Universidade de Aveiro Departamento de Química
Ano 2016/2017

Eliana Janine de Paiva Soares

Glicoperfil do CD44: Estabelecimento Molecular de alvos terapêuticos no Cancro da bexiga

CD44-glycoprofilling: Establishing the molecular basis for targeted therapeutics in bladder cancer

DECLARAÇÃO

Declaro que esta dissertação é integralmente da minha autoria, estando devidamente referenciadas as fontes e obras consultadas, bem como as citações estão identificadas de modo claro. Não contém, por isso, qualquer tipo de plágio quer de textos publicados, qualquer que seja o meio dessa publicação, incluindo meios eletrónicos, quer de trabalhos académicos.

Universidade de Aveiro Departamento de Química
Ano 2016/2017

Eliana Janine de Paiva Soares

Glicoperfil do CD44: Estabelecimento Molecular de alvos terapêuticos no Cancro da bexiga

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Bioquímica Clínica, realizada sob a orientação científica do Doutor José Alexandre Ferreira, Investigador do grupo Patologia e Terapêutica Experimental do Centro de Investigação do Instituto de Oncologia do Porto, coorientação do Doutor Luís Carlos Lima, Investigador do grupo Patologia e Terapêutica Experimental do Centro de Investigação do Instituto de Oncologia do Porto e do Professor Doutor Francisco Manuel Amado, Professor Associado com Agregação do Departamento de Química da Universidade de Aveiro.

“Everything you do has some effect, some impact”
Dalai Lama

Aos meus pais que tudo proporcionaram para chegar até aqui

o júri

presidente

Professora Doutora Rita Maria Pinho Ferreira
Professora auxiliar da Universidade de Aveiro

Doutor André Moreira Neto da Silva
Investigador externo da Faculdade de Ciências da Universidade do Porto

Doutor José Alexandre Ribeiro de Castro Ferreira
Investigador do grupo Patologia e Terapêutica Experimental do Centro de Investigação do Instituto de Oncologia do Porto

O autor declara que participou na conceção e na execução dos seguintes trabalhos publicados e em publicação durante o ano de desenvolvimento da Dissertação de Mestrado.

I. Lima L, Neves M, Oliveira MI, Dieguez L, Freitas R, Azevedo R, et al. Sialyl-Tn identifies muscle-invasive bladder cancer basal and luminal subtypes facing decreased survival, being expressed by circulating tumour cells and metastases. Urologic Oncology: Seminars and Original Investigations. 2017 Dec 13;35(12):675.e1-675.e8.

II. Azevedo R, Soares J, Gaiteiro C, Peixoto A, Lima L, Ferreira D, et al. Glycan affinity magnetic nanoplatforms for urinary glycobiomarkers discovery in bladder cancer. (submitted to Talanta)

agradecimentos

Ao Doutor José Alexandre Ferreira, orientador deste trabalho, pela sua dedicação, paciência, ensinamento, por ter confiado e acreditado nas minhas capacidades e em especial por toda a ajuda, motivação e amizade.
Ao Doutor Luís Carlos Lima, coorientador deste trabalho, por permitir a realização do estágio, por me ter inserido no seu grupo de trabalho e por me ter dado esta grande oportunidade.
Ao Doutor Francisco Manuel Amado pelo apoio dado no desenvolvimento deste trabalho
Aos meus colegas do Grupo de Patologia e Terapêutica Experimental por todos os bons momentos passados e pela grande ajuda, que sem ela não teria conseguido estar aqui hoje.
Ao meu pai e à minha mãe pelo esfoço que fazem para eu puder estar aqui hoje, por sempre acreditarem em mim e nos meus sonhos, por me aturarem e me darem força e mesmo quando há desilusões me motivarem a fazer melhor. Sem eles não seria a pessoa que sou hoje.

palavras-chave

Cancro da bexiga, isoformas de CD44, glicosilação em proteínas, quimioresistência,

resumo

O cancro da bexiga (BC) apresenta uma das maiores taxas de recorrência entre os tumoures sólidos e é a segunda causa de morte, relativamente a doenças do trato geniturinário. A introdução de modelos moleculares para um melhor prognóstico e desenvolvimento de terapias dirigidas efetivas, continua a ser um aspeto desafiador devido à significativa heterogeneidade molecular inter e intratumoral. No entanto, a CD44, uma proteína de membrana fortemente *O*-glicosilada e envolvida nas interações célula-célula, adesão celular e migração, parece desempenhar um papel crítico na progressão e disseminação do cancro da bexiga, abrindo portas para potenciais terapêuticas dirigidas. No entanto, o gene que codifica esta proteína geralmente sofre *splicing* alterativo, o que resulta em diversas isoformas funcionalmente distintas, de pesos moleculares variáveis e com vários locais de glicosilação. No entanto, a natureza dessas isoformas no contexto do cancro da bexiga ainda não está bem esclarecida. Com base nestas ideias, este trabalho tem como objetivo determinar as isoformas da CD44 mais clinicamente relevantes e com potencial de direcionar para clones mais agressivos. É dado particular ênfase à identificação de *O*-glicanos associados ao cancro, que visam aumentar o entendimento molecular para o desenho de ligandos altamente específicos. Consequentemente observou-se que a CD44 está aumentada na urina de doentes com cancro de bexiga, comparativamente com urinas controlo de indivíduos saudáveis. Esse efeito é mais pronunciado em estadios avançados da doença, particularmente após a invasão muscular, o mesmo se verifica com expressão da CD44 nos tumores de bexiga.

Além disso, uma abordagem direcionada por RT-PCR demonstrou que o modelo celular de tumores superficiais de cancro da bexiga, a linha celular 5637 e os tumores de bexiga não invasivos sobre-expressam isoformas da CD44 de alto peso molecular (CD44v3-10^{high}, CD44v8-10^{high}, CD44s^{low}). Por outro lado, as linhas celulares T24 e HT1376 derivadas de tumores musculo-invasivos e estes mesmos tumores sobre-expressam predominantemente CD44s, uma isoforma de menor peso molecular (CD44v3-10^{low}, CD44v8-10^{low}, CD44s^{high}). Além disso, os clones quimiorresistentes das células T24, tratadas com cisplatina, também sobre-expressam CD44s. Da mesma forma, os tumores invasores apresentaram um fenótipo semelhante, apoiando a associação da CD44 com fenótipos mais agressivos. Os estudos de glicómica e glicoproteómica envolvendo a linha celular T24 demonstraram ainda a expressão de CD44 glicosilado com antigénios sialil-Tn (CD44-STn) e di-sialil-T (dST), anteriormente associados a um pior prognóstico. Em paralelo, ensaios de imuno-histoquímica e de ligação de proximidade *in situ* confirmaram a existência de CD44-STn e CD44-dST em tumores musculo-invasivos.

Em conclusão, CD44s, possivelmente modificada com STn e dST, tem o potencial de direcionar selectivamente para células mais agressivas de tumores de bexiga e clones quimiorresistentes, estabelecendo assim as bases moleculares para o desenho de ligandos. Estudos futuros devem-se concentrar em avaliar o impacto funcional da remodelação da CD44 para isoformas de menor peso molecular, acompanhando a transição de tumores superficiais para invasores.

keywords Bladder cancer, CD44 isoforms, glycosylation, protein glycosylation

abstract

Bladder Cancer (BC) presents one of the highest recurrence rates amongst solid tumours, and constitutes the second deadliest disease of the genitourinary track. The introduction of molecular models for disease management and effective targeted therapeutics remains a challenging aspect due to significant inter and intra-tumour molecular heterogeneity. Nevertheless, CD44, a heavily *O*-glycosylated membrane protein involved in cell-cell interactions, cell adhesion and migration has been suggested to play a critical role in bladder cancer progression and dissemination, holding potential for targeted therapeutics. However, the gene encoding for CD44 generally undergoes significant alterative splicing, which results in many functionally distinct isoforms of variable molecular weights and glycosylation sites. Nevertheless, the nature of these isoforms in bladder cancer are yet to be fully disclosed. Building on these insights, this work aims to highlight clinically relevant CD44 isoforms with potential for targeting more aggressive clones. Particular emphasis is also given to the identification of cancer-associated *O*-glycans envisaging the molecular rational for designing highly specific cancer ligands. Accordingly, it was observed that CD44 is increased in the urine of bladder cancer patients in relation to healthy controls. This effect is more pronounced for advanced stages of the disease, particularly upon muscle invasion, mimicking CD44 expression in bladder tumours.

Moreover, a targeted approach by RT-PCR demonstrated that superficial bladder cancer cell model 5637 and non-invasive bladder tumours overexpress high molecular weight CD44 isoforms (CD44v3-10^{high}, CD44v8-10^{high}, CD44s^{low} phenotype). Conversely, T24 and HT1376 cell lines derived from muscle invasive tumours and invasive lesions predominantly overexpress lower molecular weight isoform CD44s (CD44v3-10^{low}, CD44v8-10^{low}, CD44s^{high} phenotype). In addition, chemoresistant clones from T24 cells challenged with cisplatin also overexpressed CD44s. Likewise, bladder tumours from patients with invasive tumours presented a similar phenotype, supporting CD44s association with more aggressive phenotypes. Glycomics and glycoproteomics studies involving T24 cell line further demonstrated the expression of CD44 glycosylated with sialyl-Tn (CD44-STn) and di-sialyl-T (dST) antigens, previously associated with poor prognosis. In parallel, immunohistochemistry and *in situ* proximity ligation assays confirmed the existence of CD44-STn and CD44-dST in muscle invasive tumours.

In conclusion, CD44s, possibly modified with cancer-associated STn and dST glycans, holds potential to selectively target more aggressive bladder cancer lesions and chemoresistant clones, setting the molecular rational for ligands design. Future studies should now focus on disclosing the functional impact of CD44 remodelling towards lower molecular weight isoforms, accompanying transition from superficial to invasive lesions.

Index

List of figures

List of tables

Abbreviations

BC	Bladder cancer
BCG	*Bacillus* Calmette-Guérin
BCICs	Bladder cancer initiating cells
CAMs	Cell adhesion molecules
CIS	Carcinoma in situ
C1GalT-1	Core 1, β1-3 galactosyltransferase
COX-2	Cyclooxygenase-2
CSC	Cancer stem cell
Ct	Cycle threshold
dST	Disialyl-T antigen
ECM	Extracellular matrix
EMA	Epithelial membrane antigen
ER	Endoplasmic reticulum
FFPE	Formaline-fixed paraffin embedded
FGFR3	Fibroblast Growth Factor 3
GA	Golgi apparatus
GAG	Glycosaminoglycan
Gal	Galactose
GalNAc	*N*-Acetylgalactosamine
GC	Gemcitabine and cisplatin chemotherapy
GCH	Glucocorticoid hormones
GlcNAc	*N*-Acetylglucosamine
HA	Hyaluronic acid
HABD	Hyaluronan binding domain
HRAS	Harvey rat sarcoma viral oncogene homolog
IL-1β	Interleukin 1β
MIBC	Muscle invasive bladder cancer
MVAC	Methotrexate, vinblastine, adriamycin, and cisplatin chemotherapy
MoAbs	Monoclonal antibodies
NMIBC	Non-muscle invasive bladder cancer
OSTase	Oligosacharide transferase
PI3KCA	Phosphatidylinositol 3-kinase

PGE-2	Prostaglandin E2
PLA	*in situ* Proximity ligantion assay
ppGalNAc	UDPGalNAc-polypeptide *N*- acetylgalactosaminyl- transferases
PTEN	Phosphatase and tensin homolog
PTM	Post-translational modification
RB	Retinoblastoma
Rg	Receptor globulins
Ser	Serine
S6T	Sialyl-6-T antigen
SLe^a	Sialyl Lewis a
SLe^x	Sialyl Lewis x
ST	Sialyl-T antigen
STn	Sialyl-Tn antigen
T antigen	Thomsen-Friedenreich antigen
TCC	Transitional cell carcinoma
Thr	Threonine
T-ICs	Tumour initiating cells
Tis	Tumour in situ
TUR	Transurethral resection
TNF-α	Tumour necrosis factor alpha
UDP-GalNAc	Uridine diphosphate –*N*-acetylgalactosamine

Chapter 1

INTRODUCTION

1. Introduction

1.1. Bladder cancer

1.1.1. Epidemiology

Bladder cancer (BC) is the most common malignancy of the urinary tract, the seventh most common cancer in men, and the seventeenth in women (1). Also, it is the ninth most common cancer worldwide, with 429,000 new cases and 165,000 deaths in 2012 (2). In Portugal, BC is the eighth most common cancer, with 1829 new estimated cases and 900 deaths in 2010 (3). The disease is three times more iterated in men than in women and its incidence increases with age, reaching a peak between 50 and 70 years (4). Moreover, at diagnosis, women present more advanced stages of the disease and have less favourable prognosis (5).

Environmental risk factors have a significant role in BC initiation. Chemical or environmental exposures and chronic inflammation are well known risk factors for the development of BC and may lead to genetic and molecular changes (6). Critical exposures include aromatic amines, aniline dyes, nitrites and nitrates, acrolein, coal, and arsenic. Other casual factors include *Schistosoma spp* (Schistosoma haematobium) infection, and pelvic irradiation (7). However, tobacco smoke is the predominant risk factor associated with BC, accounting for 50% of BC cases (1). Genetic predisposition also has a significant influence on BC development, especially on the susceptibility to other risk factors (8).

Haematuria is the most frequent BC symptom, occurring in approximately 80% of patients. Other symptoms include urgency and dysuria, or, in more advanced tumours, pelvic pain, and symptoms related to urinary tract obstruction (1).

1.1.2. Pathophysiology and disease progression

Urothelial cancers arise from two distinct but overlapping pathways: papillary and non-papillary. Approximately, 80% to 85% of urothelial cancers are papillary lesions, stemming from hyperplastic epithelium or minimal dysplasia of a preneoplastic clone. The continuous growth of this clone results in the development of low-grade superficial papillary tumours that present loss of heterozygosity of chromosome 9, frequent mutations in the Fibroblast Growth Factor 3 (FGFR3), Harvey rat sarcoma viral oncogene homolog (HRAS) and phosphatidylinositol 3-kinase (PI3KCA) genes (9). Non-papillary and invasive tumours are thought to arise from severe dysplasia or carcinoma in situ (CIS). In a non-papillary pathway, cells from the initial hyperplasia develop genetic instability with frequent loss of tumour suppressor genes, such as Retinoblastoma (RB) and p53 (10,11). Interestingly, most muscle-invasive lesions emerge without a prior history of disease (primary invasive bladder cancer) (12,13).

Urothelial carcinoma (transitional cell carcinoma or TCC), squamous cell carcinoma, and adenocarcinoma comprise the three main bladder cancerous lesions, with TCC accounting for approximately 90% of cases (14,15). At diagnosis, approximately 75% of patients present superficial lesions and low grade tumours, (stages Ta, T1, or tumours in situ [Tis]), 10% of which progress to recurrent muscle-invasive and metastatic disease (Figure 1) (12,13). Consequently, non-muscle invasive bladder cancer (NMIBC) is a chronic disease with varying oncologic outcomes requiring frequent follow-up and repeated treatments, making the cost per patient (from diagnosis to death) one of the highest of all cancers (16,17).

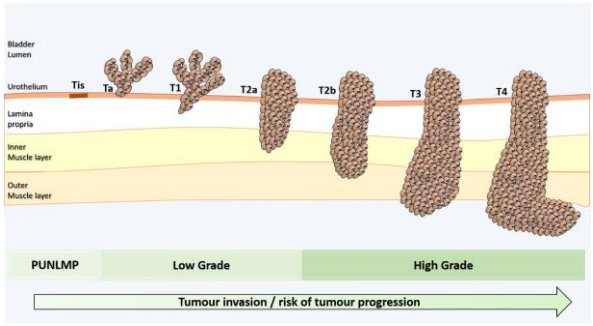

Figure 3 Schematic representation of bladder cancer staging and grading (18). The stage of the primary tumour (T) is based on the extent of penetration or invasion into the bladder wall.Tis, Tumour in situ: ‘‘flat tumour’’; Ta, Non-invasive papillary carcinoma; T1, Tumour invades subepithelial connective tissue; T2, Tumour invades muscle; T2a, Tumour invades superficial muscle (inner half); T2b, Tumour invades deep muscle (outer half); T3, Tumour invades perivesical tissue; T4, Tumour invades any of the following: prostate, uterus, vagina, pelvic or abdominal wall. Regarding tumour grading, bladder lesions can be classified as urothelial papilloma (a benign lesion), papillary urothelial neoplasm of low malignant potential (PUNLMP), low-grade papillary urothelial carcinoma and high-grade papillary urothelial carcinoma.

1.1.2.1. Diagnosis and therapeutics of bladder cancer

The most common clinical presentation of BC is asymptomatic hematuria, which prompts evaluation with cystoscopy, renal function testing, and upper urinary tract imaging in adults over 35 years with irritative voiding symptoms, and risk factors for bladder cancer (1,14). Moreover, tremendous efforts have been put in the development of biomarker panels for early diagnosis, with promising results regarding several glycans, lectins and proteoglycans (18)

The standard therapy for low-grade NMIBC is transurethral resection (TUR) of the tumour, allowing staging and primary treatment design. Further management is based upon risk factors, stage, grade, multimodality and recurrence history. Accordingly, a dose of perioperative chemotherapy can be used to reduce the risk of recurrence, and adjuvant chemotherapy is reserved for patients with multiple or multifocal recurrence. Moreover, there are subgroups of high-risk NMIBC patients which endure TUR and perioperative chemotherapy, followed by intravesical *Bacillus* Calmette-Guérin (BCG) therapy (19). Of note, intravesical BCG therapy is highly effective in comparison to chemotherapy alone (20). However, only two thirds of patients respond to BCG and one third of those responders will have recurrent disease (21), urging biomarkers able to predict therapy response and stratify patients who benefit the most from treatment.

Upon tumour invasion of the bladder wall (T2), perivesical tissue (T3), or adjacent pelvic organs (T4), endoscopic tumour resection and intravesical therapy become insufficient and the standard approach is radical cystectomy with pelvic lymphadenectomy and adjuvant or neoadjuvant chemotherapy (cisplatin-based combinations: methotrexate, vinblastine, cisplatin and doxorubicin - MVAC, or gemcitabine and cisplatin - GC) (14,22). Radical cystectomy may achieve a good local control; however, the relapse rate after radical cystectomy is as high as 50%, depending on the pathological stage of the primary tumour and the presence of loco-regional or distant metastasis (23). Importantly, the introduction of cisplatin-based chemotherapy in recent therapeutic schemes has allowed improving survival in comparison to cystectomy alone (24). Moreover, cisplatin-based chemotherapy results in nearly 15.2 and 14.0 months overall survival for MVAC and GC, respectively, compared to monotherapy and other combinations (25). Data from retrospective surgical series show that despite the early use of such aggressive approaches, many patients with muscle invasive bladder cancer (MIBC) still have high risk of recurrence and death from bladder cancer. The majority of recurrences occur within 3 years, with approximately 75% of patients already presenting distant metastasis (23).

Currently, there is a lack of specific biomarkers for targeting aggressive cell phenotypes. Additionally, the high molecular heterogeneity of MIBC is responsible for significant variations in disease course, as well as elevated recurrence and progression rates, hampering patient stratification regarding treatment response (BCG, chemotherapy). As such, to improve current clinical practices, bladder tumours of different stages have been widely screened for disease-specific glycosylation biomarkers capable of improving diagnose, surveillance, prognosis and offer novel therapeutic options.

Lately, there have been some evidences linking the expression of CD44 and the isolation of bladder tumour-initiating cells. Chan K, *et al.*, (26) isolated tumour-initiating cells (T-ICs) from bladder tumours based on the expression of basal cell markers. Particularly, $CD44^{+,}$ $CK5^{+}$, $CK20^{-}$ (T-ICs) cells could be successively transplanted up to 3 passages to induce tumour formation *in vivo*, with the tumours retaining the heterogeneity of the primary tumour. $CD44^{+}$, $CK5^{+}$, $CK20^{-}$ cells could yield either $CD44^{+}$, $CK5^{+}$, $CK20^{-}$ or $CD44^{-}$, $CK5^{-}$, $CK20^{+}$ (differentiated progeny) tumours. In contrast to T-ICs, the differentiated subpopulation could not be successively transplanted, revealing its limited self-renewal and/or proliferative capacities. Later, the same authors stratified bladder cancer into subtypes regarding cellular differentiation states based on keratin profiles, demonstrating that undifferentiated (basal-like) cells were $CD44^{+}$ (27). These results suggest that CD44 expression characterizes undifferentiated BC cells with sustained self-renewal, thereby being a cancer stem cell (CSC) marker.

1.2. Glycoprotein CD44: potential biomarker of poor prognosis in bladder cancer

1.2.1. Molecular structure and biological role

The CD44 antigen is a multifunctional and densely glycosylated transmembrane glycoprotein that together with selectins, integrins and cadherins composes the cell adhesion molecules (CAMs) family, mediating cell-matrix and cell-cell interactions, cell adhesion, migration, and tissue integrity (28,29).

The CD44 gene comprises 19 exons (1-19), of which only exons 1-5 and 15-19 are conserved, while the remaining 9 exons (6-14 exons also known as v2-v10) undergo alternative splicing (Figure 2). Particularly, the canonical CD44v2-v10 isoform contains all

variable exons (v2-v10) without the conserved exon 18, constituting the bulkiest CD44 isoform (250 kDa) (30,31). In turn, the standard CD44 (CD44s or CD44h) constitutes a much shorter CD44 isoform by losing all variable exons and exon 18, yielding an 85 kDa protein (32,33) (Figure 2). Moreover, the CD44v3-v10 isoform loses both v2 and 18 exons, being primarily found in keratinocytes, while the isoform CD44v8-v10 (CD44E) loses both v2-v7 and 18 exons and is preferentially expressed in epithelial cells (34). In addition, CD44v10 isoform loses all variable exons except v10, as well as the constitutive exon 18. Finally, the CD44st isoform, also known as short-tail CD44, loses all variable exons and exon 19. More recently, two more CD44 isoforms, namely CD44 NCBI Transcript 7 and CD44RC were described in UniProt and NCBI databases (35,36). Particularly, CD44 NCBI transcript 7 loses all variable exons, exon 15 and exon 18, while CD44RC is solely constituted by the first four conserved exons (Figure 2) (35,36). Of note, CD44 has several predicted isoforms that remain to be described in research models or humans, urging more in-depth studies regarding these markers.

Moreover, CD44 presents multiple post-translational modification sites, with at least five *N*-glycosylation sites in its ectodomain and several potential *O*-glycosylation sites in the membrane proximal extracellular region (37,38); thereby affecting its ligand binding characteristics, modulating its functions and further diversifying its structure (39). CD44 can also be considered a “part-time” proteoglycan, since some of its alternative splicing variants present glycosaminoglycan (GAG)-initiation sites (40). Further modifications can occur through tyrosine sulphation (41).

Hyaluronic acid (HA), an important component of the extracellular matrix (ECM), is the most common, but not the only, ligand of CD44. HA is a linear polymeric GAG and there are at least three sites for its binding on the CD44 molecule, one of which is the “link” domain encoded by exon 2 (42), and the other two overlap in the region encoded by exon 5 (43). Although all CD44 isoforms contain HA recognition sites, not all cells expressing CD44 constitutively bind the HA ligand. Cells can express CD44 in an active, an inducible, or an inactive state with respect to HA binding. The differences in the HA binding state of CD44 are cell specific and have been related to post-translational modification patterns (44). This interaction between CD44 and HA has been reported in chondrocytes, being an important process to direct the assembly of pericellular matrices (45) and chondrosarcoma cell lines adhesion (46). In addition, cells undergoing repair appear to upregulate both CD44 and HA. Mukesh Jain, *et al.,* demonstrated that CD44 and HA are expressed in minimal amounts on smooth muscle cells in normal arteries, in

opposition to injured arteries where the expression of these two molecules was much higher (47).

Several lymphocyte functions also appear to be CD44-dependent. Namely, there is an increase in cell surface levels of CD44 in activated T-cells (48). Additionally, CD44 expression seems to mediate the adhesion of lymphocytes to vascular endothelial cells via its binding to HA, and this interaction seems to be associated to T-cell extravasation into inflammation sites in mice (49) and humans (50). The homing of lymphocytes to inflammatory sites through CD44–HA binding is enhanced by the induction of HA synthesis in vascular endothelium by the proinflammatory cytokines tumour necrosis factor α (TNF-α) and interleukin 1β (IL-1β) (51). Accordingly, the presence of CD44 splicing variants appears to be obligatory for the migration and function of Langerhans and dendritic cells from peripheral organs to lymph nodes for antigen presentation (52). As it was previously discussed CD44 presents different ligands beside hyaluronic acid, such as osteopontin. Osteopontin (OPN) is highly expressed in bone. It is also known as, early T-lymphocyte activation (ETA-I) (53). This protein has revealed to be associated with regulation of immune cells, namely, promoting cell-mediated immune responses, playing a role in chronic inflammatory diseases, infiltration of macrophages during inflammation and induction of pro-inflammatory cytokines expression (54,55).

As it was described above, CD44 is involved in cell-cell and cell-matrix interactions, proliferation, migration and adhesion, which are directly involved in tumour progression and metastasis (56,57). As such, the clinical relevance of CD44 will be discussed in the next sections.

Figure 2 Human CD44 structure and examples of alternatively spliced mRNA transcripts. Blue filled boxes represent constitutive exons, white filled boxes represent alternative exons. The blue line represents the alternative region that is missing in each CD44 isoform. The exon box width is not proportional to the bp number. Exon 18, filled black, is not coding and contains early 3′UTR. Its inclusion gives rise to a short cytoplasmic tail mRNA that translates into CD44st. Seventy amino acids of cytoplasmic tail are encoded by exon 19. Cytoplasmic tail of CD44 contains intracellular signalling motifs and mediates interaction with cytoskeleton. Variable region of CD44 is a site of heavy O-glycosylation. The globular amino terminal domain of CD44 contains 3 disulphide bonds and two hyaluronic binding motifs.

1.2.2. Clinical relevance in solid tumours

The expression of CD44 isoforms and its binding to HA is associated with tumour growth and development (58). Moreover, some studies describe unusual CD44 transcripts not found in the corresponding normal tissues (30,59). Furthermore, it was demonstrated that *in vivo* tumour formation by human lymphoma Namalwa cells, stably transfected with CD44s, can be suppressed by a soluble human CD44s-immunoglobulin fusion protein disrupting the interaction between CD44s and its physiologic ligands. Moreover, no

tumour growth was detected even 100 days after the initial tumour cell injection. These results may provide new insides for tumour growth control *in vivo* (60). Both CD44s and its isoform CD44v10 expression are highly associated with malignant melanoma, and CD44v10 appears to facilitate local invasion. Accordingly, treatment of B16F10-bearing C57BL/6 mice either with a CD44s-/ CD44v10-specific antibody, or with receptor globulins (Rg) containing the extracellular part of CD44s or CD44v10 linked to the constant region of the immunoglobulin kappa light chain resulted in a reduction of lung metastasis. However, only CD44 Rgs prevented spread and settlement of melanoma cells in distant organs. These findings confirm the involvement of both CD44s and CD44v10 in melanoma progression, and is suggestive for the use of Rgs as therapeutic reagents (61). A similar study in malignant melanoma also demonstrated that CD44 blockage by specific antibodies inhibits tumour growth and metastisation *in vivo* by disrupting CD44-HA interactions (62). In line with these findings, tumour cell lines with high levels of CD44 protein were shown to be capable of forming more aggressive tumours in animal models (63,64).

CD44 also seems to be involved in lymphocytes preservation. Particularly, CD44 appears to inhibit DNA fragmentation and apoptosis of immature lymphocytes and peripheral T lymphocytes induced by anti-CD3 monoclonal antibodies (MoAbs) and glucocorticoid hormones (GCH), therefore modulating T-cell survival (65). In addition, CD44 provided resistance to anti-integrin antibody induced apoptosis of colon carcinoma cell lines, demonstrating the possibility of this molecule involvement in poor prognosis of colon carcinoma and metastasis development (66).

Furthermore, osteopontin and CD44 intereractions have been associated to tumour dissemination and metastization in brain tumours (67). Through the interaction with integrin receptors on macrophages and CD44, OPN seems to promote a delayed immune system response, inducing cellular immunity and an anti-oxidant effect preventing cell damage (67). Moreover, tumour dissemination is associated to neovascularization. CD44 isoforms and integrin aVβ3 are highly associated with angiogenesis process in tumours. Expression of CD44, osteopontin and integrin aVβ3 have been demonstrated to be involved in endothelial cell migration by vascular permeability factor/vascular endothelial growth factor (68,69).

1.2.3. CD44 in bladder cancer

There have been some studies relating the expression of CD44 proteins and several biologic and clinical roles in bladder cancer. Kuncová J, et al. (70) evaluated the expression of CD44s and CD44v6 in 5637 and HT1197 BC cell lines, reporting that expression of both isoforms was dependent on the differentiation state of cells. Particularly, undifferentiated 5637 cells presented a more regular positive staining for CD44 proteins, while the more differentiated HT1197 displayed variations in CD44 staining. Also, CD44 proteins expression was correlated with higher proliferative activity and stem-like phenotypes of HT1197 and 5637, as well as with cellular atypia in HT1197. Moreover, it was suggested that variations in CD44 expression might be connected to the derangement of differentiation in tumour cells, manifested by prominent cytological atypia in high-grade tumours.

It is known that CSCs can be isolated from several solid tumours, such as breast (71) and brain tumours (72); however the lack of bladder cancer stem (initial) cells (BCICs) markers has hampered the same progress in bladder cancer. Yang Y, et al. (73) reported that BCICs might be among the epithelial membrane antigen (EMA)- CD44v6+ subset isolated from bladder tumours by magnetic cell sorting, since both EMA- cells and CD44v6+ cells possess the ability for colony-forming, self-renewal and proliferation. Importantly, these cells are usually located in the basal layer of normal urothelium, the potential location of BCICs, and would infiltrate from basal layer towards superficial layer in tumourous urothelium.

Inflammation events in normal and CSCs might be associated to prostaglandin E2 (PGE2) signalling (74), which has been reported to regulate hematopoietic stem cells homeostasis (75). Cyclooxygenase-2 (COX-2), the PGE2-generating enzyme, is determinant to inflammation-related carcinogenesis (76) and COX inhibitors present antitumour activity (77). Having the above in consideration, Thanan R, et al. (78) suggested that COX-2 may play an important role in tumour initiation, regulation of stem cell proliferation and differentiation in inflammation-related bladder cancer carcinogenesis. Particularly, *S. haematobium*-induced BC was correlated with stemness marker Oct3/4 expression, while BC without associated infection correlates with CD44v6 expression. Moreover, the expression of Oct 3/4 and CD44v6 was associated with nuclear localization of COX-2, suggesting that inflammation could mediate stem cell proliferation and differentiation in bladder cancer.

