Glycoproteomics-based signatures of cancer cell lines

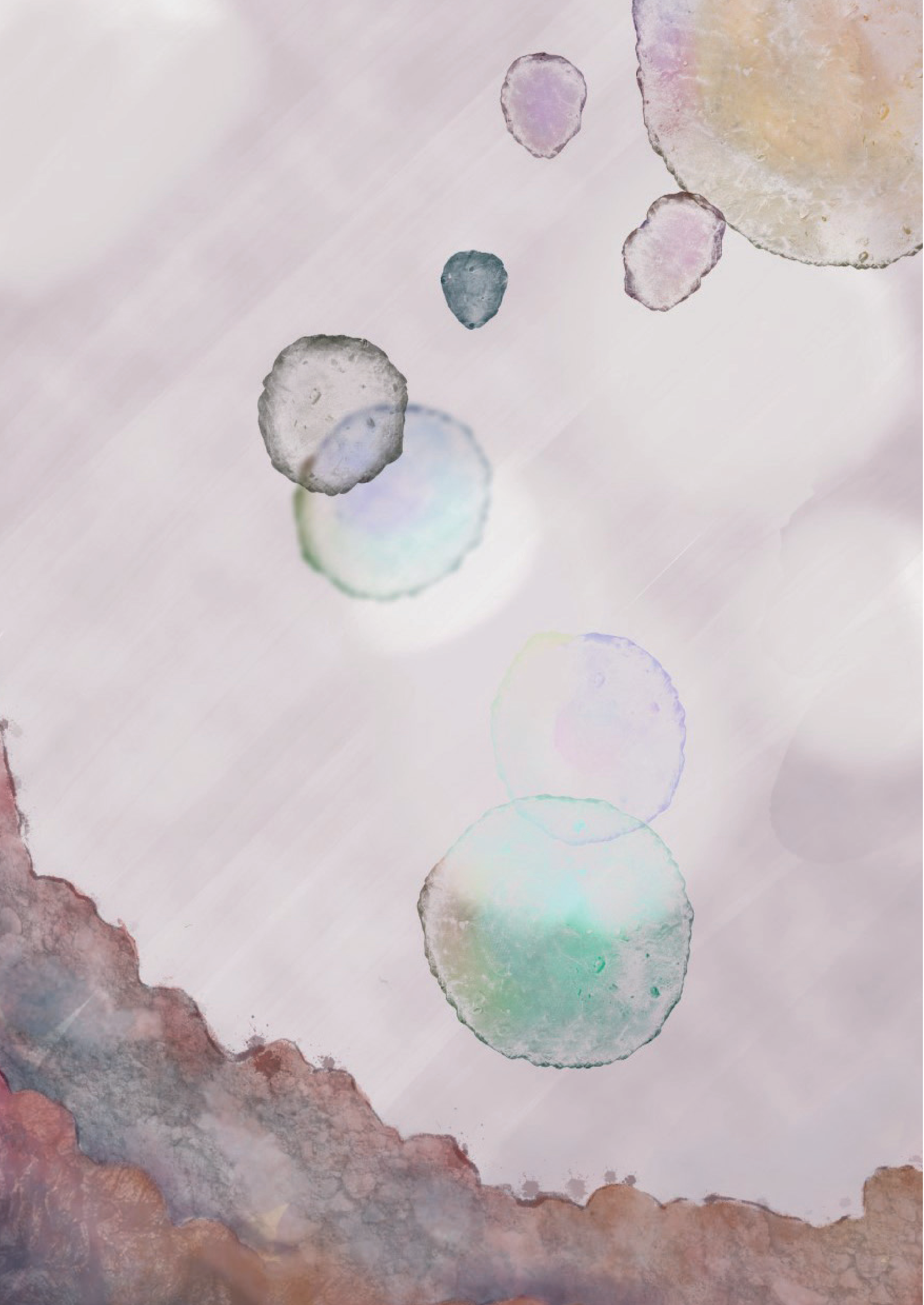
Pirro, M.

Citation

Pirro, M. (2021, September 21). *Glycoproteomics-based signatures of cancer cell lines*. Retrieved from https://hdl.handle.net/1887/3212951

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3212951

Note: To cite this publication please use the final published version (if applicable).



