kim, K.-J.; Jin, J.M.; Kim, Y.H.; Yang, Y.-H.; Lee, K.H.; Kim, Y.-G. Mass Spectrometry-Based N-Linked Glycomic Profiling as a Means for Tracking Pancreatic Cancer Metastasis. *Carbohydrate Research* **2015**, *413*, 5–11, doi:10.1016/j.carres.2015.04.019.

308.   Nouso, K.; Amano, M.; Ito, Y.M.; Miyahara, K.; Morimoto, Y.; Kato, H.; Tsutsumi, K.; Tomoda, T.; Yamamoto, N.; Nakamura, S.; et al. Clinical Utility of High-Throughput Glycome Analysis in Patients with Pancreatic Cancer. *J Gastroenterol* **2013**, *48*, 1171–1179, doi:10.1007/s00535-012-0732-7.

309.   Pan, S.; Chen, R.; Tamura, Y.; Crispin, D.A.; Lai, L.A.; May, D.H.; McIntosh, M.W.; Goodlett, D.R.; Brentnall, T.A. Quantitative Glycoproteomics Analysis Reveals Changes in N-Glycosylation Level Associated with Pancreatic Ductal Adenocarcinoma. *Journal of Proteome Research* **2014**, *13*, 1293–1306, doi:10.1021/pr4010184.

310.   Lu, H.; Xiao, K.; Tian, Z. Benchmark of Site- and Structure-Specific Quantitative Tissue N-Glycoproteomics for Discovery of Potential N-Glycoprotein Markers: A Case Study of Pancreatic Cancer. *Glycoconj J* **2021**, *38*, 213–231, doi:10.1007/s10719-021-09994-8.

311.   Barrabés, S.; Pagès-Pons, L.; Radcliffe, C.M.; Tabarés, G.; Fort, E.; Royle, L.; Harvey, D.J.; Moenner, M.; Dwek, R.A.; Rudd, P.M.; et al. Glycosylation of Serum Ribonuclease 1 Indicates a Major Endothelial Origin and Reveals an Increase in Core Fucosylation in Pancreatic Cancer. *Glycobiology* **2007**, *17*, 388–400, doi:10.1093/glycob/cwm002.

312.   Okuyama, N.; Ide, Y.; Nakano, M.; Nakagawa, T.; Yamanaka, K.; Moriwaki, K.; Murata, K.; Ohigashi, H.; Yokoyama, S.; Eguchi, H.; et al. Fucosylated Haptoglobin Is a Novel Marker for Pancreatic Cancer: A Detailed Analysis of the Oligosaccharide Structure and a Possible Mechanism for Fucosylation. *International Journal of Cancer* **2006**, *118*, 2803–2808, doi:10.1002/ijc.21728.

313.   Matsumoto, H.; Shinzaki, S.; Narisada, M.; Kawamoto, S.; Kuwamoto, K.; Moriwaki, K.; Kanke, F.; Satomura, S.; Kumada, T.; Miyoshi, E. Clinical Application of a Lectin-Antibody ELISA to Measure Fucosylated Haptoglobin in Sera of Patients with Pancreatic Cancer. *Clinical Chemistry and Laboratory Medicine* **2010**, *48*, 505–512, doi:10.1515/CCLM.2010.095.

314.   Shimomura, M.; Nakayama, K.; Azuma, K.; Terao, N.; Nishino, K.; Takamatsu, S.; Nakano, M.; Takahashi, S.; Kobayashi, Y.; Murata, K.; et al. Establishment of a Novel Lectin–Antibody ELISA System to Determine Core-Fucosylated Haptoglobin. *Clinica Chimica Acta* **2015**, *446*, 30–36, doi:10.1016/j.cca.2015.03.037.

315.   Balmaña, M.; Giménez, E.; Puerta, A.; Llop, E.; Figueras, J. Increased A1-3 Fucosylation of α-1-Acid Glycoprotein (AGP) in Pancreatic Cancer. *Journal of Proteomics* **2016**, *132*, 144–154, doi:10.1016/j.jprot.2015.11.006.

316.   Li, C.; Simeone, D.M.; Brenner, D.E.; Anderson, M.A.; Shedden, K.A.; Ruffin, M.T.; Lubman, D.M. Pancreatic Cancer Serum Detection Using A Lectin/Glyco-Antibody Array Method. *J Proteome Res* **2009**, *8*, 483–492, doi:10.1021/pr8007013.

317.   Li, C.; Zolotarevsky, E.; Thompson, I.; Anderson, M.A.; Simeone, D.M.; Casper, J.M.; Mullenix, M.C.; Lubman, D.M. A Multiplexed Bead Assay for Profiling Glycosylation Patterns on Serum Protein Biomarkers of Pancreatic Cancer. *ELECTROPHORESIS* **2011**, *32*, 2028–2035, doi:10.1002/elps.201000693.

318.   Zhao, J.; Patwa, T.H.; Qiu, W.; Shedden, K.; Hinderer, R.; Misek, D.E.; Anderson, M.A.; Simeone, D.M.; Lubman, D.M. Glycoprotein Microarrays with Multi-Lectin Detection: Unique Lectin Binding Patterns as a Tool for Classifying Normal, Chronic Pancreatitis and Pancreatic Cancer Sera. *J. Proteome Res.* **2007**, *6*, 1864–1874, doi:10.1021/pr070062p.

319.   Yue, T.; Goldstein, I.J.; Hollingsworth, M.A.; Kaul, K.; Brand, R.E.; Haab, B.B. The Prevalence and Nature of Glycan Alterations on Specific Proteins in Pancreatic Cancer Patients Revealed Using Antibody-Lectin Sandwich Arrays. *Molecular & Cellular Proteomics* **2009**, *8*, 1697–1707, doi:10.1074/mcp.M900135-MCP200.

320.   Yue, T.; Maupin, K.A.; Fallon, B.; Li, L.; Partyka, K.; Anderson, M.A.; Brenner, D.E.; Kaul, K.; Zeh, H.; Moser, A.J.; et al. Enhanced Discrimination of Malignant from Benign Pancreatic Disease by Measuring the CA 19-9 Antigen on Specific Protein Carriers. *PLoS ONE* **2011**, *6*, e29180, doi:10.1371/journal.pone.0029180.

321.   Balmaña, M.; Sarrats, A.; Llop, E.; Barrabés, S.; Saldova, R.; Ferri, M.J.; Figueras, J.; Fort, E.; Llorens, R.D.; Rudd, P.M.; et al. Identification of Potential Pancreatic Cancer Serum Markers: Increased Sialyl-Lewis X on Ceruloplasmin. *Clinica Chimica Acta* **2015**, *442*, 56–62, doi:10.1016/j.cca.2015.01.007.

322.   Balmaña, M.; Duran, A.; Gomes, C.; Llop, E.; López-Martos, R.; Ortiz, M.R.; Barrabés, S.; Reis, C.A.; Peracaula, R. Analysis of Sialyl-Lewis x on MUC5AC and MUC1 Mucins in Pancreatic Cancer Tissues. *International Journal of Biological Macromolecules* **2018**, *112*, 33–45, doi:10.1016/j.ijbiomac.2018.01.148.

323.   Guerrero, P.E.; Duran, A.; Ortiz, M.R.; Castro, E.; Garcia-Velasco, A.; Llop, E.; Peracaula, R. Microfibril Associated Protein 4 (MFAP4) Is a Carrier of the Tumor Associated Carbohydrate Sialyl-Lewis x (sLex) in Pancreatic Adenocarcinoma. *Journal of Proteomics* **2021**, *231*, 104004, doi:10.1016/j.jprot.2020.104004.

324.   Nichetti, F.; Marra, A.; Corti, F.; Guidi, A.; Raimondi, A.; Prinzi, N.; Braud, F. de; Pusceddu, S. The Role of Mesothelin as a Diagnostic and Therapeutic Target in Pancreatic Ductal Adenocarcinoma: A Comprehensive Review. *Targeted Oncology* **2018**, *13*, 333–351, doi:10.1007/s11523-018-0567-0.

325.   Chang, K.; Pastan, I. Molecular Cloning of Mesothelin, a Differentiation Antigen Present on Mesothelium, Mesotheliomas, and Ovarian Cancers. *Proceedings of the National Academy of Sciences of the United States of America* **1996**, *93*, 136–140, doi:10.1073/pnas.93.1.136.

