Targeted glycoproteomic identification of cancer cell glycosylation

Alex S Powlesland[3], Paul G Hitchen[3], Simon Parry[2,3], Sarah A Graham[3], Maria Marcela Barrio[4], Maria Teresa Elola[5], José Mordoh[4,5], Anne Dell[3], Kurt Drickamer[3], and Maureen E Taylor[1,3]

[3]Division of Molecular Biosciences, Department of Life Sciences, Imperial College, London SW7 2AZ, UK; [4]Centro de Investigaciones Oncológicas FUCA, Zabala 2836, 1426 Buenos Aires, Argentina; and [5]Fundación Instituto Leloir-IIBBA CONICET, Patricias Argentinas 435, 1405 Buenos Aires, Argentina

GalMBP is a fragment of serum mannose-binding protein that has been modified to create a probe for galactose-containing ligands. Glycan array screening demonstrated that the carbohydrate-recognition domain of GalMBP selectively binds common groups of tumor-associated glycans, including Lewis-type structures and T antigen, suggesting that engineered glycan-binding proteins such as GalMBP represent novel tools for the characterization of glycoproteins bearing tumor-associated glycans. Blotting of cell extracts and membranes from MCF7 breast cancer cells with radiolabeled GalMBP was used to demonstrate that it binds to a selected set of high molecular weight glycoproteins that could be purified from MCF7 cells on an affinity column constructed with GalMBP. Proteomic and glycomic analysis of these glycoproteins by mass spectrometry showed that they are forms of CD98hc that bear glycans displaying heavily fucosylated termini, including Lewis[x] and Lewis[y] structures. The pool of ligands was found to include the target ligands for anti-CD15 antibodies, which are commonly used to detect Lewis[x] antigen on tumors, and for the endothelial scavenger receptor C-type lectin, which may be involved in tumor metastasis through interactions with this antigen. A survey of additional breast cancer cell lines reveals that there is wide variation in the types of glycosylation that lead to binding of GalMBP. Higher levels of binding are associated either with the presence of outer-arm fucosylated structures carried on a variety of different cell surface glycoproteins or with the presence of high levels of the mucin MUC1 bearing T antigen.

Keywords: breast cancer cells/lectin/proteomics/protein engineering/tumor glycosylation

[1]To whom correspondence should be addressed: Tel: +44-207-594-5281; Fax: +44-207-594-3057; e-mail: m.taylor@imperial.ac.uk
[2]Current address: AgResearch Limited, Ruakura Research Centre, East Street, Private Bag 3123, Hamilton 3240, New Zealand.

Introduction

The normal processes of glycosylation are often severely altered during oncogenic transformation of animal cells (Hakomori 1996). These changes can affect interactions between tumor cell-surface glycans and endogenous lectins, which may determine the metastatic potential of the tumor cell. Several common structural changes occur in tumor glycans, including increases in the level of truncation and branching of structures as well as an increased expression of unusual terminal sequences (Kim and Varki 1997). These three changes often result in increased exposure of terminal galactose residues, such as those found in the cancer-associated T antigen (Gal β1-3 GalNAc) and the Lewis[x] trisaccharide (Galβ1-4(Fucα1-3)GlcNAc). One potential role of galactose ligands in tumor cell metastasis is highlighted by the correlation between increased expression of the galectin family of galactose-binding proteins and the ability of many tumor cell types to metastasize (Takenaka et al. 2004). Galectins are thought to facilitate binding of circulating tumor cells to cells of the vascular endothelium by binding to galactose-containing carbohydrate ligands on tumor cell surfaces (Glinsky 2000).