In a clinical approach, CD44v9 was reported to be an important prognostic biomarker for disease progression and cancer-specific death in MIBC and high-grade NMIBC (79). Its localization in invasion fronts demonstrated the clinical value of CD44v9

expression, which could be used as stratification marker to select patients who might progress to muscle-invasive bladder cancer. In turn, it was demonstrated that CD44s and CD44v6 correlate with tumour stage and grade (80). Particularly, low-grade papillary TCC exhibited staining of basal layer cells, with a decreasing intensity of staining towards the urothelium surface, while high-grade TCC displayed strong staining of CD44s and CD44v6 throughout the entire neoplastic urothelium. Tumour grade and depth of invasion also correlated with positive staining of CD44s and CD44v6, confirming the diagnostic and prognostic value of both markers. It was also evidenced that CD44v6 is an indicator of metastatic potential with an increased expression in higher stages and in tumours with infiltrating borders (80). Another study also associates CD44s expression with higher grade, stage, and density of tumour infiltrating lymphocytes, as well as shorter overall survival, while CD44v6 expression was associated with better prognosis, lower grade and increased overall survival (81). Contrastingly, another study reported that the loss of CD44v6 expression is an independent adverse predictor of recurrence and overall survival, further suggesting that routine evaluation of CD44v6 could be important in identifying high-risk patients which would benefit from more aggressive therapeutic approaches. Also it was hypothesised that CD44v6 expression could be a much better predictor of recurrence than BC grade (82).

The knowledge regarding the clinical significance of CD44 in BC is still scares, which allied to conflicting reports has hampered the true development of targeted therapies involving this marker. These discrepancies can be explained by the lack of standard immunohistochemical assays, the use of antibodies with different specificities, and differences in the clinicopathological status of bladder tumours used in the different studies. Therefore, integrative and standardized studies are necessary to elucidate the role of CD44 in BC as it holds an important biological and clinical value and may serve as therapeutic target. Moreover, the structural complexity of this glycoprotein, including the coexistence of different splicing variants in the same tumour (83) and significant heterogeneity in glycosylation patterns (39) also may explain the existence of conflicting data. Therefore, it is crucial to determine the nature of the splicing variants expressed by bladder tumour of different histopathological natures, as well as their biological and clinical significances. Moreover, since CD44 is a highly *O*-glycosylated protein, exploiting its *O*-glycosylation patterns may allow narrowing down its isoforms to clinically relevant glycospecies capable of improve BC management (detection, prognosis, follow-up, and therapy selection).

1.3. Protein glycosylation and cancer

1.3.1. Patterns of glycosylation

Protein glycosylation is the most frequent post-translational modification (PTM) of membrane-bound and secreted proteins. This non-templated process involves the highly coordinated action of several glycosyltransferases, glycosidases and nucleotide sugar transporters in the endoplasmic reticulum (ER) and Golgi apparatus (GA) to generate carbohydrate-associated proteins. Glycosylation increases proteome complexity due to the diversity in sugar compositions, glycosidic linkages (*N*- , *O*- and C-linked glycosylation, glypiation (GPI anchor attachment) and phosphoglycosylation), chain length and substitution patterns (84–88). Given its structural diversity, glycans are important mediators of a plethora of biological events, such as cell-cell adhesion, cell differentiation, migration, signalling transduction, immune recognition, host-pathogen interactions, protein folding, traffic and stability (89–91).

Two main classes of glycans can be found at the cell surface, namely *N*- and *O*-glycans (Figure 3). Although glycosylation has been characterized as a post-translational modification, *N*-glycosylation often occurs during the translation and transport of proteins into the ER, promoting the proper folding of newly synthesized polypeptides. Precursor *N*-glycan synthesis begins on the cytosolic face of the ER and is further elongated after the initial structure is flipped into the ER lumen (91). During *N*-glycan synthesis, a 14-sacharide “core” unit (Glc3Man9GlcNAc2-) is transferred by oligosaccharide transferase (OSTase) to asparagine (Asn) residues in an Asn-X-Ser/Thr sequence of the nascent protein, with X being any amino acid except proline. The diversification of the glycans occurs in the GA to yield either partially unprocessed oligomannose antenna or, more frequently, complex or hybrid type structures (92). Mature *N*-glycans often yield highly relevant terminal structures, such as Lewis (Le) blood group related antigens and ABO(H) blood group determinants (93). Further glycome complexity is added by sugar phosphorylation, *O*-acetylation of sialic acids, and *O*-sulfation of galactose and *N*-acetylgalactosamines (94).

N-glycosylation does not prevent *O*-glycosylation from happening, as *O*-glycosylation commonly takes place on glycoproteins previously *N*-glycosylated in the ER. *O*-glycosylation occurs post-translationally by covalently α-linking a *N*-acetylgalactosamine (GalNAc) moiety to protein serine (Ser) or threonine (Thr) residues

by UDPGalNAc-polypeptide *N*- acetylgalactosaminyl- transferases (ppGalNAc-Ts), forming the simplest *O*-glycan Tn antigen. Subsequently, T antigen (core 1 or Thomsen-Friedenreich antigen) results from the attachment of a galactose (Gal) to GalNAc by T synthase (β1-3 galactosyltransferase or C1GalT-1) (95). Moreover, T antigen can be the precursor of several more complex core structures (from core 2 to 8) bearing similar terminal structures than mature *N*-glycans (96). In addition, Tn and T antigens can be sialylated by sialyltransferases, forming the sialyl-Tn (STn), sialyl-T (ST) and disialyl-T (dST) antigens. Upon formation of the STn antigen, any further processing of the oligosaccharide chain stops (97).

In cancer tissues, glycosylation patterns are profoundly altered, being characterized by the expression of highly branched and heavily sialylated/fucosylated glycans, alterations in glycan terminal structures, and overexpression of truncated glycans (18,98). Glycome alterations at the cell surface derive from the synergism of different events that go beyond altered glycosyltransferases and/or glycosidases expression. These may include alterations in peptide backbone or nascent glycan structures (99), glycosidases or glycosyltransferases activity (100,101), chaperone functions (102), glycosyltransferases mislocation in secretory organelles (103), and the bioavailability of sugar nucleotide donors and cofactors (104). Importantly, many glycoepitopes constitute tumour-associated antigens. Those found at the cell surface are easily accessible to antibodies and lectins to selectively target specific tumour cells, while those secreted or shed into bodily fluids can be explored for non-invasive cancer diagnosis through serological assays. As such, cancer specific alterations to protein glycosylation provide a unique opportunity for clinical intervention.

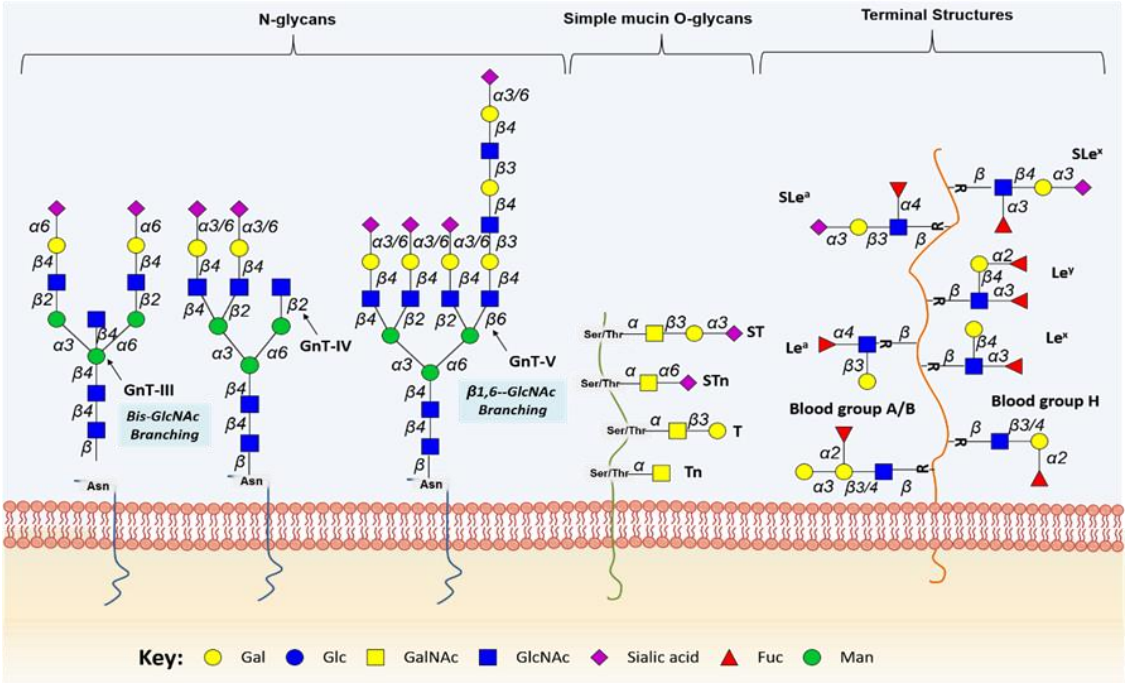

Figure 4 Schematic representation of protein-associated glycan structures relevant in bladder cancer (18). The figure represents specific N-linked and *O*-linked glycan structures, as well as terminal Lewis and sialylated Lewis structures that have biological significance in bladder cancer. Key enzymes mediating the addition of specific sugars are also shown. Protein *N*-glycan alterations include the β1-6 branching of *N*-glycans in result of GlcNAcT-V (GnT-V) overexpression, and the addition of bisecting GlcNAc branches by GlcNAcT-III (GnT-III) glycosyltransferases. Alterations in *O*-glycosylation pathways are also a common hallmark of malignant transformations of the bladder. Herein, we represent the overexpression of simple mucin-type O-glycans and their sialylated counterparts, T, sialyl T (ST), Tn and sialyl Tn (STn) antigens. Altered expression of terminal structures is also a common feature of bladder tumours. Namely, the abnormally low or absent expression of ABO(H) blood group determinants is frequently found in high grade and invasive disease. Carbohydrate terminal Lewis antigens are significantly under-expressed in healthy urothelium when compared to bladder tumours and are also highlighted here. Lewis type 1 antigens includes Lewis^a (Le^a), and sialyl Lewis^a (SLe^a), while the type 2 group includes Lewis^x (Le^x) and sialyl Lewis^x (Sle^x).

1.3.2. Altered glycosylation in bladder cancer

It has been long known that advanced stage bladder tumours present severe dysregulations in glycosylation pathways, translated by the loss of terminal ABO blood group determinants at the cell-surface of ABH secretor individuals (105,106), over- and/or de novo expression of short-chained *O*-GalNAc glycans (107), Lewis blood group related antigens, and their sialylated counterparts (108,109), as well as oversialylation and fucosylation of glycan chains (110). Particularly, ABO(H) antigens in the initial biopsy of bladder carcinomas is predictive of a much higher chance of subsequent invasion than in those tumours in which the ABO(H) antigens are detectable (111). However, a significant number of patients whose initial tumours were reported as blood group antigen negative failed to develop an invasive tumour (111). It is possible that these conflicting results may, at least in part, be explained by differences in methodology, interpretation, or both. These antigens are present on normal bladder epithelium but not on some low-grade and early-

stage papillary transitional cell carcinomas of the bladder (112). In bladder urothelium the most studied change has been the deletion of blood group A antigens from A individuals and H antigen from O individuals. Moreover, the loss of activity of the A and B gene-encoded transferases (112), and ABO(H) gene and/or its promoter hypermethylation (105) have been suggested to be amongst the events driving these, explaining the deletion of these antigens in bladder tumours.

The A, B, H antigens have biosynthetic and structural similarities with the Lewis antigens, including the type 1 Lewis^a and type 2 Lewis^x antigens and their sialylated counterparts, namely sialyl-Lewis a (SLe^a) and x (SLe^x). Several authors have associated Lewis^a expression patterns with malignant transformations, reporting significantly lower expression of this antigen in healthy urothelium when compared to invasive tumours (113,114). As such, the expression of Lewis^a can be associated with worse bladder cancer phenotypes. Moreover, Lewis^a antigen expression patterns change at an early neoplastic stage, suggesting that Lewis^a determination might be useful in the diagnosis of very early premalignant changes in the urothelium (115). In addition, the employment of Lewis^a staining scores allows the sub-classification of histologically identical tumours into prognostically different groups, pointing to a relationship between the pathological grade and stage of the evaluated tumours and a morphological and functional de-differentiation (115). Given this, Lewis^a antigen is a valuable functional marker of the malignant potential in superficial bladder cancer. In turn, the Lewis^x antigen is not expressed in normal urothelium, except for occasional umbrella cells, but is demonstrated in most invasive tumours, regardless of blood type and secretor status of the individuals studied (106). In cancer cells, both SLe^a and SLe^x mimic their normal expression on blood cells, allowing cancer cell binding to endothelial selectins and extravasation of cancer cells (116). Minimal structural alterations in SLe terminal motifs may significantly alter their biological behaviour, allowing them to be explored has biomarkers of disease. Namely, a glycan epitope that is very similar to the selectin ligand SLe^x, the sialyl-6-sulfo Lewis^x, is preferentially expressed in normal epithelial cells compared to cancer cells (117). Essentially all genes involved in the synthesis of SLe^x are predicted to be the same as those governing the synthesis of sialyl-6-sulfo Lewis^x, except for the genes engaged in its sulfation. This finding supports the hypothesis that reduced expression of sialyl-6-sulfo Lewis^x induces SLe^x expression in cancers(118). Sialyl-6-sulfo Lewis^x antigens are expressed in bladder urothelial carcinomas playing divergent roles in its progression, since it promotes E-selectin-mediated tumour cell adhesion to vascular endothelial cells, which is potentially associated with metastasis, while aiding lymphocyte recruitment to

enhance anti-tumour immune responses (119). A structure very similar to SLe^a termed disialyl-$Lewis^a$ is preferentially expressed in non-malignant cells, and is useful as a marker for tissue injuries occurring in benign diseases (120). The structural difference between SLe^a and disialyl-$Lewis^a$ is the presence of one extra sialic acid residue attached to the C-6 position of βGlcNAc in the carbohydrate structure of the latter glycan epitope. This implies that the α-2-6 sialylation of the GlcNAc moiety is impaired in cancer cells compared to non-malignant cells. Moreover, this event has been associated with the epigenetic silencing of the *ST6GalNAcIV* gene responsible for the α2-6 sialylation of the βGlcNAc moiety in cancer cells (120,121). The sialylated forms of Lewis antigens have also been associated with BC malignant potential. Particularly, it has been demonstrated that loss/reduction of SLe^a expression was associated with higher atypia grade (122), while SLe^x has been closely linked to invasive and metastatic potential of primary bladder tumours (123). Yet, another study demonstrated no associations between SLe^x, grade or stage in urothelial carcinoma of the renal pelvis, ureter, and urinary bladder (122). Notwithstanding, the overall increase in cell-surface sialic acid content was shown to reduce the attachment of metastatic tumour cells to the extracellular matrix, possibly protecting them from recognition by the alternative pathway of complement activation and favouring metastatic spread (124).

ABO(H) loss may also stem from premature stop in protein *O*-GalNAc glycosylation, ultimately translating in the accumulation of Tn, T and their sialylated counterparts STn and ST antigens. In particular, the overexpression of *ST6GalNAc-I* has been found to promote the premature sialylation of the Tn antigen and consequent formation of the STn antigen in bladder cancer (125,126). Specifically, the STn antigen is absent in the healthy urothelium, while being present in more than 70% of high-grade NMIBC and MIBC, denoting a cancer specific nature (125). This posttranslational modification of cell surface proteins is mostly expressed in non-proliferative tumour areas, known for their high resistance to cytostatic agents currently used to improve the overall survival of advanced stage bladder cancer patients (125). Recently, a novel STn-dependent mechanism for chemotherapeutic resistance of gastric cancer cells to cisplatin has been described, in which STn protects cancer cells against chemotherapeutic-induced cell death by decreasing the interaction of cell surface glycan receptors with galectin-3 and increasing its intracellular accumulation (113). The relationship between chemoresistance and STn overexpression remains to be fully explored in bladder cancer. Furthermore, STn expression is significantly higher in MIBC when compared to NMIBC, denoting its association with muscle invasion and poor prognosis (127). Studies *in vitro*

have further demonstrated that this antigen plays an important role in bladder cancer cell migration and invasion through mechanisms so far unexplored (125,128). However, glycoproteomics studies of bladder cancer cell models highlighted that STn was mainly present in integrins and cadherins, further reinforcing a possible role for this glycan in adhesion, cell motility and invasion (128). Also, recent work from our group has demonstrated the presence of STn in lymph node and distant metastasis, strengthening the notion that STn expression may influence cancer cell motility and matastization capability (129). Moreover, STn inclusion improved the predictive capacity of a molecular model proposed for stratification and prognostication of bladder tumours based on keratin (KRT14, 5, and 20) expressions (129). Accordingly, the STn antigen has been associated with basal-like phenotypes (KRT14+ and/or KRT5+, KRT20-) facing worst prognosis (129). Furthermore, STn-expressing BC cells have shown the ability to induce a tolerogenic microenvironment by impairing dendritic cells maturation, allowing cancer cells to evade innate and adaptive immune system responses (130). Interestingly, the tolerogenic effect of short-chained *O*-glycans has also been correlated with bladder tumour metastasis through a mechanism in which MUC1 carrying core 2 *O*-glycans functions as a molecular shield against NK cells attack, thereby promoting metastization (131). In addition, STn expression in bladder cancer tissues has been used in combination with other surrogate markers of tumour aggressiveness envisaging patient stratification regarding disease stage and therapeutic benefit. Specifically, STn and sialyl-6-T (S6T), an STn-related antigen, expression are independent predictive markers of BCG treatment response and were found useful in the identification of patients who could benefit more from this immunotherapy (132). Moreover, STn was found to be a marker of poor prognosis in bladder cancer and, in combination with PI3K/Akt/mTOR pathway evaluation, holds potential to improve disease stage stratification (127).

In turn, several reports associated the presence of Tn and T antigens with recurrence and metastization in bladder cancer, suggesting that these antigens may be surrogate markers of profound cellular alterations (133,134). Also, there is growing evidences linking the overexpression of the sialyl-T antigen and ST3Gal.I, the enzyme responsible by T antigen sialylation, with bladder cancer aggressiveness, recurrence (135), poor prognosis, and tumour grade (133,136). Moreover, the expression of T antigen is significantly associated with higher risk for subsequent recurrences with deep muscle invasion and metastatic involvement of regional lymph nodes (133,136,137). The fact that these simple glycans are absent, significantly under-expressed or restricted to some cell

types in healthy tissues, makes them ideal diagnostic and therapeutic targets for bladder cancer therapy (138).

1.3.3. CD44 glycosylation in solid tumours

As previously described, CD44 undergoes extensive post-translational modification, including *N*- and *O*-linked glycosylation and substitution with high molecular weight glycosaminoglycans (139), thereby modulating its biological functions. Particularly, Dasgupta et al. demonstrated that CD44s *O*-linked glycosylation decreased CD44s-mediated adhesion to hyaluronate (HA), while *N*-linked glycosylation had minimal influence on CD44 function. These findings suggest that *O*-linked glycosylation may be as important as alternative splicing in the regulation of CD44 function and the broad spectrum of biological processes attributed to it, including normal development, tumour metastases, and lymphocyte function (38). Contrastingly, Bartolazzi et al. showed that treatment of a panel of human cell lines which constitutively express CD44 with the inhibitor of *N*-linked glycosylation tunicamycin results in the loss of attachment of these cells to HA-coated substrate (140) The authors further demonstrated that treatment of the same cells with deoxymannojirimycin, which inhibits the conversion of high mannose oligosaccharides to complex *N*-linked carbohydrates, results in either no change or an increase in CD44-mediated adhesion to hyaluronate, suggesting that complex *N*-linked oligosaccharides may not be required or even inhibit CD44-HA interaction. Using human melanoma cells stably transfected with CD44 *N*-linked glycosylation site-specific mutants it was shown that integrity of five potential *N*-linked glycosylation sites within the HA recognition domain of CD44 is critical for HA binding. Mutation of any one of these potential *N*-linked glycosylation sites abrogates CD44-mediated melanoma cell attachment to HA-coated surfaces, suggesting that all five sites are necessary to maintain the HA-recognition domain in the appropriate conformation. They also demonstrated that mutation of serine residues, which constitute the four Ser-Gly motifs in the membrane proximal domain and provide potential sites for glycosaminoglycan side chain attachment, impairs HA binding (140). These observations suggest that the relevance of *N*-glycosylation sites for CD44 functions is site specific. Olgun Guvench developed molecular dynamics simulations that provide atomic-resolution mechanistic understanding of CD44-HA interaction to help bridge gaps between existing experimental binding and structural biology data. Findings from these simulations included that Tyr42 may function

as a molecular switch that converts the HA-binding site from a low affinity to a high affinity state; in the partially disordered form of hyaluronan binding domain (HABD), basic amino acids in the C-terminal region can gain sufficient mobility to form direct contacts with bound HA to further stabilize binding; and terminal sialic acids on covalently attached *N*-glycans can form charge-paired hydrogen bonding interactions with basic amino acids that could otherwise bind to HA, thereby blocking HA binding to glycosylated CD44 HABD (141). Glycosylation modulates CD44 biological role not only by affecting its interaction with ligands. Namely, Hu et al. suggest that the Lewis y antigen, as an important component of the molecular structure of CD44, promotes proliferation and inhibits apoptosis of ovarian cancer cells, leading to drug resistance via activation of PI3K/AKT signalling pathways. Moreover, it is proposed that the Lewis y antigen, as a structural component of CD44, integrins a5b1 and avb3, as well as EGFR, play a role in various cell adhesion processes that mediate both cell adhesion and drug resistance (142).These results suggested that fucosylation of CD44 is related to drug resistance in ovarian cancer. Campos et al. uncovered CD44-STn glycoforms in gastric cancer patients serum, further validating its expression in gastric cancer tissue, suggesting that aberrantly expressed CD44 as potential biomarkers in gastric cancer (143). Particularly, the CD44v6 expression levels were associated with pre-malignant and malignant lesions of the stomach, providing a potential biomarker for gastric mucosa transformation (143). Singh et al. also reported the coexpression of oncofetal carbohydrate antigens T and STn on CD44 splice variants providing a link between cancer-associated changes in glycosylation and CD44 splicing, both of which correlate with increased metastatic potential of colon cancer (144).

Regarding BC, Carrascal et al. reported that STn^+-CD44 cancer cells impair dendritic cells (DC) maturation and endow DCs with a tolerogenic function, limiting their capacity to trigger protective anti-tumour T cell responses, suggesting that STn antigens and, in particular, STn+ glycoproteins, such as CD44, are potential targets for circumventing tumour-induced tolerogenic mechanisms (145). More recently, glycoproteomic analyses of advanced bladder tumours based on enzymatic treatments, Vicia villosa lectin-affinity chromatography enrichment and nanoLC-ESI-MS/MS analysis resulted in the identification of several key cancer-associated glycoproteins carrying altered glycosylation, including STn^+-CD44 (146). These observations suggest that exploring CD44 glycosylation may improve its biomarker potential. Nevertheless, the comprehensive glycomics mapping of CD44, including *N*- and *O*-glycan patterns, in bladder cancer as yet to be performed.

Chapter 2

AIMS AND SCOPES

2. Aims and Scopes

Bladder Cancer presents one of the highest recurrence rates amongst solid tumours, and constitutes the second deadliest disease of the genitourinary track. Moreover, the only therapeutic approach for managing advanced patients are cisplatin-based regimens that often fail to prevent disease relapse and dissemination. The high molecular heterogeneity presented by advanced bladder tumours and the lack of specific biomarkers for targeting more aggressive cancer cells remain challenging topics in bladder cancer management.

According to the current state-of-the art, CD44 presents an opportunity to address more aggressive bladder cancer cells and consequently improve these patients management. However, the gene encoding for CD44 generally undergoes significant alterative splicing, which results in many functionally distinct isoforms of variable molecular weights and glycosylation sites. Nevertheless, at the moment there are no consensus regarding the nature of CD44 isoforms associated to muscle invasion, which would be crucial for designing targeted therapeutics.

Having as starting point preliminary associations between increased CD44 levels in urine of cancer patients with more aggressive forms of the disease (Appendix 1), the first part of this work will be devoted to confirming these observations. The second part of the work will focus in disclosing the nature of CD44 isoforms associated with muscle invasion and ultimately chemoresistance using molecular models and patient samples, envisaging the molecular rationale for intervention. Emphasis will also be set on CD44 O-glycosylation patterns with the objective of narrowing down the isoforms to clinical relevant species.

Based on these observations, this work comprehends the following specific aims:

i. Identification of clinically relevant CD44 isoforms for targeted therapeutics;
ii. *O*-glycome characterization of CD44 isoforms envisaging highly cancer-specific target;
iii. Analysis of CD44 isoform expression in context of chemotherapy

Chapter 3

MATERIAL AND METHODS

3. Material and methods

3.1. Patient and sampling

Between 2010 and 2011, thirty-one first void urine samples without visible signs of haematuria were prospectively collected before surgery of bladder cancer male patients with mean age of 70 (range, 45–89) years, attending the Portuguese Institute of Oncology of Porto (IPO-Porto; Portugal). Only patients that had not been previously submitted to neoadjuvant therapy were included. Corresponding formalin-fixed paraffin embedded (FFPE) tumours were also obtained for this study. Based on the World Health Organization urothelial carcinoma grading and staging criteria, three different groups were considered in this study, namely low-grade (n=15) and high grade (n=9) non-muscle-invasive bladder tumours (NMIBC) and muscle-invasive (n=7) tumours (MIBC). An additional 15 urines were obtained from healthy control male volunteers, mean age of 68 (range 41-82) years. A representative set spanning all stages of disease was also used for protein extraction. This included 3 Ta, 3 T1 and 3 MIBC tumours. Furthermore, a subset of 10 patients treated with neoadjuvant chemotherapy were included. A total of 20 formalin-fixed paraffin-embedded (FFPE) tissues sections were analysed, 10 from prior chemotherapy transurethral resections (TUR) and 10 from post-chemotherapy cystectomies. All procedures were performed under the approval of the institution ethics committee and upon patients' informed consent. All clinicopathological information was obtained from patients' clinical records.

3.2. Cell lines and cell culture conditions

The T24 (grade III), HT1376 (grade III) and 5637 (grade II) bladder cancer cell lines used in this work were acquired from DSMZ (Düsseldorf, Germany) and recently characterized from the genetic standpoint by our group (147). Accordingly, the T24 cell line is representative of the FGFR3/CCND1 pathway, presenting a mutated HRAS and overexpression of CCND1. The HT1376 and 5637 cell lines correspond to the E2F3/RB1 pathway with loss of one copy of RB1 and mutation of the remaining copy. Additionally, HT1376 cells exhibit deletion of the Phosphatase and tensin homolog (PTEN) gene and

no alteration of Phosphatidylinositol 3-kinase catalytic subunit alpha (PIK3CA), which in combination with the inactivation of p53, translates into a more invasive and metastatic potential. In contrast, the 5637 cell line does not present any loss of PTEN and loses PIK3CA gene, which gives it a less-invasive phenotype.

The cells were cultured in RPMI 1640+GlutaMAXTM-I medium (Gibco, Life Technologies) supplemented with 10% heat-inactivated FBS (Gibco, Life Technologies) and 1% penicillin-streptomycin (10,000 Units/mL P; 10,000 µg/mL S; Gibco, Life Technologies). Cell lines were cultured as a monolayer at 37˚C in a 5% CO2 humidified atmosphere, and were routinely subcultured after trypsinization.

Chemoresistance assays were performed by exposing $1x10^6$ cells of the above-mentioned cell lines to 3µM cisplatin (Teva Pharma, Portugal) for 72h (ic50 doses stablished by Arantes-Rodrigues et al. (148) In parallel, cells were cultured using complete media alone and used as controls. Immediately after the incubation periods, remaining cells were lysed for protein or RNA extraction. All experiments were performed in triplicates and the presented results reflect the average of three independent replicates.

3.3. Isolation of Urine Proteins

Five to forty millilitres of collected urine were centrifuged at 5000x*g* for 40 min at 4 °C to remove cells and debris, desalted on Amicon Ultra 10 kDa centrifugal filters (Merck KGaA, Darmstadt, Germany) and proteins were resuspended on 50 mM Ammonium bicarbonate (pH 7.8, Sigma-Aldrich, St. Louis, MO, EUA). Protein quantification was accessed using the DC Protein assay (Bio-Rad, Hercules, CA, USA).

3.4. CD44 analysis by Slot-Blot

Urine proteins (10 μg) were slot-blotted on a nitrocellulose membrane (Whatman, Protan; pore size 0.45 µm) using the Hybri-slot apparatus (21052-014; Gibco BRL, Life Technologies, Waltham, MA, USA). Fetuin from fetal calf serum and deglycosylated bovine serum albumin (both purchased at Sigma-Aldrich) were used as positive and negative controls, respectively. The amount of protein in the slots was confirmed by Ponceau S (Sigma-Aldrich). Nonspecific binding was then blocked with Carbo-Free Blocking Solution (SP-5040, Vector Laboratories) for 30 minutes at room temperature.

Samples were screened for CD44 using a recombinant polyclonal antibody (anti-CD44, 1:150 in PBS; ab157107; Abcam, Cambridge, UK) and a goat anti-rabbit IgG (H+L) Secondary Antibody, HRP (Invitrogen, Carlsbad, CA, USA). Reactive bands were detected by enhanced chemiluminescence ECL (Amersham Pharmacia Biotech Inc, Piscataway, NJ, USA) according to the manufacturer's instructions. Images were recorded using X-ray films (Kodak Biomax light Film, Sigma-Aldrich). The films were scanned in Molecular Imager Gel Doc XR+ System (Bio-Rad) and analyzed with QuantityOne software (v 4.6.3, Bio-Rad). CD44 evaluation was performed in triplicates and the presented results reflect the average of three independent measurements.

3.5. RNA extraction and mRNA expression analysis

RNA was isolated from FFPE tissue samples using “Absolutely RNA FFPE Kit” (Stratagene, La Jolla, CA), according to the vendor’s instructions. Regarding cultured cells, total RNA was extracted using TriPure isolation Reagent (Roche Diagnostic GmbH, Mannheim, Germany). The RNA quantity and its purity were determined based on the A260/A280 using a Nanodrop Lite Spectrophotometer (Fisher Scientific). Only samples with ratios between 1.9 and 2.1 were considered for downstream molecular studies.