326.   Hassan, R.; Bera, T.; Pastan, I. Mesothelin: A New Target for Immunotherapy. *Clinical Cancer Research* **2004**, *10*, 3937–3942, doi:10.1158/1078-0432.CCR-03-0801.

327.   Liu, X.; Chan, A.; Tai, C.-H.; Andresson, T.; Pastan, I. Multiple Proteases Are Involved in Mesothelin Shedding by Cancer Cells. *Communications Biology* **2020**, *3*, 2–10, doi:10.1038/s42003-020-01464-5.

328. Sapede, C.; Gauvrit, A.; Barbieux, I.; Padieu, M.; Cellerin, L.; Sagan, C.; Scherpereel, A.; Dabouis, G.; Grégoire, M. Aberrant Splicing and Protease Involvement in Mesothelin Release from Epithelioid Mesothelioma Cells. *Cancer Science* **2008**, *99*, 590–594, doi:10.1111/j.1349-7006.2007.00715.x.

329. Muminova, Z.E.; Strong, T.V.; Shaw, D.R. Characterization of Human Mesothelin Transcripts in Ovarian and Pancreatic Cancer. *BMC cancer* **2004**, *4*, doi:10.1186/1471-2407-4-19.

330. Scholler, N.; Fu, N.; Yang, Y.; Ye, Z.; Goodman, G.E.; Hellstrom, K.E.; Hellstrom, I. Soluble Member(s) of the Mesothelin/Megakaryocyte Potentiating Factor Family Are Detectable in Sera from Patients with Ovarian Carcinoma. *Proceedings of the National Academy of Sciences* **1999**, *96*, 11531–11536, doi:10.1073/pnas.96.20.11531.

331. Sathyanarayana, B.K.; Hahn, Y.; Patankar, M.S.; Pastan, I.; Lee, B. Mesothelin, Stereocilin, and Otoancorin Are Predicted to Have Superhelical Structures with ARM-Type Repeats. *BMC Struct Biol* **2009**, *9*, 1, doi:10.1186/1472-6807-9-1.

332. Ho, M.; Onda, M.; Wang, Q.C.; Hassan, R.; Pastan, I.; Lively, M.O. Mesothelin Is Shed from Tumor Cells. *Cancer Epidemiology Biomarkers and Prevention* **2006**, *15*, 1751, doi:10.1158/1055-9965.EPI-06-0479.

333. Fujihira, H.; Takakura, D.; Matsuda, A.; Abe, M.; Miyazaki, M.; Nakagawa, T.; Kajino, K.; Denda-Nagai, K.; Noji, M.; Hino, O.; et al. Bisecting-GlcNAc on Asn388 Is Characteristic to ERC/Mesothelin Expressed on Epithelioid Mesothelioma Cells. *The Journal of Biochemistry* **2021**, *170*, 317–326, doi:10.1093/jb/mvab044.

334. Yang, J.; Zhang, Y. I-TASSER Server: New Development for Protein Structure and Function Predictions. *Nucleic Acids Res* **2015**, *43*, W174–W181, doi:10.1093/nar/gkv342.

335. Rump, A.; Morikawa, Y.; Tanaka, M.; Minami, S.; Umesaki, N.; Takeuchi, M.; Miyajima, A. Binding of Ovarian Cancer Antigen CA125/MUC16 to Mesothelin Mediates Cell Adhesion. *Journal of Biological Chemistry* **2004**, *279*, 9190–9198, doi:10.1074/jbc.M312372200.

336. Gubbels, J.A.A.; Belisle, J.; Onda, M.; Rancourt, C.; Migneault, M.; Ho, M.; Bera, T.K.; Connor, J.P.; Sathyanarayana, B.K.; Lee, B.; et al. Mesothelin-MUC16 Binding Is a High Affinity, N-Glycan Dependent Interaction That Facilitates Peritoneal Metastasis of Ovarian Tumors. *Molecular Cancer* **2006**, *5*, 50, doi:10.1186/1476-4598-5-50.

337. Hassan, R.; Thomas, A.; Alewine, C.; Le, D.T.; Jaffee, E.M.; Pastan, I. Mesothelin Immunotherapy for Cancer: Ready for Prime Time? *Journal of Clinical Oncology* **2016**, *34*, 4171–4179, doi:10.1200/JCO.2016.68.3672.

338. Montemagno, C.; Cassim, S.; Pouyssegur, J.; Broisat, A.; Pagès, G. From Malignant Progression to Therapeutic Targeting: Current Insights of Mesothelin in Pancreatic Ductal Adenocarcinoma. *International Journal of Molecular Sciences* **2020**, *21*, 1–18, doi:10.3390/ijms21114067.

339.   Zheng, C.; Jia, W.; Tang, Y.; Zhao, H.; Jiang, Y.; Sun, S. Mesothelin Regulates Growth and Apoptosis in Pancreatic Cancer Cells through P53-Dependent and -Independent Signal Pathway. *Journal of experimental & clinical cancer research : CR* **2012**, *31*, 84, doi:10.1186/1756-9966-31-84.

340.   Bharadwaj, U.; Marin-Muller, C.; Li, M.; Chen, C.; Yao, Q. Mesothelin Confers Pancreatic Cancer Cell Resistance to TNF-α-Induced Apoptosis through Akt/PI3K/NF-κB Activation and IL-6/Mcl-1 Overexpression. *Molecular Cancer* **2011**, *10*, 106, doi:10.1186/1476-4598-10-106.

341.   He, X.; Wang, L.; Riedel, H.; Wang, K.; Yang, Y.; Dinu, C.Z.; Rojanasakul, Y. Mesothelin Promotes Epithelial-to-Mesenchymal Transition and Tumorigenicity of Human Lung Cancer and Mesothelioma Cells. *Molecular Cancer* **2017**, *16*, 63, doi:10.1186/s12943-017-0633-8.

342.   Hassan, R.; Remaley, A.T.; Sampson, M.L.; Zhang, J.; Cox, D.D.; Pingpank, J.; Alexander, R.; Willingham, M.; Pastan, I.; Onda, M. Detection and Quantitation of Serum Mesothelin, a Tumor Marker for Patients with Mesothelioma and Ovarian Cancer. *Clinical Cancer Research* **2006**, *12*, 447–453, doi:10.1158/1078-0432.CCR-05-1477.

343.   Hanaoka, T.; Hasegawa, K.; Kato, T.; Sato, S.; Kurosaki, A.; Miyara, A.; Nagao, S.; Seki, H.; Yasuda, M.; Fujiwara, K. Correlation Between Tumor Mesothelin Expression and Serum Mesothelin in Patients with Epithelial Ovarian Carcinoma: A Potential Noninvasive Biomarker for Mesothelin-Targeted Therapy. *Molecular Diagnosis & Therapy* **2017**, *21*, 187–198, doi:10.1007/s40291-017-0255-2.

344.   Lv, J.; Li, P. Mesothelin as a Biomarker for Targeted Therapy. *Biomarker Research* **2019**, *7*, 18, doi:10.1186/s40364-019-0169-8.

345.   Chu, Q. Targeting Mesothelin in Solid Tumours: Anti-Mesothelin Antibody and Drug Conjugates. *Curr Oncol Rep* **2023**, *25*, 309–323, doi:10.1007/s11912-023-01367-8.

346.   Argani, P.; Iacobuzio-Donahue, C.; Ryu, B.; Rosty, C.; Goggins, M.; Wilentz, R.E.; Murugesan, S.R.; Leach, S.D.; Jaffee, E.; Yeo, C.J.; et al. Mesothelin Is Overexpressed in the Vast Majority of Ductal Adenocarcinomas of the Pancreas: Identification of a New Pancreatic Cancer Marker by Serial Analysis of Gene Expression (SAGE). *Clinical cancer research* **2001**, *7*, 3862–3868.

347.   Hassan, R.; Laszik, Z.G.; Lerner, M.; Raffeld, M.; Postier, R.; Brackett, D. Mesothelin Is Overexpressed in Pancreaticobiliary Adenocarcinomas But Not in Normal Pancreas and Chronic Pancreatitis. *American Journal of Clinical Pathology* **2005**, *124*, 838–845, doi:10.1309/F1B6-4CL7-H8VJ-KEAF.