Monoclonal and polyclonal antibodies have proved to be useful tools in the identification of tumor-specific glycosylation (Dube and Bertozzi 2005). For example, the CD15 epitope, which corresponds to the Lewis[x] trisaccharide, is a useful diagnostic marker for Reed-Sternberg cells of Hodgkin's lymphoma (Dorfman et al. 1986) and an antibody specific to sialyl-Lewis[a] is used in diagnostic tests for progression of colon cancer (Macdonald 1999). Characterization of human cell glycosylation with antibodies has been complemented by structural comparison of glycans on normal and tumor cells and by analysis of changes in the complement of enzymes involved in glycan biosynthesis in tumor cells (Kobata and Amano 2005). However, establishing how changes in glycosylation affect the properties of tumor cells requires a more detailed understanding of how glycans are presented at the cell surface. Specifically, identification of proteins that carry tumor glycans is vital to describing the cellular adhesion and subsequent processes that enable the tumor cell to adhere to the endothelium (Hakomori 1996; Kim and Varki 1997; Takenaka et al. 2004). New tools are needed because it is often difficult to identify protein carriers of glycan epitopes. For individual cell-surface glycoproteins, the nature of the glycans bearing particular epitopes is largely unknown.

Sugar-binding receptors of the C-type lectin family contain structurally related carbohydrate recognition domains (CRDs) but display a diversity of carbohydrate-binding specificities (Weis et al. 1998). Some receptors that contain C-type CRDs, such as the selectins, interact with tumor cells and facilitate tumor dissemination (Kim et al. 1998). Structural knowledge of how C-type CRDs bind selectively to specific sugars has made
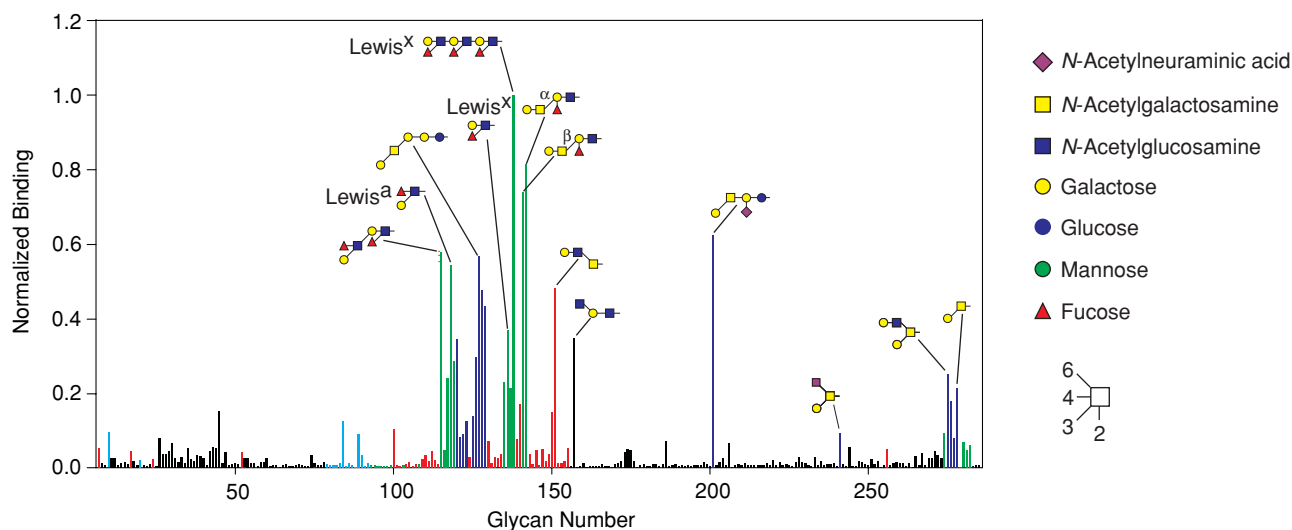
Fig. 1. Screening of a glycan array with GalMBP. Immobilized glycans were probed with fluorescein-labeled GalMBP. Glycans bearing terminal galactose are highlighted in red, except those that also contain fucose, which are colored green, and those linked via a β1,3 linkage to GalNAc, which are colored dark blue. Glycans bearing terminal GalNAc are indicated in light blue. The level of fluorescence was normalized to glycan 138. A full list of glycans present on the array is included in supplementary Table I.

it possible to modify the sugar-binding characteristics of these domains (Iobst and Drickamer 1994), raising the possibility of creating CRDs with novel specificities for tumor glycans.