Up to 2 mg of total RNA from tissues sections and adherent cells were reverse transcribed with random primers, using the “High Capacity cDNA Reverse Transcription Kit” (Applied Biosystems, Foster City, CA). The amplification conditions were the following: 25°C for 10 minutes, 37°C for 120 minutes, and reverse transcriptase inactivation at 85°C for 5minutes. The products were amplified in a StepOne Real-Time PCR System (Applied Biosystems) using TaqMan Gene Expression Master Mix, primers, and probes provided by Applied Biosystems. TaqMan expression assays were obtained from Applied Biosystems CD44: Hs01075864_m1; CD44v3-v10: Hs01081480_m1; CDv8-v10: Hs01081475_m1; and CD44s: Hs01081473_m1. Schematic representation of PCR probes recognition zones is shown on Figure 4. Thermal cycling conditions were: 10 minutes at 95°C followed by 45 cycles of 15 seconds at 95°C and 1 minute at 60°C. All reactions were run in duplicate. GAPDH gene was selected for normalization from a set of 4 housekeeping genes, ACTB, GAPDH, HPRT and 18S, since it presents higher stability among the bladder tumour samples (data not shown). Concerning cultured cells, HPRT and B2M were the housekeeping genes selected for normalization. The raw −ΔCt was

used to analyze the CD44 expression and therefore used as an estimate of the mRNA relative levels. ΔCt stands for the difference between the cycle threshold (Ct) of the amplification curve of the target gene and that of the GAPDH or the HPRT/B2M. The efficiency of the amplification reaction for each primer-probe is more than 95% (as determined by the manufacturer)

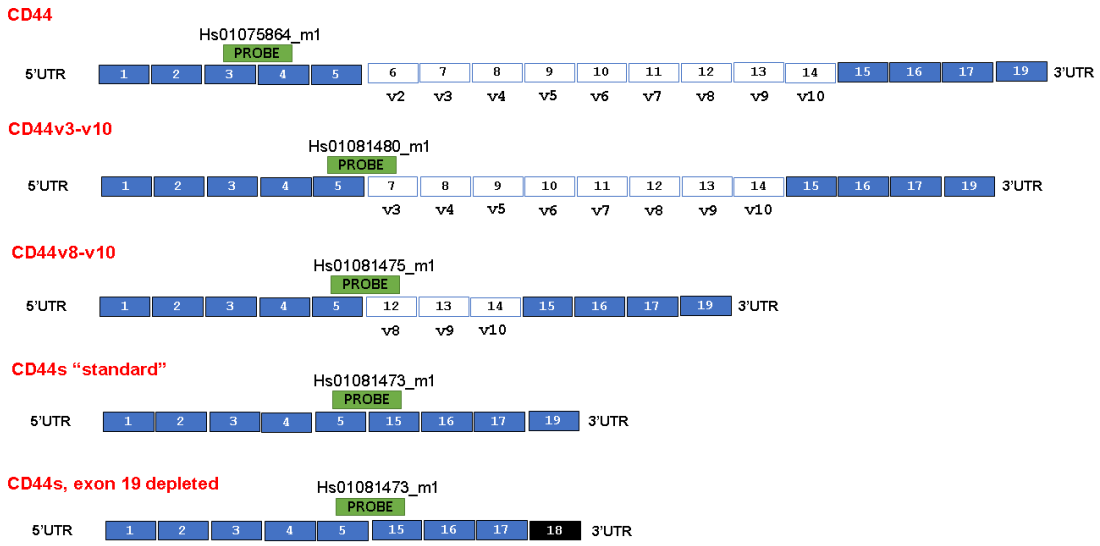

Figure 4 Schematic representation of PCR probes recognition zones. CD44 probe recognizes a sequence in the constitutive zone, amplifying any CD44 isoform. CD44v3-v10 is a probe designed to recognize exon-exon boundaries containing exon 5 and exon v3., while CD44v8-v10 recognizes exon-exon boundaries containing exon 5 and exon v8. CD44s is recognized by exon-exon boundaries containing exon 5 and 15 of the constitutive zone.

3.6. CD44 analysis by Western Blot

Proteins were extracted from T24, HT1376 and 5637 bladder whole cell lysates with RIPA lysis buffer (50mM Tris [pH=8], 150mM NaCl, 1%NP 40, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with protease and phosphatase inhibitor cocktail (Halt™ Protease & Phosphatase inhibitor cocktail; Thermo Scientific). Protein quantification was done using the DC protein assay kit (Bio Rad) according with the manufacturer instructions. Protein samples were separated in reducing SDS/PAGE gels, transferred to 0.45 mm nitrocellulose membranes (GE Healthcare Life Sciences, Uppsala, Sweden) and blotted for CD44 and β-actin. Prior to blotting protein loads were confirmed by Ponceau s (Sigma-Aldrich) staining. The membrane for CD44 and β-actin antigens was blocked with Carbo-free Blocking Solution 1X (Vector Laboratories) for 1h at room

temperature (RT) and incubated according to Table 1 specifications. After final washing, the antibodies were revealed by chemiluminescence using Amersham ECL Prime Western Blotting Detection Reagents (GE Healthcare). CD44 expression was normalized in relation to β-actin. Blots for three independent experiments were performed and the presented results reflect the average for three independent replicates.

Table 1 - Specification of primary and secondary antibodies conditions

PRIMARY ANTIBODY	SECONDARY ANTIBODY	CONDITIONS
Rabbit polyclonal to CD44 (abcam 157107)	Goat anti-Rabbit IgG (H+L) Secondary antibody, HRP conjugate (Invitrogen)	1º Ab dilution:1:10000, 1h, RT; 2º Ab dilution 1:60000, 30', RT
Anti-Beta-Actin Antibody Mouse Mono (Sigma-Aldrich)	Peroxidase-conjugated AffinPure Goat Anti-Mouse IgG (H+L) (Jackson ImmunoResearch Laboratories, Inc)	1º Ab dilution:1,4µg/µL, 1h, RT; 2º Ab dilution:1:70000, 30mins, RT;

3.7. Cellular *O*-glycome analysis by MALDI-TOF-MS

The analysis of the *O*-glycome in living cells using a chemical *O*-glycan precursor was performed using a modified Cellular *O*-Glycome Reporter/Amplification (CORA) method described by Kudelka et al. (149). Briefly, the Benzyl-α-D-GalNAc (Sigma-Aldrich, St. Louis, MO, USA) was acetylated to generate the Benzyl-α-D-GalNAcAc3 by incubation with 1:6 pyridine:acetic anhydride for 30 min at 30°C. The compound was purified using chloroform-water extraction and dried in a speedvac. Then, the Benzyl-α-D-GalNAcAc3 was dissolved in 50µL of DMSO and further diluted to 50µM, 150µM and 250 µM in complete media with 10% FBS and 1% penicillin-streptomycin. Media with compound was administered to approximately $10x10^6$ adherent T24 and HT1317 cells following a 24h, 48h and 72h incubation period after which conditioned medium was collected. The conditioned medium was filtered in 10kDa centrifugal filters (Amicon – Ultra 4, EMD Millipore, Burlington, MA, USA) upon 2500×g centrifugation for 1 h at 4ºC and the flow through was collected. Subsequently, the Benzyl-α-D-GalNAcAc3-O-glycans was purified from the flow through using a Sep-Pak 3 cc C18 cartridge (Waters, Milford, MA, USA) by gravity chromatography. The column was equilibrated with 2 × 2 ml acetonitrile, then 4 × 2

ml 0.1% TFA in water. The flow through was added to the column following washing with with 4 × 2 ml 0.1% TFA in water. Benzyl-α-D-GalNAcAc3-*O*-glycans were then eluted with 2 × 1.5 ml 50% acetonitrile/0.1% TFA in water and dried in a speedvac. The dried Benzyl-α-D-GalNAcAc3-*O*-glycans were permethylated using a standard protocol (150) with 500 µl DMSO/NaOH slurry followed by 200 µl methyl iodide. Samples were shaken for 30 min, another 50 µl methyl iodide was added and samples were shaken again for 20 min. The permethylated Benzyl-α-D-GalNAcAc3-*O*-glycans was purified using chloroform-water extraction: 1 ml chloroform and 1 ml water were added to supernatant, mixed, and centrifuged at 430xg for 2 minutes. Then the aqueous phase is successively removed and added for five washing cycles until chloroform is evaporated using a speedvac.

Benzyl-α-D-GalNAcAc3-*O*-glycans were resuspended in 50 µl 50% methanol. Then, 0.5 µl matrix (10 mg/ml 2,5-dihydrobenzoic acid, 50% acetonitrile, 0.1% TFA) and 0.5 µl sample were spotted on target plate, air dried, and analysed by a MALDI-TOF/TOF mass spectrometer (4800 Proteomics Analyzer, Applied Biosystems, Europe) in the positive ion reflector mode. Peak masses were identified and structures assigned by composition and knowledge of glycan biosynthetic pathways, or MS/MS where indicated. For the acquisition of tandem mass spectra (MS/MS) of each sample spot, a data dependent acquisition method was created to select the 10 most intense peaks, excluding those from the matrix. Fragmentation was promoted by collision induced dissociation (CID) with air at a pressure of 1.2 e-6 torr in the collision cell.

3.8. Isolation of CD44 isoforms by immunoprecipitation

Proteins were extracted from whole cells using cell Fractionation buffer (20mM HEPES [pH=7.4], 10mM KCl, $MgCl_2$ 2mM, EDTA 1mM, EGTA 1mM) containing protease and phosphatase inhibitors (Halt™ Protease & Phosphatase inhibitor cocktail; Thermo Scientific) and by passing the cell suspension through a 25-gauge needle. Whole cell lysates were centrifuged at 720xg for 5 min at 4ºC (cellular debris and nucleus exclusion), following a 10000xg centrifugation of the resultant supernatant for 5 minutes at 4ºC (mitochondria and zymogen granules exclusion). The subsequent supernatant was then ultracentrifuged at 100000xg for 1h at 4ºC to obtain membrane-bound proteins, following resuspension by up and down pipetting and passage through a 25-gauge needle. Utracentrifugation at 100000xg for 45 min at 4ºC follows. The obtained pellet was

resuspended in TBS/0.1% SDS buffer and protein quantified with the DC protein assay kit (Bio Rad) according to the manufacturer's instructions.

CD44 isoforms were immunoprecipitated from membrane-bound protein extracts with polyclonal antibody CD44 (ab157107; Abcam, Cambridge, UK) using the Pierce Direct Magnetic IP/CO-IP Kit (Thermo Scientific) according to the supplier's instructions.

3.9. CD44 isoforms identification by nanoLC-MS/MS

For proteomics analysis, CD44 isoforms isolated by immunoprecipitation were separated by 4–16% gradient SDS–PAGE under reducing conditions, the bands were excised from the gels and proteins reduced with 5 mM DTT (Sigma-Aldrich, Missouri, USA) for 40 min at 60˚C, alkylated with 10 mM iodoacetamide (Sigma-Aldrich, Missouri, USA) for 45 min in the dark, digested with trypsin (Promega, Wisconsin, USA) in situ, and identified by nanoLC-MS/MS. A nanoLC system (Dionex, 3000 Ultimate nano-LC) was coupled on-line to a LTQ-Orbitrap XL mass spectrometer (Thermo Scientific, Bremen, Germany) equipped with a nano-electrospray ion source (EASY-Spray source; Thermo Scientific). Eluent A was aqueous formic acid (0.2%) and eluent B was formic acid (0.2%) in acetonitrile. Samples (10 µl) were injected directly into a trapping column (C18 PepMap 100, 5 µm particle size) and washed with an isocratic flux of 95% eluent A and 5% eluent B at a flow rate of 30 µl/min. After 3 min, the flux was redirected to the analytical column (EASY-Spray C18 PepMap, 100 Å, 150 mm x 75µm ID and 3 µm particle size) at a flow rate of 0.3 µl/min. Column temperature was set at 35º C. Peptide separation occurred using a linear gradient of 5–50% eluent B over 90 min., 50–90% eluent B over 5 min. and 5 min. with 90% eluent B. The mass spectrometer was operated in the positive ion mode, with a spray voltage of 1.9 kV and a transfer capillary temperature of 250˚C. Tube lens voltage was set to 120 V. MS survey scans were acquired at an Orbitrap resolution of 60,000 for an m/z range from 300 to 2000. Tandem MS (MS/MS) data were acquired in the linear ion trap using a data dependent method with dynamic exclusion: The top 6 most intense ions were selected for collision induced dissociation (CID). CID settings were 35% normalized collision energy, 2 Da isolation window, 30 ms activation time and an activation Q of 0.250. A window of 45 s was used for dynamic exclusion. Automatic Gain Control (AGC) was enabled and target values were $1.00e^{+6}$ for the Orbitrap and $1.00e^{+4}$ for LTQ MSn analysis. Data were recorded with Xcalibur software version 2.1 and analysed automatically using the SequestHT search engine with the Decoy algorithm for validation

of protein identifications (Proteome Discoverer 1.4; Thermo Scientific, Massachusetts, USA). Data were searched against an internal database consisting of CD44 isoforms encoded by human proteome obtained from the SwissProt database on 4/12/2017. Trypsin was selected as the enzyme, allowing for up to 2 missed cleavage sites, a precursor ion mass tolerance of 10 ppm, and 0.6 Da for product ions. Carbamidomethylcysteine (+57.02146) was selected as a fixed modification while oxidation of methionine (+15.99491), modification of serine and threonine with HexNac (+203.07937), HexNacNeuNac (STn) (+494.17479), HexHexNac (T) (+365.13220), HexHexNacNeuNac (+656.22761) and HexHexNacNeuNacNeuNac (+947.32303) were defined as variable modifications, taking into account our previous publications in bladder cancer and *O*-glycomics analysis (146,151). CD44 is a heavily glycosylated protein and the identification of glycosylated peptides is often difficult, therefore, protein grouping filters were set to consider peptide identifications with low confidence to avoid loss of relevant peptides as false negatives. Identification of specific isoforms was based on the presence of isoform specific tryptic peptides and was only considered if the peptide spectrum match could be validated manually and presented unambiguous sequence coverage. For manual validation, typical carbohydrate fragmentation patterns were also considered, including sugar neutral loss, and multiple dehydrations.

3.10. Immunohistochemistry

Formalin-fixed, paraffin-embedded tissue sections (FFPE) were screened for CD44 expression by immunohistochemistry using the strepavidin/biotin peroxidase method. Briefly, 3 µm sections were deparaffinized with xylene, rehydrated with graded ethanol series, microwaved for 15 min in boiling citrate buffer (10mM Citric Acid, 0.05% Tween 20, pH 6.0, Sigma-Aldrich), and exposed to 3% hydrogen peroxide for 25 min. CD44 was detected using a recombinant polyclonal antibody (anti-CD44, 1:4000 in PBS; ab157107; Abcam, Cambridge, UK) after incubation overnight at 4°C. Upon digestion with an α-neuraminidase from Clostridium perfringens (Sigma Aldrich, St. Louis, MO, USA) for 4 h at 37°C, ST and dST antigens were detected using an hybridoma-derived 3C9 antibody 1:2 overnight at 4°C. Also, STn antigen was evaluated using a B72.3 monoclonal antibody (Abcam) 0.5ug/mL overnight at 4°C. The antigens were identified with UltraVision HRP Detection System Kit (Thermo Fisher Scientific, Waltham, MO, USA) followed by incubation with 3,3-diaminobenzidine tetrahydrochloride (Impact Dab, Vector

Laboratories) for chromogenic development. Finally, the slides were counterstained with Harris's haematoxylin for 1 min. Negative control sections were performed by adding BSA (5% in PBS) devoided of primary antibody. The immunostained sections were blindly assessed using light microscopy by two independent observers and validated by an experienced pathologist. Briefly, a semi-quantitative approach was established to score immunoreactivity based on the intensity and extension of the staining. The extension of staining was rated in cutoffs of 10%, and staining intensity was rated as follows: negative-0, weak-1, moderate-2, strong-3. The tumours were then classified based on the multiplication of extension evaluation and intensity. Disaccording readings were re-analysed using a double-headed microscope and consensus was reached.

3.11. *In situ* proximity ligation assays on tissue sections

The simultaneous detection of CD44 and STn+, ST+/dST+- glycoforms was achieved made by in situ proximity ligation assays (PLA), using the Duolink in situ detection reagents Red (Olink Bioscience, Uppsala, Sweden) according to the manufacturer's instructions and based on previous reports (143,146,152) FFPE tissues were deparaffinized, rehydrated and subjected to acid- and heat-induced antigen retrieval, followed by incubation with 3% hydrogen peroxide. Then, tissue sections were incubated with blocking solution in aunder humidified atmosphere chamber. CD44, STn and ST/dST antigens were detected by indirect PLA using the polyclonal antibody anti-CD44 1:2000 overnight at 4°C (clone ab157107; Abcam, Cambridge, UK), hybridoma-derived 3F1 antibody 1:4 overnight at 4°C, and hybridoma-derived 3C9 antibody 1:2 overnight at 4°C, respectively. For ST/dST detection, tissues were previously treated with an α-neuraminidase from Clostridium perfringens (Sigma Aldrich, St. Louis, MO, USA) for 4 h at 37°C. The PLA probes anti-rabbit MINUS and anti-mouse PLUS were both added and sections were incubated at 37°C for 1 h. Next, ligation was performed for 30 min at 37 °C, and amplification was carried out for 120 min at 37 °C to produce rolling circle products, followed by incubation with 4',6-diamidino-2-phenylindole for 10 min at room temperature and sample mountinged for fluorescence microscopy. PLA results were evaluated by two distinct observers and validated by an experienced pathologist, who independently registered cytolocalizationcolocalization of staining.

3.12. Statistical analysis

Statistical analysis was performed using Graphpad prism7 by GraphPad Software, Inc. Differences between continuous variables among the evaluated groups were accessed by Mann-Whitney non- parametric test for independent samples. Differences were considered significant when $p<0.05$.

Pair Wise Fixed Reallocation Randomization Test©, calculated by REST software, was used to determine the relative gene expression between pre- and post-chemotherapy tumours in non- responsive patients. This tool integrates randomization and bootstrapping methods that test the statistical significance of calculated expression ratios by the hypothesis test p (H1), representing the probability of the alternate hypothesis thatwhere the difference between groups is due only to chance (153).

Chapter 4

RESULTS AND DISCUSSION

4. Results

Previous work from our group has demonstrated the potential of glycan-affinity glycoproteomics nanoplatforms for urinary biomarkers discovery in bladder cancer (Appendix 1). Magnetic nanoprobes (MNP) coated with three broad-spectrum lectins (Concanavalin A, Wheat Germ Agglutinin, and *Sambucus nigra*) have been used to selectively capture glycoproteins from the urine of low-grade and high-grade non-muscle invasive and muscle-invasive BC patients. Glycoproteins were identified by nano-LC MALDI-TOF/TOF and data was curated using bioinformatics tools (UniProt, NetOGlyc, NetNGlyc, Cytoscape and Oncomine) to highlight clinically relevant species, which included the BC stem-cell marker CD44.

4.1. CD44 expression in urine and bladder tumours

This work starts with the orthogonal validation of previous MS-based results (Appendix 1) by determining CD44 levels in the urine of BC patients by slot-blot, as well as in the corresponding bladder tumours by immunohistochemistry. According to Figure 5, CD44 urinary levels showed a trend increase with the severity of the lesions; however, this effect is only significant for high-grade tumours in comparison to controls and low-grade NMIBC ($p<0.05$).

In parallel, immunohistochemical screening of the corresponding bladder cancer tissues for CD44 expression demonstrated that CD44 antigen was mainly found at the cellular membrane of cancer cells, in accordance with expected cellular location. A significant increase in CD44 expression with the severity of the lesions was seen, especially in high-grade tumours ($p<0,05$), thus mimicking urine analysis (Figure 5). More importantly, CD44 was not observed in the tumour-adjacent histologically normal urothelium of the one case with normal urothelium representation, neither in the most superficial lesions, supporting its overexpression in cancer. In superficial tumours, CD44 was predominantly found in basal layer cells, which frequently harbour more malignant clones. Contrastingly, MIBC cells presented an extensive and intense CD44 staining without a defined pattern throughout the tumour (Figure 5). These findings agree with MNP@ lectins-based glycoproteomics and reinforces the role of CD44 in the context of advanced disease (high-grade NMIBC and MIBC). This project further explores the expression of CD44 and its splicing variants, including its glycoforms, in BC cell lines aiming for CD44-associated signatures associated with more malignant phenotypes.

These results were subsequently validated in patient samples envisaging clinical translation.

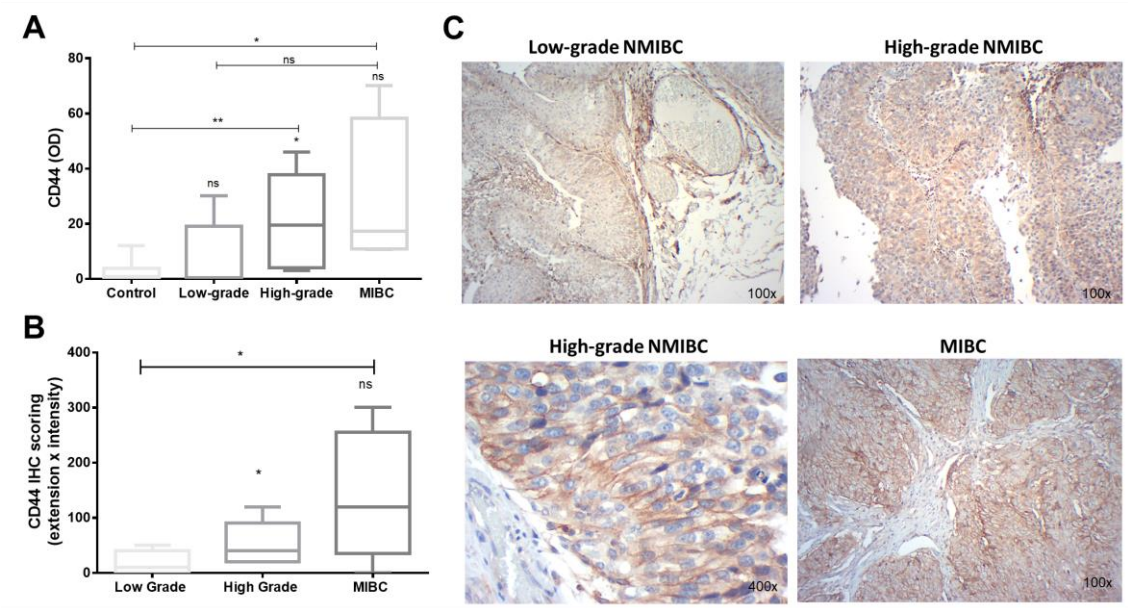

Figure 5 Validation of CD44 expression in urine samples and immunohistochemistry of bladder cancer patients. A. CD44 urinary levels showed a trend increase with the severity of the lesions; however this effect is only significant for high-grade tumours in comparison to the control and low-grade NMIBC groups. **B.** CD44 antigen was mainly found at membrane of cancer cells, in accordance with expected cellular location. Again, we observed a significant overexpression of CD44 with the severity of the lesions, which is more pronounced for high-grade tumours, thus, mimicking urine analyses. **C.** In superficial tumours, CD44 was predominantly found in cells underlining the basal layer, which have been described to harbor more malignant clones. Contrastingly, MIBC cells presented an extensive and intense CD44 expression without a defined pattern throughout the tumour. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

4.2. CD44 expression in bladder cancer cell lines

CD44 expression was evaluated in three bladder cancer cell models, two highly undifferentiated cells lines isolated from invasive urothelial carcinomas (T24 and HT1376) and a more differentiated cell line (5637), isolated from a non-muscle invasive tumour. A preliminary screening of these cell lines by western blot highlighted key differences in CD44 major bands, demonstrating a different CD44 expression profile between invasive (T24 and HT1376) and non-invasive carcinomas derived (5637) cell lines (Figure 6). Namely, T24 and HT1376 presented a major band between 75-100 kDa, whereas the 5637 cell line presented a major band between 100 and 150 kDa. Interestingly, all cell lines presented vestigial bands bellow 50 kDa, suggesting the expression of lower molecular weight CD44 isoforms (data not shown).

Complementary analysis by RT PCR, using commercially available probes for all forms of CD44 (total CD44), CD44s, CD44v3-v10, and CD44v8-v10 (Figure 4), confirmed

the existence of distinct CD44 profiles between invasive and non-invasive cell lines. Particularly, invasive T24 and HT1376 exhibit high expression of CD44s, while under-expressing CD44v3-10 and CD44v8-10. The 5637 cell line presents the exact opposite profile (Figure 7). Interestingly, CD44s (39.4 kDa) presents a lower molecular weight when compared to CD44v8-10 (53.4 kDa) and CD44v3-10 (76.7 kDa). These observations suggest that that more aggressive bladder cancer cells are enriched for lower molecular weight CD44 isoforms in comparison to the less aggressive cell line (5637), thus in accordance with western blot analysis. Differences between the predicted molecular weight and western blot bands may results from marked difference in the number of *O*-glycosylation sites (CD44s: 32 putative *O*-glycosylation sites; CD44v8-10: 79 putative *O*-glycosylation sites; CD44v3-10: 140 putative *O*-glycosylation sites) and possibly *O*-glycan structures.

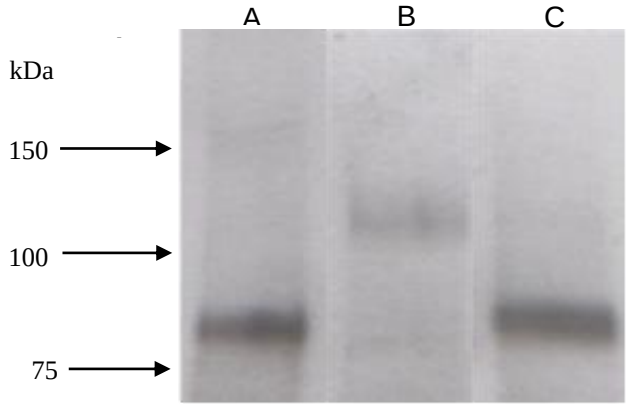

Figure 6 Western blot of T24 (A), 5637 (B) and HT1376 (C) presenting the most intense band for CD44. T24 present a predominant band, with high intensity, between 75-100 kDa. HT1376 exhibits a band pattern similar to T24, with a predominant and high intensity band between 75-100 kDa. 5637 shows a predominant band between 150-100 kDa and a much less intense band between 75-100 kDa.

Bladder cancer cell lines were submitted to ic50 doses of cisplatin chemotherapy and CD44 was evaluated by by western blot in untreated cells (controls) and the remaing population after chemotherapy challenge (ic50). According to Figure 8, T24 presented higher levels of CD44 in relation to HT1376 and 5637 cell lines. Nevertheless, both the CD44 profile and expression levels remained unaltered in all cell lines upon chemotherapeutic challenge. Given the higher levels of CD44 in T24 cell line, further studies were only conducted using this model. Accordingly, RT PCR analysis of ic50 resistant T24 subpopulation confirmed the WB results of unchanged CD44 profile, since upon chemotherapeutic challenge maintained high expression of CD44s, while under-expressing CD44v3-v10 and CD44v8-v10.

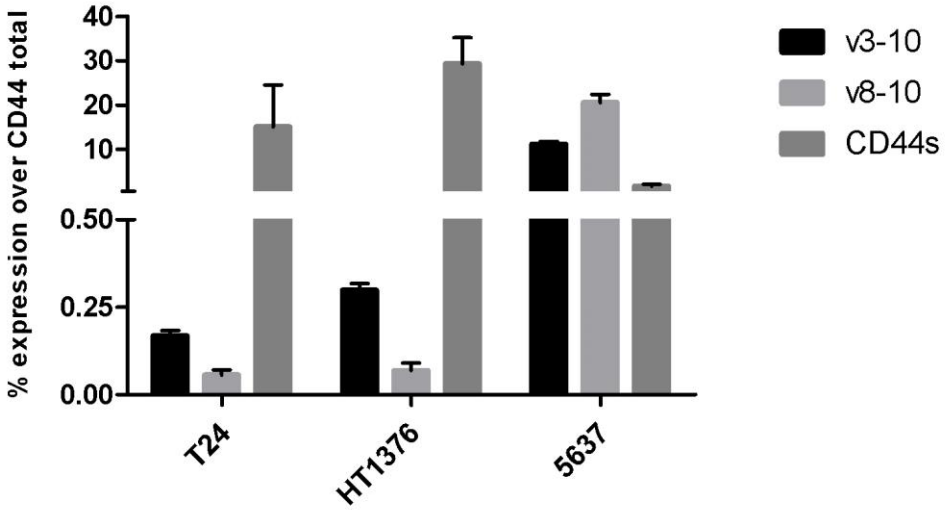

Figure 7 Graphical representation of CD44 isoforms expression in BC cell lines. T24 and HT1376 express high ratios of CD44s, while under-expressing CD44v3-v10 and CD44v8-v10. 5637 cell line exhibits an opposite expression profile.

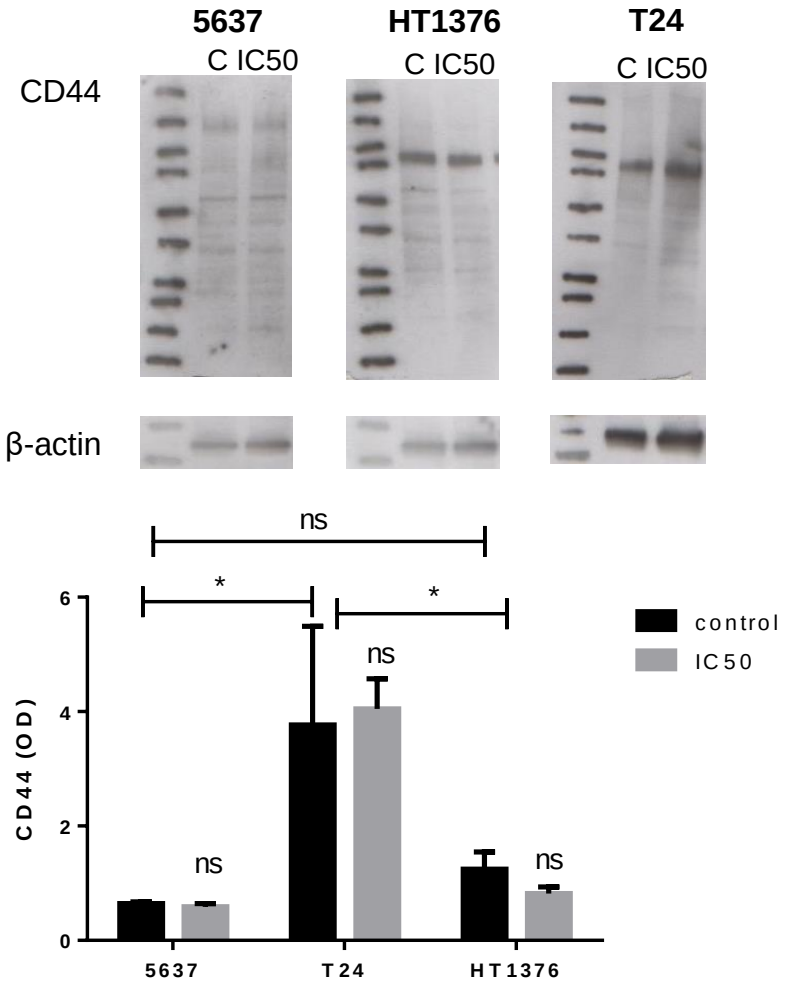

Figure 8 CD44 expression in superficial tumour derived cell line 5637 and muscle-invasive bladder cancer derived T24 and HT1376 cell lines. CD44 levels are significantly higher in T24 cell line in comparison to 5637 and HT1376 cell lines. Nevertheless, CD44 levels did not vary for any cell line after exposure to IC50 doses of Cisplatin.