348.   Frank, R.; Li, S.; Ahmad, N.A.; Sepulveda, A.R.; Jhala, N.C. Mesothelin Expression in Pancreatic Mucinous Cysts. *American Journal of Clinical Pathology* **2014**, *142*, 313–319, doi:10.1309/AJCPDTTL2I5ECMFG.

349.   Einama, T.; Kamachi, H.; Nishihara, H.; Homma, S.; Kanno, H.; Ishikawa, M.; Kawamata, F.; Konishi, Y.; Sato, M.; Tahara, M.; et al. Importance of Luminal Membrane Mesothelin Expression in Intraductal Papillary Mucinous Neoplasms. *Oncology Letters* **2015**, *9*, 1583–1589, doi:10.3892/ol.2015.2969.

350.   Ordóñez, N.G. Application of Mesothelin Immunostaining in Tumor Diagnosis. *The American Journal of Surgical Pathology* **2003**, *27*, 1418–1428, doi:10.1097/00000478-200311000-00003.

351.   Weidemann, S.; Perez, D.; Izbicki, J.R.; Neipp, M.; Mofid, H.; Daniels, T.; Nahrstedt, U.; Jacobsen, F.; Bernreuther, C.; Simon, R.; et al. Mesothelin Is Commonly Expressed in Pancreatic Adenocarcinoma but Unrelated to Cancer Aggressiveness. *Cancer Investigation* **2021**, 1–20, doi:10.1080/07357907.2021.1943747.

352.   Jhala, N.; Jhala, D.; Vickers, S.M.; Eltoum, I.; Batra, S.K.; Manne, U.; Eloubeidi, M.; Jones, J.J.; Grizzle, W.E. Biomarkers in Diagnosis of Pancreatic Carcinoma in Fine-Needle Aspirates: A Translational Research Application. *Am J Clin Pathol* **2006**, *126*, 572–579, doi:10.1309/CEV30BE088CBDQD9.

353.   Sato, N.; Fukushima, N.; Maitra, A.; Iacobuzio-Donahue, C.A.; van Heek, N.T.; Cameron, J.L.; Yeo, C.J.; Hruban, R.H.; Goggins, M. Gene Expression Profiling Identifies Genes Associated with Invasive Intraductal Papillary Mucinous Neoplasms of the Pancreas. *The American Journal of Pathology* **2004**, *164*, 903–914, doi:10.1016/S0002-9440(10)63178-1.

354.   Kendrick, Z.W.; Firpo, M.A.; Repko, R.C.; Scaife, C.L.; Adler, D.G.; Boucher, K.M.; Mulvihill, S.J. Serum IGFBP2 and MSLN as Diagnostic and Prognostic Biomarkers for Pancreatic Cancer. *HPB (Oxford)* **2014**, *16*, 670–676, doi:10.1111/hpb.12199.

355.   Beyer, H.L.; Geschwindt, R.D.; Glover, C.L.; Tran, L.; Hellstrom, I.; Hellstrom, K.E.; Miller, M.C.; Verch, T.; Allard, W.J.; Pass, H.I.; et al. MESOMARK: A Potential Test for Malignant Pleural Mesothelioma. *Clinical Chemistry* **2007**, *53*, 666–672, doi:10.1373/clinchem.2006.079327.

356.   Sharon, E.; Zhang, J.; Hollevoet, K.; Steinberg, S.M.; Pastan, I.; Onda, M.; Gaedcke, J.; Ghadimi, B.M.; Ried, T.; Hassan, R. Serum Mesothelin and Megakaryocyte Potentiating Factor in Pancreatic and Biliary Cancers. *Clinical Chemistry and Laboratory Medicine* **2012**, *50*, 721–725, doi:10.1515/cclm.2011.816.

357.   Mehta, S.; Bhimani, N.; Gill, A.J.; Samra, J.S.; Sahni, S.; Mittal, A. Serum Biomarker Panel for Diagnosis and Prognosis of Pancreatic Ductal Adenocarcinomas. *Front. Oncol.* **2021**, *11*, 708963, doi:10.3389/fonc.2021.708963.

358.   Zhang, X.; Yu, Y.; Peer, C.J.; Landsman, R.; Skorupan, N.; Cao, L.; Alewine, C. Low Serum Mesothelin in Pancreatic Cancer Patients Results from Retention of Shed Mesothelin in the Tumor Microenvironment. *Translational Oncology* **2022**, *21*, 101440, doi:10.1016/j.tranon.2022.101440.

359.   Chen, Z.; Downing, S.; Tzanakakis, E.S. Four Decades After the Discovery of Regenerating Islet-Derived (Reg) Proteins: Current Understanding and Challenges. *Frontiers in Cell and Developmental Biology* **2019**, *7*, 1–16, doi:10.3389/fcell.2019.00235.

360.   Jin, C.X.; Hayakawa, T.; Ko, S.B.H.; Ishiguro, H.; Kitagawa, M. Pancreatic Stone Protein/Regenerating Protein Family in Pancreatic and Gastrointestinal Diseases. *Intern. Med.* **2011**, *50*, 1507–1516, doi:10.2169/internalmedicine.50.5362.

361.   Parikh, A.; Stephan, A.-F.; Tzanakakis, E.S. Regenerating Proteins and Their Expression, Regulation and Signaling. *Biomol Concepts* **2012**, *3*, 57–70, doi:10.1515/bmc.2011.055.

362.   Watanabe, T.; Yonekura, H.; Terazono, K.; Yamamoto, H.; Okamoto, H. Complete Nucleotide Sequence of Human Reg Gene and Its Expression in Normal and Tumoral Tissues. The Reg Protein, Pancreatic Stone Protein, and Pancreatic Thread Protein Are One and the Same Product of the Gene. *Journal of Biological Chemistry* **1990**, *265*, 7432–7439, doi:10.1016/S0021-9258(19)39132-X.

363.   Sanchez, D.; Figarella, C.; Marchand-Pinatel, S.; Bruneau, N.; Guy-Crotte, O. Preferential Expression of Reg Iβ Gene in Human Adult Pancreas. *Biochemical and Biophysical Research Communications* **2001**, *284*, 729–737, doi:10.1006/bbrc.2001.5033.

364.   Kimura, N.; Yonekura, H.; Okamofo, H.; Nagura, H. Expression of Human Regenerating Gene mRNA and Its Product in Normal and Neoplastic Human Pancreas. *Cancer* **1992**, *70*, 1857–1863, doi:10.1002/1097-0142(19921001)70:7<1857::AID-CNCR2820700708>3.0.CO;2-8.

365.   Satomura, Y.; Sawabu, N.; Ohta, H.; Watanabe, H.; Yamakawa, O.; Motoo, Y.; Okai, T.; Toya, D.; Makino, H.; Okamoto, H. The Immunohistochemical Evaluation of PSP/Reg-Protein in Normal and Diseased Human Pancreatic Tissues. *Int J Pancreatol* **1993**, *13*, 59–67, doi:10.1007/BF02795200.

366.   Said, S.; Kurtin, P.J.; Nasr, S.H.; Graham, R.P.; Dasari, S.; Vrana, J.A.; Yasir, S.; Torbenson, M.S.; Zhang, L.; Mounajjed, T.; et al. Carboxypeptidase A1 and Regenerating Islet-Derived 1α as New Markers for Pancreatic Acinar Cell Carcinoma. *Human Pathology* **2020**, *103*, 120–126, doi:10.1016/j.humpath.2020.07.019.

367.   Uhlén, M.; Fagerberg, L.; Hallström, B.M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, Å.; Kampf, C.; Sjöstedt, E.; Asplund, A.; et al. Tissue-Based Map of the Human Proteome. *Science* **2015**, *347*, 1260419, doi:10.1126/science.1260419.

368.   The Human Protein Atlas Tissue Expression of REG1A - Staining in Pancreas Available online: https://www.proteinatlas.org/ENSG00000115386-REG1A/tissue/pancreas (accessed on 30 June 2023).

369.   The Human Protein Atlas Tissue Expression of REG1B - Staining in Pancreas Available online: https://www.proteinatlas.org/ENSG00000172023-REG1B/tissue/pancreas (accessed on 30 June 2023).