The results reported in this manuscript demonstrate that GalMBP, an engineered variant of serum mannose-binding protein (MBP-A) that selectively binds to glycans bearing terminal galactose, is an effective tool for identifying and purifying tumor-specific glycans. This glycoproteomic approach has allowed identification of the heavy chain of the cell-surface receptor CD98 as the primary bearer of Lewis[x] and Lewis[y/b] tumor glycans on the MCF7 breast cancer cell line, including a large population of novel, heavily fucosylated, multi-antennary tumor-associated glycans. The ability of tumor cell lines of different origins to bind GalMBP varies widely and reflects heterogeneity in the levels of terminally fucosylated N-linked glycans and mucin MUC1 bearing O-linked glycans with exposed terminal galactose residues.

Results

GalMBP binding to tumor-associated glycans

A trimeric C-terminal fragment of rat MBP-A, consisting of the CRD and an adjacent α-helical region, can be modified to bind to galactose-related sugars by three single amino acid changes and the insertion of a five-amino-acid glycine-rich loop to create GalMBP (mutant QPDWG in Iobst and Drickamer (1994)). In order to assess the ability of this engineered CRD to bind specific sets of galactose-containing ligands, the carbohydrate-binding characteristics of fluorescently labeled GalMBP were examined by probing a glycan array consisting of 285 natural and synthetic oligosaccharides covalently linked to a chemically modified glass slide. The results show that GalMBP selectively binds to glycans bearing terminal galactose residues, with little binding to other glycans, even those containing terminal N-acetylgalactosamine residues (Figure 1). Glycan 157, the

one ligand that does not contain a terminal galactose residue, ends with a galactose that is modified with an α1,6-linked N-acetylglucosamine residue. Thus, GalMBP can tolerate limited substitution of the 6-position of galactose, although the presence of 2–6-linked sialic acid prevents binding.

GalMBP shows preference for two distinct subsets of the 78 glycans bearing terminal galactose residues. One subset consists of glycans in which a terminal α1,3- or α1,4-linked fucose residue is adjacent to a terminal galactose residue, including 8 of the 20 ligands that give the strongest signals and contain either the Lewis[a] (Galβ1-3(Fucα1-4)GlcNAc) or the Lewis[x] (Galβ1-4(Fucα1-3)GlcNAc) trisaccharides. Little binding is observed to glycans containing terminal fucose in α1,2 linkage (glycans 95–99 and 108). This observation is consistent with the fact that the primary binding site in GalMBP binds galactose and not fucose (Iobst and Drickamer 1994), so although both native MBP and GalMBP can bind the Lewis[x] structure, they must do so in very different ways, with fucose in the primary binding site of MBP and galactose in the primary binding site of GalMBP. Another 8 of the 20 best ligands contain Galβ1-3GalNAc. A requirement for GalNAc as the subterminal residue in this disaccharide is indicated by the lack of binding to glycans containing Galβ1-3GlcNAc (glycans 131–134) or Galβ1-3Gal (glycan 130). The disaccharide Galβ1-3GalNAc is the core 1 structure of mucin-type O-linked glycans, which in an unmodified form is commonly referred to as the cancer-associated Thomsen–Friedenreich antigen (T antigen). The ability of GalMBP to bind to the T antigen is also indicated by its binding to glycan 278, where Galβ1-3GalNAc is presented on a threonine linker.