4.3. *O*-glycome and glycoproteomics profile of CD44

A glycoproteomics approach was used to validate CD44 isoforms previously accessed by real-time PCR in T24 cells (CD44 v3-v10, CD44 v8-v10 and CD44s) and to broaden the current knowledge on this field. T24 cell line was elected for this preliminary glycoproteomics analysis based on its higher CD44 expression in comparison to other cell lines (Figure 8). Nevertheless, the high glycosylation degree of this protein (approximately 150 glycosylation sites for the canonical isoform) poses a significant hurdle for isoforms identification by mass spectrometry (154). Therefore, an *O*-glycomics approach was first performed to highlight the most abundant glycoforms produced by T24 cells. Accordingly, the most predominant forms included the mono- and di-sialylated T antigens (m/z=954 and m/z=1315, respectively) that have been previously found overexpressed in advanced bladder tumours (146,155) In addition, this experiment also demonstrate the presence of STn antigen (m/z=783), including its monoacetylated form (m/z=825), which are important for disease progression and dissemination (126,129,156,157) Dissialylated core 2 (m/z=1764) and fucosylated core 1 (m/z=767) glycans were also detected, but their biological and clinical significance remains unknown in bladder cancer. In addition there was a minor expression of more elongated, sialylated and fucosylated *O*-glycoforms (Figure 9), whose role in bladder cancer also warrants future clarification.

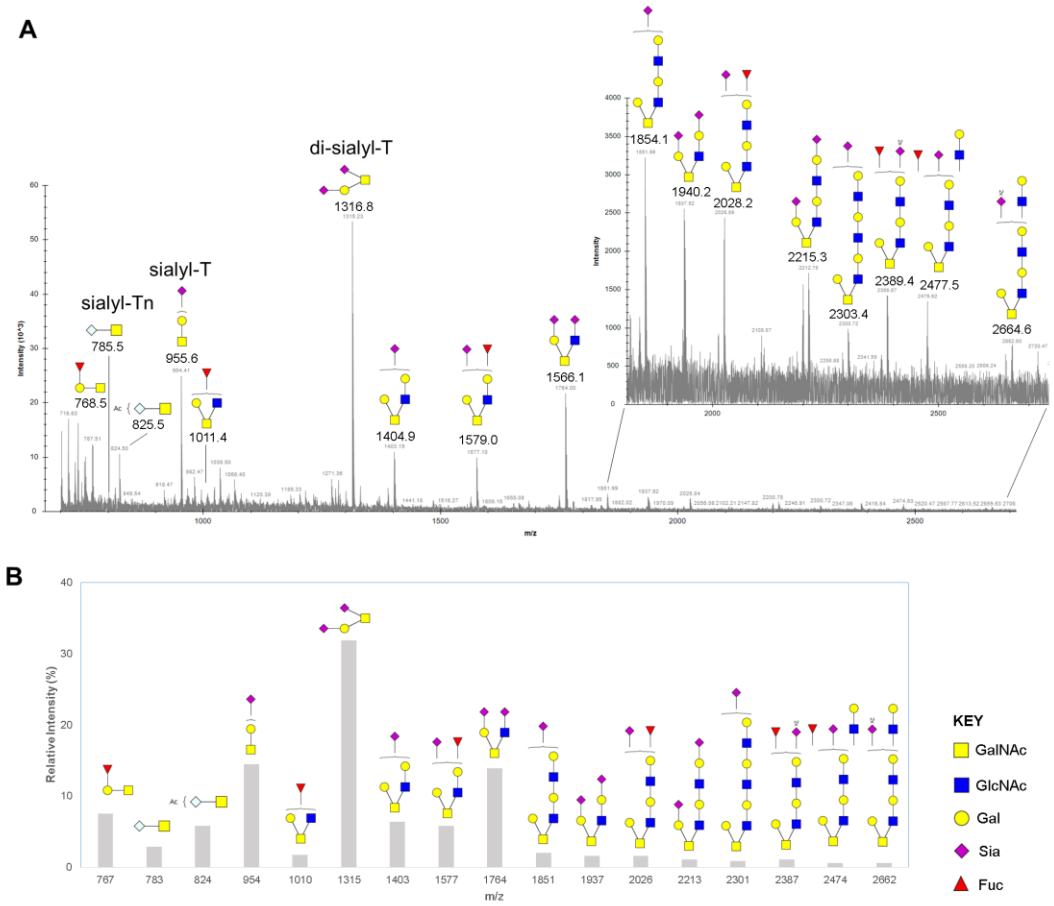

Figure 9 Mass spectrometry spectra of identified *O*-glycoforms in T24 cells. The most predominant forms included the mono- and di-sialylated T antigens (*m/z*=954 and *m/z*=1315, respectively), dissialylated core 2 (*m/z*=1764) and fucosylated core 1 (*m/z*=767). In addition, there is minor expression of more elongated, sialylated and fucosylated *O*-glycoforms.

In parallel, CD44 and its isoforms were immunoprecipitated from membrane proteins that were isolated from protein lysates by differential ultracentrifugation. CD44 immunoprecipitates were first analyzed by western blot (Figure 10A), which confirmed the main band between 75 and 100 kDa but also several bands at both higher and, particularly, lower molecular weights, suggesting a broader array of CD44 iso- and glycoforms. Then CD44 isoforms were separated by SDS-PAGE, bands were excised from gel, digested with trypsin and identified by nano-LC MS/MS. Take into account that the CD44 isoforms definition is governed by variable exons encoding for a significant number of *O*-glycosylation sites, information on the glycome was comprehensively included for protein identification. This approach led to the unequivocal identification of 12 isoforms. In addition to confirming the expression of CD44v3-v10, CD44v8-10 and CD44s, previously accessed by RT-PCR targeted approach, it included CD44v10, CD44 NCBI

Transcript 7 and CD44RC and also Uniprot isoforms 2, 3, 5, 7, 8, 9, 13 and 14 (predicted isoforms) (Figure 10A), information summarized in table 2. Interestingly, the most predominant band between 75 and 100 kDa in western blots, comprehended isoforms 3, 4 (CD44v8-10 also analyzed by RT-PCR), isoform 7, and isoform 12/15 (CD44s also analyzed by RT-PCR). More importantly, many CD44 isoforms spanned different molecular weights most likely reflecting distinct *O*-glycosylation patterns. In addition, glycopeptides carrying previously described cancer-associated glycans could be identified for all isoforms. Specifically, the short isoform CD44s, associated with muscle invasive phenotypes and chemoresistance, has demonstrated specific *O*-glycosylation sites exhibiting di-sialyl-T and STn glycoforms (Figure 10B and C).

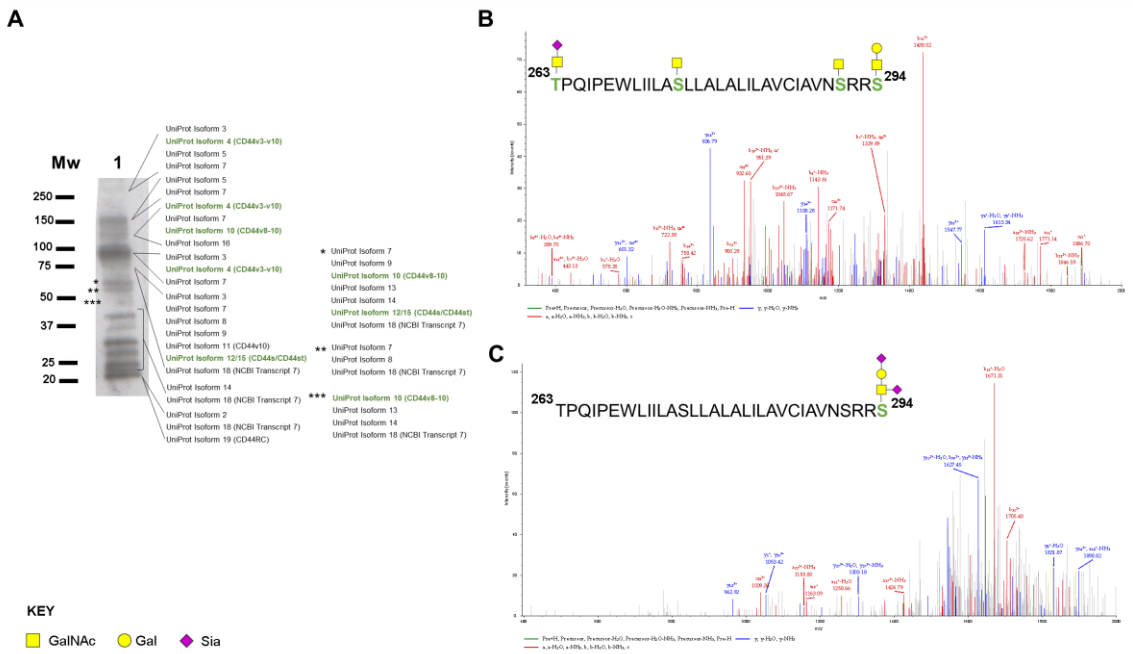

Figure 10 Identification of CD44 isoforms by mass spectrometry. A. A targeted proteomics approach led to the unequivocal identification of 12 CD44 isoforms; namely, CD44v3-v10, CD44v8v-v10, CD44s, CD44v10, CD44 NCBI Transcript 7, CD44RC, and Uniprot isoforms 2, 3, 5, 7, 8, 9, 13 and 14 (predicted isoforms). **B.** MS/MS spectrum of CD44s specific glycopeptide glycosylated with the STn antigen as well as Tn and T. **C.** MS/MS spectrum of CD44s specific glycopeptide glycosylated with the ST antigen.

Table 3 - Representation of CD44 isoforms analysed compared to decribed and predicted isoforms

Predicted CD44 isoforms	CD44 described isoforms	Identified by MS/MS	Identified by PCR
Isoform 1	Isoform 1		
Isoform 2		Isoform 2	
Isoform 3		Isoform 3	
Isoform 4	Isoform 4 (CD44v3-v10)	Isoform 4	Isoform 4
Isoform 5		Isoform 5	
Isoform 6			
Isoform 7		Isoform 7	
Isoform 8		Isoform 8	
Isoform 9		Isoform 9	
Isoform 10	Isoform 10 (CD44v8-v10)	Isoform 10	Isoform 10
Isoform 11	Isoform 11 (CD44v10)	Isoform 11	
Isoform 12	Isoform 12 (CD44s)	Isoform 12	Isoform 12/15
Isoform 13		Isoform 13	
Isoform 14		Isoform 14	
Isoform 15	Isoform 15 (CD44st)	Isoform 15	Isoform 12/15
Isoform 16			
Isoform 17			
Isoform 18	Isoform 18 CD44 NCBI Transcript 7	Isoform 18	
Isoform 19	Isoform 19 (CD44RC)	Isoform 19	

To validate the existence of abnormally glycosylated CD44, muscle-invasive bladder tumour tissues were then screened for CD44 and mono- and di-sialyl-T glycoforms, which are the most abundant glycans in *O*-glycomics analysis. In addition, the sialyl-Tn antigen, which has been previously associated with more aggressive phenotypes, was also evaluated. Two complementary validation strategies were applied, first consecutive tumour tissues were screened for CD44 and cancer associated *O*-glycans by immunohistochemistry. In parallel, the same tumour sections were screened for CD44 and the targeted glycans by in situ proximity ligand assays (PLA), which allowed the simultaneous detection of both targets when in close proximity. Since there are no monoclonal antibodies for mono- and di-sialyl-T antigens, the tissues were first digested with an alpha-neuraminidase to expose the T antigen. Subsequently, detection was conducted using a commercial available against the T antigen. Importantly, prior to analyses, the tumour tissues were found to be negative or present just vestigial T antigen expression, in accordance to previous observations from our group (146). Upon neuraminidase digestion, T antigen presented an extensive and intense expression consistent with the presence of ST and/or dST antigens. Immunohistochemistry analyses further demonstrated the coexpression of the glycans and CD44 in the same tumour areas, which was confirmed by PLA analyses (Figure 11). These findings support the existence of CD44 carrying cancer-associated glycoforms in bladder tumours and validate the results of the proteomics approach.

In summary, glycomics and glycoproteomics analysis confirmed the existence of CD44 variants previously observed by RT-PCR, which have allowed distinction between more aggressive and superficial cancer cells, In particular, it confirmed the presence of CD44s, which has been associated with cell lines derived from muscle invasive tumours and chemoresistance. Nevertheless, it also highlighted the presence of several other isoforms, which warrant comprehensive evaluation in future studies. Moreover, it demonstrated that CD44s was glycosylated with sialylated short-chain *O*-glycans frequently present in invasive tumours, providing key insights for designing highly specific targeted therapeutics.

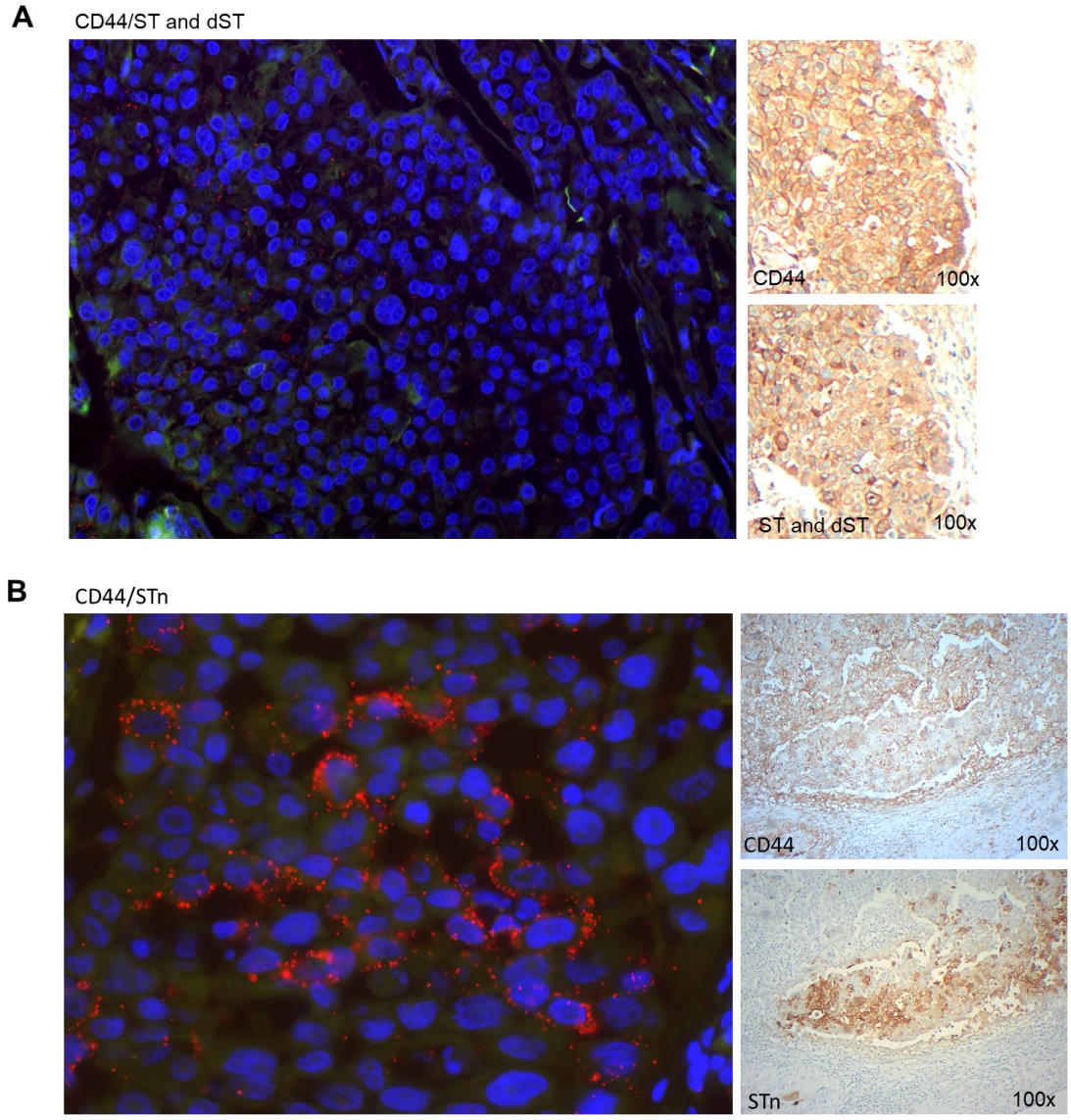

Figure 11 *In situ* proximity ligand assay and immunohistochemistry for the simultaneous detection of CD44/ST and dST as well as CD44/STn glycoforms. Two complementary validation strategies were applied: first consecutive tumour tissues were screened for CD44 and cancer associated *O*-glycans by immunohistochemistry (right); second, the same tumour sections were screened for CD44 and the targeted glycans by *in situ* proximity ligand assays (left), which allowed the simultaneous detection of both targets when in close proximity. Results showed the co-expression of CD44, ST, and dST (**A**) as well as STn and CD44 (**B**) in the same tumour areas and support the existence of CD44 carrying cancer-associated glycosylation in bladder tumours.

4.4. CD44 expression in bladder cancer tumours

To explore the CD44 expression profile in bladder tumours, western blot analysis was performed using extracted proteins from tumours representative of high grade tumours (more aggressive) from all stages of the disease . In agreement with T24 and HT1376 cell lines, bladder tumours also presented a main band between 75 and 100 kDa, denoting a common ground at the CD44 level (Figure 6 and 12). A major band at 250 kDa

was also observed in bladder tissues that may correspond to CD44 canonical form, which presents approximately 150 putative *O*-glycosylation sites. Interestingly, bladder tumours also exhibit low molecular weight forms (bellow 37 kDa) of CD44, as also previously observed for T24 cell line (Figure 10 and 12). These may correspond to truncated and/or poorly glycosylated CD44 isoforms, whose nature warrant comprehensive elucidation in future studies. Nevertheless, these protein were isolated from FFPE tissues, which may ultimately lead to proteolysis and the accumulation of low molecular weight species. Therefore, studies using freshly collected surgical specimens will be required to validate these findings.

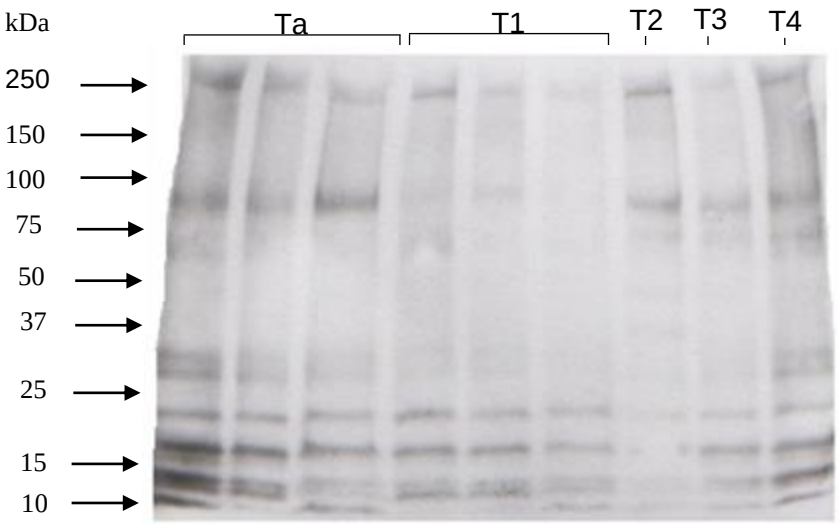

Figure 12 Western blot analysis of bladder cancer tumours. CD44 expression patterns are unaltered by disease stage. However, all major CD44 bands in bladder tumour samples are of low molecular weights.

Nevertheless, particular emphasis was given to confirming by RT-PCR the associations between CD44v3-10, CD44v8-10 and CD44s with the severity of the lesions, suggested upon the analysis of cell models. Accordingly, NMIBC presented high levels of CD44v3-v10, while under-expressing CD44v8-v10 and CD44s. In turn, MIBC presents significantly higher levels of CD44s ($p<0.05$) compared with NMIBC (Figure 13). Moreover, no differences were found regarding CD44v3-v10 and CD44v8-v10 between NMIBC and MIBC. These observations support the existence of a CD44v8-10^{Low}, CD44v3-10^{Low} and CD44s^{high} phenotype for muscle invasive tumours, which contrasts with the CD44v8-10^{high}, CD44v3-10^{high} and CD44s^{low} phenotype presented by NMIBC.

A group of 10 patients non-responsive to neo-adjuvant chemotherapy with cisplating were selected for CD44 isoform characterization, and analysis was performed in

tumours before and after chemotherapy. Upon chemotherapeutic challenge, total CD44 and CD44s expression showed approximately a 3-fold-increase, while CD44v3-v10 and CD44v8-v10 demonstrated a 50% decrease (Figure 14). These results suggests that CD44s might have a role in chemoresistance to cisplatin in bladder cancer patients, supporting previous observations using T24 cell model.

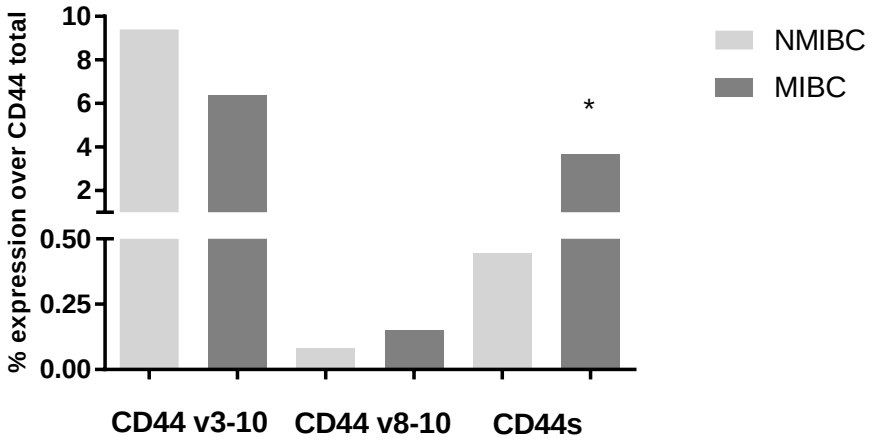

Figure 13 Graphic schematization of CD44 isoforms expression in NMIBC and MIBC. NMIBC exhibited high expressions of CD44v3-v10 and low CD44s, while MIBC low expressions of CD44v3-v10 and high expression of CD44s. MIBC expressed higher levels of CD44s compared to NMIBC *p<0.05.

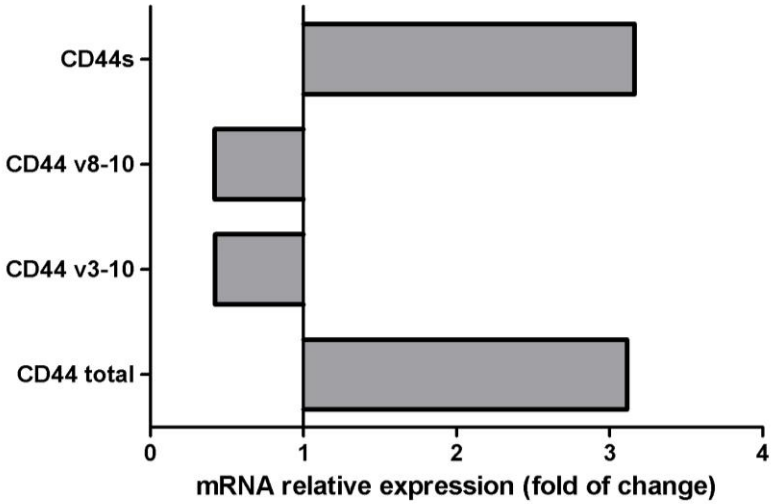

Figure 14 Graphic schematization of CD44 expression in pre- and post-chemotherapy tumour samples. Both total CD44 and CD44s show a 3-fold increased expression upon chemotherapeutic challenge, while CD44v3-v10 and CD44v8-v10 had a 50% decrease.

In summary, patient samples analysis demonstrates a CD44 remodelling accompanying muscle invasion, translated by the transition from high-molecular weight to low molecular weight isoforms, as previously suggested using cell models. More importantly, data suggests that CD44s low molecular weight isoforms may be a marker of chemoresistant clones.

Chapter 5

DISCUSSION, CONCLUDING REMARKS AND FUTURE PERSPECTIVES

5. Discussion, concluding remarks and future perspectives

Bladder Cancer (BC) presents one of the highest recurrence rates amongst solid tumours, and constitutes the second deadliest disease of the genitourinary track. The introduction of molecular models for disease management and effective targeted therapeutics remains a challenging aspect due to significant inter and intra-tumour molecular heterogeneity. Therefore, the management of BC patients presents significant hurdles due to high recurrence rates, rapid progression, dissemination, and poor response to chemotherapy. As previously described, several studies have demonstrated that CD44 has an important role in carcinogenesis, being a biomarker of undifferentiated stem-like cell phenotypes (26,158). As such, CD44 offers the opportunity for specifically target more malignant tumour cell phenotypes, including chemoresistant clones (159). Importantly, CD44 gene suffers extensive alternative splicing, generating several isoforms whose biological and clinical role in BC is still poorly understood. These observations allied to conflicting reports has hampered the true development of targeted therapies involving this marker. These discrepancies can be explained by the lack of standard immunohistochemical assays, the use of antibodies with different specificities, and differences in the clinicopathological status of bladder tumours used in the different studies. For instance, CD44s expression was associated with higher disease grade, stage, and density of tumour infiltrating lymphocytes, as well as with shorter overall survival, while CD44v6 expression was associated with better prognosis, lower grade and increased overall survival (81). Contrastingly, other authors reported that the loss of CD44v6 expression is an independent adverse predictor of recurrence and overall survival, further suggesting that routine evaluation of CD44v6 could be important for identifying high-risk patients which would benefit from more aggressive therapeutic approaches. Also, it was hypothesised that CD44v6 expression could be a much better predictor of recurrence than BC grade (82).

In order to gain knowledge regarding CD44 and its isoforms relevance in BC, a preliminary study concerning CD44 expression in urine and bladder tumours demonstrated that CD44 urinary levels increased with the severity of the lesions; however, this effect is only significant for high-grade tumours in comparison to controls and low-grade NMIBC. Furthermore, tumour expression of CD44 also increases with the severity

of the lesions, which is more pronounced for high-grade tumours, thus, mimicking urine analysis. More importantly, CD44 was not observed in the tumour-adjacent histologically normal urothelium of the one case with normal urothelium representation, neither in the most superficial lesions, supporting its overexpression in cancer. In superficial tumours, CD44 was predominantly found in basal layer cells, which has been described to harbour more malignant clones (80,81). Contrastingly, MIBC cells presented an extensive and intense CD44 staining without a defined pattern throughout the tumour. Given these insights, this work brings together a targeted-analysis of CD44 isoforms present in bladder cancer cell lines and bladder tumours, as well as a non-target, more explorative approach to broaden our knowledge on the isoforms present in both cell lines and patient samples. Accordingly, a short CD44 isoform, specifically CD44s, is predominantly expressed by more aggressive cell lines and tumours. On the other hand, more elongated isoforms (CD44v3-10 and CD44v8-10) have been found associated with more superficial lesions, in both cell models and in human tumours. These observations strongly suggest that the acquisition of invasive and/or other more aggressive traits may be accompanied by a remodelling of CD44 expression patterns towards shorter isoforms. Moreover, upon chemotherapeutic challenge, CD44s becomes overexpressed or unchanged in BC cell lines and in muscle-invasive bladder tumours that do not respond to chemotherapy, highlighting its relevance in the acquisition of chemoresistance traits in bladder cancer. Of note, CD44s has been previously associated with undifferentiated bladder cancer cells, higher disease stage and grade, depth of invasion, density of tumour infiltrating lymphocytes and short overall survival (80,81,160). However, other reports suggest that increased CD44s is related to low grade tumours and increased overall survival (83,161). Conflicting results are also found in other tumours models as hepatocellular, breast and colorectal cancers. Particularly, some studies indicate that CD44s overexpression is related to invasion, lymph node metastasis, aggressive disease, and expression of mesenchymal markers (162,163), while other studies suggest that CD44s is related to low grade tumours and higher overall survival (164). Interestingly, we observed that T24 cells that overexpress CD44s after chemotherapy challenge concomitantly activate a series of epithelial-to-mesenchymal (EMT)-related genes (165), thereby acquiring more motile phenotypes. These observations support that CD44s overexpression may be associated with invasive traits, at least for certain bladder cancer cells subpopulation. In agreement with these observations, the functional role for CD44 alternative splicing during EMT has already been demonstrated in breast cancer (166). Particularly, upon EMT programs activation a switch in CD44 isoform expression from CD44 variants to CD44s occurs.

Importantly, the overall level of CD44 protein did not significantly change during this process. Furthermore, CD44s expression, but not CD44 variants expression, was functionally linked to Akt activation, upregulation of anti-apoptotic molecules and apoptosis resistance. Given these insights, it is possible that a chemotherapy-induced EMT activation could lead to CD44s overexpression in bladder cancer. As such, CD44s seems to hold tremendous potential to target more aggressive and chemoresistant BC cells. Future studies should now focus on disclosing a link between CD44s overexpression and both invasion and chemoresistance. Moreover, a glycoproteomics approach has disclosed a wider array of isoforms for T24 cells line, including lower molecular weight protein species that should also be addressed in other cell lines and bladder tumours to provide a more comprehensively addressed envisaging clinically relevant and biomarkers.

In addition to isoform remodelling CD44 may also experience glycome remodelling accompanying significant alterations at this level experienced by bladder cancer cells (18,145,146). As such, the combined knowledge of clinically relevant isoforms of CD44 in BC with its associated glycosylation can generate highly specific targets for aggressive BC clones. Supporting this hypothesis, previous work from our group has also highlighted that CD44 presents the cancer-associated STn antigen as posttranslational modifications (145,146). The *O*-glycome characterization of aggressive BC cell line T24 has now shown the overexpression of sialylated short-chain of *O*-glycans, as ST, dST and STn, mimicking its expression in advanced stage bladder tumours (126,146). Downstream glycoproteomics studies confirmed the existence of CD44 isoforms carrying these glycans, including CD44s. The presence of CD44-STn and CD44-ST and/or dST was also confirmed in muscle invasive bladder tumours. As such, future studies should devout to explore the functional role of CD44s in BC progression, as well as glycomapping its structure, envisaging highly specific targets. Silencing of long CD44 isoforms in superficial cell lines overexpressing these glycans is currently being evaluated as a starting point for functional studies. Based on this molecular rational, specific antibodies against CD44s glycoforms could be developed. Such ligands may either be used to block oncogenic signalling transduction pathways potentially mediated by CD44, deliver drugs to more aggressive and chemoresistant clones and/or elicit immune responses against these cells. Furthermore, the generated knowledge could be employed for addressing metastatic disease, since CD44 could be functionally associated with metastatic propagation by stem-like cancer cells in bladder cancer.