CHAPTER 5

OPTIMIZATION OF PROTEOMICS WORKFLOWS ALLOWS FOR SPECIFIC IDENTIFICATION OF SECRETED MGL BINDING PROTEINS FROM COLORECTAL CANCER CELL LINES

Martina Pirro[1], Yoann Rombouts[2], Yassene Mohammed[1], Sandra J. van Vliet[3], Arnoud H. de Ru[1], George M. C. Janssen[1], Manfred Wuhrer[1], Peter A. van Veelen[1], Paul J. Hensbergen[1*]

[1] Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, The Netherlands

[2] Institut de Pharmacologie et de Biologie Structurale, Université de Toulouse, CNRS, UPS, Toulouse, France

[3] Amsterdam UMC, Vrije Universiteit Amsterdam, dept. of Molecular Cell Biology and Immunology, Cancer Center Amsterdam, Amsterdam Infection & Immunity Institute, Amsterdam, the Netherlands

Abstract

Colorectal cancer (CRC) is the third leading cause of death worldwide, and currently lacks clinically relevant biomarkers with high sensitivity and specificity, in both diagnosis and prognosis. Ideal tumor markers are tissue specific molecules produced and secreted or shed into the surrounding tissue or in the circulation. Hence, the secretome of cancer cells has emerged as a promising source for biomarker discovery. In this study, we applied lectin enrichment in combination with LC-MS/MS-based glycoproteomics , to study the secretomes of three CRC cell lines, HCT116, HT29 and LS174T. By performing comparative analysis between the conditional media and whole cell lysate, as well as between conditional media obtained at different timings of incubation, our method allowed us to optimize the enrichment of secreted proteins and reduce the intracellular contaminants. Using MGL lectin, we then captured the secreted ligands of the CRC cell lines and analyzed them with mass spectrometry. The analysis revealed that all three cell lines secrete glycoproteins carrying the specific MGL epitope, which could have a role in tumor-host cells interaction and subsequent immune response to CRC.

Introduction

Colorectal cancer is the third most deadly and the fourth most common cancer worldwide [47]. CRC incidence have been increasing in developing countries due to a more "western" lifestyle, where obesity, sedentarism, meat and alcohol consumption are driving factors behind CRC progression [47]. At the same time, CRC-related mortality has reduced thanks to advances especially in early detection screening and treatment options [48]. However, current CRC serum biomarkers, such as carcinoembryonic antigen (CEA) and carbohydrate antigen 19-9 (CA 19-9) have low specificity and sensitivity for CRC diagnosis [48]. Hence, there is a clinical need for more sensitive and specific serum biomarkers for CRC.

Proteins secreted by tumors found in circulation or other body fluids are involved in extracellular matrix (ECM) degradation, cell proliferation, invasion, angiogenesis, evasion of host immune responses and many other processes influencing tumor progression and metastasis. For this reason, human blood and body fluids represent a good source for the discovery of potential biomarkers or therapeutic targets. However, biomarker discovery from human blood serum is challenging due to the high abundance serum proteins (such as albumin, immunoglobulins, transferrin, fibrinogen, and haptoglobin) that can mask the low abundant potential biomarkers. A good alternative used in proteomic studies is the analysis of the cancer cells "secretome". The term secretome refers to the overall proteins secreted by or shed from the surface of tumor cells into the microenvironment, which can, eventually, go into circulation [120]. Cancer cell lines secretomes are usually obtained incubating cell lines in serum-free media for a certain period of time, typically 24 hours (h). Subsequently, the so-called conditioned media (CM) is harvested and concentrated using centrifugal filters [121]. This is important in order to avoid masking of low abundance secreted proteins by the fetal bovine serum proteins. However, starving conditions are likely to induce spontaneous lysis, causing intracellular proteins release into the CM [122]. Since secretomes are inevitably contaminated by intracellular proteins, it is vital to establish the optimal incubation time and a filtering criterion to select *bona fide* secreted proteins in the subsequent analysis. Another technical obstacle of using cell lines comes from the absence of host-tumor microenvironment. On the other hand, cell lines secretomes have the advantage of being immediately available from various cell lines mimicking different tumor stages/histotypes and allowing the identification of low abundant tumor-specific biomarkers compared to the more

complex blood source [123]. In the current study, we provide an optimized and adapted method to obtain secretome from three colorectal cancer (CRC) cell lines (HCT116, HT29 and LS174T), as reservoir of glycoproteins for lectin affinity purification.

Recently, we characterized the same cell lines for the cell surface expression of glycoproteins binding to the C-type macrophage galactose-type lectin (MGL/CLEC10/CD301) [110]. MGL, expressed by tolerogenic dendritic cells (DCs) and macrophages, exclusively recognizes terminal GalNAc residues on *N*- and/or *O*-glycans, such as LacdiNAc epitope (GalNAcβ1,4GlcNAcβ1-) and (S)Tn ((NeuAcα2,6)GalNAcα1-Ser/Thr) antigens [38], well-known tumor-associated antigens (TACA) [11]. In CRC, their higher expression represents a hallmark of poor prognosis in patients with stage III of the disease [45]. Moreover, MGL can be used as a tool to discriminate healthy tissues from CRC ones [45] as well as CRC cell lines with different molecular characteristics (e.g. BRAF mutation involved in MGL ligands expression [83]) and subtypes (colon-like versus undifferentiated [52]) [110]. These CRC cell lines have been fully characterized for their MGL binding profile (high for HCT116 and HT29, or low for LS174T), for the identity of MGL binding proteins [110], the contribution of *N*- and *O*-glycosylation involved in lectin binding and mechanisms behind their expression [124]. However, primary tumors comprehend not only the malignant cells but also a wide variety of stromal cells, which contribute to the development and progression of cancer [123]. Notwithstanding, we hypothesized that also the cancer secretome, including growth factors, enzymes and shed receptors, could have a role in the creation of a favorable *milieu* for the progression of the disease, interacting with host-cells, such as MGL expressing DCs. For this reason, after enriching the secretome of the above-mentioned CRC cell lines, our aim was first to evaluate MGL binding to the secreted glycoproteins. Next, we determined the identity of MGL binding proteins using our established MGL enrichment method combined with (glyco)proteomics.