370.   Eggimann, P.; Que, Y.-A.; Rebeaud, F. Measurement of Pancreatic Stone Protein in the Identification and Management of Sepsis. *Biomarkers in Medicine* **2019**, *13*, 135–145, doi:10.2217/bmm-2018-0194.

371.   de Hond, T.A.P.; Oosterheert, J.J.; van Hemert-Glaubitz, S.J.M.; Musson, R.E.A.; Kaasjager, K.A.H. Pancreatic Stone Protein as a Biomarker for Sepsis at the Emergency Department of a Large Tertiary Hospital. *Pathogens* **2022**, *11*, 559, doi:10.3390/pathogens11050559.

372.   De Reggi, M.; Gharib, B. Protein-X, Pancreatic Stone-, Pancreatic Thread-, Reg-Protein, P19, Lithostathine, and Now What. *Current Protein & Peptide Science* **2001**, *2*, 19–42, doi:10.2174/1389203013381233.

373.   De Caro, A.M.; Adricil, Z.; Fournet, B.; Capon, C.; Bonicel, J.J.; De Caro, J.D.; Rovery, M. N-Terminal Sequence Extension in the Glycosylated Forms of Human Pancreatic Stone Protein. The 5-Oxoproline N-Terminal Chain Is O-Glycosylated on the 5th Amino Acid Residue. *Biochimica et biophysica acta* **1989**, *994*, 281–284, doi:10.1016/0167-4838(89)90305-1.

374.   De Reggi, M.; Capon, C.; Gharib, B.; Wieruszeski, J.-M.; Michel, R.; Fournet, B. The Glycan Moiety of Human Pancreatic Lithostathine. Structure Characterization and Possible Pathophysiological Implications. *Eur J Biochem* **1995**, *230*, 503–510, doi:10.1111/j.1432-1033.1995.tb20589.x.

375.   Gerbaud, V.; Pignol, D.; Loret, E.; Bertrand, J.A.; Berland, Y.; Fontecilla-Camps, J.-C.; Canselier, J.-P.; Gabas, N.; Verdier, J.-M. Mechanism of Calcite Crystal Growth Inhibition by the N-Terminal Undecapeptide of Lithostathine. *Journal of Biological Chemistry* **2000**, *275*, 1057–1064, doi:10.1074/jbc.275.2.1057.

376.   Zhou, L.; Zhang, R.; Wang, L.; Shen, S.; Okamoto, H.; Sugawara, A.; Xia, L.; Wang, X.; Noguchi, N.; Yoshikawa, T.; et al. Upregulation of REG Iα Accelerates Tumor Progression in Pancreatic Cancer with Diabetes. *Int. J. Cancer* **2010**, *127*, 1795–1803, doi:10.1002/ijc.25188.

377.   Liu, Z.; Zhang, Y.; Xie, J.; Li, C.; Wang, X.; Shen, J.; Zhang, Y.; Wang, S.; Cheng, N. Regenerating Gene 1B Silencing Inhibits Colon Cancer Cell HCT116 Proliferation and Invasion. *Int J Biol Markers* **2015**, *30*, 217–225, doi:10.5301/jbm.5000133.

378.   Li, Q.; Wang, H.; Zogopoulos, G.; Shao, Q.; Dong, K.; Lv, F.; Nwilati, K.; Gui, X.; Cuggia, A.; Liu, J.-L.; et al. Reg Proteins Promote Acinar-to-Ductal Metaplasia and Act as Novel Diagnostic and Prognostic Markers in Pancreatic Ductal Adenocarcinoma. *Oncotarget* **2016**, *7*, 77838–77853, doi:10.18632/oncotarget.12834.

379.   Radon, T.P.; Massat, N.J.; Jones, R.; Alrawashdeh, W.; Dumartin, L.; Ennis, D.; Duffy, S.W.; Kocher, H.M.; Pereira, S.P.; Guarner (posthumous), L.; et al. Identification of a Three-Biomarker Panel in Urine for Early Detection of Pancreatic Adenocarcinoma. *Clinical Cancer Research* **2015**, *21*, 3512–3521, doi:10.1158/1078-0432.CCR-14-2467.

380. Debernardi, S.; O'Brien, H.; Algahmdi, A.S.; Malats, N.; Stewart, G.D.; Plješa-Ercegovac, M.; Costello, E.; Greenhalf, W.; Saad, A.; Roberts, R.; et al. A Combination of Urinary Biomarker Panel and PancRISK Score for Earlier Detection of Pancreatic Cancer: A Case–Control Study. *PLoS Med* **2020**, *17*, e1003489, doi:10.1371/journal.pmed.1003489.

381. Debernardi, S.; Blyuss, O.; Rycyk, D.; Srivastava, K.; Jeon, C.Y.; Cai, H.; Cai, Q.; Shu, X.-O.; Crnogorac-Jurcevic, T. Urine Biomarkers Enable Pancreatic Cancer Detection up to 2 Years before Diagnosis. *International Journal of Cancer* **2023**, *152*, 769–780, doi:10.1002/ijc.34287.

382. Porterfield, M.; Zhao, P.; Han, H.; Cunningham, J.; Aoki, K.; Von Hoff, D.D.; Demeure, M.J.; Pierce, J.M.; Tiemeyer, M.; Wells, L. Discrimination between Adenocarcinoma and Normal Pancreatic Ductal Fluid by Proteomic and Glycomic Analysis. *J. Proteome Res.* **2014**, *13*, 395–407, doi:10.1021/pr400422g.

383. Zhang, Y.; Yuan, X.; Zhu, X.; Wang, Q.; Yu, X.; Wei, Q.; Li, L. Serum REG Iα as a Potential Novel Biomarker in Cancer: An Observational Study. *Medicine* **2020**, *99*, e22281, doi:10.1097/MD.0000000000022281.

384. Bacon, S.; Kyithar, M.P.; Schmid, J.; Rizvi, S.R.; Bonner, C.; Graf, R.; Prehn, J.H.; Byrne, M.M. Serum Levels of Pancreatic Stone Protein (PSP)/reg1A as an Indicator of Beta-Cell Apoptosis Suggest an Increased Apoptosis Rate in Hepatocyte Nuclear Factor 1 Alpha (HNF1A-MODY) Carriers from the Third Decade of Life Onward. *BMC Endocr Disord* **2012**, *12*, 13, doi:10.1186/1472-6823-12-13.

385. Kobayashi, Y.; Tateno, H.; Dohra, H.; Moriwaki, K.; Miyoshi, E.; Hirabayashi, J.; Kawagishi, H. A Novel Core Fucose-Specific Lectin from the Mushroom Pholiota Squarrosa. *Journal of Biological Chemistry* **2012**, *287*, 33973–33982, doi:10.1074/jbc.M111.327692.

386. Rebelo, A.L.; Gubinelli, F.; Roost, P.; Jan, C.; Brouillet, E.; Van Camp, N.; Drake, R.R.; Saldova, R.; Pandit, A. Complete Spatial Characterisation of N-Glycosylation upon Striatal Neuroinflammation in the Rodent Brain. *J Neuroinflammation* **2021**, *18*, 116, doi:10.1186/s12974-021-02163-6.

387. DeLong, E.R.; DeLong, D.M.; Clarke-Pearson, D.L. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics* **1988**, *44*, 837, doi:10.2307/2531595.

388. Robin, X.; Turck, N.; Hainard, A.; Tiberti, N.; Lisacek, F.; Sanchez, J.-C.; Müller, M. pROC: An Open-Source Package for R and S+ to Analyze and Compare ROC Curves. *BMC Bioinformatics* **2011**, *12*, 77, doi:10.1186/1471-2105-12-77.

389. Campbell, M.P.; Royle, L.; Radcliffe, C.M.; Dwek, R.A.; Rudd, P.M. GlycoBase and autoGU: Tools for HPLC-Based Glycan Analysis. *Bioinformatics* **2008**, *24*, 1214–1216, doi:10.1093/bioinformatics/btn090.

390.   Abrahams, J.L.; Campbell, M.P.; Packer, N.H. Building a PGC-LC-MS N-Glycan Retention Library and Elution Mapping Resource. *Glycoconj J* **2018**, *35*, 15–29, doi:10.1007/s10719-017-9793-4.