References

References

1. Babjuk M, Oosterlinck W, Sylvester R, Kaasinen E, Böhle A, Palou-Redorta J, Rouprêt M. EAU guidelines on non-muscle-invasive Urothelial carcinoma of the bladder: Update 2013. European Urology. 2013;64(4):639–53.
2. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. International Journal of Cancer. 2010;127(12):2893–917.
3. Goodison S, Rosser CJ, Urquidi V. Bladder cancer detection and monitoring: Assessment of urine- and blood-based marker tests. Vol. 17, Molecular Diagnosis and Therapy. 2013. p. 71–84.
4. Zhang Y. Understanding the gender disparity in bladder cancer risk: the impact of sex hormones and liver on bladder susceptibility to carcinogens. Journal of environmental science and health Part C, Environmental carcinogenesis & ecotoxicology reviews [Internet]. 2013;31(March 2015):287–304.
5. Knowles MA, Hurst CD. Molecular biology of bladder cancer : new insights into pathogenesis and clinical diversity. Nature Publishing Group. 2015;15(1):25–41.
6. Burger M, Catto JWF, Dalbagni G, Grossman HB, Herr H, Karakiewicz P, Kassouf W, Kiemeney LA, La Vecchia C, Shariat S, Lotan Y. Epidemiology and risk factors of urothelial bladder cancer. European Urology. 2013;63(2):234–41.
7. Kaufman DS, Shipley WU, Feldman AS. Bladder cancer. Vol. 374, The Lancet. 2009. p. 239–49.
8. Mason RA, Morlock E V., Karagas MR, Kelsey KT, Marsit CJ, Schned AR, Andrew AS. EGFR pathway polymorphisms and bladder cancer susceptibility and prognosis. Carcinogenesis. 2009;30(7):1155–60.
9. Dovedi SJ, Davies BR. Emerging targeted therapies for bladder cancer: A disease waiting for a drug. Vol. 28, Cancer and Metastasis Reviews. 2009. p. 355–67.
10. Markl IDC, Jones PA. Presence and location of TP53 mutation determines pattern of CDKN2A/ARF pathway inactivation in bladder cancer. Cancer Research. 1998;58(23):5348–53.
11. Shariat SF, Tokunaga H, Zhou J, Kim J, Ayala GE, Benedict WF, Lerner SP. P53, p21, pRB, and p16 expression predict clinical outcome in cystectomy with bladder cancer. Journal of Clinical Oncology. 2004;22(6):1014–24.
12. Cote RJ, Datar RH. Therapeutic approaches to bladder cancer: Identifying targets and mechanisms. In: Critical Reviews in Oncology/Hematology. 2003.
13. Wu X-R. Urothelial tumorigenesis: a tale of divergent pathways. Nature reviews Cancer. 2005;5(9):713–25.
14. Witjes JA, Compérat E, Cowan NC, De Santis M, Gakis G, Lebret T, Ribal MJ, Van Der

Heijden AG, Sherif A. EAU guidelines on muscle-invasive and metastatic bladder cancer: Summary of the 2013 guidelines. Vol. 65, European Urology. 2014. p. 778–92.

15. Humphrey PA, Moch H, Cubilla AL, Ulbright TM, Reuter VE. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs—Part B: Prostate and Bladder Tumours. European Urology. 2016;70:106–19.
16. Botteman MF, Pashos CL, Redaelli A, Laskin B, Hauser R. The health economics of bladder cancer: a comprehensive review of the published literature. PharmacoEconomics. 2003;21(18):1315–30.
17. van Rhijn BW, Burger M, Lotan Y, Solsona E, Stief CG, Sylvester RJ, Witjes JA, Zlotta AR. Recurrence and progression of disease in non-muscle-invasive bladder cancer: from epidemiology to treatment strategy. Eur Urol. 2009;56(3):430–42.
18. Azevedo R, Peixoto A, Gaiteiro C, Fernandes E, Neves M, Lima L, Santos LL, Ferreira JA. Over forty years of bladder cancer glycobiology: Where do glycans stand facing precision oncology? Oncotarget. 2017 Oct 31;8(53):91734–64.
19. Amling CL. Diagnosis and management of superficial bladder cancer. Current Problems in Cancer. 2017 Jan 29;25(4):219–78.
20. Hussain MHA, Wood DP, Bajorin DF, Bochner BH, Dreicer R, Lamm DL, O'Donnell MA, Siefker-Radtke AO, Theodorescu D, Dinney CP. Bladder Cancer: Narrowing the Gap Between Evidence and Practice. Journal of Clinical Oncology. 2009 Dec 1;27(34):5680–4.
21. Zlotta AR, Fleshner NE, Jewett MA. The management of BCG failure in non-muscle-invasive bladder cancer: An update. Vol. 3, Journal of the Canadian Urological Association. 2009.
22. Dall'Era MA, Cheng L, Pan C-X. Contemporary management of muscle-invasive bladder cancer. Expert review of anticancer therapy. 2012;12(7):941–50.
23. Stein JP, Lieskovsky G, Cote R, Groshen S, Feng a C, Boyd S, Skinner E, Bochner B, Thangathurai D, Mikhail M, Raghavan D, Skinner DG. Radical cystectomy in the treatment of invasive bladder cancer: long-term results in 1,054 patients. J Clin Oncol. 2001;19(3):666–75.
24. Zargar H, Espiritu PN, Fairey AS, Mertens LS, Dinney CP, Mir MC, Krabbe L-M, Cookson MS, Jacobsen N-E, Gandhi N, Griffin J, Montgomery JS, Vasdev N, Yu EY, Youssef D, Xylinas E, Campain NJ, Kassouf W, Dall'Era MA, Seah J-A, Ercole CE, Horenblas S, Sridhar SS, McGrath JS, Aning J, Shariat SF, Wright JL, Thorpe AC, Morgan TM, Holzbeierlein JM, Bivalacqua TJ, North S, Barocas DA, Lotan Y, Garcia JA, Stephenson AJ, Shah JB, van Rhijn BW, Daneshmand S, Spiess PE, Black PC. Multicenter Assessment of Neoadjuvant Chemotherapy for Muscle-invasive Bladder Cancer. European urology. 2015 Feb 23;67(2):241–9.
25. von der Maase H, Sengelov L, Roberts JT, Ricci S, Dogliotti L, Oliver T, Moore MJ, Zimmermann A, Arning M. Long-Term Survival Results of a Randomized Trial Comparing Gemcitabine Plus Cisplatin, With Methotrexate, Vinblastine, Doxorubicin, Plus Cisplatin in

Patients With Bladder Cancer. Journal of Clinical Oncology. 2005 Jul 20;23(21):4602–8.

26. Chan KS, Espinosa I, Chao M, Wong D, Ailles L, Diehn M, Gill H, Presti J, Chang HY, van de Rijn M, Shortliffe L, Weissman IL. Identification, molecular characterization, clinical prognosis, and therapeutic targeting of human bladder tumor-initiating cells. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(33):14016–21.
27. Volkmer J-P, Sahoo D, Chin RK, Ho PL, Tang C, Kurtova A V, Willingham SB, Pazhanisamy SK, Contreras-Trujillo H, Storm TA, Lotan Y, Beck AH, Chung BI, Alizadeh AA, Godoy G, Lerner SP, van de Rijn M, Shortliffe LD, Weissman IL, Chan KS. Three differentiation states risk-stratify bladder cancer into distinct subtypes. Proceedings of the National Academy of Sciences of the United States of America. 2012 Feb 7;109(6):2078–83.
28. Orian-Rousseau V. CD44, a therapeutic target for metastasising tumours. European Journal of Cancer. 2010;46(7):1271–7.
29. Ponta H, Sherman L, Herrlich PA. Cd44 : From adhesion molecules to signalling regulators. 2003;4(January).
30. Stamenkovic I, Aruffo A, Amiot M, Seed B. The hematopoietic and epithelial forms of CD44 are distinct polypeptides with different adhesion potentials for hyaluronate-bearing cells. The EMBO Journal. 1991 Feb;10(2):343–8.
31. Fox SB, Fawcett J, Jackson DG, Collins I, Gatter KC, Harris AL, Gearing A, Simmons DL. Normal Human Tissues, in Addition to Some Tumors, Express Multiple Different CD44 Isoforms. Cancer Research. 1994 Aug 15;54(16):4539 LP-4546.
32. Screaton GR, Bell M V, Jackson DG, Cornelis FB, Gerth U, Bell JI. Genomic structure of DNA encoding the lymphocyte homing receptor CD44 reveals at least 12 alternatively spliced exons. Proceedings of the National Academy of Sciences of the United States of America. 1992 Dec 15;89(24):12160–4.
33. Screaton GR, Bell M V, Bell JI, Jackson DG. The identification of a new alternative exon with highly restricted tissue expression in transcripts encoding the mouse Pgp-1 (CD44) homing receptor. Comparison of all 10 variable exons between mouse, human, and rat. Journal of Biological Chemistry . 1993 Jun 15;268(17):12235–8.
34. Naor D, Sionov R V, Ish-Shalom D. CD44: structure, function, and association with the malignant process. Advances in cancer research. 1997;71(September):241–319.
35. The UniProt Consortium. UniProt: the universal protein knowledgebase. Nucleic Acids Research. 2017 Jan 4;45(Database issue):D158–69.
36. Coordinators NR. Database resources of the National Center for Biotechnology Information. Nucleic Acids Research. 2016 Jan 4;44(Database issue):D7–19.
37. Bennett KL, Modrell B, Greenfield B, Bartolazzi A, Stamenkovic I, Peach R, Jackson DG, Spring F, Aruffo A. Regulation of CD44 binding to hyaluronan by glycosylation of variably spliced exons. The Journal of Cell Biology. 1995 Dec 15;131(6):1623 LP-1633.

38. Dasgupta A, Takahashi K, Cutler M, Tanabe KK. O-linked Glycosylation Modifies CD44 Adhesion to Hyaluronate in Colon Carcinoma Cells. Biochemical and Biophysical Research Communications. 1996;227(1):110–7.
39. Skelton TP, Zeng C, Nocks A, Stamenkovic I. Glycosylation Provides Both Stimulatory and Inhibitory Effects on Cell Surface and Soluble CD44 Binding to Hyaluronan . The Journal of Cell Biology. 1998 Jan 26;140(2):431–46.
40. David G. Jackson,* John I. Bell,* Richard Dickinson, Jackie Timans JS, Whittle and N. Proteoglycan forms of the lymphocyte homing receptor CD44 are alternatively spliced variants containing the v3 exon. The Journal of Cell Biology. 1995 Feb 2;128(4):673–85.
41. Sleeman JP, Rahmsdorf U, Steffen A, Ponta H, Herrlich P. CD44 variant exon v5 encodes a tyrosine that is sulphated. European Journal of Biochemistry. 1998;255(1):74–80.
42. Yang B, Yang BL, Savani RC, Turley EA. Identification of a common hyaluronan binding motif in the hyaluronan binding proteins RHAMM, CD44 and link protein. The EMBO Journal. 1994 Jan 15;13(2):286–96.
43. Liao HX, Lee DM, Levesque MC, Haynes BF. N-terminal and central regions of the human CD44 extracellular domain participate in cell surface hyaluronan binding. The Journal of Immunology. 1995 Oct 15;155(8):3938 LP-3945.
44. Lesley J, English N, Perschl A et al. Variant cell lines selected for alterations in the function of the hyaluronan receptor CD44 show differences in glycosylation. The Journal of Experimental Medicine. 1995 Aug 1;182(2):431–7.
45. Hua Q, Knudson CB, Knudson W. Internalization of hyaluronan by chondrocytes occurs via receptor-mediated endocytosis. Journal of Cell Science. 1993 Sep 1;106(1):365 LP-375.
46. Ishida O, Tanaka Y, Morimoto I, Takigawa M, Eto S. Chondrocytes Are Regulated by Cellular Adhesion Through CD44 and Hyaluronic Acid Pathway. Journal of Bone and Mineral Research. 1997;12(10):1657–63.
47. Jain M, He Q, Lee W Sen, Kashiki S, Foster LC, Tsai JC, Lee ME, Haber E. Role of CD44 in the reaction of vascular smooth muscle cells to arterial wall injury. Journal of Clinical Investigation. 1996;97(3):596–603.
48. DeGrendele HC, Kosfiszer M, Estess P, Siegelman MH. CD44 activation and associated primary adhesion is inducible via T cell receptor stimulation. The Journal of Immunology. 1997 Sep 15;159(6):2549 LP-2553.
49. DeGrendele HC, Estess P, Siegelman MH. Requirement for CD44 in Activated T Cell Extravasation into an Inflammatory Site. Science. 1997 Oct 24;278(5338):672 LP-675.
50. Estess P, DeGrendele HC, Pascual V, Siegelman MH. Functional activation of lymphocyte CD44 in peripheral blood is a marker of autoimmune disease activity. Journal of Clinical Investigation. 1998 Sep 15;102(6):1173–82.
51. Mohamadzadeh M, DeGrendele H, Arizpe H, Estess P, Siegelman M. Proinflammatory stimuli regulate endothelial hyaluronan expression and CD44/HA-dependent primary adhesion. Journal of Clinical Investigation. 1998 Jan 1;101(1):97–108.

52. Weiss JM, Renkl AC, Sleeman J, Dittmar H, Termeer CC, Taxis S, Howells N, Schöpf E, Ponta H, Herrlich P, Simon JC. CD44 variant isoforms are essential for the function of epidermal Langerhans cells and dendritic cells. Cell adhesion and communication. 1998;6(2–3):157–60.
53. Kahles F, Findeisen HM, Bruemmer D. Osteopontin: A novel regulator at the cross roads of inflammation, obesity and diabetes. Molecular Metabolism. 2014 Jul 22;3(4):384–93.
54. Crawford HC, Matrisian LM, Liaw L. Distinct Roles of Osteopontin in Host Defense Activity and Tumor Survival during Squamous Cell Carcinoma Progression in Vivo. Cancer Research. 1998 Nov 15;58(22):5206 LP-5215.
55. Zhu B, Suzuki K, Goldberg HA, Rittling SR, Denhardt DT, McCulloch CAG, Sodek J. Osteopontin modulates CD44-dependent chemotaxis of peritoneal macrophages through G-protein-coupled receptors: Evidence of a role for an intracellular form of osteopontin. Journal of Cellular Physiology. 2004;198(1):155–67.
56. Basakran NS. CD44 as a potential diagnostic tumor marker. Saudi Medical Journal. 2015;36(3):273–9.
57. Farahani E, Patra HK, Jangamreddy JR, Rashedi I, Kawalec M, Rao Pariti RK, Batakis P, Wiechec E. Cell adhesion molecules and their relation to (cancer) cell stemness. Carcinogenesis. 2014 Apr 1;35(4):747–59.
58. Misra S, Hascall VC, Markwald RR, Ghatak S. Interactions between Hyaluronan and Its Receptors (CD44, RHAMM) Regulate the Activities of Inflammation and Cancer. Frontiers in Immunology. 2015 May 6;6:201.
59. Stamenkovic I, Amiot M, Pesando JM, Seed B. A lymphocyte molecule implicated in lymph node homing is a member of the cartilage link protein family. Cell. 1989;56(6):1057–62.
60. Sy M-S, Guo Y-J, Stamenkovic I. Inhibition of tumor growth in vivo with a soluble CD44-immunoglobulin fusion protein . Journal of Experimental Medicine. 1992;176(August):623–7.
61. Zawadzki V, Perschl A, Rösel M, Hekele A, Zöller M. Blockade of metastasis formation by CD44-receptor globulin. International Journal of Cancer. 1998;75(6):919–24.
62. Guo Y, Ma J, Wang J, Che X, Narula J, Bigby M, Wu M, Sy M-S. Inhibition of Human Melanoma Growth and Metastasis in Vivo by Anti-CD44 Monoclonal Antibody. Cancer Research. 1994 Mar 1;54(6):1561 LP-1565.
63. Birch M, Mitchell S, Hart IR. Isolation and Characterization of Human Melanoma Cell Variants Expressing High and Low Levels of CD44. Cancer Research. 1991 Dec 1;51(24):6660 LP-6667.
64. Sy MS, Guo YJ, Stamenkovic I. Distinct effects of two CD44 isoforms on tumor growth in vivo. The Journal of Experimental Medicine. 1991 Oct 1;174(4):859 LP-866.
65. Ayroldi E, Cannarile L, Migliorati G, Bartoli A, Nicoletti I, Riccardi C. CD44 (Pgp-1) inhibits CD3 and dexamethasone-induced apoptosis. Blood. 1995 Oct 1;86(7):2672 LP-2678.
66. Bates RC, Elith CA, Thorne RF, Burns GF. Engagement of Variant CD44 Confers

Resistance to Anti-Integrin Antibody-Mediated Apoptosis in a Colon Carcinoma Cell Line. Cell Communication & Adhesion. 1998 Jan 1;6(1):21–38.

67. Weber GF, Ashkar S. Molecular mechanisms of tumor dissemination in primary and metastatic brain cancers. Brain Research Bulletin. 2000;53(4):421–4.
68. Senger DR, Ledbetter SR, Claffey KP, Papadopoulos-Sergiou A, Peruzzi CA, Detmar M. Stimulation of endothelial cell migration by vascular permeability factor/vascular endothelial growth factor through cooperative mechanisms involving the alphavbeta3 integrin, osteopontin, and thrombin. The American Journal of Pathology. 1996 Jul;149(1):293–305.
69. Trochon V, Mabilat C, Bertrand P, Legrand Y, Smadja-Joffe F, Soria C, Delpech B, Lu H. Evidence of involvement of CD44 in endothelial cell proliferation, migration and angiogenesis in vitro. International Journal of Cancer. 1996;66(5):664–8.
70. Kuncov J, Kostrouch Z, Viale M, Revoltella R, Mandys V. Expression of CD44v6 correlates with cell proliferation and cellular atypia in urothelial carcinoma cell lines 5637 and HT1197. Folia Biologica. 2005;51(1):3–11.
71. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. Proceedings of the National Academy of Sciences. 2003;100(7):3983–8.
72. Singh SK, Clarke ID, Terasaki M, Bonn VE, Hawkins C, Squire J, Dirks PB. Identification of a Cancer Stem Cell in Human Brain Tumors. Cancer Research. 2003 Sep 15;63(18):5821 LP-5828.
73. Yang YM, Chang JW. Bladder cancer initiating cells (BCICs) are among EMA-CD44v6+ subset: Novel methods for isolating undetermined cancer stem (initiating) cells. Cancer Investigation. 2008;26(7):725–33.
74. Logan CM, Giordano A, Puca A, Cassone M. Prostaglandin E2: at the crossroads between stem cell development, inflammation and cancer. Cancer Biol Ther. 2007;6(10):1517–20.
75. North TE, Goessling W, Walkley CR, Lengerke C, Kopani KR, Lord AM, Weber GJ, Bowman T V, Jang I-H, Grosser T, FitzGerald GA, Daley GQ, Orkin SH, Zon LI. Prostaglandin E2 regulates vertebrate haematopoietic stem cell homeostasis. Nature. 2007 Jun 21;447:1007.
76. Kitayama W, Denda A, Okajima E, Tsujiuchi T, Konishi Y. Increased expression of cyclooxygenase-2 protein in rat urinary bladder tumors induced by N-butyl-N-(4-hydroxybutyl) nitrosamine. Carcinogenesis. 1999;20(12):2305–10.
77. Okajima E, Denda A, Ozono S, Takahama M, Akai H, Sasaki Y, Kitayama W, Wakabayashi K, Konishi Y. Chemopreventive effects of nimesulide, a selective cyclooxygenase-2 inhibitor, on the development of rat urinary bladder carcinomas initiated by N-butyl-N-(4-hydroxybutyl)nitrosamine. Cancer Res. 1998;58(14):3028–31.
78. Thanan R, Murata M, Ma N, Hammam O, Wishahi M, El Leithy T, Hiraku Y, Oikawa S, Kawanishi S. Nuclear localization of COX-2 in relation to the expression of stemness markers in urinary bladder cancer. Mediators of Inflammation. 2012;2012.

79. Kobayashi K, Matsumoto H, Matsuyama H, Fujii N, Inoue R, Yamamoto Y, Nagao K. Clinical significance of CD44 variant 9 expression as a prognostic indicator in bladder cancer. Oncology Reports. 2016;36(5):2852–60.
80. Omran OM, Ata HS. CD44s and CD44v6 in Diagnosis and Prognosis of Human Bladder Cancer. Ultrastructural Pathology. 2012 May 1;36(3):145–52.
81. Lipponen P, Aaltoma S, Kosma VM, Ala-Opas M, Eskelinen M. Expression of CD44 standard and variant-v6 proteins in transitional cell bladder tumours and their relation to prognosis during a long-term follow-up. Journal of Pathology. 1998;186(2):157–64.
82. Klatte T, Seligson DB, Rao JY, Yu H, de Martino M, Garraway I, Wong SG, Belldegrun AS, Pantuck AJ. Absent CD44v6 Expression is an Independent Predictor of Poor Urothelial Bladder Cancer Outcome. The Journal of Urology. 2010;183(6):2403–8.
83. Sugino T, Gorham H, Yoshida K, Bolodeoku J, Nargund V, Cranston D, Goodison S, Tarin D. Progressive loss of CD44 gene expression in invasive bladder cancer. The American Journal of Pathology. 1996 Sep;149(3):873–82.
84. Haynes PA. Phosphoglycosylation: A new structural class of glycosylation? Vol. 8, Glycobiology. 1998. p. 1–5.
85. Trombetta ES. The contribution of N-glycans and their processing in the endoplasmic reticulum to glycoprotein biosynthesis. Vol. 13, Glycobiology. 2003.
86. Bektas M, Rubenstein DS. The role of intracellular protein O-glycosylation in cell adhesion and disease. Vol. 25, Journal of Biomedical Research. 2011. p. 227–36.
87. Doucey M a, Hess D, Cacan R, Hofsteenge J. Protein C-mannosylation is enzyme-catalysed and uses dolichyl-phosphate-mannose as a precursor. Molecular biology of the cell. 1998;9(2):291–300.
88. Pittet M, Conzelmann A. Biosynthesis and function of GPI proteins in the yeast Saccharomyces cerevisiae. Vol. 1771, Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids. 2007. p. 405–20.
89. Rabinovich G a, Toscano M a. Turning “sweet” on immunity: galectin-glycan interactions in immune tolerance and inflammation. Nature reviews Immunology. 2009;9(5):338–52.
90. Häuselmann I, Borsig L. Altered tumor-cell glycosylation promotes metastasis. Frontiers in oncology. 2014;4(February):28.
91. Helenius A, Aebi M. Intracellular Functions of N-Linked Glycans. Science. 2001;291(5512):2364–9.
92. Helenius A. How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. Molecular Biology of the Cell. 1994 Mar;5(3):253–65.
93. Noro E, Togayachi A, Sato T, Tomioka A, Fujita M, Sukegawa M, Suzuki N, Kaji H, Narimatsu H. Large-Scale Identification of N-Glycan Glycoproteins Carrying Lewis x and Site-Specific N-Glycan Alterations in Fut9 Knockout Mice. Journal of Proteome Research. 2015 Sep 4;14(9):3823–34.
94. Muthana SM, Campbell CT, Gildersleeve JC. Modifications of Glycans: Biological

Significance and Therapeutic Opportunities. ACS Chemical Biology. 2012 Jan 20;7(1):31–43.

95. Ju T, Brewer K, Souza AD, Cummings RD, Canfield WM. Cloning and expression of human core 1 beta 1,3-galactosyltransferase. Journal of Biological Chemistry. 2002;277(1):178–86.
96. Irazoqui FJ, Sendra VG, Lardone RD, Nores GA. Immune response to Thomsen–Friedenreich disaccharide and glycan engineering. Immunology And Cell Biology. 2005 Aug 1;83:405.
97. Marcos NT, Pinho S, Grandela C, Cruz A, Samyn-Petit B, Harduin-Lepers A, Almeida R, Silva F, Morais V, Costa J, Kihlberg J, Clausen H, Reis CA. Role of the human ST6GalNAc-I and ST6GalNAc-II in the synthesis of the cancer-associated Sialyl-Tn antigen. Cancer Research. 2004;64(19):7050–7.
98. Varki A, Kannagi R, Toole BP. Chapter 44 Glycosylation Changes in Cancer. In: Varki, A. et al., editor. Essentials of Glycobiology. Cold Spring Harbor (NY); 2009. p. 1–11.
99. Kumamoto K, Goto Y, Sekikawa K, Takenoshita S, Ishida N, Kawakita M, Kannagi R. Increased Expression of UDP-Galactose Transporter Messenger RNA in Human Colon Cancer Tissues and Its Implication in Synthesis of Thomsen-Friedenreich Antigen and Sialyl Lewis A/X Determinants. Cancer Research. 2001 Jun 1;61(11):4620 LP-4627.
100. Dennis JW, Laferte S, Waghorne C, Breitman ML, Kerbel RS. Beta 1-6 branching of Asn-linked oligosaccharides is directly associated with metastasis. Science. 1987 May 1;236(4801):582 LP-585.
101. Zhao Y, Sato Y, Isaji T, Fukuda T, Matsumoto A, Miyoshi E, Gu J, Taniguchi N. Branched N-glycans regulate the biological functions of integrins and cadherins. FEBS Journal. 2008 May 1;275(9):1939–48.
102. Voss M, Künzel U, Higel F, Kuhn P-H, Colombo A, Fukumori A, Haug-Kröper M, Klier B, Grammer G, Seidl A, Schröder B, Obst R, Steiner H, Lichtenthaler SF, Haass C, Fluhrer R. Shedding of glycan-modifying enzymes by signal peptide peptidase-like 3 (SPPL3) regulates cellular N-glycosylation. The EMBO Journal. 2014 Dec 17;33(24):2890–905.
103. Nakahara S, Saito T, Kondo N, Moriwaki K, Noda K, Ihara S, Takahashi M, Ide Y, Gu J, Inohara H, Katayama T, Tohyama M, Kubo T, Taniguchi N, Miyoshi E. A secreted type of β1,6 N-acetylglucosaminyltransferase V (GnT-V), a novel angiogenesis inducer, is regulated by γ-secretase. The FASEB Journal. 2006 Dec 1;20(14):2451–9.
104. Hakomori S. Tumor Malignancy Defined by Aberrant Glycosylation and Sphingo(glyco)lipid Metabolism. Cancer Research. 1996 Dec 1;56(23):5309 LP-5318.
105. Chihara Y, Sugano K, Kobayashi A, Kanai Y, Yamamoto H, Nakazono M, Fujimoto H, Kakizoe T, Fujimoto K, Hirohashi S, Hirao Y. Loss of blood group A antigen expression in bladder cancer caused by allelic loss and//or methylation of the ABO gene. Lab Invest. 2005 May 9;85(7):895–907.
106. Sheinfeld J, Reuter VE, Fair WR, Cordon-Cardo C. Expression of blood group antigens in bladder cancer: current concepts. Semin Surg Oncol. 1992;8(5):308–15.

107. Gruszewska E, Chrostek L. The alterations of glycosylation in malignant diseases. Pol Merkur Lekarski. 2013;34(199):58–61.
108. Numahata K, Satoh M, Handa K, Saito S, Ohyama C, Ito A, Takahashi T, Hoshi S, Orikasa S, Hakomori SI. Sialosyl-Lex expression defines invasive and metastatic properties of bladder carcinoma. Cancer. 2002;94(3):673–85.
109. Sheinfeld J, Reuter VE, Melamed MR, Fair WR, Morse M, Sogani PC, Herr HW, Whitmore WF, Cordon-Cardo C. Enhanced Bladder Cancer Detection with the Lewis X Antigen as a Marker of Neoplastic Transformation. The Journal of Urology. 2017 Dec 14;143(2):285–8.
110. Ferreira JA, Magalhães A, Gomes J, Peixoto A, Gaiteiro C, Fernandes E, Santos LL, Reis CA. Protein glycosylation in gastric and colorectal cancers: Toward cancer detection and targeted therapeutics. Cancer Letters. 2017 Dec 14;387:32–45.
111. Thorpe SJ, Abel P, Slavin G, Feizi T. Blood group antigens in the normal and neoplastic bladder epithelium. Journal of Clinical Pathology. 1983 Aug;36(8):873–82.
112. Ørntoft TF, Wolf H, Watkins WM. Activity of the Human Blood Group ABO, Se, H, Le, and X Gene-encoded Glycosyltransferases in Normal and Malignant Bladder Urothelium. Cancer Research. 1988 Aug 1;48(15):4427 LP-4433.
113. Santos SN, Junqueira MS, Francisco G, Vilanova M, Magalhães A, Baruffi MD, Chammas R, Harris AL, Reis CA, Bernardes ES. O-glycan sialylation alters galectin-3 subcellular localization and decreases chemotherapy sensitivity in gastric cancer. Oncotarget. 2016 Dec 13;7(50):83570–87.
114. Tran DT, Ten Hagen KG. Mucin-type O-Glycosylation during Development. The Journal of Biological Chemistry. 2013 Mar 8;288(10):6921–9.
115. Susuki Y, Sutoh M, Hatakeyama S, Mori K, Yamamoto H, Koie T, Saitoh H, Yamaya K, Funyu T, Habuchi T, Arai Y, Fukuda M, Ohyama C, Tsuboi S. MUC1 carrying core 2 O-glycans functions as a molecular shield against NK cell attack, promoting bladder tumor metastasis. International Journal of Oncology. 2012 Jun 23;40(6):1831–8.
116. Gomes C, Osório H, Pinto MT, Campos D, Oliveira MJ, Reis CA. Expression of ST3GAL4 Leads to SLe(x) Expression and Induces c-Met Activation and an Invasive Phenotype in Gastric Carcinoma Cells. Williams BO, editor. PLoS ONE. 2013 Jun 14;8(6):e66737.
117. Krüger S, Weitsch G, Büttner H, Matthiensen A, Böhmer T, Marquardt T, Sayk F, Feller AC, Böhle A. HER2 overexpression in muscle-invasive urothelial carcinoma of the bladder: Prognostic implications. International Journal of Cancer. 2002;102(5):514–8.
118. Yan M, Schwaederle M, Arguello D, Millis SZ, Gatalica Z, Kurzrock R. HER2 expression status in diverse cancers: review of results from 37,992 patients. Cancer Metastasis Reviews. 2015 Feb 25;34:157–64.
119. Zhao J, Xu W, Zhang Z, Song R, Zeng S, Sun Y, Xu C. Prognostic role of HER2 expression in bladder cancer: a systematic review and meta-analysis. International Urology and Nephrology. 2015;47(1):87–94.
120. Balzar M, J Winter M, de Boer C, Litvinov S. The biology of the 17-1a antigen (Ep-CAM).

Vol. 77, Journal of molecular medicine (Berlin, Germany). 1999. 699-712 p.