Material and Methods

CRC cell lines culture

In order to mimic tumor heterogeneity, we selected three colorectal cancer cell lines with different molecular characteristics and stage as in Pirro et al. [110]. In brief, HT29 contains the BRAFV600E mutation and belongs to colon-like subtype and grade I adenocarcinoma [110]; LS174T originates from same subtype and grade as HT29 but lacks of the BRAF mutation; HCT116 has undifferentiated classification and originates from a grade IV carcinoma [110] without the BRAF mutation. HCT116 and HT29 were provided by the Department of Surgery of the Leiden University Medical Center (Leiden, The Netherlands), whereas LS174T was obtained from Amsterdam UMC (Amsterdam, the Netherlands). The forensic laboratory for DNA-research (ISO 17025) performed cell line authentication, using short-tandem repeat (STR) profiling. All cell lines matched for 100% with the known profile [125]. For optimal culture conditions, HCT116 and HT29 were cultivated in RPMI-1640 and LS174T in DMEM media containing L-glutamine, 10% fetal bovine serum (FBS) (Invitrogen) and streptomycin/penicillin (Sigma-Aldrich) at 5% CO2 and 37°C.

Secretome collection and cell lysates

Cells were cultured upon 70% of confluence in a T175 cm^2 flask in medium containing 10% FBS and antibiotics. Cell were washed 3 times with 1x PBS solution and once with serum-free media, followed by 6 h or 1 h pre-incubation with serum-free media. Subsequently, the media was discarded, and cells incubated with serum-free media for 6 hours or 24 hours until conditioned media (CM) collection. The CMs were centrifuged at 300 g for 10 min and subsequently passed through a 0.45 µm filter (Millex-HV, Millipore, Amsterdam, The Netherlands) to remove cell debris. Finally, the CM was concentrated from 20 ml to approximately 300 µl using a 10 KDa MWCO centrifugal concentrator (Amicon ultra, Millipore, Amsterdam, The Netherlands), according to product information. Protein quantification was performed using Bradford assay (Thermo Fisher Scientific, Massachusetts, United States) and samples were frozen at -80°C until use. Meanwhile, whole cellular proteins were obtained after CM harvesting. Cells were gently rinsed with 1× PBS and incubated for approximately 5 min in 1× trypsin/EDTA solution in 1× PBS, then inhibited by the addition of serum containing

medium. Then, cells were harvested and lysed as described before [125], for 20 min on ice in lysis buffer (10 mM triethanolamine pH 8.2, 150 mM NaCl, 1 mM MgCl2, 1 mM CaCl2 and 1% (volume/volume) Triton X-100, containing EDTA-free protease inhibitor (Roche Diagnostics)). BCA assay (BCA Protein Assay Kit, Pierce™) was used for protein quantification, following the manufacturer's instructions.

Lectins and Antibodies

Chimeric MGL-Fc and mutant-MGL-Fc (MGL short H259T-Fc, obtained by site-directed mutagenesis) were prepared as previously described [38] and provided by Amsterdam UMC, Amsterdam, the Netherlands. Lectins were targeted with a peroxidase-conjugated goat anti-human IgG-Fcγ secondary antibody (Jackson Immuno Research, West Grove, PA, USA, dilution 1:1500).

SDS-PAGE and Lectin blot

Twenty µg of concentrated CRC secretomes were separated by SDS-PAGE (4–15% Mini-PROTEAN® TGX Stain-Free™ Protein Gels, Bio-Rad) and transferred to a PVDF membrane (Trans-Blot Turbo Mini PVDF Transfer Packs, Bio-Rad, Hercules, CA, USA). Five % bovine serum albumin (BSA, Sigma-Aldrich) in TSM buffer (20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM CaCl2 and 2 mM MgCl2) was used as blocking buffer. Lectin staining was performed in 1% BSA TSM, followed by washes and incubation with peroxidase-conjugated secondary antibody in TSM with 0.1% Tween-20. Detection was obtained by enhanced chemiluminescence (ECL) using Clarity Western ECL substrate (Bio-Rad) and an Amersham Imager 600 (Cytiva, Marlborough, United States).