391.   Rahib, L.; Smith, B.D.; Aizenberg, R.; Rosenzweig, A.B.; Fleshman, J.M.; Matrisian, L.M. Projecting Cancer Incidence and Deaths to 2030: The Unexpected Burden of Thyroid, Liver, and Pancreas Cancers in the United States. *Cancer Research* **2014**, *74*, 2913–2921, doi:10.1158/0008-5472.CAN-14-0155.

392.   Füzéry, A.K.; Levin, J.; Chan, M.M.; Chan, D.W. Translation of Proteomic Biomarkers into FDA Approved Cancer Diagnostics: Issues and Challenges. *Clinical Proteomics* **2013**, *10*, 13, doi:10.1186/1559-0275-10-13.

393.   Diamandis, E.P. The Failure of Protein Cancer Biomarkers to Reach the Clinic: Why, and What Can Be Done to Address the Problem? *BMC Med* **2012**, *10*, 87, doi:10.1186/1741-7015-10-87.

394.   Kulasingam, V.; Diamandis, E.P. Strategies for Discovering Novel Cancer Biomarkers through Utilization of Emerging Technologies. *Nat Rev Clin Oncol* **2008**, *5*, 588–599, doi:10.1038/ncponc1187.

395.   Badr, H.A.; AlSadek, D.M.; Darwish, A.A.; ElSayed, A.I.; Bekmanov, B.O.; Khussainova, E.M.; Zhang, X.; Cho, W.C.; Djansugurova, L.B.; Li, C.-Z. Lectin Approaches for Glycoproteomics in FDA-Approved Cancer Biomarkers. *Expert Review of Proteomics* **2014**, *11*, 227–236, doi:10.1586/14789450.2014.897611.

396.   Ueda, K. Glycoproteomic Strategies: From Discovery to Clinical Application of Cancer Carbohydrate Biomarkers. *PROTEOMICS – Clinical Applications* **2013**, *7*, 607–617, doi:10.1002/prca.201200123.

397.   Kuzmanov, U.; Kosanam, H.; Diamandis, E.P. The Sweet and Sour of Serological Glycoprotein Tumor Biomarker Quantification. *BMC Med* **2013**, *11*, 31, doi:10.1186/1741-7015-11-31.

398.   Wang, W.; Wei, C. Advances in the Early Diagnosis of Hepatocellular Carcinoma. *Genes Dis* **2020**, *7*, 308–319, doi:10.1016/j.gendis.2020.01.014.

399.   Gupta, S.; Bent, S.; Kohlwes, J. Test Characteristics of α-Fetoprotein for Detecting Hepatocellular Carcinoma in Patients with Hepatitis C. *Ann Intern Med* **2003**, *139*, 46–50, doi:10.7326/0003-4819-139-1-200307010-00012.

400.   Yi, X.; Yu, S.; Bao, Y. Alpha-Fetoprotein-L3 in Hepatocellular Carcinoma: A Meta-Analysis. *Clinica Chimica Acta* **2013**, *425*, 212–220, doi:10.1016/j.cca.2013.08.005.

401.   Best, J.; Bechmann, L.P.; Sowa, J.-P.; Sydor, S.; Dechêne, A.; Pflanz, K.; Bedreli, S.; Schotten, C.; Geier, A.; Berg, T.; et al. GALAD Score Detects Early Hepatocellular Carcinoma in an International Cohort of Patients With Nonalcoholic Steatohepatitis. *Clinical Gastroenterology and Hepatology* **2020**, *18*, 728-735.e4, doi:10.1016/j.cgh.2019.11.012.

402. Gratacós-Mulleras, A.; Duran, A.; Asadi Shehni, A.; Ferrer-Batallé, M.; Ramírez, M.; Comet, J.; de Llorens, R.; Saldova, R.; Llop, E.; Peracaula, R. Characterisation of the Main PSA Glycoforms in Aggressive Prostate Cancer. *Sci Rep* **2020**, *10*, 18974, doi:10.1038/s41598-020-75526-3.

403. Llop, E.; Ferrer-Batallé, M.; Barrabés, S.; Guerrero, P.E.; Ramírez, M.; Saldova, R.; Rudd, P.M.; Aleixandre, R.N.; Comet, J.; Llorens, R. de; et al. Improvement of Prostate Cancer Diagnosis by Detecting PSA Glycosylation-Specific Changes. *Theranostics* **2016**, *6*, 1190–1204, doi:10.7150/thno.15226.

404. Haga, Y.; Uemura, M.; Baba, S.; Inamura, K.; Takeuchi, K.; Nonomura, N.; Ueda, K. Identification of Multisialylated LacdiNAc Structures as Highly Prostate Cancer Specific Glycan Signatures on PSA. *Anal. Chem.* **2019**, *91*, 2247–2254, doi:10.1021/acs.analchem.8b04829.

405. Inoue, T.; Kaneko, T.; Muramatsu, S.; Kimura, H.; Yoshino, T.; Goto, T.; Sawada, A.; Akamatsu, S.; Kobayashi, T.; Yamasaki, T.; et al. LacdiNAc-Glycosylated Prostate-Specific Antigen Density Is a Potential Biomarker of Prostate Cancer. *Clinical Genitourinary Cancer* **2020**, *18*, e28–e36, doi:10.1016/j.clgc.2019.10.011.

406. Hagiwara, K.; Tobisawa, Y.; Kaya, T.; Kaneko, T.; Hatakeyama, S.; Mori, K.; Hashimoto, Y.; Koie, T.; Suda, Y.; Ohyama, C.; et al. Wisteria Floribunda Agglutinin and Its Reactive-Glycan-Carrying Prostate-Specific Antigen as a Novel Diagnostic and Prognostic Marker of Prostate Cancer. *Int J Mol Sci* **2017**, *18*, 261, doi:10.3390/ijms18020261.

407. Hatano, K.; Yoneyama, T.; Hatakeyama, S.; Tomiyama, E.; Tsuchiya, M.; Nishimoto, M.; Yoshimura, K.; Miyoshi, E.; Uemura, H.; Ohyama, C.; et al. Simultaneous Analysis of Serum A2,3-Linked Sialylation and Core-Type Fucosylation of Prostate-Specific Antigen for the Detection of High-Grade Prostate Cancer. *Br J Cancer* **2022**, *126*, 764–770, doi:10.1038/s41416-021-01637-x.

408. Pihikova, D.; Kasak, P.; Kubanikova, P.; Sokol, R.; Tkac, J. Aberrant Sialylation of a Prostate-Specific Antigen: Electrochemical Label-Free Glycoprofiling in Prostate Cancer Serum Samples. *Anal Chim Acta* **2016**, *934*, 72–79, doi:10.1016/j.aca.2016.06.043.

409. Ishikawa, T.; Yoneyama, T.; Tobisawa, Y.; Hatakeyama, S.; Kurosawa, T.; Nakamura, K.; Narita, S.; Mitsuzuka, K.; Duivenvoorden, W.; Pinthus, J.H.; et al. An Automated Micro-Total Immunoassay System for Measuring Cancer-Associated A2,3-Linked Sialyl N-Glycan-Carrying Prostate-Specific Antigen May Improve the Accuracy of Prostate Cancer Diagnosis. *Int J Mol Sci* **2017**, *18*, 470, doi:10.3390/ijms18020470.

410. Li, Q.K.; Chen, L.; Ao, M.-H.; Chiu, J.H.; Zhang, Z.; Zhang, H.; Chan, D.W. Serum Fucosylated Prostate-Specific Antigen (PSA) Improves the Differentiation of Aggressive from Non-Aggressive Prostate Cancers. *Theranostics* **2015**, *5*, 267–276, doi:10.7150/thno.10349.

411.   Salminen, L.; Nadeem, N.; Jain, S.; Grènman, S.; Carpén, O.; Hietanen, S.; Oksa, S.; Lamminmäki, U.; Pettersson, K.; Gidwani, K.; et al. A Longitudinal Analysis of CA125 Glycoforms in the Monitoring and Follow up of High Grade Serous Ovarian Cancer. *Gynecologic Oncology* **2020**, *156*, 689–694, doi:10.1016/j.ygyno.2019.12.025.

412.   Labrou, N.E. Protein Purification: An Overview. In *Protein Downstream Processing*; Labrou, N.E., Ed.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2014; Vol. 1129, pp. 3–10 ISBN 978-1-62703-976-5.