Detection of GalMBP ligands on MCF7 cell surfaces

The selectivity of GalMBP for structures such as the T antigen and Lewis[x] trisaccharide, which are expressed at increased levels in a variety of cancerous tissues such as carcinomas, indicated that the modifications made to create GalMBP might
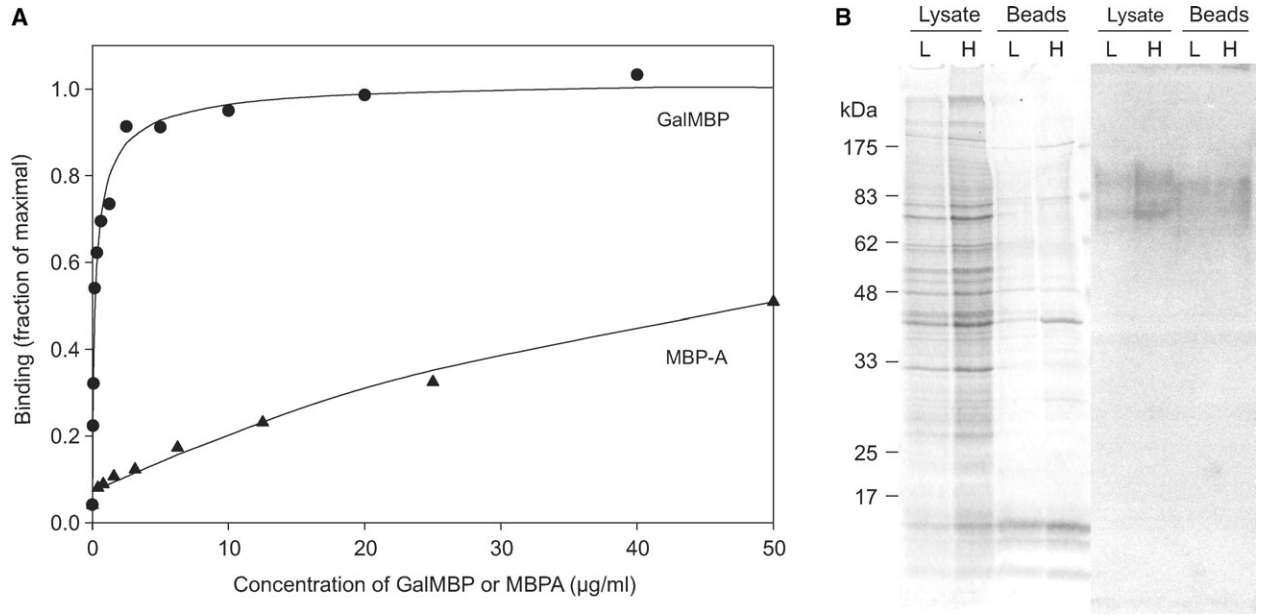
Fig. 2. GalMBP binding to MCF7 cells and glycoproteins. (**A**) Binding of GalMBP or wild-type MBP-A to MCF7 cells was detected using a rabbit polyclonal antibody to MBP-A, followed by a goat anti-rabbit secondary antibody conjugated to horseradish peroxidase. Experimental data (symbols) are shown together with curves fitted to the data using a direct binding equation. The concentration of protein required for 50% of maximal binding was approximately 0.1 μg/mL for GalMBP and 20 μg/mL for wild-type MBP-A. (**B**) Total cell lysates and biotinylated cell-surface proteins purified on avidin-conjugated beads were analyzed by SDS–polyacrylamide gel electrophoresis followed by Coomassie blue staining or blotting onto nitrocellulose followed by probing with ^{125}I-GalMBP. The experiment was performed with low (L) and high (H) densities of MCF7 cells.

target it selectively to cancer cells. Binding studies with the human breast cancer cell line MCF7, which is known to express Lewis[x] (Elola et al. 2007), showed that GalMBP binds to MCF7 cells with 200-fold greater affinity than wild-type MBP-A does, confirming that GalMBP is able to detect high affinity ligands on these cells (Figure 2A). Specific glycoprotein ligands in the MCF7 cell membranes were detected on nitrocellulose blots probed with radioiodinated GalMBP (Figure 2B). The two major glycoprotein ligands, migrating at approximately 75 kDa and 100 kDa, were also detected when MCF7 cell surface proteins were biotinylated and purified on avidin-conjugated agarose beads. Thus, GalMBP interacts specifically with a restricted set of cell surface glycoproteins that are potential carriers of tumor-specific glycans.

Identification of glycoprotein ligands for GalMBP
Columns of immobilized GalMBP were used to enrich GalMBP ligands in detergent extracts of MCF7 cells for further characterization. SDS–polyacrylamide gels of ligands bound to GalMBP and released with EDTA revealed that the major Coomassie blue-stained proteins retained on the column correspond to the two bands of 75 kDa and 100 kDa identified in the blotting experiments (Figure 3A). Probing with ^{125}I-GalMBP confirmed that these bands are ligands for GalMBP, while the absence of significant binding of ^{125}I-GalMBP to proteins in the flow-through and wash fractions demonstrated that the majority of ligand was bound to the column. When the elution fractions were pooled and repurified on the GalMBP column, which was washed with high and low salt buffers before ligands were again eluted with EDTA, only two Coomassie-blue stained bands that

were not recognized by ^{125}I-GalMBP remained, at 35 kDa and 25 kDa (Figure 3B).