Pull-down assay

MGL ligands were pulled down from 175 µg of secreted proteins with 2 µg of chimeric MGL-Fc or MGL short H259T mutant-Fc coupled to 50 µl Dynabeads protein G (Invitrogen) as previously described [110]. Specific ligands were, subsequently, eluted by adding 100 mM EDTA. Elutions were dried under vacuum for 1 h.

SDS-PAGE and LC-MS/MS analysis

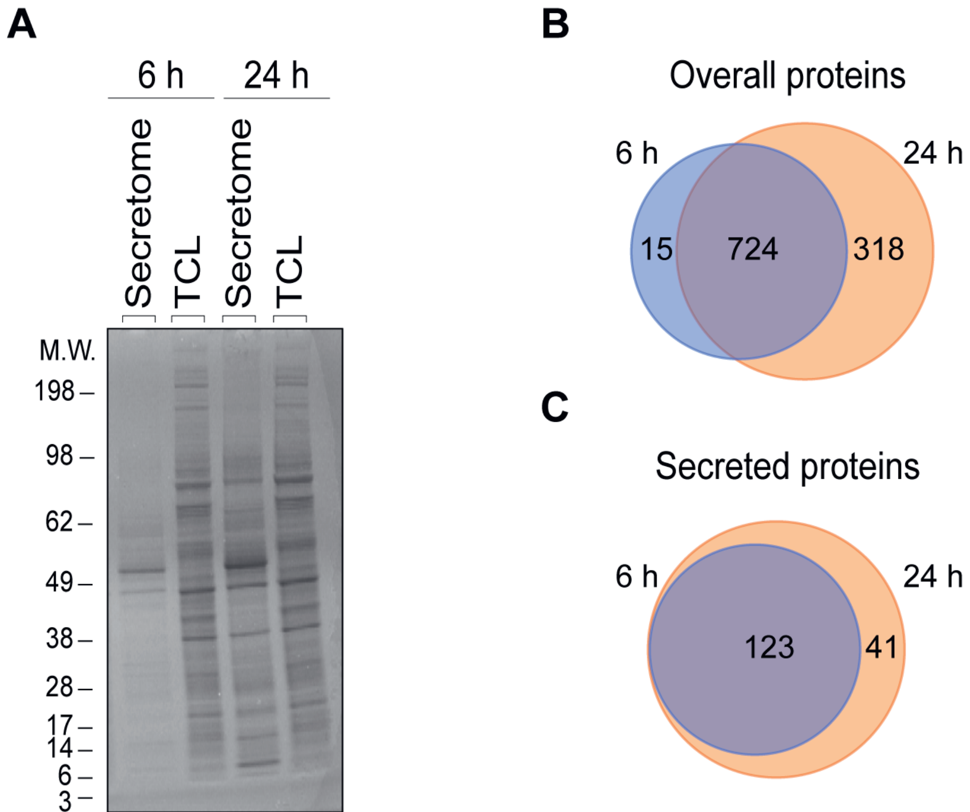
MGL binding proteins from CRC secretomes were cleaned up by a short SDS-PAGE run (NuPAGE™ 4–12% Bis-Tris Protein Gels, Thermo Fisher Scientific). Gel staining was performed with SimplyBlue™ Safe Stain (Invitrogen) 1 h incubation at room temperature (RT), followed by distilled water washes for 3 h. Gel bands corresponding to each sample lane were cut and subjected to dithiothreitol (10 mM) reduction, iodoacetamide (50 mM) alkylation and in-gel trypsin (Worthington Enzymes Lakewood, NJ, US) digestion, using a Proteineer DP digestion robot (Bruker, Billerica, MA, US). Tryptic peptides were analyzed by C18 nano-High Performance Liquid Chromatography (HPLC)-MS/MS with an Easy nLC 1000 gradient HPLC system (Thermo, Bremen, Germany) coupled to an Orbitrap Fusion LUMOS mass spectrometer (Thermo), as previously described [110]. Briefly, fractions, loaded onto a homemade precolumn, were eluted via a homemade analytical nano-HPLC column for 20 min and subjected to electrospray injection into the mass spectrometer. MS/MS spectra were acquired in data-dependent top-10 mode with a normalized collision energy of 32%, followed by MS/MS spectrum recording in the Orbitrap. Protein identification was obtained converting raw data in to mzXML using Proteowizard software. Peptide/protein identification and subsequent statistical validation was performed using the Trans Proteomics Pipeline version 5.1.0 with X! Tandem Jackhammer TPP (version 2013.06.15.1-LabKey, Insilicos, ISB) as search engine. The parameters used were: 10 ppm and 0.04 Da for, respectively, precursor and fragment mass tolerance, carbamidomethyl (Cys) as fixed modification and oxidation (Met) as variable one. A false discovery rate (FDR) threshold of 1% as well as a minimum of two peptides per protein were used. R version 3.4.4 was used for data extraction and table generation.