413.   Fitzgerald, J.; Leonard, P.; Darcy, E.; O'Kennedy, R. Immunoaffinity Chromatography. In *Protein Chromatography*; Walls, D., Loughran, S.T., Eds.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2011; Vol. 681, pp. 35–59 ISBN 978-1-60761-912-3.

414.   Harlow, E.; Lane, D. Immunoprecipitation: Purifying the Immune Complexes. *Cold Spring Harbor protocols* **2006**, *2006*, doi:10.1101/pdb.prot4536.

415.   Kaboord, B.; Perr, M. Isolation of Proteins and Protein Complexes by Immunoprecipitation. In *2D PAGE: Sample Preparation and Fractionation*; Posch, A., Ed.; Methods in Molecular Biology™; Humana Press: Totowa, NJ, 2008; Vol. 424, pp. 349–364 ISBN 978-1-58829-722-8.

416.   Le, D.T.; Brockstedt, D.G.; Nir-Paz, R.; Hampl, J.; Mathur, S.; Nemunaitis, J.; Sterman, D.H.; Hassan, R.; Lutz, E.; Moyer, B.; et al. A Live-Attenuated Listeria Vaccine (ANZ-100) and a Live-Attenuated Listeria Vaccine Expressing Mesothelin (CRS-207) for Advanced Cancers: Phase I Studies of Safety and Immune Induction. *Clinical Cancer Research* **2012**, *18*, 858–868, doi:10.1158/1078-0432.CCR-11-2121.

417.   Hassan, R.; Sharon, E.; Thomas, A.; Zhang, J.; Ling, A.; Miettinen, M.; Kreitman, R.J.; Steinberg, S.M.; Hollevoet, K.; Pastan, I. Phase 1 Study of the Antimesothelin Immunotoxin SS1P in Combination with Pemetrexed and Cisplatin for Front-Line Therapy of Pleural Mesothelioma and Correlation of Tumor Response with Serum Mesothelin, Megakaryocyte Potentiating Factor, and Cancer Antigen: SS1P Plus Chemotherapy for Mesothelioma. *Cancer* **2014**, *120*, 3311–3319, doi:10.1002/cncr.28875.

418.   Kreitman, R.J.; Hassan, R.; FitzGerald, D.J.; Pastan, I. Phase I Trial of Continuous Infusion Anti-Mesothelin Recombinant Immunotoxin SS1P. *Clin Cancer Res* **2009**, *15*, 5274–5279, doi:10.1158/1078-0432.CCR-09-0062.

419.   Hassan, R.; Alewine, C.; Pastan, I. New Life for Immunotoxin Cancer Therapy. *Clin Cancer Res* **2016**, *22*, 1055–1058, doi:10.1158/1078-0432.CCR-15-1623.

420.   Hollevoet, K.; Mason-Osann, E.; Liu, X.; Imhof-Jung, S.; Niederfellner, G.; Pastan, I. In Vitro and In Vivo Activity of the Low-Immunogenic Antimesothelin Immunotoxin RG7787 in Pancreatic Cancer. *Mol Cancer Ther* **2014**, *13*, 2040–2049, doi:10.1158/1535-7163.MCT-14-0089-T.

421.   Matsuzawa, F.; Kamachi, H.; Mizukami, T.; Einama, T.; Kawamata, F.; Fujii, Y.; Fukai, M.; Kobayashi, N.; Hatanaka, Y.; Taketomi, A. Mesothelin Blockage by Ama-tuximab Suppresses Cell Invasiveness, Enhances Gemcitabine Sensitivity and Regulates Cancer Cell Stemness in Mesothelin-Positive Pancreatic Cancer Cells. *BMC Cancer* **2021**, *21*, 200, doi:10.1186/s12885-020-07722-3.

422.   Hassan, R.; Ebel, W.; Routhier, E.L.; Patel, R.; Kline, J.B.; Zhang, J.; Chao, Q.; Jacob, S.; Turchin, H.; Gibbs, L.; et al. Preclinical Evaluation of MORAb-009, a Chimeric Antibody Targeting Tumor-Associated Mesothelin. *Cancer Immunity* **2007**, *7*, 20.

423.   Weekes, C.D.; Lamberts, L.E.; Borad, M.J.; Voortman, J.; McWilliams, R.R.; Diamond, J.R.; de Vries, E.G.E.; Verheul, H.M.; Lieu, C.H.; Kim, G.P.; et al. Phase I Study of DMOT4039A, an Antibody–Drug Conjugate Targeting Mesothelin, in Patients with Unresectable Pancreatic or Platinum-Resistant Ovarian Cancer. *Mol Cancer Ther* **2016**, *15*, 439–447, doi:10.1158/1535-7163.MCT-15-0693.

424.   Morello, A.; Sadelain, M.; Adusumilli, P.S. Mesothelin-Targeted CARs: Driving T Cells to Solid Tumors. *Cancer Discov* **2016**, *6*, 133–146, doi:10.1158/2159-8290.CD-15-0583.

425.   Hollevoet, K.; Reitsma, J.B.; Creaney, J.; Grigoriu, B.D.; Robinson, B.W.; Scher-pereel, A.; Cristaudo, A.; Pass, H.I.; Nackaerts, K.; Rodríguez Portal, J.A.; et al. Serum Mesothelin for Diagnosing Malignant Pleural Mesothelioma: An Individual Patient Data Meta-Analysis. *JCO* **2012**, *30*, 1541–1549, doi:10.1200/JCO.2011.39.6671.

426.   Ibrahim, D.A.; Abouhashem, N.S. Diagnostic Value of IMP3 and Mesothelin in Differentiating Pancreatic Ductal Adenocarcinoma from Chronic Pancreatitis. *Pathology - Research and Practice* **2016**, *212*, 288–293, doi:10.1016/j.prp.2016.01.007.

427.   Wang, Y.-N.; Lee, H.-H.; Hsu, J.L.; Yu, D.; Hung, M.-C. The Impact of PD-L1 N-Linked Glycosylation on Cancer Therapy and Clinical Diagnosis. *J Biomed Sci* **2020**, *27*, 77, doi:10.1186/s12929-020-00670-x.

428.   Uray, K.; Mizuno, M.; Inazu, T.; Goto, K.; Hudecz, F. The Effect of Glycosyla-tion on the Antibody Recognition of a MUC2 Mucin Epitope: Effect of Glycosylation on MUC2 Mucin Epitope. *Biopolymers* **2014**, *102*, 390–395, doi:10.1002/bip.22526.

429.   Ma, J.; Tang, W.K.; Esser, L.; Pastan, I.; Xia, D. Recognition of Mesothelin by the Therapeutic Antibody MORAb-009. *Journal of Biological Chemistry* **2012**, *287*, 33123–33131, doi:10.1074/jbc.M112.381756.

430.   Zhang, Y.-F.; Phung, Y.; Gao, W.; Kawa, S.; Hassan, R.; Pastan, I.; Ho, M. New High Affinity Monoclonal Antibodies Recognize Non-Overlapping Epitopes on Mesothe-lin for Monitoring and Treating Mesothelioma. *Scientific reports* **2015**, *5*, 9928, doi:10.1038/srep09928.

431.   Onda, M.; Willingham, M.; Nagata, S.; Bera, T.K.; Beers, R.; Ho, M.; Hassan, R.; Kreitman, R.J.; Pastan, I. New Monoclonal Antibodies to Mesothelin Useful for Im-munohistochemistry, Fluorescence-Activated Cell Sorting, Western Blotting, and ELISA. *Clinical cancer research* **2005**, *11*, 5840–5846, doi:10.1158/1078-0432.CCR-05-0578.

432. Inami, K.; Kajino, K.; Abe, M.; Hagiwara, Y.; Maeda, M.; Suyama, M.; Watanabe, S.; Hino, O. Secretion of N-ERC/Mesothelin and Expression of C-ERC/Mesothelin in Human Pancreatic Ductal Carcinoma. *Oncol Rep* **2008**, *20*, 1375–1380, doi:10.3892/or_00000155.

433. Argani, P.; Iacobuzio-Donahue, C.; Ryu, B.; Rosty, C.; Goggins, M.; Wilentz, R.E.; Murugesan, S.R.; Leach, S.D.; Jaffee, E.; Yeo, C.J.; et al. Mesothelin Is Overexpressed in the Vast Majority of Ductal Adenocarcinomas of the Pancreas: Identification of a New Pancreatic Cancer Marker by Serial Analysis of Gene Expression (SAGE). *Clinical cancer research* **2001**, *7*, 3862–3868.