121. Munz M, Baeuerle PA, Gires O. The Emerging Role of EpCAM in Cancer and Stem Cell Signaling. Cancer Research. 2009 Jul 15;69(14):5627 LP-5629.
122. Dollé L, Theise ND, Schmelzer E, Boulter L, Gires O, van Grunsven LA. EpCAM and the biology of hepatic stem/progenitor cells. American Journal of Physiology - Gastrointestinal and Liver Physiology. 2015 Feb 15;308(4):G233–50.
123. Bryan RT, Shimwell NJ, Wei W, Devall AJ, Pirrie SJ, James ND, Zeegers MP, Cheng KK, Martin A, Ward DG. Urinary EpCAM in urothelial bladder cancer patients: characterisation and evaluation of biomarker potential. British Journal of Cancer. 2014 Feb 4;110(3):679–85.
124. Brunner A, Prelog M, Verdorfer I, Tzankov A, Mikuz G, Ensinger C. EpCAM is predominantly expressed in high grade and advanced stage urothelial carcinoma of the bladder. Journal of Clinical Pathology. 2008 Mar 1;61(3):307 LP-310.
125. Sanli O, Dobruch J, Knowles MA, Burger M, Alemozaffar M, Nielsen ME, Lotan Y. Bladder cancer. Nature Reviews Disease Primers. 2017 Apr 13;3:17022.
126. Ferreira JA, Videira PA, Lima L, Pereira S, Silva M, Carrascal M, Severino PF, Fernandes E, Almeida A, Costa C, Vitorino R, Amaro T, Oliveira MJ, Reis CA, Dall'Olio F, Amado F, Santos LL. Overexpression of tumour-ssociated carbohydrate antigen sialyl-Tn in advanced bladder tumours. Molecular Oncology. 2013 Jun 21;7(3):719–31.
127. Chang JT, Mani SA. Sheep, Wolf, or Werewolf: Cancer Stem Cells and the Epithelial-to-Mesenchymal Transition. Cancer letters. 2013 Nov 28;341(1):10.1016/j.canlet.2013.03.004.
128. Grossman HB, Lee C, Bromberg J, Liebert M. Expression of the alpha6beta4 integrin provides prognostic information in bladder cancer. Oncol Rep. 2000;7(1):13–6.
129. Lima L, Neves M, Oliveira MI, Dieguez L, Freitas R, Azevedo R, Gaiteiro C, Soares J, Ferreira D, Peixoto A, Fernandes E, Montezuma D, Tavares A, Ribeiro R, Castro A, Oliveira M, Fraga A, Reis CA, Santos LL, Ferreira JA. Sialyl-Tn identifies muscle-invasive bladder cancer basal and luminal subtypes facing decreased survival, being expressed by circulating tumor cells and metastases. Urologic Oncology: Seminars and Original Investigations. 2017 Dec 13;35(12):675.e1-675.e8.
130. Juhl BR, Hartzen SH, Hainau BO. Lewis a antigen in transitional cell tumors of the urinary bladder. Cancer. 1986;58(2):222–8.
131. van der Horst G, van den Hoogen C, Buijs JT, Cheung H, Bloys H, Pelger RCM, Lorenzon G, Heckmann B, Feyen J, Pujuguet P, Blanque R, Clément-Lacroix P, van der Pluijm G. Targeting of α(v)-Integrins in Stem/Progenitor Cells and Supportive Microenvironment Impairs Bone Metastasis in Human Prostate Cancer. Neoplasia (New York, NY). 2011 Jun 7;13(6):516–25.
132. Chen Q, Manning CD, Millar H, McCabe FL, Ferrante C, Sharp C, Shahied-Arruda L, Doshi P, Nakada MT, Anderson GM. CNTO 95, a fully human anti αv integrin antibody, inhibits cell signaling, migration, invasion, and spontaneous metastasis of human breast cancer cells. Clinical & Experimental Metastasis. 2008;25(2):139–48.

133. Yoo NJ, Kim MS, Lee SH. Absence of COSMC gene mutations in breast and colorectal carcinomas. APMIS. 2008;116(2):154–5.
134. Stojnev S, Ristic-Petrovic A, Velickovic LJ, Krstic M, Bogdanovic D, Khanh DT, Ristic A, Conic I, Stefanovic V. Prognostic significance of mucin expression in urothelial bladder cancer. International Journal of Clinical and Experimental Pathology. 2014 Jul 15;7(8):4945–58.
135. Simms, Hughes, Limb, Price, Bishop. MUC1 mucin as a tumour marker in bladder cancer. BJU International. 1999;84(3):350–2.
136. Kaur S, Momi N, Chakraborty S, Wagner DG, Horn AJ, Lele SM, Theodorescu D, Batra SK. Altered Expression of Transmembrane Mucins, MUC1 and MUC4, in Bladder Cancer: Pathological Implications in Diagnosis. PLOS ONE. 2014 Mar 26;9(3):e92742.
137. Trikha M, Zhou Z, Nemeth JA, Chen Q, Sharp C, Emmell E, Giles-Komar J, Nakada MT. CNTO 95, a fully human monoclonal antibody that inhibits αv integrins, has antitumor and antiangiogenic activity in vivo. International Journal of Cancer. 2004;110(3):326–35.
138. Sachs MD, Rauen KA, Ramamurthy M, Dodson JL, De Marzo AM, Putzi MJ, Schoenberg MP, Rodriguez R. Integrin α$_v$ and coxsackie adenovirus receptor expression in clinical bladder cancer. Urology. 2017 Dec 13;60(3):531–6.
139. Rodgers AK, Nair A, Binkley PA, Tekmal R, Schenken RS. Inhibition of CD44 N- and O-linked Glycosylation Decreases Endometrial Cell Lines Attachment to Peritoneal Mesothelial Cells. Fertility and sterility. 2011 Feb;95(2):823–5.
140. Bartolazzi A, Nocks A, Aruffo A, Spring F, Stamenkovic I. Glycosylation of CD44 is implicated in CD44-mediated cell adhesion to hyaluronan. Journal of Cell Biology. 1996;132(6):1199–208.
141. Guvench O. Revealing the mechanisms of protein disorder and N-glycosylation in CD44-hyaluronan binding using molecular simulation. Vol. 6, Frontiers in Immunology. 2015.
142. Hu Z, Gao J, Zhang D, Liu Q, Yan L, Gao L, Liu J, Liu D, Zhang S, Lin B. High Expression of Lewis y Antigen and CD44 Is Correlated with Resistance to Chemotherapy in Epithelial Ovarian Cancers. Vol. 8, PLoS ONE. 2013.
143. Campos D, Freitas D, Gomes J, Magalhães A, Steentoft C, Gomes C, Vester-Christensen MB, Ferreira JA, Afonso LP, Santos LL, Pinto de Sousa J, Mandel U, Clausen H, Vakhrushev SY, Reis CA. Probing the O-Glycoproteome of Gastric Cancer Cell Lines for Biomarker Discovery. Molecular & Cellular Proteomics : MCP. 2015 Jun 26;14(6):1616–29.
144. Singh R, Campbell BJ, Yu LG, Fernig DG, Milton JD, Goodlad RA, FitzGerald AJ, and Rhodes JM. Cell surface-expressed Thomsen-Friedenreich antigen in colon cancer is predominantly carried on high molecular weight splice variants of CD44. Glycobiology. 2001;11(7):587–92.
145. Carrascal MA, Severino PF, Guadalupe Cabral M, Silva M, Ferreira JA, Calais F, Quinto H, Pen C, Ligeiro D, Santos LL, Dall'Olio F, Videira PA. Sialyl Tn-expressing bladder cancer cells induce a tolerogenic phenotype in innate and adaptive immune cells. Molecular

Oncology. 2014;8(3):753–65.

146. Cotton S, Azevedo R, Gaiteiro C, Ferreira D, Lima L, Peixoto A, Fernandes E, Neves M, Neves D, Amaro T, Cruz R, Tavares A, Rangel M, Silva AMN, Santos LL, Ferreira JA. Targeted O-glycoproteomics explored increased sialylation and identified MUC16 as a poor prognosis biomarker in advanced-stage bladder tumours. Molecular Oncology. 2017 Aug 2;11(8):895–912.
147. Pinto-Leite R, Carreira I, Melo J, Ferreira SI, Ribeiro I, Ferreira J, Filipe M, Bernardo C, Arantes-Rodrigues R, Oliveira P, Santos L. Genomic characterization of three urinary bladder cancer cell lines: understanding genomic types of urinary bladder cancer. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine. 2014;35(5):4599–617.
148. Arantes-Rodrigues R, Pinto-Leite R, Fidalgo-Gonçalves L, Palmeira C, Santos L, Colaço A, Oliveira P. Synergistic Effect between Cisplatin and Sunitinib Malate on Human Urinary Bladder-Cancer Cell Lines. BioMed Research International. 2013 Nov 28;2013:791406.
149. Kudelka MR, Antonopoulos A, Wang Y, Duong DM, Song X, Seyfried NT, Dell A, Haslam SM, Cummings RD, Ju T. Cellular O-Glycome Reporter/Amplification to explore O-glycans of living cells. Nature Methods. 2015;
150. Jang-Lee J, North SJ, Sutton-Smith M, Goldberg D, Panico M, Morris H, Haslam S, Dell A. Glycomic Profiling of Cells and Tissues by Mass Spectrometry: Fingerprinting and Sequencing Methodologies. Vol. 415, Methods in Enzymology. 2006. p. 59–86.
151. Azevedo R, Silva AMN, Reis CA, Santos LL, Ferreira JA. In silico approaches for unveiling novel glycobiomarkers in cancer. Journal of Proteomics. 2017;
152. Ricardo S, Marcos-Silva L, Pereira D, Pinto R, Almeida R, Söderberg O, Mandel U, Clausen H, Felix A, Lunet N, David L. Detection of glyco-mucin profiles improves specificity of MUC16 and MUC1 biomarkers in ovarian serous tumours. Molecular Oncology. 2015;9(2):503–12.
153. Pfaffl MW, Horgan GW, Dempfle L. Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Research. 2002 May 1;30(9):e36–e36.
154. Kolli V, Schumacher KN, Dodds ED. Engaging challenges in glycoproteomics: recent advances in MS-based glycopeptide analysis. Bioanalysis. 2015 Jan 1;7(1):113–31.
155. Lima L, Severino PF, Silva M, Miranda A, Tavares A, Pereira S, Fernandes E, Cruz R, Amaro T, Reis CA, Dall'Olio F, Amado F, Videira PA, Santos L, Ferreira JA. Response of high-risk of recurrence/progression bladder tumours expressing sialyl-Tn and sialyl-6-T to BCG immunotherapy. British Journal of Cancer. 2013 Oct 15;109(8):2106–14.
156. Costa C, Pereira S, Lima L, Peixoto A, Fernandes E, Neves D, Neves M, Gaiteiro C, Tavares A, Gil da Costa RM, Cruz R, Amaro T, Oliveira PA, Ferreira JA, Santos LL. Abnormal Protein Glycosylation and Activated PI3K/Akt/mTOR Pathway: Role in Bladder Cancer Prognosis and Targeted Therapeutics. Real FX, editor. PLoS ONE. 2015 Nov

16;10(11):e0141253.

157. Peixoto A, Fernandes E, Gaiteiro C, Lima L, Azevedo R, Soares J, Cotton S, Parreira B, Neves M, Amaro T, Tavares A, Teixeira F, Palmeira C, Rangel M, Silva AMN, Reis CA, Santos LL, Oliveira MJ, Ferreira JA, Peixoto A, Fernandes E, Gaiteiro C, Lima L, Azevedo R, Soares J, Cotton S, Parreira B, Neves M, Amaro T, Tavares A, Teixeira F, Palmeira C, Rangel M, Silva AMN, Reis CA, Santos LL, Oliveira MJ, Alexandre Ferreira J. Hypoxia enhances the malignant nature of bladder cancer cells and concomitantly antagonizes protein O -glycosylation extension. Oncotarget. 2016;5(0).
158. Bànkfalvi À, Terpe H-J, Breukelmann D, Bier B, Rempe D, Pschadka G, Krech R, Böcker W. Gains and losses of CD44 expression during breast carcinogenesis and tumour progression. Histopathology. 1998;33(2):107–16.
159. Hoofd C, Wang X, Lam S, Jenkins C, Wood B, Giambra V, Weng AP. CD44 promotes chemoresistance in T-ALL by increased drug efflux. Experimental Hematology. 2017 Dec 17;44(3):166–171.e17.
160. Southgate J, Trejdosiewicz LK, Smith B, Selby PJ. Patterns of splice variant CD44 expression by normal human urothelium in situ and in vitro and by bladder-carcinoma cell lines. International Journal of Cancer. 1995;62(4):449–56.
161. Stavropoulos NE, Filliadis I, Ioachim E, Michael M, Mermiga E, Hastazeris K, Nseyo UO. CD44 standard form expression as a predictor of progression in high risk superficial bladder tumors. International Urology and Nephrology. 2001;33(3):479–83.
162. Mima K, Okabe H, Ishimoto T, Hayashi H, Nakagawa S, Kuroki H, Watanabe M, Beppu T, Tamada M, Nagano O, Saya H, Baba H. CD44s regulates the TGF-b-mediated mesenchymal phenotype and is associated with poor prognosis in patients with hepatocellular carcinoma. Cancer Research. 2012;72(13):3414–23.
163. Xu Y, Gao XD, Lee J-H, Huang H, Tan H, Ahn J, Reinke LM, Peter ME, Feng Y, Gius D, Siziopikou KP, Peng J, Xiao X, Cheng C. Cell type-restricted activity of hnRNPM promotes breast cancer metastasis via regulating alternative splicing. Genes & Development. 2014 Jun 1;28(11):1191–203.
164. Bendardaf R, Elzagheid A, Lamlum H, Ristamäki R, Collan Y, Pyrhönen S. E-cadherin, CD44s and CD44v6 correlate with tumour differentiation in colorectal cancer. Oncology Reports. 2005;13(5):831–5.
165. Neves D. Glycoproteomic Characterization of Bladder cancer chemoresistant cells. Faculdade de Ciências da Universidade do Porto; 2015.
166. Brown RL, Reinke LM, Damerow MS, Perez D, Chodosh LA, Yang J, Cheng C. CD44 splice isoform switching in human and mouse epithelium is essential for epithelial-mesenchymal transition and breast cancer progression. The Journal of Clinical Investigation. 2011 Mar 1;121(3):1064–74.

Appendix

Glycan affinity magnetic nanoplatforms for urinary glycobiomarkersdiscovery in bladder cancer (submitted to Talanta)

Appendix 1

Glycan affinity magnetic nanoplatforms for urinary glycobiomarkers discovery in bladder cancer

Rita Azevedo[1,2*], Janine Soares[1*], Cristiana Gaiteiro[1,2], Andreia Peixoto[1,2,3], Luís Lima[1,3,4], Dylan Ferreira[1], Marta Relvas-Santos[1], Elisabete Fernandes[1,2,3], Ana Tavares[1], Sofia Cotton[1], Daniel-da-Silva AL[6], Lúcio Lara Santos[1,2], Rui Vitorino[7,8], Francisco Amado[9] and José Alexandre Ferreira[1,2,3,4,10]

[1]Experimental Pathology and Therapeutics Group, Research Center of Portuguese Institute of Oncology (CI-IPOP), Porto, Portugal; [2]Institute of Biomedical Sciences Abel Salazar, University of Porto, Porto, Portugal; [3]Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Portugal; [4]Glycobiology in Cancer, Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP), Porto, Portugal; [5]Department of Urology, Portuguese Institute of Oncology, Porto, Portugal; [6]CICECO - Aveiro Institute of Materials, Department of Chemistry, University of Aveiro, Aveiro, Portugal; [7]iBiMED - Institute of Biomedicine, Department of Medical Sciences, University of Aveiro, Aveiro, Portugal; [8]Research Unit of the Physiology and Cardiothoracic Surgery Department of the Faculty of Medicine, University of Porto, Porto, Portugal; [9]Mass Spectrometry Centre, Organic Chemistry and Natural Products Unit, Department of Chemistry, University of Aveiro, Aveiro, Portugal; [10]International Iberian Nanotechnology Laboratory, Braga, Portugal.

Corresponding authors: José Alexandre Ferreira (jose.a.ferreira@ipoporto.min-saude.pt)
*Equal contribution

Keywords:
Glycoproteomics, urine biomarkers, bladder cancer, glycosylation, nanoprobes

Abstract

Bladder Cancer (BC) presents one of the highest recurrence rates amongst solid tumours and constitutes the second deadliest disease of the genitourinary track. Non-invasive identification of patients facing disease recurrence and/or progression remains one of the most critical and challenging aspects in disease management. To contribute to this goal, we demonstrate the potential of glycan-affinity glycoproteomics nanoplatforms for urinary biomarkers discovery in bladder cancer. Briefly, magnetic nanoprobes (MNP) coated with three broad-spectrum lectins, namely *Concanavalin A* (ConA; MNP@ConA), Wheat Germ Agglutinin (WGA; MNP@WGA), and *Sambucus nigra* (SNA; MNP@SNA), were used to selectively capture glycoproteins from the urine of low-grade and high-grade non-muscle invasive as well as muscle-invasive BC patients. Proteins were identified by nano-LC MALDI-TOF/TOF and data was curated using bioinformatics tools (UniProt, NetOGlyc, NetNGlyc, Cytoscape and Oncomine) to highlight clinically relevant species. Accordingly, 63 glycoproteins were exclusively identified in cancer samples compared with healthy controls matching in age and gender. Specific glycoprotein sets exclusively found in low-grade non-muscle invasive bladder tumours may aid early diagnosis, while those only found in high-grade non-invasive and muscle-invasive tumours hold potential for accessing progression. Amongst these proteins is bladder cancer stem-cell marker CD44, which has been associated with poor prognosis in bladder tumours. Orthogonal validation studies by slot-blotting demonstrated an elevation in urine CD44 levels of high-grade patients, which became more pronounced upon muscle-invasion, in mimicry of the primary tumour. These observations demonstrate the potential of MNP@lectins for identification of clinically relevant glycoproteomics signatures in bladder cancer. Future clinical validation in a larger and well characterized patient subset is required envisaging clinical translation of the results.

Introduction

Bladder cancer (BC) is the second most deadly malignancy of the urinary tract [1]. Patients diagnosed with low-grade non-muscle invasive BC (NMIBC) generally face better prognosis, whereas those with high-grade lesions are frequently burdened by significant recurrences, generally accompanied by progression to muscle invasion (MIBC) and metastasis [2]. As such, these patients are subjected to several post-operative cycles of intravesical therapy proceeded by intensive and invasive follow-up interventions [2]. The development of non-invasive tools supporting the differentiation between low and high-grade lesions, as well as the early identification of patients at risk of progression, remains a critical aspect for disease management [3].

Urine is easily accessible and, in bladder cancer cases, is in direct contact with the tumour, thereby providing a key source of biomarkers for addressing the above mentioned difficulties [4]. Moreover, the human urine proteome is well characterized, supporting proteomics-based biomarker studies [5]. Interestingly, the most relevant targets arising from these studies include several glycoproteins, such as members of the apolipoprotein family, fibrinogen chains, alpha-1-antitrypsin, alpha-2-macroglobulin, and uromodulin [5, 6]. Nevertheless, these biomarkers may be reflective of haematuria and have also been associated with kidney disease (e.g. apolipoproteins, uromodulin, or alpha-1-antitrypsin) and other pathologies [6]. As such, current urine biomarkers lack the necessary sensitivity and specificity for clinical use; however, their incorporation in multi-biomarker panels may provide the necessary molecular context for more effective applications. However, the lack of efficient enrichment strategies remains one of the most critical aspects underlying the identification of cancer-specific biomarkers, generally present in minute amounts (nano-femtomolar range) in bodily fluids such as urine [7].

Targeting glycan moieties in cancer-associated glycoproteins holds tremendous potential to overcome these limitations, while providing an important source of clinically relevant biomarkers [8]. Moreover, the identification of disease-associated glycoforms may further improve the biomarker potential of already proposed urine glycobiomarkers. Accordingly, several decades of research, mostly using lectins and antibodies for tissue screening, have disclosed a plethora of alterations in membrane proteins glycosylation associated with bladder cancer [3]. Namely, we have reported that disease progression to invasion and

dissemination is accompanied by the overexpression of short-chain sialylated *O*-glycans, associated with poor prognosis [9-11]. Lectin-enrichment targeting these glycans further enabled the identification of clinically relevant glycoproteins in patient samples, including several integrins, cadherins, the stem cell marker CD44 and MUC16 [12, 13]. Taken together, our studies demonstrated the importance of pre-enrichment strategies for downstream comprehensive glycoproteomics approaches envisaging biomarker discovery [12-14]. Concomitantly to *O*-glycan alterations, advanced stage bladder cancer cells also experience *N*-glycome remodelling, including changes in mannose core, branching, elongation and degree of fucosylation and sialylation [3]. However, the nature of the glycoproteins exhibiting these alterations is still poorly understood. Nevertheless, we hypothesize that abnormally glycosylated proteins may be secreted or shed from the primary tumour into bodily fluids, constituting an important source of potentially relevant urine glycobiomarkers.

Building on these insights, our group has developed lectin functionalized magnetic nanoprobes for glycoproteomics research, which have demonstrated enhanced sensitivity in comparison to conventional enrichment strategies [15, 16]. Herein, we aim to explore the potential of these nanoplatforms to highlight the existence of glycoprotein signatures associated with bladder cancer stage and grade. We envisage that this preliminary approach may provide the necessary analytical rationale for more in depth translational studies.

Material and Methods

Patient Samples

Between 2010 and 2011, thirty-one first void urine samples without visible signs of haematuria were prospectively collected before surgery of bladder cancer male patients with mean age of 70 (range, 45–89) years, attending the Portuguese Institute of Oncology of Porto (IPO-Porto; Portugal). Only patients that had not been previously submitted to neoadjuvant therapy were included. Corresponding formalin-fixed paraffin embedded (FFPE) tumours were also obtained for this study. Based on the World Health Organization urothelial carcinoma grading and staging criteria, three different groups were considered in this study, namely low-grade (n=15) and high grade (n=9) non-muscle-invasive bladder tumours (NMIBC) and muscle-invasive (n=7) tumours (MIBC). An additional 15 urines

were obtained from healthy control male volunteers, mean age of 68 (range 41-82) years. All procedures were performed under the approval of the institution ethics committee and upon patients' informed consent. All clinicopathological information was obtained from patients' clinical records.

Isolation of Urine Proteins

Five to forty millilitres of collected urine were centrifuged at 5000*g* for 40 min at 4 °C to remove cells and debris, desalted on Amicon Ultra 10 kDa centrifugal filters (Merck KGaA, Darmstadt, Germany) and proteins were resuspended on 50 mM Ammonium bicarbonate (pH 7.8, Sigma-Aldrich, St. Louis, MO, EUA). Protein quantification was accessed using the DC Protein assay (Bio-Rad, Hercules, CA, USA) and proteins content was normalized in relation to creatinine.

Total Sialic acids

Total sialic acids (TSA) were determined by fluorimetry as previously described [15, 16]. All solutions were precooled in an ice bath. Twenty microliters of sodium periodate solution (10 mM, Sigma-Aldrich) were added to 30 μl of glycoconjugate sample (10–200 μg) placed in a 2 mL polypropylene test tube. The solution was chilled in the ice bath for 45 min. The reaction was stopped by adding 100 μl of 50 mM sodium thiosulfate (Sigma-Aldrich), 500 μl of 4.0 M ammonium acetate (pH 7.5, Sigma-Aldrich) and 400 μl of an ethanolic solution of 100 mM acetoacetanilide (Sigma-Aldrich), followed by incubation for 10 min at room temperature. The fluorescence intensities of the solution were measured at 471 nm with an excitation wavelength of 388 nm. Detection limits of the determination were obtained according to the relative fluorescence of sample solution to a sample blank prepared under identical conditions. TSA content was normalized in relation to the creatinine content in the samples.

Synthesis of lectin functionalized magnetic nanoparticles (MNP@ConA; MNP@WGA; MNP@SNA)

Lectin functionalized ferromagnetic nanoparticles were synthetized as previously described by Ferreira *et al.* (2011) [15] and Cova M *et al.* (2015) [16]. Briefly, the iron oxide magnetic core of the MNPs was synthesized by coprecipitation of $FeCl_2$ and $FeCl_3$ (Sigma-

Aldrich) under alkaline conditions and coated with amorphous silica to prevent particle clustering and ensure their chemical stability. Derivatization was achieved by reaction with tetraethyl orthosilicate (TEOS, Sigma-Aldrich) under alkaline conditions with trimethylamine (Sigma-Aldrich) as catalyst. Subsequently, the silica-coated nanoparticle surface was functionalized with amine groups with 3-aminopropyltrimethoxysilane (APS, Sigma-Aldrich) to reduce nonspecific interactions with the target proteins. Successful synthesis and surface modification of the MNP matrix was ensured by step-by-step monitorization with Fourier transform infrared (FT-IR) spectroscopy. The size and shape of the MNPs was determined by transmission electron microscopy. The MNPs exhibited relatively narrow size distribution and an average diameter of 14±3 nm, thus confirming their nanoscale dimensions. The MNPs were then dispersed in dimethyl sulfoxide (DMSO, Sigma-Aldrich) and their surface was activated with suberic acid bis-N-hydroxysuccinimide ester (DSS, Sigma-Aldrich), washed thoroughly with DMSO, and incubated with *Concanavalin A* (ConA; Vector Laboratories, Burlingame, CA, USA), Wheat Germ Agglutinin (WGA; Vector Laboratories), and *Sambucus nigra* (SNA; Vector Laboratories) (1 mg/mL in binding buffer; Table 1S) at 4 °C for 12 h to produce MNP@ConA, MNP@WGA, and MNP@SNA particles, respectively. Binding of the lectins at the cells surface was confirmed by FT-IR and UV spectroscopy at 280 nm. The amount of bounded lectin was indirectly estimated from the conjugation supernatant.

Urine Glycoprotein enrichment with MNP@lectins

Proteins isolated from the urine of male individuals (n=5) were polled together according to the following groups: i) Controls; ii) low-grade NMIBC; iii) high-grade NMIBC; iv) MIBC. Protein extracts (100 μg) were diluted in 500 μL of lectin binding buffers and incubated with 1 mg of MNP@lectins at 4°C for 30 min. Specifically, ConA binding buffer was 20 mM Tris, 500 mM NaCl, 5 mM $MgCl_2$, 5 mM $MnCl_2$, and 5 mM $CaCl_2$, pH 7.45. WGA binding buffer was 20 mM Tris, 150 mM NaCl, and 0.05% Tween-20, pH 7.45. SNA binding buffer was 20 mM Tris, 150 mM NaCl, and 0.05% Tween-20, pH 7.45. The MNP were then pulled to the tube walls under the influence of a strong magnetic field, allowing an easy and complete removal of the supernatant. The unbound proteins were removed by washing three times with 200 μL of wash buffer (0.05% Tween-20 in binding

buffer). The glycoproteins were recovered with specific elution buffers for each ligand as previously reported [15, 16].

Protein identification and data curation

The proteins recovered from MNP@lectins were first reduced upon incubation with 5 mM dithiothreitol (DTT, Sigma-Aldrich) at 75ºC for 1 h, then alkylated with 20 mM iodoacetamide (Sigma-Aldrich) at room temperature in the dark for 2 h and digested with trypsin (Promega, Madison, WI, USA) for 37.0°C overnight. Protein identification was performed as previously described by our group [16]. Tryptic digests were separated with a C18 Pepmap (Dionex) column on an Ultimate 3000 (Dionex/LC Packings, Sunnyvale, CA) nano-HPLC, and fractions were collected with a Probot (Dionex/LC Packings, Sunnyvale, CA) directly onto a matrix-assisted laser desorption ionization (MALDI) plate. The MALDI-TOF/TOF (time-of-flight) mass spectrometry (MS) analysis was performed on a 4800 MALDI-TOF/TOF Analyzer (Applied Biosystems, Foster City, CA). The MS and MS/MS spectra acquired were processed and analyzed by the Global Protein Server Workstation (Applied Biosystems). LC-MALDI-MS/MS runs were done in triplicates for each MNP@lectin (ConA, WGA, SNA) and experimental condition (control, low-grade NMIBC, high-grade NMIBC; MIBC). Protein identification was achieved with a search performed against the Swiss-Prot protein database (March 2009, 428 650 entries) for *Homo sapiens*. The final list includes proteins common to all three independent LC-MALDI-MS/MS runs, which were then queried using the "Retrieve ID/mapping" tool of UniProtKB [17] for membrane glycoproteins with extracellular domain prone to exhibit or secreted glycoproteins [14]. The identified glycoproteins were then screened for putative *N*-glycosylation sites using NetNglyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc), an artificial neural network that examines the sequence context of Asn-X-Ser/Thr (where X is not Pro) sequons [18]. *O*-glycosylation was predicted using the NetOglyc 4.0 server (http:// www.cbs.dtu.dk/services/NetOGlyc) that produces neural network predictions of mucin-type GalNAc *O*-glycosylation sites in mammalian proteins [19]. Online bioinformatics tools Oncomine [20] and the ClueGO and CluePedia apps for Cytoscape (http://www.cytoscape.org/) [21, 22] allowed narrowing the identified species to glycobiomarkers with non-invasive clinical potential, as previously described by us [14].

Slot-Blot analysis

Urine proteins were slot-blotted on a nitrocellulose membrane (Whatman, Protan; pore size 0.45 μm) using the Hybri-slot apparatus (21052-014; Gibco BRL, Life Technologies, Waltham, MA, USA). Fetuin from fetal calf serum and deglycosylated bovine serum albumin (both purchased at Sigma-Aldrich) were used as positive and negative controls, respectively. Nonspecific binding was blocked with Carbo-Free Blocking Solution (SP-5040, Vector Laboratories) for 30 minutes at room temperature. The samples were incubated with biotinylated ConA, WGA and SNA lectins (Vector Laboratories) for 30 minutes at room temperature. Membranes were then washed with TBS-T (TBS with 0.5 % Tween 20) and incubated with VECTASTAIN® ABC (Peroxidase, PK-6100, Vector Laboratories) for 30 minutes at room temperature. Lectin-affinity was determined using biotin/avidin interaction. Samples were also screened for CD44 using a recombinant monoclonal antibody (anti-CD44, 1:150 in PBS; EPR1013Y; Abcam, Cambridge, UK) and a goat anti-rabbit IgG (H+L) Secondary Antibody, HRP (Invitrogen, Carlsbad, CA, USA). Reactive bands were detected by enhanced chemiluminescence ECL (Amersham Pharmacia Biotech Inc, Piscataway, NJ, USA) according to the manufacturer's instructions. Images were recorded using X-ray films (Kodak Biomax light Film, Sigma-Aldrich). The films were scanned in Molecular Imager Gel Doc XR+ System (Bio-Rad) and analyzed with QuantityOne software (v 4.6.3, Bio-Rad). The results reflected the average of at least three independent replicates.

Immunohistochemistry

Formalin-fixed, paraffin-embedded tissue sections (FFPE) were screened for CD44 expression by immunohistochemistry using the strepavidin/biotin peroxidase method. Briefly, 3 μm sections were deparaffinized with xylene, rehydrated with graded ethanol series, microwaved for 15 min in boiling citrate buffer (10mM Citric Acid, 0.05% Tween 20, pH 6.0, Sigma-Aldrich), and exposed to 3% hydrogen peroxide for 25 min. CD44 was detected using a recombinant monoclonal antibody (anti-CD44, 1:4000 in PBS; ab157107; Abcam, Cambridge, UK) after incubation overnight at 4°C. Upon digestion with an α-neuraminidase from *Clostridium perfringens* (Sigma Aldrich, St. Louis, MO, USA) for 4 h at 37ºC, ST and dST antigens were detected using an hybridoma-derived 3C9 antibody

1:2 overnight at 4ºC. The antigens were identified with UltraVision HRP Detection System Kit (Thermo Fisher Scientific, Waltham, MO, USA) followed by incubation with 3,3-diaminobenzidine tetrahydrochloride (Impact Dab, Vector Laboratories) for chromogenic development. Finally, the slides were counterstained with Harris's haematoxylin for 1 min. Negative control sections were performed by adding BSA (5% in PBS) devoided of primary antibody. The immunostained sections were blindly assessed using light microscopy by two independent observers and validated by an experienced pathologist. Briefly, a semi-quantitative approach was established to score immunoreactivity based on the intensity and extension of the staining. The extension of staining was rated in cutoffs of 10%, and staining intensity was rated as follows: negative-0, weak-1, moderate-2, strong-3. The tumours were then classified based on the multiplication of extension evaluation and intensity. Disaccording readings were re-analyzed using a double-headed microscope and consensus was reached.