Results

Optimization of secretome collection in CRC cell line

First, we optimized a suitable method to collect secreted proteins from the media of colorectal cancer cell line cultures. For this purpose, we used the CRC cell line HCT116. We focused on different challenges of sample preparation: reducing the contamination from the fetal bovine serum (FBS) as well as intracellular protein and detection of low abundant

proteins. After reaching 70% confluency of HCT116 cell culture in presence of FBS, we rinsed and incubated the cells for 6 h in serum-free media. This media was discarded and fresh media was added for 24 h for CM collection. Since the proteins are diluted in the media, the CM was concentrated through a 10 kDa filter to obtain 300 µl from 20 ml of media. To screen for major contamination of intracellular proteins in our samples, we analyzed the CM by SDS-PAGE and compared this with the total proteome obtained by cell lysis of the same cells (Supplemental figure S1). This analysis revealed a substantial contamination corresponding to intracellular proteins as shown by the similar bands pattern of CM and total proteome. This result suggested that the starvation induced excessive cell death. To prevent this effect, we decreased the pre-incubation time from 6 h to 1 h. After discarding this sample, cells were then incubated for 6 or 24 h and analyzed again by SDS-PAGE, compared with the corresponding total cell lysates (Figure 1A). This analysis shows a characteristic pattern for the CM samples, which differs from the corresponding cell lysate, suggesting a lower contamination from intracellular proteins using 1 h of pre-incubation. Moreover, as shown by the intensity of the gel staining, 24 h of incubation results in higher abundance of secreted proteins compared to the 6 h sample. To gain more insight in the proteins present in our samples, we used LC-MS/MS, after on-filter tryptic digestion. We were able to identify 739 and 1042 total proteins after 6 and 24 h of incubation, respectively (Figure 1B). Within these, 724 were in common between the 2 samples, 15 exclusive of the 6 h CM and 318 identified only in the 24 h CM. To analyze the proteins that are truly secreted by HCT116, we compared only proteins annotated as “extracellular” in the Gene Ontology database. Overall, 123 secreted proteins were identified in both samples, with additional 41 secreted proteins found only after 24 h of incubation. In conclusion, these results show a similar contamination with intracellular proteins in both samples but higher abundance and number of protein IDs after 24 h of incubation. For this reason, we decided to select 1 h and 24 h as pre- and post-incubation time, respectively, for the collection of secreted proteins from HCT116 and other two CRC cell lines (HT29 and LS174T).



5

Figure 1: Analysis of secreted proteins from HCT116 cells. *A) SDS-PAGE of HCT116 conditioned medium (secretome), collected after 6 or 24 hours (h) and total cell lysate (TCL). Prior to the collection of conditioned medium, cells were incubated in serum-free media for 1h. B) Venn-diagram presenting the overall number of proteins identified in samples of conditioned medium collected after 6 h (in blue) and 24 h (in orange), respectively. C) Venn-diagram of the annotated "secreted" proteins in the two samples. M.W.: molecular weight.*

<u>MGL binding to secreted proteins by CRC cell lines</u>

Next, we evaluated MGL binding to the secretome of CRC cell lines by WB (Figure 2), using a recombinant Fc-coupled human MGL or mutant-MGL and secondary antibody (Ab) alone as negative controls. Secretomes from HCT116 and HT29 show specific bands stained by MGL, at different molecular weights. LS174T showed only one specific band around 62 kDa and a broader smear in the higher molecular weight, which appeared also in the negative control with mutant-MGL. This suggests that MGL staining to the higher molecular weight proteins

secreted by LS174T might be aspecific. On the other hand, aspecific binding by the secondary Ab alone was negligible in all three cell lines.

Knowing that MGL ligands can also be secreted by CRC cell lines, we, then, investigated their identity. CM from the three cell lines was used for MGL pull-downs with our previously established method [110]. Mutant-MGL was used as negative control for the pull-downs. After EDTA-elution, the specific ligands were loaded and run in a short SDS PAGE, followed by in-gel tryptic digestion and LC-MS/MS (Supplemental figure 1B). Altogether, these experiments resulted in the identification 707 proteins from all cell lines and triplicates combined. Within these, we filtered for proteins specifically binding to MGL by their identification at least twice in the MGL pull-down and at most once in the negative control. In order to avoid the selection of intracellular or cell surface contaminants, we next selected only for those proteins annotated as "extracellular" or "secreted" in the Gene Ontology database. As a result, 81, 102 and 57 secreted proteins were identified as MGL binders in HCT116, HT29 and LS174T respectively (Table S1). In line with the WB (Figure 2), LS174T was found to secrete proteins carrying the binding epitope, in contrast to what we previously found for cell surface proteins [110].