Proteomic analysis of the major bands in the repurified ligand pool by trypsin digestion and mass spectrometry demonstrated that both high-molecular-weight GalMBP ligands comprise the heavy chain of the cell-surface glycoprotein CD98 (CD98hc) (Figure 3 and supplementary Table II). No other proteins were identified across a range of search conditions. CD98hc consists of a 58-kDa polypeptide backbone with four potential *N*-glycosylation sites (Deves and Boyd 2000). The heterogeneity of CD98hc purified from MCF7 cells is presumably due to variation in the number and the size of *N*-linked glycans, with two dominant populations being evident. The 35-kDa band was identified as the large neutral amino-acid transporter LAT1, one of six light chains (CD98lc) that form disulfide-linked dimers with CD98hc (Deves and Boyd 2000). The molecular weight of the LAT1 polypeptide is 55 kDa, but the protein migrates anomalously on SDS gels due to its hydrophobic composition (Deves and Boyd 2000). The 25-kDa band was derived from GalMBP, presumably as a result of limited proteolytic degradation of the column.

CD98 was originally identified as a surface antigen in lymphocytes and subsequently was found to be expressed in most tissues and virtually all tumor cells (Deves and Boyd 2000). CD98hc is required for cell-surface expression of the CD98lc amino acid transporter (Nakamura et al. 1999). High-level expression of CD98hc therefore may be required for MCF7 cells to produce large amounts of amino-acid transporter necessary for sustained growth. At the same time, the CD98hc projects an extensive array of glycans that could mediate adhesion to the vascular endothelium and facilitate invasion of tissues. A
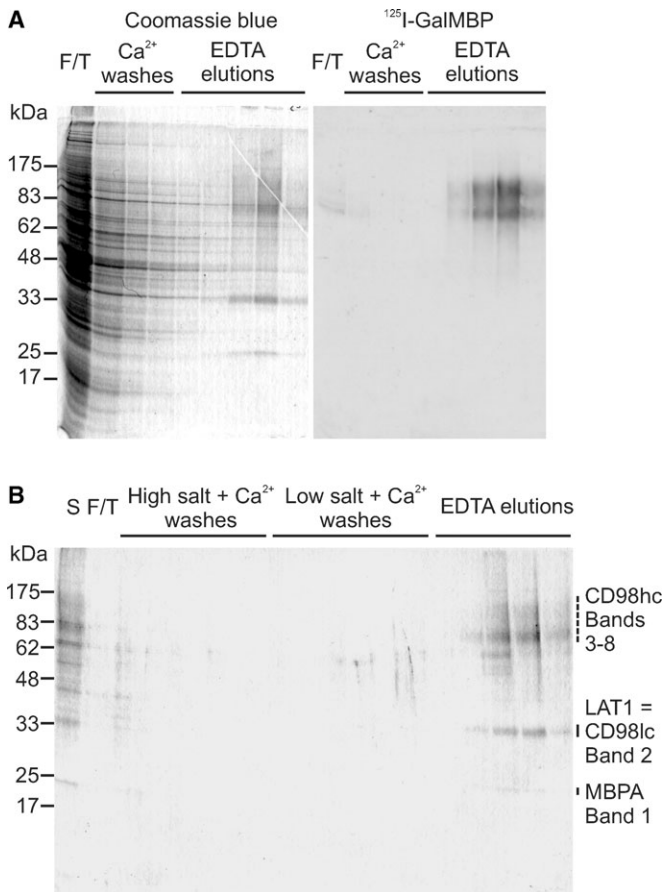
Fig. 3. Purification of GalMBP ligands from MCF7 cell lysates by affinity chromatography on immobilized GalMBP. (**A**) Solubilized MCF7 cells were passed over a column containing immobilized GalMBP, which was washed in the presence of calcium and eluted with EDTA. Fractions were analyzed by SDS–polyacrylamide gel electrophoresis followed by staining with Coomassie blue or blotting onto nitrocellulose followed by probing with ^{125}I-GalMBP. (**B**) GalMBP ligands were repurified on the GalMBP column and analyzed by SDS–polyacrylamide gel electrophoresis followed by staining with Coomassie blue. Bands from the elution fractions were excised and digested with trypsin. Proteins identified by mass spectrometry in each band are indicated.