Statistical analysis and data mining

Statistical analysis was performed using Graphpad prism7 by GraphPad Software, Inc. Differences between continuous variables among the evaluated groups were accessed by Mann-Whitney non- parametric test for independent samples. Differences were considered significant when $p<0.05$.

Results

To highlight the potential of glycan-affinity glycoproteomics nanoplatforms for biomarker discovery in BC, we have prospectively collected urine samples from low-grade and high-grade NMIBC and MIBC patients. A significant increase in protein content was observed in these urines in comparison to healthy controls matching in age and gender (**Figure S1A**-Supporting Information). Moreover, proteinuria increased with the severity of the lesions, in accordance with previous reports [23]. Concomitantly, the urines of cancer patients presented higher sialic acids content in comparison to the controls, which also was more pronounced in high-grade NMIBC and MIBC (**Figure S1B** – Supporting Information), suggesting profound alterations in the urine glycome. Urine glycoproteins were then isolated after dialysis and pooled together in accordance with the histopathological natures of the tumours.

As a proof-of-concept, we have incubated urine proteins from each group with MNP functionalized with three broad spectra lectins: i) ConA, targeting α-linked mannose present in *N*-glycan core oligosaccharides; ii) WGA, targeting terminal GlcNAc units common to many glycans found in membrane proteins; iii) SNA, targeting preferentially *O*-6 linked and, to less extent, *O*-3 linked sialic acids. Proteins showing affinity for these lectins were recovered from the beads and identified based on a bottom-up shotgun proteomics by nanoLC-MALDI-TOF/TOF. The resulting data was sorted for membrane or secreted glycoproteins showing putative *N*- and/or *O*-glycosylation sites, based on bioinformatics predictions using UniProt, NetNGlyc and NetOGlyc, respectively.

Overall, we have identified 114 glycoproteins, 63 of which were exclusively found in cancer patients' urines and 21 solely in the controls (**Table S1 and S2**-Supporting Information). Amongst bladder cancer patient urine samples, commonly observed proteins were apolipoprotein family members (APOA1, APOE), fibrinogen chains (FGA, FGB, FGG), alpha-1-antitrypsin (SERPINA1) and alpha-2-macroglobulin (A2M) (**Table 1S**-Supporting Information). A higher number of proteins were identified in high-grade NMIBC (n=80) and MIBC (n=96) in comparison to low-grade NMIBC (n=52) and control urines (n=55), considering all three lectins and accounting the repetitive capture of some proteins by the three lectins (**Figure 1A**). Nevertheless, cancer-specific glycoproteins could be observed in all cancer groups, including low-grade tumours (**Figure 1B**). Moreover, supporting the alterations observed in the glycome, slot-blot analysis of individual samples demonstrated that all three lectins presented higher affinity for proteins isolated from high-grade lesions, particularly in patients facing muscle-invasion, when compared to the controls (**Figure S2**-Supporting Information). Taken together, these findings suggest the existence of significant alterations in the nature of urine glycome and glycoproteome, accompanying the transition from low-grade to high-grade lesions; however, these alterations become more pronounced upon muscle invasion.

Considering cancer-associated glycoproteins, 8 protein species were exclusively found in low-grade tumours, suggesting potential for early diagnosis of patients with superficial lesions. Conversely, 13 were exclusively found in high-grade NMIBC and may aid non-invasive detection of progression of low-grade lesions. Moreover, 19 glycoproteins have been exclusively found in MIBC, holding potential to determine further progression to invasion. On the other hand, 12 glycoproteins have been detected in all groups and lack

potential for patient discrimination (**Figure 2 and Table S2**-Supporting Information). We have then comprehensively matched these findings against previous reports in bladder cancer tissues using Oncomine considering "p=0.001, 2-fold upregulation, All rank" thresholds. Interestingly, 81% of the cancer-associated urine glycoproteins had been previously detected in both healthy human urothelium and bladder cancer tissues (**Figure 2 and Table S2**-Supporting Information). It is possible that the net loss of integrity of the urothelium accompanying malignant transformations may ultimately dictate their release into urine in pathological situations. Moreover, many of these glycoproteins may experience glycome remodeling, favoring affinity for the lectins chosen for enrichment. Moreover, 19% of these glycoproteins have been found overexpressed in cancer tissues (IGHG3, SERPINA1, CD44, IGHM, HSPG2, GGH, APOE, MUC16, ATP13A2, ADCK2, PRSS2, and LAMA5) (**Figure 2**). More importantly, we note that SERPINA1, IGHM and CD44, which are significantly overexpressed in MIBC (according to Oncomine), can be detected in the urine of both high-grade NMIBC as well as MIBC patients. These observations suggest that these proteins may further aid on non-invasive differentiation between low-grade and high-grade tumours and ultimately be useful in risk stratification. In accordance, literature demonstrated that SERPINA1 is frequently found in higher stage and grade patients and was associated with tumour progression and shorter disease-free survival in both urines and tumour tissues from bladder cancer patient's analysis [24, 25]. Regarding IGHM, this is the first report of this protein in bladder cancer as a non-invasive differentiator between low-grade and high-grade tumours. Conversely, CD44 has been frequently reported as a bladder cancer stem-cell biomarker expressed by more aggressive subpopulations [13, 26]. Previous studies have demonstrated the presence of CD44 in exfoliated cancer cells in urine and its potential as a urinary marker for bladder cancer detection [27, 28]. This is the first report identifying CD44 as a protein that can be released into urine and suggesting its potential as a non-invasive stratification marker between low and high-grade bladder tumours.

A comprehensive integration of cancer-associated urine glycobiomarkers using Cytoscape's ClueGo and Cluepedia plugins further demonstrated an overrepresentation of MIBC-associated pathways, including regulation of immune responses (humoral and complement), platelet degranulation, which has been considered a key event for tumour growth and metastasis [29], and cholesterol transport, which may play a critical role in

cancer progression [30]. However, there are also common biological features to both NMIBC and MIBC, concerning the promotion of heterotypic regulation of cell-cell adhesion, mostly mediated by CD44. Given the key role played by CD44, its presence in urine samples and overexpression in more aggressive and invasive bladder tumours (Oncomine), further emphasis was devoted to validate these results and its potential biomarker value in the context of advanced disease.

CD44 expression in urine and bladder tumours

Orthogonal validation of CD44 expression in urine samples was performed by slot blotting. According to **Figure 4A**, CD44 urinary levels showed a trend increase with the severity of the lesions; however, this effect is only significant for high-grade tumours in comparison to controls and low-grade NMIBC.

In parallel, we have screened the corresponding bladder tissues for CD44 expression by immunohistochemistry. CD44 antigen was mainly found at the cancer cells membrane, in accordance with expected cellular location. Again, we observed a significant overexpression of CD44 with the severity of the lesions, which is more pronounced for high-grade tumours, thus, mimicking urine analysis. More importantly, CD44 was not observed in the tumour-adjacent histologically normal urothelium of the one case with normal urothelium representation, neither in the most superficial lesions, supporting its overexpression in cancer. In superficial tumours, CD44 was predominantly found in basal layer cells, which have been described to harbor more malignant clones. Contrastingly, MIBC cells presented an extensive and intense CD44 staining without a defined pattern throughout the tumour.

These findings agree with MNP@ lectins-based glycoproteomics and reinforce the role of CD44 in the context of advanced disease (high-grade NMIBC and MIBC).

Discussion and concluding remarks

The establishment of proteomics workflows to address clinically relevant urine glycoproteins in bladder cancer remains challenging due to the lack of efficient enrichment strategies. Herein, we have explored the potential of targeting glycans using MNP@lectins and/or MNP@moAbs. Nevertheless, we have previously demonstrated that MNP@lectins significantly increase glycoprotein yields when compared to conventional immobilization

strategies (sepharose/agarose) and microspheres [15]. This is most likely associated to superior dynamics provided by large surface volume ratios, while displaying low unspecific binding. Moreover, we have used this approach to address the serum, saliva, and urine glycoproteomes, which resulted in the identification of proteins spanning large dynamic ranges, including low abundant species [16]. Likewise, we have hypothesized that this approach could be ideal for addressing cancer-biomarkers, generally present in minute amounts (nano-fentomolar range) in urine samples. Accordingly, we have combined information from MNP functionalized with three large spectra lectins (ConA - *N*-glycans structures; WGA - GlcNAc in *N*- and *O*-glycans; and SNA - sialic acids also present in in N- and O-glycans) envisaging biomarker panels associated with different aspects of the disease.

Even though explorative, this study has highlighted significant alterations in the urine glycome and glycoproteome, warranting comprehensive and more in-depth evaluation in future studies. These changes were significantly more pronounced in MIBC in comparison to initial stages of the disease and healthy controls, supporting a disease-associated nature of these findings. Amongst the identified proteins were known bladder cancer-associated biomarkers apolipoprotein family, fibrinogen chains, alpha-1-antitrypsin, alpha-2-macroglobulin and uromodulin, previously reported as urinary biomarkers for bladder cancer [31, 32]. However, it has been extensively discussed that, individually, these proteins lack the necessary sensitivity and specificity to face the main clinical challenges raised by bladder cancer management, i.e. early identification of patients facing disease progression to high-grade disease and invasion. Giving this challenge, our strategy has generated a broader panel of 42 glycoproteins associated to high-grade disease (NMIBC and MIBC) that may hold clinical potential, 19 of which exclusively found in MIBC. In addition, the comprehensive integration of data using protein-protein interaction networks (Cytoscape) demonstrated the existence of common pathways to both NMIBC and MIBC (humoral response, complement activation, cell-adhesion) even though driven by different protein players. Amongst these glycoproteins is CD44, which has been found to modulate cell-cell adhesion in both NMIBC and MIBC (oncomine) and has been previously reported to drive invasion and metastasis [33, 34]. In agreement with these observations, we found high levels of CD44 in the urine of patients diagnosed with high-grade NMIBC and MIBC, possibly reflecting the molecular nature of the tumour. The CD44 antigen has been

previously found overexpressed in more aggressive bladder tumours and bladder cancer cells exfoliated into urine, and is frequently associated to poor prognosis [27]. Nevertheless, this is the first report demonstrating the shedding of this glycoprotein into urine, supporting its potential as biomarkers of disease progression. However, CD44 may present several isoforms that may change according to the origin of the tumour as well as the severity of the lesions. In bladder cancer, the most described CD44 isoforms are CD44s, CD44v6 and CD44v9, both associated with poor prognosis [35-37]. Variations in CD44 expression might be connected to the derangement of differentiation in bladder tumour cells [38]. Moreover, these cells are usually located in the basal layer of normal urothelium, the potential location of initiation tumour bladder cells, thereby reinforcing the value of CD44 as bladder cancer stem cell marker [39].

Future studies should therefore devote to a comprehensive clarification of the nature of CD44 isoforms patterns as well as its glycoforms in bladder cancer envisaging highly specific biomarkers. The identification of clinically relevant CD44 glycoforms may further enhance the biomarker value of this glycoprotein, has previously demonstrated for other glycoproteins, namely PSA in prostate cancer [40]. Accordingly, we believe that the MNP@lectins-based enrichment strategy may also constitute a key starting point for glycome characterization by narrowing down identified glycoproteins to its clinically relevant glycoforms.

In summary, we believe that the future of biomarker discovery in bladder cancer should include a comprehensive interrogation of the glycome and glycoproteome, as well as the integration of relevant glycobiomarkers into broader panomics data. The generalization of glycan-based enrichment nanoplatforms may provide the necessary means to achieve this goal.

Acknowledgments

The authors wish to acknowledge the Portuguese Foundation for Science and Technology (FCT) for the human resources grants: PhD grant SFRH/BD/105355/2014 (Rita Azevedo), SFRH/BD/111242/2015 (Andreia Peixoto), SFRH/BD/103571/2014 (Elisabete Fernandes), SFRH/BD/127327/2016 Cristiana Gaiteiro and Postdoctoral grants SFRH/BPD/101827/2014 (Luis Lima) and SFRH/BPD/111048/2015 (José Alexandre Ferreira). FCT is co-financed by European Social Fund (ESF) under Human Potential

Operation Programme (POPH) from National Strategic Reference Framework (NSRF). The authors also acknowledge the Portuguese Oncology Institute of Porto Research Centre (CI-IPOP-29-2014; CI-IPOP-58-2015) and PhD Programs in Biomedicine and Pathology and Molecular Pathology of ICBAS-University of Porto.

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D and Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015; 136(5):E359-386.

2. Babjuk M, Bohle A, Burger M, Capoun O, Cohen D, Comperat EM, Hernandez V, Kaasinen E, Palou J, Roupret M, van Rhijn BW, Shariat SF, Soukup V, Sylvester RJ and Zigeuner R. EAU Guidelines on Non-Muscle-invasive Urothelial Carcinoma of the Bladder: Update 2016. Eur Urol. 2017; 71(3):447-461.

3. Azevedo R, Peixoto A, Gaiteiro C, Fernandes E, Neves M, Lima L, Santos LL and Ferreira JA. Over forty years of bladder cancer glycobiology: Where do glycans stand facing precision oncology? Oncotarget. 2017; 8:91734-91764.

4. Bassi P, De Marco V, De Lisa A, Mancini M, Pinto F, Bertoloni R and Longo F. Non-invasive diagnostic tests for bladder cancer: a review of the literature. Urol Int. 2005; 75(3):193-200.

5. Zhao M, Li M, Yang Y, Guo Z, Sun Y, Shao C, Li M, Sun W and Gao Y. A comprehensive analysis and annotation of human normal urinary proteome. Sci Rep. 2017; 7(1):3024.

6. Chen YT, Chen HW, Domanski D, Smith DS, Liang KH, Wu CC, Chen CL, Chung T, Chen MC, Chang YS, Parker CE, Borchers CH and Yu JS. Multiplexed quantification of 63 proteins in human urine by multiple reaction monitoring-based mass spectrometry for discovery of potential bladder cancer biomarkers. J Proteomics. 2012; 75(12):3529-3545.

7. Frantzi M, Latosinska A, Fluhe L, Hupe MC, Critselis E, Kramer MW, Merseburger AS, Mischak H and Vlahou A. Developing proteomic biomarkers for bladder cancer: towards clinical application. Nat Rev Urol. 2015; 12(6):317-330.

8. Hashim OH, Jayapalan JJ and Lee CS. Lectins: an effective tool for screening of potential cancer biomarkers. PeerJ. 2017; 5:e3784.

9. Lima L, Neves M, Oliveira MI, Dieguez L, Freitas R, Azevedo R, Gaiteiro C, Soares J, Ferreira D, Peixoto A, Fernandes E, Montezuma D, Tavares A, Ribeiro R, Castro

A, Oliveira M, et al. Sialyl-Tn identifies muscle-invasive bladder cancer basal and luminal subtypes facing decreased survival, being expressed by circulating tumour cells and metastases. Urol Oncol. 2017.

10. Lima L, Severino PF, Silva M, Miranda A, Tavares A, Pereira S, Fernandes E, Cruz R, Amaro T, Reis CA, Dall'Olio F, Amado F, Videira PA, Santos L and Ferreira JA. Response of high-risk of recurrence/progression bladder tumours expressing sialyl-Tn and sialyl-6-T to BCG immunotherapy. Br J Cancer. 2013; 109(8):2106-2114.

11. Costa C, Pereira S, Lima L, Peixoto A, Fernandes E, Neves D, Neves M, Gaiteiro C, Tavares A, Gil da Costa RM, Cruz R, Amaro T, Oliveira PA, Ferreira JA and Santos LL. Abnormal Protein Glycosylation and Activated PI3K/Akt/mTOR Pathway: Role in Bladder Cancer Prognosis and Targeted Therapeutics. PLoS One. 2015; 10(11):e0141253.

12. Peixoto A, Fernandes E, Gaiteiro C, Lima L, Azevedo R, Soares J, Cotton S, Parreira B, Neves M, Amaro T, Tavares A, Teixeira F, Palmeira C, Rangel M, Silva AM, Reis CA, et al. Hypoxia enhances the malignant nature of bladder cancer cells and concomitantly antagonizes protein O-glycosylation extension. Oncotarget. 2016; 7(39):63138-63157.

13. Cotton S, Azevedo R, Gaiteiro C, Ferreira D, Lima L, Peixoto A, Fernandes E, Neves M, Neves D, Amaro T, Cruz R, Tavares A, Rangel M, Silva AMN, Santos LL and Ferreira JA. Targeted O-glycoproteomics explored increased sialylation and identified MUC16 as a poor prognosis biomarker in advanced-stage bladder tumours. Mol Oncol. 2017; 11(8):895-912.

14. Azevedo R, Silva AMN, Reis CA, Santos LL and Ferreira JA. In silico approaches for unveiling novel glycobiomarkers in cancer. J Proteomics. 2018; 171:95-106.

15. Ferreira JA, Daniel-da-Silva AL, Alves RM, Duarte D, Vieira I, Santos LL, Vitorino R and Amado F. Synthesis and optimization of lectin functionalized nanoprobes for the selective recovery of glycoproteins from human body fluids. Anal Chem. 2011; 83(18):7035-7043.

16. Cova M, Oliveira-Silva R, Ferreira JA, Ferreira R, Amado F, Daniel-da-Silva AL and Vitorino R. Glycoprotein enrichment method using a selective magnetic nano-probe platform (MNP) functionalized with lectins. Methods Mol Biol. 2015; 1243:83-100.

17. The UniProt C. UniProt: the universal protein knowledgebase. Nucleic Acids Res. 2017; 45(D1):D158-D169.

18. Gupta R, Jung E and Brunak S. Prediction of N-glycosylation sites in human proteins. In preparation, 2004.

19. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT, Lavrsen K, Dabelsteen S, Pedersen NB, Marcos-Silva L, Gupta R, Bennett EP, Mandel U, Brunak S, Wandall HH, Levery SB, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. The EMBO journal. 2013; 32(10):1478-1488.

20. Rhodes DR, Kalyana-Sundaram S, Mahavisno V, Varambally R, Yu J, Briggs BB, Barrette TR, Anstet MJ, Kincead-Beal C, Kulkarni P, Varambally S, Ghosh D and Chinnaiyan AM. Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. Neoplasia. 2007; 9(2):166-180.

21. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003; 13(11):2498-2504.

22. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pages F, Trajanoski Z and Galon J. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics. 2009; 25(8):1091-1093.

23. Yang DY, Thompson RH, Zaid HB, Lohse CM, Rule AD, Boorjian SA, Leibovich BC, Cheville JC and Tollefson MK. Severity of Preoperative Proteinuria is a Risk Factor for Overall Mortality in Patients Undergoing Nephrectomy. J Urol. 2017; 198(4):795-802.

24. Linden M, Segersten U, Runeson M, Wester K, Busch C, Pettersson U, Lind SB and Malmstrom PU. Tumour expression of bladder cancer-associated urinary proteins. BJU Int. 2013; 112(3):407-415.

25. Zhang G, Gomes-Giacoia E, Dai Y, Lawton A, Miyake M, Furuya H, Goodison S and Rosser CJ. Validation and clinicopathologic associations of a urine-based bladder cancer biomarker signature. Diagn Pathol. 2014; 9:200.

26. Hofner T, Macher-Goeppinger S, Klein C, Schillert A, Eisen C, Wagner S, Rigo-Watermeier T, Baccelli I, Vogel V, Trumpp A and Sprick MR. Expression and prognostic significance of cancer stem cell markers CD24 and CD44 in urothelial bladder cancer xenografts and patients undergoing radical cystectomy. Urol Oncol. 2014; 32(5):678-686.

27. Sugiyama M, Woodman A, Sugino T, Crowley S, Ho K, Smith J, Matsumura Y and Tarin D. Non-invasive detection of bladder cancer by identification of abnormal CD44 proteins in exfoliated cancer cells in urine. Clin Mol Pathol. 1995; 48(3):M142-147.

28. Woodman AC, Goodison S, Drake M, Noble J and Tarin D. Noninvasive diagnosis of bladder carcinoma by enzyme-linked immunosorbent assay detection of CD44 isoforms in exfoliated urothelia. Clin Cancer Res. 2000; 6(6):2381-2392.

29. Yan M and Jurasz P. The role of platelets in the tumour microenvironment: From solid tumours to leukemia. Biochim Biophys Acta. 2016; 1863(3):392-400.

30. Silvente-Poirot S and Poirot M. Cholesterol metabolism and cancer: the good, the bad and the ugly. Curr Opin Pharmacol. 2012; 12(6):673-676.

31. Frantzi M and Vlahou A. Ten Years of Proteomics in Bladder Cancer: Progress and Future Directions. Bladder Cancer. 2017; 3(1):1-18.

32. Sandim V, Pereira Dde A, Kalume DE, Oliveira-Carvalho AL, Ornellas AA, Soares MR, Alves G and Zingali RB. Proteomic analysis reveals differentially secreted proteins in the urine from patients with clear cell renal cell carcinoma. Urol Oncol. 2016; 34(1):5 e11-25.

33. Yu G, Yao W, Xiao W, Li H, Xu H and Lang B. MicroRNA-34a functions as an anti-metastatic microRNA and suppresses angiogenesis in bladder cancer by directly targeting CD44. J Exp Clin Cancer Res. 2014; 33:779.

34. Wu K, Ning Z, Zeng J, Fan J, Zhou J, Zhang T, Zhang L, Chen Y, Gao Y, Wang B, Guo P, Li L, Wang X and He D. Silibinin inhibits beta-catenin/ZEB1 signaling and suppresses bladder cancer metastasis via dual-blocking epithelial-mesenchymal transition and stemness. Cell Signal. 2013; 25(12):2625-2633.

35. Omran OM and Ata HS. CD44s and CD44v6 in diagnosis and prognosis of human bladder cancer. Ultrastruct Pathol. 2012; 36(3):145-152.

36. Klatte T, Seligson DB, Rao JY, Yu H, de Martino M, Garraway I, Wong SG, Belldegrun AS and Pantuck AJ. Absent CD44v6 expression is an independent predictor of poor urothelial bladder cancer outcome. J Urol. 2010; 183(6):2403-2408.

37. Kobayashi K, Matsumoto H, Matsuyama H, Fujii N, Inoue R, Yamamoto Y and Nagao K. Clinical significance of CD44 variant 9 expression as a prognostic indicator in bladder cancer. Oncol Rep. 2016; 36(5):2852-2860.

38. Kuncova J, Kostrouch Z, Viale M, Revoltella R and Mandys V. Expression of CD44v6 correlates with cell proliferation and cellular atypia in urothelial carcinoma cell lines 5637 and HT1197. Folia Biol (Praha). 2005; 51(1):3-11.

39. Yang YM and Chang JW. Bladder cancer initiating cells (BCICs) are among EMA-CD44v6+ subset: novel methods for isolating undetermined cancer stem (initiating) cells. Cancer Invest. 2008; 26(7):725-733.

40. Llop E, Ferrer-Batalle M, Barrabes S, Guerrero PE, Ramirez M, Saldova R, Rudd PM, Aleixandre RN, Comet J, de Llorens R and Peracaula R. Improvement of Prostate Cancer Diagnosis by Detecting PSA Glycosylation-Specific Changes. Theranostics. 2016; 6(8):1190-1204.

Figures

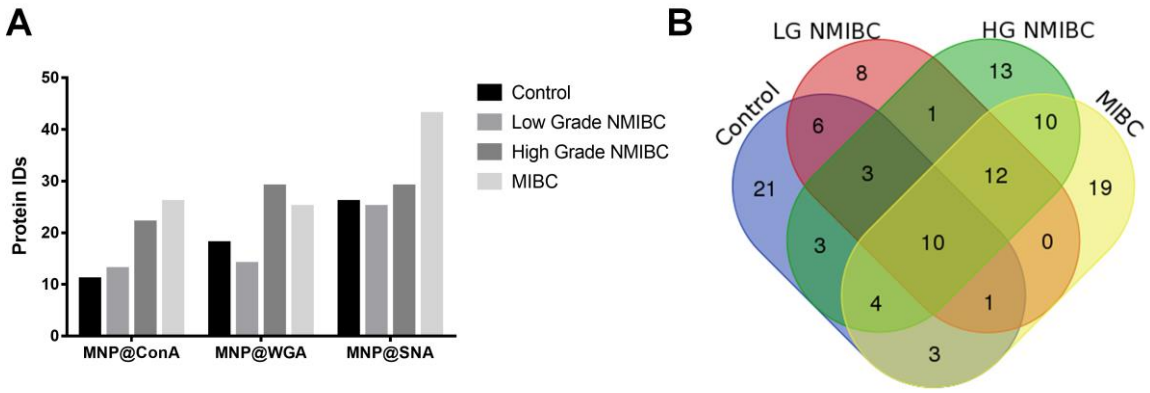

Figure 1. Urine glycoproteins identified in controls and low-grade NMIBC, high-grade NMIBC and MIBC using MNP@ConA, MNP@WGA and MNP@SNA lectins. **A.** A higher number of proteins were identified for high-grade NMIBC (n=80) and MIBC (n=96) in comparison to low-grade NMIBC (n=52) and control urines (n=55), considering all three lectins. **B.** Venn diagram highlighting the existence of distinct lectin-based glycoprotein signatures for each group, holding potential for early detection (low-grade tumours group); disease progression to undifferentiating (high-grade NMIBC group); progression to invasion (MIBC) and poor prognosis (high-grade NMIBC+MIBC group).

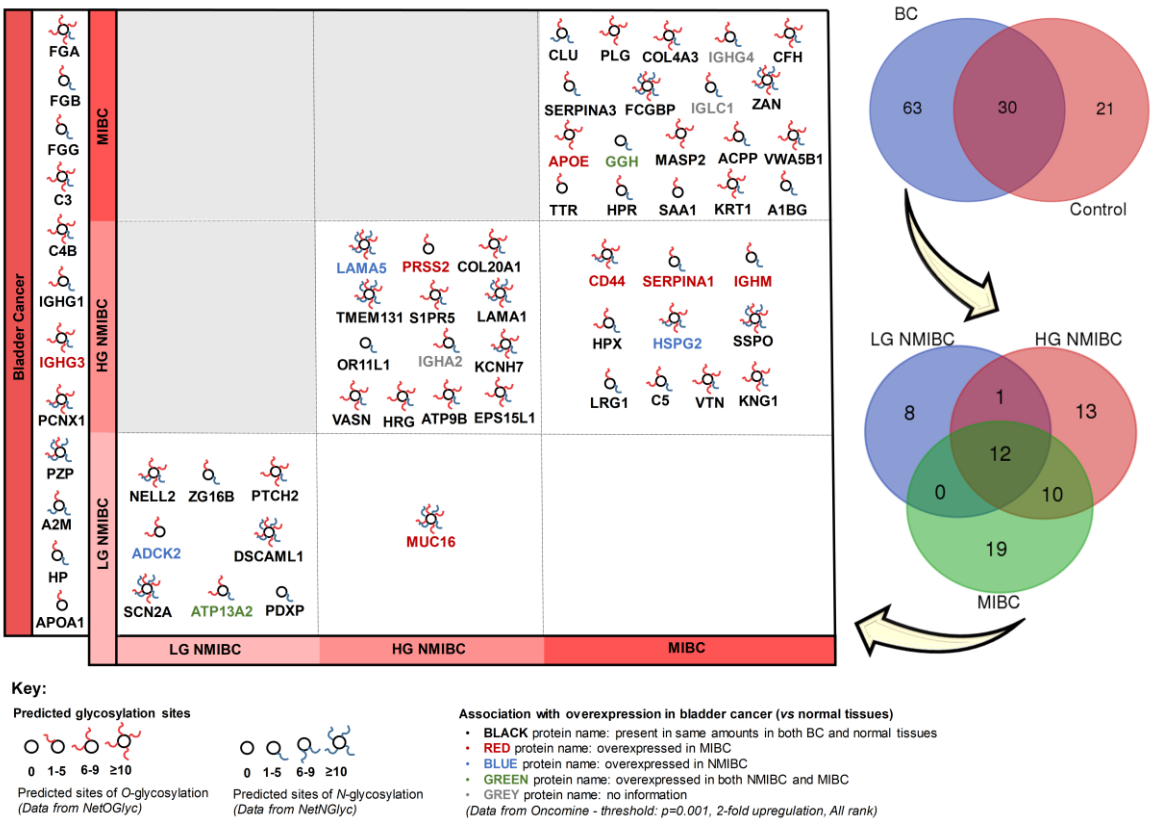

Figure 2. Graphical representation of cancer-specific urine glycobiomarkers with affinity to MNP@ConA, MNP@WGA and MNP@SNA lectin nanoprobes distributed according to their expression in bladder cancer. Accordingly 63 glycoproteins were exclusively found in cancer samples, (top right Venn diagram). The left panel represents the distribution of these glycoproteins amongst the cancer groups (low-grade and high-grade NMIBC and MIBC). Significant associations of these glycoproteins with different aspects of the disease according to an *in silico* analysis with Oncomine was represented by different protein name colours. The number of predicted *O*- and *N*-glycosylation sites were also represented according to results from NetOGlyc and NetNGlyc bioinformatics tools, respectively.

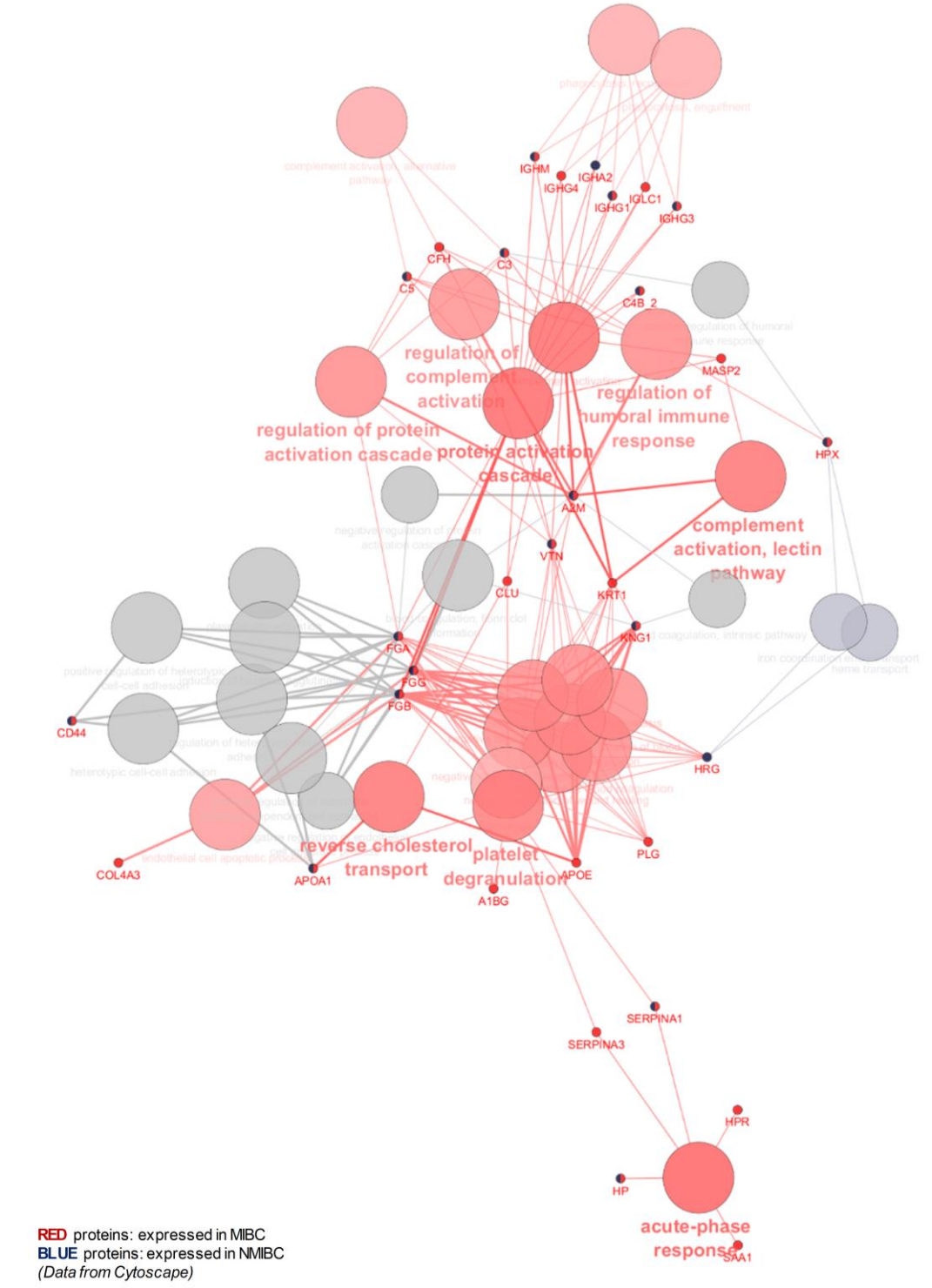

Figure 3. Network presenting the most significant biological functions of the candidate glycobiomarkers in bladder cancer. This network mainly reflects biological functions of immune system regulation (humoral response and complement), platelet degranulation and cholesterol transport as also heterotypic cell-cell adhesion, which is essentially played by CD44 glycoprotein in both NMIBC and MIBC.