Interestingly, only few proteins appeared to be in common between two cell lines. Of note, HT29 shows ligands in common with both HCT116 and LS174T. However, HCT116 and LS174T show only one common MGL binding protein, the secreted isoform of c-Met receptor, already found in HCT116 and HT29 as MGL binding protein within the cell surface MGL binders [110].

In order to investigate the glycosylated epitopes carried by secreted proteins captured by MGL, we used an automated software, Byonic to annotate *N*- and *O*-glycans of glycopeptides. The overall analysis resulted in the confident identification of 86 *N*-glycopeptides and 32 *O*-glycopeptides in the three cell lines (Supplemental Table S2 and S3). Three of those glycopeptides belong to identified secreted MGL ligands, MET, DSC2 and CSPG4. However, none of these carried the specific terminal GalNAc, required for MGL recognition. Interestingly, one tryptic glycopeptide obtained from the secreted metallopeptidase inhibitor 1 (TIMP1) carried the LacdiNAc epitope on *N*-glycans (HexNAc(6)Hex(3)), fucosylated (in HT29) or not (in HCT116). However, TIMP1 was not selected as specific MGL binder in both cell lines, because it was also found in the pull-downs using mutant-MGL.

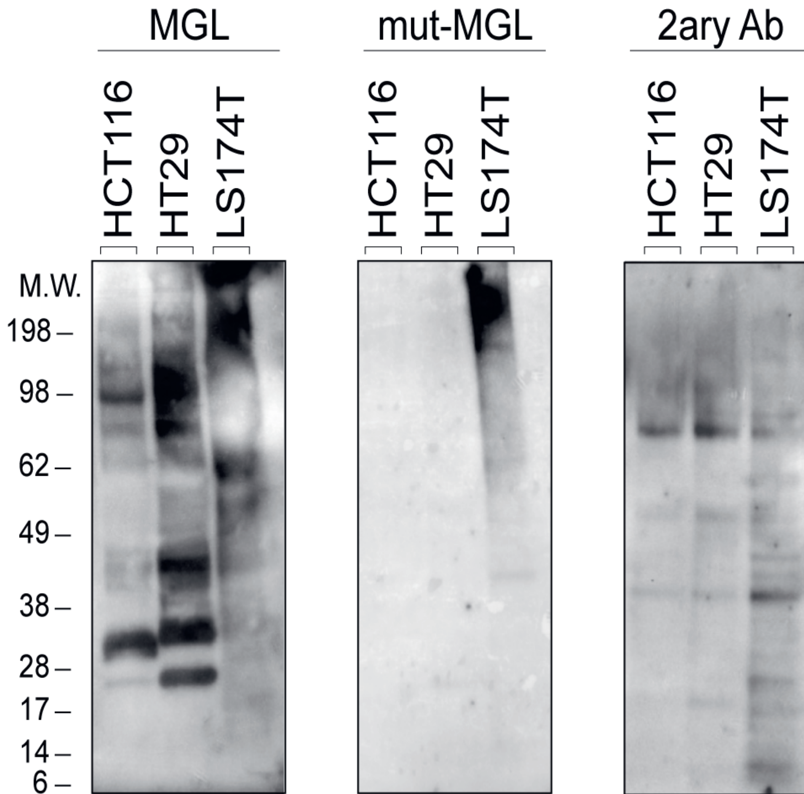


Figure 2: MGL binds to secreted glycoproteins of colorectal cancer cell lines. *MGL-Fc lectin blot (first panel) of secretome collected from HCT116, HT29 and LS174T. Mutant-MGL (second panel) and secondary antibody (2ary Ab) alone (third panel) were used as negative controls. The lectin blots are representative of 3 biological replicates. M.W.: molecular weight.*

Discussion

The cancer secretome released by cultured cancer cells in the CM *in vitro* reflects, to a certain extent, proteins secreted by cancers *in vivo* [126]. As such, secretome analysis might be a straightforward approach to identify putative markers, responsible for cancer-host cell interactions and consequent immune responses dampening or promoting cancer progression. In the current study, we adapted a method for secretome collection in the CRC cell line HCT116, with the aim of reducing not only serum contaminants used in cell culture but also intracellular proteins released due to cell starvation during collection. We achieved our goal by performing a series of washing steps and short incubation times of 1 h with serum-free medium, prior to CM collection after 24 h.