434. Hassan, R.; Laszik, Z.G.; Lerner, M.; Raffeld, M.; Postier, R.; Brackett, D. Mesothelin Is Overexpressed in Pancreaticobiliary Adenocarcinomas But Not in Normal Pancreas and Chronic Pancreatitis. *American Journal of Clinical Pathology* **2005**, *124*, 838–845, doi:10.1309/F1B6-4CL7-H8VJ-KEAF.

435. Frank, R.; Li, S.; Ahmad, N.A.; Sepulveda, A.R.; Jhala, N.C. Mesothelin Expression in Pancreatic Mucinous Cysts. *American Journal of Clinical Pathology* **2014**, *142*, 313–319, doi:10.1309/AJCPDTTL2I5ECMFG.

436. Zhu, L.; Liu, Y.; Chen, G. Diagnostic Value of Mesothelin in Pancreatic Cancer: A Meta-Analysis. *International Journal of Clinical and Experimental Medicine* **2014**, *7*, 4000–4007.

437. Le, K.; Wang, J.; Zhang, T.; Guo, Y.; Chang, H.; Wang, S.; Zhu, B. Overexpression of Mesothelin in Pancreatic Ductal Adenocarcinoma (PDAC). *Int. J. Med. Sci.* **2020**, *17*, 422–427, doi:10.7150/ijms.39012.

438. Ho, M.; Onda, M.; Wang, Q.C.; Hassan, R.; Pastan, I.; Lively, M.O. Mesothelin Is Shed from Tumor Cells. *Cancer Epidemiology Biomarkers and Prevention* **2006**, *15*, 1751, doi:10.1158/1055-9965.EPI-06-0479.

439. Mann, B.F.; Goetz, J.A.; House, M.G.; Schmidt, C.M.; Novotny, M.V. Glycomic and Proteomic Profiling of Pancreatic Cyst Fluids Identifies Hyperfucosylated Lactosamines on the N-Linked Glycans of Overexpressed Glycoproteins*. *Molecular & Cellular Proteomics* **2012**, *11*, M111.015792, doi:10.1074/mcp.M111.015792.

440. Zhao, J.; Qiu, W.; Simeone, D.M.; Lubman, D.M. *N*- Linked Glycosylation Profiling of Pancreatic Cancer Serum Using Capillary Liquid Phase Separation Coupled with Mass Spectrometric Analysis. *J. Proteome Res.* **2007**, *6*, 1126–1138, doi:10.1021/pr0604458.

441. Deer, E.L.; González-Hernández, J.; Coursen, J.D.; Shea, J.E.; Ngatia, J.; Scaife, C.L.; Firpo, M.A.; Mulvihill, S.J. Phenotype and Genotype of Pancreatic Cancer Cell Lines. *Pancreas* **2010**, *39*, 425–435, doi:10.1097/MPA.0b013e3181c15963.

442. Altman, M.O.; Gagneux, P. Absence of Neu5Gc and Presence of Anti-Neu5Gc Antibodies in Humans—An Evolutionary Perspective. *Front Immunol* **2019**, *10*, 789, doi:10.3389/fimmu.2019.00789.

443.   Irie, A.; Koyama, S.; Kozutsumi, Y.; Kawasaki, T.; Suzuki, A. The Molecular Basis for the Absence ofN-Glycolylneuraminic Acid in Humans*. *Journal of Biological Chemistry* **1998**, *273*, 15866–15871, doi:10.1074/jbc.273.25.15866.

444.   Seo, N.; Ko, J.; Lee, D.; Jeong, H.; Oh, M.J.; Kim, U.; Lee, D.H.; Kim, J.; Choi, Y.J.; An, H.J. In-Depth Characterization of Non-Human Sialic Acid (Neu5Gc) in Human Serum Using Label-Free ZIC-HILIC/MRM-MS. *Anal Bioanal Chem* **2021**, *413*, 5227–5237, doi:10.1007/s00216-021-03495-1.

445.   Galili, U. Anti-Gal: An Abundant Human Natural Antibody of Multiple Pathogeneses and Clinical Benefits. *Immunology* **2013**, *140*, 1–11, doi:10.1111/imm.12110.

446.   Koike, C.; Uddin, M.; Wildman, D.E.; Gray, E.A.; Trucco, M.; Starzl, T.E.; Goodman, M. Functionally Important Glycosyltransferase Gain and Loss during Catarrhine Primate Emergence. *Proc Natl Acad Sci U S A* **2007**, *104*, 559–564, doi:10.1073/pnas.0610012104.

447.   Galili, U.; Rachmileuitz, E.; Peleg, A.; Flechner, I. A Unique Natural Human IgG Antibody with Anti-Alpha-Galactosyl Specificity. *J Exp Med* **1984**, *160*, 1519–1531.

448.   Ghaderi, D.; Zhang, M.; Hurtado-Ziola, N.; Varki, A. Production Platforms for Biotherapeutic Glycoproteins. Occurrence, Impact, and Challenges of Non-Human Sialylation. *Biotechnology and Genetic Engineering Reviews* **2012**, *28*, 147–176, doi:10.5661/bger-28-147.

449.   Liyanage, S.H.; Yan, M. Quantification of Binding Affinity of Glyconanomaterials with Lectins. *Chem Commun (Camb)* **2020**, *56*, 13491–13505, doi:10.1039/d0cc05899h.

450.   Kamada, Y.; Kinoshita, N.; Tsuchiya, Y.; Kobayashi, K.; Fujii, H.; Terao, N.; Kamihagi, K.; Koyama, N.; Yamada, S.; Daigo, Y.; et al. Reevaluation of a Lectin Antibody ELISA Kit for Measuring Fucosylated Haptoglobin in Various Conditions. *Clinica Chimica Acta* **2013**, *417*, 48–53, doi:10.1016/j.cca.2012.12.014.

451.   Ueda, M.; Kamada, Y.; Takamatsu, S.; Shimomura, M.; Maekawa, T.; Sobajima, T.; Fujii, H.; Nakayama, K.; Nishino, K.; Yamada, M.; et al. Specific Increase in Serum Core-Fucosylated Haptoglobin in Patients with Chronic Pancreatitis. *Pancreatology* **2016**, *16*, 238–243, doi:10.1016/j.pan.2016.01.004.

452.   Kusama, K.; Okamoto, Y.; Saito, K.; Kasahara, T.; Murata, T.; Ueno, Y.; Kobayashi, Y.; Kamada, Y.; Miyoshi, E. Reevaluation of Pholiota Squarrosa Lectin-Reactive Haptoglobin as a Pancreatic Cancer Biomarker Using an Improved ELISA System. *Glycoconjugate Journal* **2017**, *34*, 537–544, doi:10.1007/s10719-017-9772-9.

453.   Bailey, S.E.R.; Abel, G.A.; Atkins, A.; Byford, R.; Davies, S.-J.; Mays, J.; McDonald, T.J.; Miller, J.; Neck, C.; Renninson, J.; et al. Diagnostic Performance of a Faecal Immunochemical Test for Patients with Low-Risk Symptoms of Colorectal Cancer in Primary Care: An Evaluation in the South West of England. *Br J Cancer* **2021**, *124*, 1231–1236, doi:10.1038/s41416-020-01221-9.

454.   Chari, S.T.; LEIBSON, C.L.; RABE, K.G.; RANSOM, J.; DE ANDRADE, M.; PETERSEN, G.M. Probability of Pancreatic Cancer Following Diabetes: A Population-Based Study. *Gastroenterology* **2005**, *129*, 504–511, doi:10.1053/j.gastro.2005.05.007.

455.   Vujasinovic, M.; Dugic, A.; Maisonneuve, P.; Aljic, A.; Berggren, R.; Panic, N.; Valente, R.; Pozzi Mucelli, R.; Waldthaler, A.; Ghorbani, P.; et al. Risk of Developing Pancreatic Cancer in Patients with Chronic Pancreatitis. *J Clin Med* **2020**, *9*, 3720, doi:10.3390/jcm9113720.