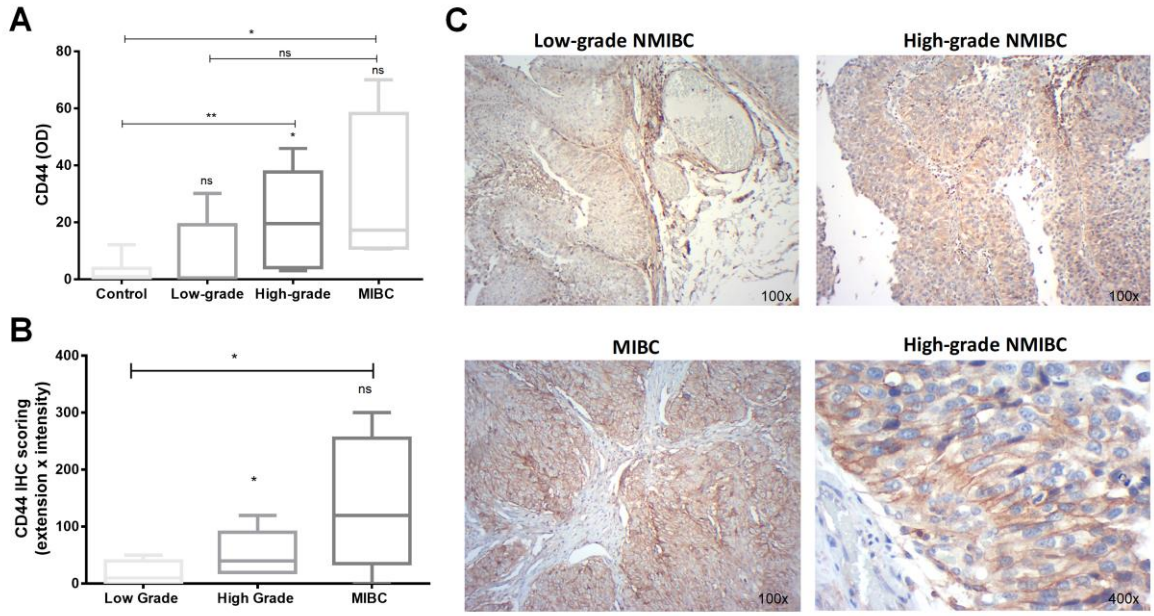

Figure 4. Validation of CD44 expression in urine samples and immunohistochemistry of bladder cancer patients. A. CD44 urinary levels showed a trend increase with the severity of the lesions; however this effect is only significant for high-grade tumours in comparison to the control and low-grade NMIBC groups. **B.** CD44 antigen was mainly found at membrane of cancer cells, in accordance with expected cellular location. Again, we observed a significant overexpression of CD44 with the severity of the lesions, which is more pronounced for high-grade tumours, thus, mimicking urine analyses. **C.** In superficial tumours, CD44 was predominantly found in cells underlining the basal layer, which have been described to harbor more malignant clones. Contrastingly, MIBC cells presented an extensive and intense CD44 expression without a defined pattern throughout the tumour. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Glycan affinity magnetic nanoplatforms for urinary glycobiomarkers discovery in bladder cancer

Rita Azevedo[1,2*], Janine Soares[1*], Cristiana Gaiteiro[1,2], Andreia Peixoto[1], Luis Lima[1,3,4], Dylan Ferreira[1], Marta Relvas-Santos[1], Elisabete Fernandes[1,2], Ana Tavares[1], Sofia Cotton[1], Daniel-da-Silva AL[6], Rui Vitorino[7,8], Lúcio Lara Santos[1,2], Francisco Amado[9] and José Alexandre Ferreira[1,2,3,4,10]

[1]Experimental Pathology and Therapeutics Group, Research Center of Portuguese Institute of Oncology (CI-IPOP), Porto, Portugal; [2]Institute of Biomedical Sciences Abel Salazar, University of Porto, Porto, Portugal; [3]Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Portugal; [4]Glycobiology in Cancer, Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP), Porto, Portugal; [5]Department of Urology, Portuguese Institute of Oncology, Porto, Portugal; [6]CICECO - Aveiro Institute of Materials, Department of Chemistry, University of Aveiro, Aveiro, Portugal; [7]iBiMED - Institute of Biomedicine, Department of Medical Sciences, University of Aveiro, Aveiro, Portugal; [8]Research Unit of the Physiology and Cardiothoracic Surgery Department of the Faculty of Medicine, University of Porto, Porto, Portugal; [9]Mass Spectrometry Centre, Organic Chemistry and Natural Products Unit, Department of Chemistry, University of Aveiro, Aveiro, Portugal; [10]International Iberian Nanotechnology Laboratory, Braga, Portugal.

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

*Equal contribution

Keywords:

Glycoproteomics, urine biomarkers, bladder cancer, glycosylation, nanoprobes

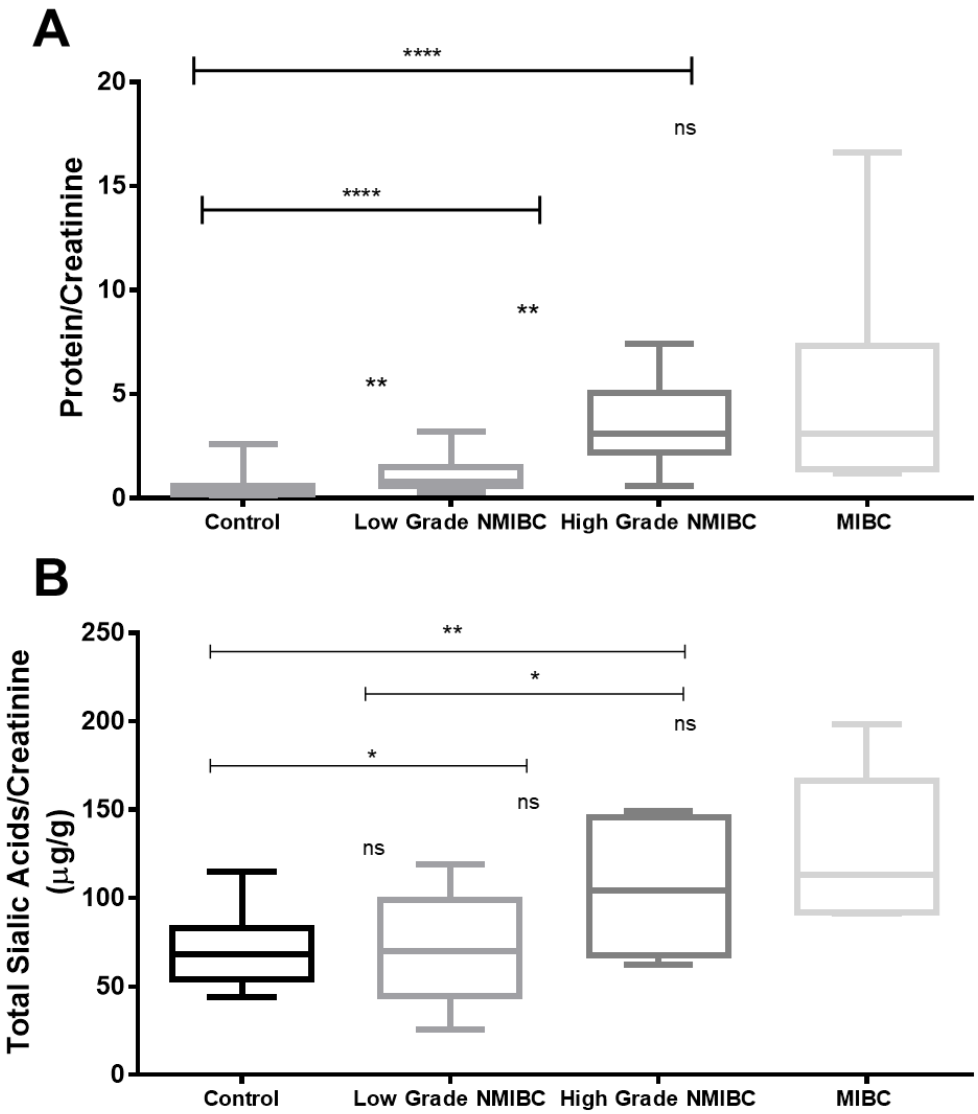

Figure S1. Protein content and total sialic acids of urines from controls and bladder cancer patients with low-grade NMIBC, high-grade NMIBC and MIBC. A significant increase in urinary proteinuria and in sialic acids content was observed along with the severity of the bladder cancer lesions in comparison to individuals without known neoplasia, matched in age and gender.

* $p \leq 0.05$, ** $p \leq 0.01$, **** $p \leq 0.0001$

Table S1. Extracellular membrane-bound or secreted glycoproteins with affinity to ConA, WGA or SNA lectins showing *N*- or *O*-glycosylation that were identified in urine samples from controls

Protein Name	MW	Entry	*N*-glycosylated	*O*-glycosylated	ConA	WGA	SNA
Alpha-1-acid glycoprotein 1 OS=Homo sapiens GN=ORM1 PE=1 SV=1	23	P02763	YES	NO			√
Alpha-2-HS-glycoprotein OS=Homo sapiens GN=AHSG PE=1 SV=1	39	P02765	YES	YES			√
Alpha-2-macroglobulin-like protein 1 OS=Homo sapiens GN=A2ML1 PE=1 SV=2	161	A8K2U0	YES	YES			√
Alpha-amylase 2B OS=Homo sapiens GN=AMY2B PE=1 SV=1	58	P19961	YES	YES			√
Apolipoprotein B-100 OS=Homo sapiens GN=APOB PE=1 SV=1	515	P04114	YES	YES			√
Cerebellin-2 OS=Homo sapiens GN=CBLN2 PE=2 SV=1	24	Q8IUK8	YES	YES		√	
Ceruloplasmin OS=Homo sapiens GN=CP PE=1 SV=1	122	P00450	YES	YES			√
Desmoglein-1 OS=Homo sapiens GN=DSG1 PE=1 SV=2	114	Q02413	YES	YES	√		
Dipeptidase 2 OS=Homo sapiens GN=DPEP2 PE=1 SV=1	53	Q9H4A9	YES	YES		√	
Fibrocystin-L OS=Homo sapiens GN=PKHD1L1 PE=2 SV=2	465	Q86WI1	YES	YES	√		
Fibronectin OS=Homo sapiens GN=FN1 PE=1 SV=3	262	P02751	YES	YES			√
Frizzled-4 OS=Homo sapiens GN=FZD4 PE=1 SV=2	60	Q9ULV1	YES	YES		√	
Galectin-3-binding protein OS=Homo sapiens GN=LGALS3BP PE=1 SV=1	65	Q08380	YES	YES			√
Gamma-interferon-inducible lysosomal thiol reductase OS=Homo sapiens GN=IFI30 PE=1 SV=2	29	P13284	YES	NO			√
G-protein coupled receptor 98 OS=Homo sapiens GN=GPR98 PE=1 SV=1	692	Q8WXG9	YES	YES			√
Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=2	38	P01876	YES	YES			√
Ig gamma-2 chain C region OS=Homo sapiens GN=IGHG2 PE=1 SV=2	36	P01859	YES	YES			√
Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1	12	P01834	NO	YES			√
Immunoglobulin superfamily member 10 OS=Homo sapiens GN=IGSF10 PE=1 SV=1	291	Q6WRI0	YES	YES			√
Indian hedgehog protein OS=Homo sapiens GN=IHH PE=1 SV=4	45	Q14623	YES	YES		√	
Integrin alpha-L OS=Homo sapiens GN=ITGAL PE=1 SV=3	129	P20701	YES	YES		√	
Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4	103	Q14624	YES	YES			√
Interleukin-17 receptor A OS=Homo sapiens GN=IL17RA PE=1 SV=2	96	Q96F46	YES	YES		√	
Junction plakoglobin OS=Homo sapiens GN=JUP PE=1 SV=3	82	P14923	YES	YES	√		
Laminin subunit alpha-2 OS=Homo sapiens GN=LAMA2 PE=1 SV=4	344	P24043	YES	YES	√	√	
Large proline-rich protein BAT3 OS=Homo sapiens GN=BAT3 PE=1 SV=2	119	P46379	YES	YES	√	√	

Lutropin-choriogonadotropic hormone receptor OS=Homo sapiens GN=LHCGR PE=1 SV=4	79	P22888	YES	YES		√	
Macrophage-stimulating protein receptor OS=Homo sapiens GN=MST1R PE=1 SV=1	152	Q04912	YES	YES	√		
Monocyte differentiation antigen CD14 OS=Homo sapiens GN=CD14 PE=1 SV=2	40	P08571	YES	YES			√
Neurotensin receptor type 2 OS=Homo sapiens GN=NTSR2 PE=2 SV=1	45	O95665	NO	YES		√	
Osteopontin OS=Homo sapiens GN=SPP1 PE=1 SV=1	35	P10451	YES	YES		√	
Otoferlin OS=Homo sapiens GN=OTOF PE=1 SV=3	227	Q9HC10	YES	YES		√	
Pancreatic alpha-amylase OS=Homo sapiens GN=AMY2A PE=1 SV=2	58	P04746	YES	YES			√
Pepsin A OS=Homo sapiens GN=PGA3 PE=1 SV=1	42	P0DJD8	NO	YES			√
Plasma serine protease inhibitor OS=Homo sapiens GN=SERPINA5 PE=1 SV=2	46	P05154	YES	YES			√
Prostaglandin-H2 D-isomerase OS=Homo sapiens GN=PTGDS PE=1 SV=1	21	P41222	YES	NO			√
Protein AMBP OS=Homo sapiens GN=AMBP PE=1 SV=1	39	P02760	YES	YES			√
Protein notum homolog OS=Homo sapiens GN=NOTUM PE=2 SV=2	56	Q6P988	YES	YES		√	
Protocadherin Fat 1 OS=Homo sapiens GN=FAT1 PE=1 SV=1	506	Q14517	YES	YES	√		
Protocadherin Fat 2 OS=Homo sapiens GN=FAT2 PE=1 SV=1	479	Q9NYQ8	YES	YES	√		
Putative maltase-glucoamylase-like protein FLJ16351 OS=Homo sapiens PE=2 SV=1	74	Q2M2H8	YES	YES			√
Serotransferrin OS=Homo sapiens GN=TF PE=1 SV=2	77	P02787	YES	YES			√
Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2	69	P02768	NO	YES			√
Tenascin-X OS=Homo sapiens GN=TNXB PE=1 SV=3	464	P22105	YES	YES		√	
Transmembrane protein 132D OS=Homo sapiens GN=TMEM132D PE=2 SV=1	122	Q14C87	YES	YES	√		
Transmembrane protein 214 OS=Homo sapiens GN=TMEM214 PE=1 SV=2	77	Q6NUQ4	YES	YES		√	
Transmembrane protein 8A OS=Homo sapiens GN=TMEM8A PE=1 SV=2	85	Q9HCN3	YES	YES		√	
Uromodulin OS=Homo sapiens GN=UMOD PE=1 SV=1	70	P07911	YES	YES	√		√
Vesicle-associated membrane protein 5 OS=Homo sapiens GN=VAMP5 PE=1 SV=1	13	O95183	YES	YES		√	
WAP, kazal, immunoglobulin, kunitz and NTR domain-containing protein 2 OS=Homo sapiens GN=WFIKKN2 P	64	Q8TEU8	YES	YES	√	√	
Zinc-alpha-2-glycoprotein OS=Homo sapiens GN=AZGP1 PE=1 SV=2	34	P25311	YES	YES			√

Table S2. Cancer-specific extracellular membrane-bound or secreted glycoproteins with affinity to ConA, WGA or SNA lectins showing *N*- or *O*-glycosylation that were identified in urine samples from bladder cancer patients

Protein Name	MW	Entry	O-glycosylation sites	N-glycosylation sites	Bladder cancer	ConA	WGA	SNA	Oncomine
Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=2	163	P01023	4	8	All	√	√	√	NS
Fibrinogen beta chain OS=Homo sapiens GN=FGB PE=1 SV=2	56	P02675	5	1	All	√	√	√	NS
Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	187	P01024	>10	2	All		√	√	NS
Fibrinogen gamma chain OS=Homo sapiens GN=FGG PE=1 SV=3	51	P02679	4	1	All		√	√	NS
Pregnancy zone protein OS=Homo sapiens GN=PZP PE=1 SV=4	164	P20742	8	>10	All		√	√	NS
Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1	31	P02647	4	NO	All	√	√	√	NS
Complement C4-B OS=Homo sapiens GN=C4B PE=1 SV=1	193	P0C0L5	>10	5	All			√	NS
Ig gamma-1 chain C region OS=Homo sapiens GN=IGHG1 PE=1 SV=1	36	P01857	7	1	All	√	√	√	NS
Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1	45	P00738	5	4	All	√	√	√	NS
Ig gamma-3 chain C region OS=Homo sapiens GN=IGHG3 PE=1 SV=2	41	P01860	>10	1	All	√	√	√	MIBC
Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	95	P02671	>10	2	All		√	√	NS
Pecanex-like protein 1 OS=Homo sapiens GN=PCNX PE=2 SV=2	259	Q96RV3	>10	7	All		√	√	NS
Leucine-rich alpha-2-glycoprotein OS=Homo sapiens GN=LRG1 PE=1 SV=2	38	P02750	5	2	HG NMIBC/MIBC	√	√		NS
Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3	47	P01009	5	4	HG NMIBC/MIBC		√	√	MIBC
CD44 antigen OS=Homo sapiens GN=CD44 PE=1 SV=2	82	P16070	>10	8	HG NMIBC/MIBC		√		MIBC
Complement C5 OS=Homo sapiens GN=C5 PE=1 SV=4	188	P01031	8	3	HG NMIBC/MIBC		√	√	NS
Hemopexin OS=Homo sapiens GN=HPX PE=1 SV=2	52	P02790	9	5	HG NMIBC/MIBC		√	√	NS
Kininogen-1 OS=Homo sapiens GN=KNG1 PE=1 SV=2	72	P01042	>10	3	HG NMIBC/MIBC		√	√	NS
Vitronectin OS=Homo sapiens GN=VTN PE=1	54	P04004	>10	2	HG NMIBC/MIBC		√	√	NS

SV=1									
Ig mu heavy chain disease protein OS=Homo sapiens PE=1 SV=1	43	P01871	2	5	HG NMIBC/MIBC			√	MIBC
Basement membrane-specific heparan sulfate proteoglycan core protein OS=Homo sapiens GN=HSPG2 PE=1 SV=3	469	P98160	>10	8	HG NMIBC/MIBC		√	√	NMIBC
SCO-spondin OS=Homo sapiens GN=SSPO PE=2 SV=1	547	A2VEC9	>10	>10	HG NMIBC/MIBC			√	NS
Ig gamma-4 chain C region OS=Homo sapiens GN=IGHG4 PE=1 SV=1	36	P01861	9	1	MIBC	√		√	No results
Ig lambda chain C regions OS=Homo sapiens GN=IGLC1 PE=1 SV=1	11	P0CG04	3	1	MIBC	√		√	No results
Prostatic acid phosphatase OS=Homo sapiens GN=ACPP PE=1 SV=3	45	P15309	6	3	MIBC	√			NS
Mannan-binding lectin serine protease 2 OS=Homo sapiens GN=MASP2 PE=1 SV=3	76	O00187	>10	NO	MIBC	√			NS
Transthyretin OS=Homo sapiens GN=TTR PE=1 SV=1	16	P02766	1	NO	MIBC	√			NS
Gamma-glutamyl hydrolase OS=Homo sapiens GN=GGH PE=1 SV=2	36	Q92820	NO	3	MIBC	√			Both
Complement factor H OS=Homo sapiens GN=CFH PE=1 SV=4	139	P08603	>10	4	MIBC		√	√	NS
Haptoglobin-related protein OS=Homo sapiens GN=HPR PE=1 SV=2	39	P00739	2	3	MIBC			√	NS
Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6	66	P04264	>10	2	MIBC			√	NS
Alpha-1B-glycoprotein OS=Homo sapiens GN=A1BG PE=1 SV=3	54	P04217	1	4	MIBC			√	NS
Alpha-1-antichymotrypsin OS=Homo sapiens GN=SERPINA3 PE=1 SV=2	48	P01011	2	4	MIBC			√	NS
Plasminogen OS=Homo sapiens GN=PLG PE=1 SV=2	91	P00747	>10	NO	MIBC			√	NS
Serum amyloid A protein OS=Homo sapiens GN=SAA1 PE=1 SV=2	14	P0DJI8	1	NO	MIBC			√	NS

Apolipoprotein E OS=Homo sapiens GN=APOE PE=1 SV=1	36	P02649	10	NO	MIBC		√	MIBC
Clusterin OS=Homo sapiens GN=CLU PE=1 SV=1	52	P10909	5	6	MIBC		√	NS
von Willebrand factor A domain-containing protein 5B1 OS=Homo sapiens GN=VWA5B1 PE=1 SV=2	134	Q5TIE3	>10	5	MIBC		√	NS
Zonadhesin OS=Homo sapiens GN=ZAN PE=2 SV=3	305	Q9Y493	>10	10	MIBC		√	NS
IgGFc-binding protein OS=Homo sapiens GN=FCGBP PE=1 SV=3	572	Q9Y6R7	>10	>10	MIBC		√	NS
Collagen alpha-3(IV) chain OS=Homo sapiens GN=COL4A3 PE=1 SV=3	162	Q01955	>10	1	MIBC		√	NS
Sodium channel protein type 2 subunit alpha OS=Homo sapiens GN=SCN2A PE=1 SV=3	228	Q99250	>10	>10	LG NMIBC	√		NS
Mucin-16 OS=Homo sapiens GN=MUC16 PE=1 SV=2	3007	Q8WXI7	>10	>10	LG/HG NMIBC	√	√	MIBC
Probable cation-transporting ATPase 13A2 OS=Homo sapiens GN=ATP13A2 PE=2 SV=2	129	Q9NQ11	6	2	LG NMIBC	√		Both
Pyridoxal phosphate phosphatase OS=Homo sapiens GN=PDXP PE=1 SV=2	32	Q96GD0	NO	2	LG NMIBC	√		NS
Uncharacterized aarF domain-containing protein kinase 2 OS=Homo sapiens GN=ADCK2 PE=2 SV=1	69	Q7Z695	6	NO	LG NMIBC	√		NMIBC
Down syndrome cell adhesion molecule-like protein 1 OS=Homo sapiens GN=DSCAML1 PE=1 SV=2	224	Q8TD84	>10	>10	LG NMIBC		√	NS
Protein kinase C-binding protein NELL2 OS=Homo sapiens GN=NELL2 PE=1 SV=1	91	Q99435	>10	7	LG NMIBC		√	NS
Protein patched homolog 2 OS=Homo sapiens GN=PTCH2 PE=2 SV=2	130	Q9Y6C5	>10	3	LG NMIBC		√	NS
Zymogen granule protein 16 homolog B OS=Homo sapiens GN=ZG16B PE=1	23	Q96DA0	5	1	LG NMIBC		√	NS

SV=3									
Trypsin-2 OS=Homo sapiens GN=PRSS2 PE=1 SV=1	26	P07478	1	NO	HG NMIBC	√			MIBC
Ig alpha-2 chain C region OS=Homo sapiens GN=IGHA2 PE=1 SV=3	37	P01877	6	5	HG NMIBC	√			No results
Collagen alpha-1(XX) chain OS=Homo sapiens GN=COL20A1 PE=1 SV=4	136	Q9P218	>10	1	HG NMIBC	√			NS
Probable phospholipid-transporting ATPase IIB OS=Homo sapiens GN=ATP9B PE=2 SV=4	129	O43861	10	1	HG NMIBC	√			NS
Epidermal growth factor receptor substrate 15-like 1 OS=Homo sapiens GN=EPS15L1 PE=1 SV=1	94	Q9UBC2	>10	4	HG NMIBC	√			NS
Olfactory receptor 11L1 OS=Homo sapiens GN=OR11L1 PE=2 SV=1	37	Q8NGX0	NO	1	HG NMIBC	√			NS
Laminin subunit alpha-1 OS=Homo sapiens GN=LAMA1 PE=1 SV=2	337	P25391	>10	>10	HG NMIBC	√			NS
Potassium voltage-gated channel subfamily H member 7 OS=Homo sapiens GN=KCNH7 PE=1 SV=2	135	Q9NS40	>10	8	HG NMIBC	√			NS
Vasorin OS=Homo sapiens GN=VASN PE=1 SV=1	72	Q6EMK4	>10	5	HG NMIBC		√		NS
Histidine-rich glycoprotein OS=Homo sapiens GN=HRG PE=1 SV=1	60	P04196	>10	3	HG NMIBC			√	NS
Laminin subunit alpha-5 OS=Homo sapiens GN=LAMA5 PE=1 SV=7	400	O15230	>10	>10	HG NMIBC			√	NMIBC
Sphingosine 1-phosphate receptor 5 OS=Homo sapiens GN=S1PR5 PE=2 SV=1	42	Q9H228	>10	1	HG NMIBC			√	NS
Transmembrane protein 131 OS=Homo sapiens GN=TMEM131 PE=1 SV=2	201	Q92545	>10	>10	HG NMIBC			√	NS

NS: not significantly associated. In Oncomine: threshold: p=0.001, 2-fold, All rank.

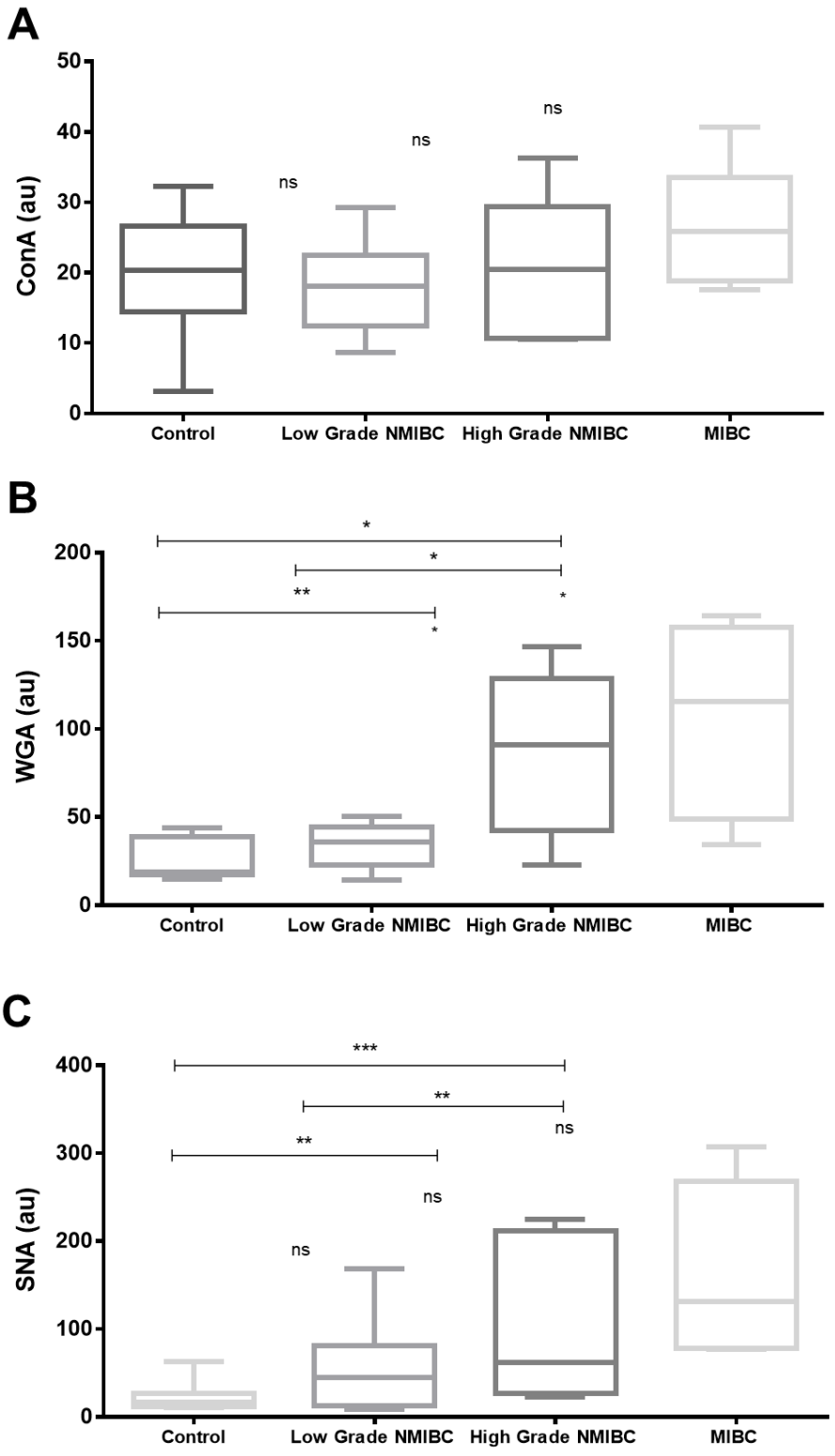

Figure S2. Slot-blot analysis of urine samples using the ConA, WGA and SNA lectins. Slot-blot analysis of individual samples demonstrated that all three lectins presented higher

affinity for proteins isolated from high-grade lesions, in particular in patients facing muscle-invasion, when compared to the controls.

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$