HCT116 and HT29, together with other CRC cell lines, have been previously studied for in-depth proteomic analysis of their secretome [121]. Even though the secretome collection was achieved following the same wash steps and conditioned media incubation time, we introduced a pre-incubation step, prior to washes, to reduce serum contamination from the cell culturing conditions. On the other hand, de Wit et al. [121], also compared the CRC cell lines secretomes with those obtained from CRC tissues of patients. In fact, protein biomarkers possibly detectable in stool or blood, can be found at higher concentrations in fluids in close proximity to the tumor. Results showed a substantial overlap of secreted proteins between tissue and cell lines secretomes, as well as selection of candidates already validated as serum/stool biomarkers for CRC [121]. Within the selected list of 76 candidates overlapping between CRC cell lines and tissues secretomes [121], we found only a few proteins and glycoproteins, such as Serpin B5 and Agrin, in our study. Notwithstanding the relevance of protein components in the tumor interstitial fluid (TIF) for cell to cell/ microenvironment cross talking, proteomic profiling of secreted proteins from body fluids was used as relevant approach for biomarkers discovery also in other cancer types such as breast cancer [127], lung cancer [128] and prostate cancer [129]. However, the availability and heterogeneity of cancer cell lines make them the most suitable source for secretome profiling studies and proteomics analysis[122].

HCT116, HT29 and LS174T show variable affinity binding for MGL, a lectin suspected to dampen adaptive immune response responsible for cancer progression [31, 34]. Knowing the identity of proteins and receptors binding to the lectin might be used to study not only the differential immune response to various MGL ligands recognition but also the cancer mechanisms activated by MGL binding. However, in CRC, the dynamic cross-talk favoring tumorigenesis sustained by extracellular mediators binding to MGL is still unknown. On the contrary, MGL affinity to Tn moieties on extracellular proteins was recently explored in ovarian cancer [130]. We sought first to investigate whether CRC cell line secretomes bind to MGL. To our surprise, the secretome of both high and low cell surface MGL ligands-expressing cell lines were found to bind to the lectin. Analyzing the identities of secreted proteins binding to MGL, using lectin enrichment followed by MS analysis, we noticed secreted isoforms of respective cell surface MGL ligands as well as new set of proteins. Within the first group, the secreted isoform of c-Met receptor drew our attention, since it was selected as secreted

binder not only in HCT116 but also in LS174T, while in the latter cell line c-Met on the cell surface was not found to bind MGL. This finding suggests that glycosylation may vary depending on protein sublocation. However, we also used a less stringent filtering routine to select for secreted proteins binding to MGL compared to our previous studies where we used total cell lysates, which could also partially explain these differences. We applied less stringent filtering routines due to technical changes in the procedure, such as pull-downs performed using culture medium.

Notwithstanding the annotation of more than one hundred glycopeptides, our glycoproteomic analysis did not result in the identification of peptides belonging to selected MGL binders, which carry the specific terminal GalNAc epitope. This might be due to intrinsic technical challenges of glycopeptide detection and identification, albeit the analytical method used was similar as described before by our group as well as others [110, 130].

Proteins can be secreted following two main mechanisms: the classical and non-classical secretory pathways [131]. The first one is driven by the activation of signaling pathways moving proteins characterized by a signal peptide at the N-terminus from the ER, through Golgi, into vesicles released in the extracellular space [131]. The non-classical one does not involve the ER-Golgi system and it involves a vesicular or non-vesicular transport [132]. While most of the proteins follow the conventional pathway, some signal peptide-containing proteins, such as CD45, a receptor protein Tyr phosphatase, expressed by T cells binding to MGL, pass through the ER, but bypass the Golgi to reach the cell surface in a vesicle. Others, such as fibroblast growth factor 2 (FGF2) and galectin 3, the latter found in LS174T MGL pull-downs, exit the cells through both ER- and Golgi- independent mechanisms [132].

During the trafficking, the proteins are further modified for example by *N*- and *O*-glycosylation, which influence the sorting of secretion [133]. Most proteins in the human circulation are heavily glycosylated, thereby decreasing undesired intermolecular interactions but increasing protection against undesired proteolysis [2]. However, changes in glycan structures occurring in cancer cells may be associated with tumor transformation and progression. Glycan structures are recognized by lectins, involved in quality control of secreted proteins, intercellular interaction and many other different biological function [134]. The role of MGL in tumor immunity and progression has been widely studied [31, 41]. MGL was recently described to bind not only to surface but also to extracellular proteins, decorated