456.   McWilliams, R.R.; Rabe, K.G.; Olswold, C.; De Andrade, M.; Petersen, G.M. Risk of Malignancy in First-Degree Relatives of Patients with Pancreatic Carcinoma. *Cancer* **2005**, *104*, 388–394, doi:10.1002/cncr.21166.

457.   Brand, R.E.; Persson, J.; Bratlie, S.O.; Chung, D.C.; Katona, B.W.; Carrato, A.; Castillo, M.; Earl, J.; Kokkola, A.; Lucas, A.L.; et al. Detection of Early-Stage Pancreatic Ductal Adenocarcinoma From Blood Samples: Results of a Multiplex Biomarker Signature Validation Study. *Clin Transl Gastroenterol* **2022**, *13*, e00468, doi:10.14309/ctg.0000000000000468.

458.   Yang, Y.; Franc, V.; Heck, A.J.R. Glycoproteomics: A Balance between High-Throughput and In-Depth Analysis. *Trends in Biotechnology* **2017**, *35*, 598–609, doi:10.1016/j.tibtech.2017.04.010.

459.   Tabang, D.N.; Ford, M.; Li, L. Recent Advances in Mass Spectrometry-Based Glycomic and Glycoproteomic Studies of Pancreatic Diseases. *Front. Chem.* **2021**, *9*, 707387, doi:10.3389/fchem.2021.707387.

460.   Sun, C.; Wang, X.; Hui, Y.; Fukui, H.; Wang, B.; Miwa, H. The Potential Role of REG Family Proteins in Inflammatory and Inflammation-Associated Diseases of the Gastrointestinal Tract. *Int J Mol Sci* **2021**, *22*, 7196, doi:10.3390/ijms22137196.

461.   Grégoire, C.; Marco, S.; Thimonier, J.; Duplan, L.; Laurine, E.; Chauvin, J.-P.; Michel, B.; Peyrot, V.; Verdier, J.-M. Three-Dimensional Structure of the Lithostathine Protofibril, a Protein Involved in Alzheimer's Disease. *EMBO J* **2001**, *20*, 3313–3321, doi:10.1093/emboj/20.13.3313.

462.   Lebart, M.-C.; Trousse, F.; Valette, G.; Torrent, J.; Denus, M.; Mestre-Frances, N.; Marcilhac, A. Reg-1α, a New Substrate of Calpain-2 Depending on Its Glycosylation Status. *IJMS* **2022**, *23*, 8591, doi:10.3390/ijms23158591.

463.   Munkley, J. Aberrant Sialylation in Cancer: Therapeutic Opportunities. *Cancers (Basel)* **2022**, *14*, 4248, doi:10.3390/cancers14174248.

464.   Astorri, E.; Guglielmi, C.; Bombardieri, M.; Alessandri, C.; Buzzetti, R.; Maggi, D.; Valesini, G.; Pitzalis, C.; Pozzilli, P. Circulating Reg1α Proteins and Autoantibodies to Reg1α Proteins as Biomarkers of β-Cell Regeneration and Damage in Type 1 Diabetes. *Horm Metab Res* **2010**, *42*, 955–960, doi:10.1055/s-0030-1267206.

465.   Huang, Y.-F.; Aoki, K.; Akase, S.; Ishihara, M.; Liu, Y.-S.; Yang, G.; Kizuka, Y.; Mizumoto, S.; Tiemeyer, M.; Gao, X.-D.; et al. Global Mapping of Glycosylation Pathways in Human-Derived Cells. *Developmental Cell* **2021**, *56*, 1195-1209.e7, doi:10.1016/j.devcel.2021.02.023.

466.   Yang, X.; Tao, S.; Orlando, R.; Brockhausen, I.; Kan, F.W.K. Structures and Biosynthesis of the N- and O-Glycans of Recombinant Human Oviduct-Specific Glyco-protein Expressed in Human Embryonic Kidney Cells. *Carbohydr Res* **2012**, *358*, 47–55, doi:10.1016/j.carres.2012.05.027.

467.   Zhang, Q.; Li, Z.; Wang, Y.; Zheng, Q.; Li, J. Mass Spectrometry for Protein Sialoglycosylation. *Mass Spectrometry Reviews* **2018**, *37*, 652–680, doi:10.1002/mas.21555.

468.   de Haan, N.; Reiding, K.R.; Haberger, M.; Reusch, D.; Falck, D.; Wuhrer, M. Linkage-Specific Sialic Acid Derivatization for MALDI-TOF-MS Profiling of IgG Glyco-peptides. *Anal. Chem.* **2015**, *87*, 8284–8291, doi:10.1021/acs.analchem.5b02426.

469.   Moran, A.B.; Gardner, R.A.; Wuhrer, M.; Lageveen-Kammeijer, G.S.M.; Spen-cer, D.I.R. Sialic Acid Derivatization of Fluorescently Labeled N-Glycans Allows Linkage Differentiation by Reversed-Phase Liquid Chromatography–Fluorescence Detection–Mass Spectrometry. *Anal Chem* **2022**, *94*, 6639–6648, doi:10.1021/acs.analchem.1c02610.

470.   Nishikaze, T.; Tsumoto, H.; Sekiya, S.; Iwamoto, S.; Miura, Y.; Tanaka, K. Dif-ferentiation of Sialyl Linkage Isomers by One-Pot Sialic Acid Derivatization for Mass Spectrometry-Based Glycan Profiling. *Anal. Chem.* **2017**, *89*, 2353–2360, doi:10.1021/acs.analchem.6b04150.

471.   Jiang, K.; Zhu, H.; Li, L.; Guo, Y.; Gashash, E.; Ma, C.; Sun, X.; Li, J.; Zhang, L.; Wang, P.G. Sialic Acid Linkage-Specific Permethylation for Improved Profiling of Pro-tein Glycosylation by MALDI-TOF MS. *Anal Chim Acta* **2017**, *981*, 53–61, doi:10.1016/j.aca.2017.05.029.

472.   Fukuyama, Y.; Funakoshi, N.; Takeyama, K.; Hioki, Y.; Nishikaze, T.; Ka-neshiro, K.; Kawabata, S.; Iwamoto, S.; Tanaka, K. 3-Aminoquinoline/p-Coumaric Acid as a MALDI Matrix for Glycopeptides, Carbohydrates, and Phosphopeptides. *Anal. Chem.* **2014**, *86*, 1937–1942, doi:10.1021/ac4037087.

473.   Watanabe, M.; Terasawa, K.; Kaneshiro, K.; Uchimura, H.; Yamamoto, R.; Fu-kuyama, Y.; Shimizu, K.; Sato, T.-A.; Tanaka, K. Improvement of Mass Spectrometry Analysis of Glycoproteins by MALDI-MS Using 3-Aminoquinoline/α-Cyano-4-Hydro-xycinnamic Acid. *Anal Bioanal Chem* **2013**, *405*, 4289–4293, doi:10.1007/s00216-013-6771-y.

474.   Gilar, M.; Yu, Y.-Q.; Ahn, J.; Xie, H.; Han, H.; Ying, W.; Qian, X. Characteri-zation of Glycoprotein Digests with Hydrophilic Interaction Chromatography and Mass Spectrometry. *Analytical Biochemistry* **2011**, *417*, 80–88, doi:10.1016/j.ab.2011.05.028.

475.   Domínguez-Vega, E.; Tengattini, S.; Peintner, C.; van Angeren, J.; Temporini, C.; Haselberg, R.; Massolini, G.; Somsen, G.W. High-Resolution Glycoform Profiling of Intact Therapeutic Proteins by Hydrophilic Interaction Chromatography-Mass Spectrometry. *Talanta* **2018**, *184*, 375–381, doi:10.1016/j.talanta.2018.03.015.

476.   Magalhães, A.; Duarte, H.O.; Reis, C.A. The Role of O-Glycosylation in Human Disease. *Molecular Aspects of Medicine* **2021**, *79*, 100964, doi:10.1016/j.mam.2021.100964.

477.   Bellis, S.L.; Reis, C.A.; Varki, A.; Kannagi, R.; Stanley, P. Glycosylation Changes in Cancer. In *Essentials of Glycobiology*; Varki, A., Cummings, R.D., Esko, J.D., Stanley, P., Hart, G.W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N.H., Prestegard, J.H., Schnaar, R.L., Seeberger, P.H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2022 ISBN 978-1-62182-421-3.