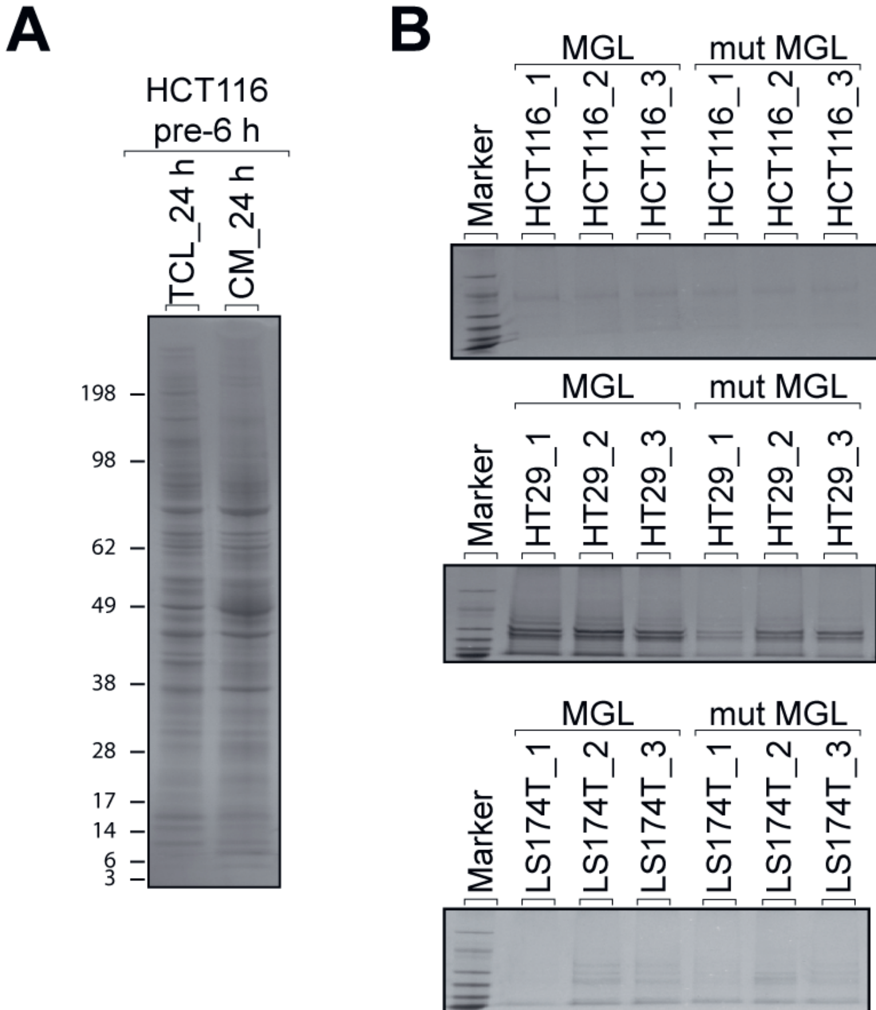
with Tn clusters, in ovarian cancer [130]. This evidence suggests that also the tumor secretome might have a contribution in modulating the tumor microenvironment. Among the HCT116 MGL binding secreted proteins identified in this study, the biggest cluster of proteins is associated with extracellular matrix (ECM) interaction and adhesion, including proteins belonging to the family of laminins and integrins. In common between HCT116 and HT29, different members of the semaphorin family were found as binders. For example, SEMA4D, found in HCT116, has been described for its immunoregulatory activity and is shed from the cell surface upon metalloprotease-mediated cleavage of the transmembrane receptor [135]. In HT29, GLG1, DAG1, GOLM and PTPR family members were identified also as cell surface ligands. The secreted mucin MUC5AC was found to bind MGL, within the CM of LS174T. Interestingly, we previously found that transmembrane mucins expressed in this cell line did not carry MGL epitopes, hence ruling out the involvement of cell surface mucins in MGL binding to CRC cell lines [110].

Overall, the recent advances in the MGL binding specificity in CRC using cell line secretome, and, previously, proteome give a better understanding of the tumor-immune cell mediators. Importantly, the current study shows that the investigation of both cell line total proteome and secretome is pivotal to have a complete picture of cancer cell interaction with the microenvironment.

Acknowledgments

We acknowledge G.W. van Pelt for providing the human CRC cells and the availability to use the cell culture facility. This research was funded by the European Commission’s Horizon 2020 programme “GlyCoCan” project, grant number 676421, and by the research programme Investment Grant NWO Medium with project number 91116004, which is (partially) financed by ZonMw.

Supporting Information



Supplemental Figure 1: Secretome of HCT116. *A) SDS-PAGE of HCT116 conditioned medium, collected after 6 h (pre-incubation, medium discarded) and 24 h (collection) of in serum-free media. For comparison, a total cell lysate (TCL) was used. B) Short SDS-PAGE of MGL/mutant-MGL pull-downs from HCT116, HT29 and LS174T secretomes.*

Supplemental table S1: secreted MGL-binding proteins from HCT116, HT29 and LS174T. *The overall list of secreted MGL-binding proteins in CRC cell lines was sorted for the Gene ontology annotation. Only proteins containing "secreted" or extracellular region" in their description were selected. Mol.Weight: molecular weight.*

5

Supplemental table S2: N- and O-glycopeptides from secreted proteins of HCT116 and HT29. *N- and O-glycopeptides we automatically annotated by Byonic. Only glycopeptides with Byonic score abore 250 were selected. Z: charge; HexNAc: N-Acetylhexosamine; Hex: Hexose; Fuc: Fucose; NeuAc: N-Acetylneuroaminic Acid.*