direct role of CD98hc in tumor-cell activation and adhesion is suggested by the finding that crosslinking of CD98hc with monoclonal antibodies specific to the CD98hc protein or carbohydrate blocks the proliferation of tumor cells (Yagita et al. 1986) and the demonstration that overexpression of CD98hc induces malignant transformation of fibroblasts (Hara et al. 1999).

Glycans on CD98hc

The proteomic experiments indicate that the heavy chain of CD98 is the predominant ligand for GalMBP in MCF7 cells and suggest that this glycoprotein may be the carrier for tumor-specific glycans. The preference of GalMBP for Lewis-type structures, coupled with previous reports of Lewis[x] expression on MCF7 cells (Elola et al. 2007), suggested that glycans on CD98 that are recognized by GalMBP may contain Lewis[x]. This idea was confirmed by matrix-assisted laser desorption-ionization time-of-flight (MALDI-TOF) mass spectrometry analysis of *N*-glycans on CD98hc. The MALDI-TOF

profile shows that the majority of molecular ions correspond to large, complex-type *N*-glycans composed of 5–11 hexose residues, 4–10 *N*-acetylhexosamine residues, and 1–6 fucose residues (Figure 4). MS/MS analysis of the major molecular ions coupled with the knowledge of the biosynthetic pathway of *N*-linked glycosylation was used to predict the structures of the main glycans present (Dell 1994). Many of the glycans identified are tetra-antennary, with bi- and tri-antennary structures also present at lower levels. There is evidence of substantial terminal fucosylation but very limited terminal sialylation and some high mannose oligosaccharides are also present. Sequences predicted to be *O*-glycosylated are not found in CD98hc and *O*-glycosylation of the protein has not been reported, so the most likely ligands for GalMBP would be the *N*-linked glycans.

MS/MS analysis provided evidence that Lewis[a/x], Lewis[y/b], and core fucose are present in all molecular ion peaks containing two or more fucose residues (Figure 5). All of the peaks analyzed produced fragment ions at m/z 660, corresponding to Lewis[a/x]. Diagnostic fragments corresponding to elimination of fucose from the 3-position of GlcNAc indicate that Lewis[x] is the predominant singly fucosylated terminal structure, although some Lewis[a] termini may also be present. For example, the dominant species at m/z 2591 consists of a bi-antennary glycan with both of the antennae terminating in Lewis[x] (Figure 5A). Similarly, a fragment at m/z 834 is present in multiply fucosylated peaks, indicating the presence of Lewis[y/b] (Figure 5A–C). Analysis of larger molecular ions revealed the presence of two- and three-repeat poly-*N*-acetyllactosamine extensions bearing one or two fucose residues, which have the potential to form repeated Lewis[x/a] structures (Figure 5C–D).

Glycans on CD98hc as targets for cell adhesion

The detection of heavily fucosylated structures on CD98hc explains the selective binding of GalMBP to this protein and indicated that it may provide a scaffold for interactions of MCF7 cells with endothelia (Elola et al. 2007). This suggestion was supported by the fact that CD98hc is bound by the scavenger receptor C-type lectin (Figure 6). This endothelial receptor recognizes Lewis[x] selectively and has been proposed to function in adhesion of tumor cells during metastasis (Coombs et al. 2005; Elola et al. 2007). In previous studies, in which whole cell extracts were blotted with scavenger receptor C-type lectin, two lower-molecular-weight bands were detected (Elola et al. 2007). These bands, which appear to be proteolytic fragments of CD98hc, were not observed when more stringent steps were taken to prevent proteolysis during sample preparation.

Survey of GalMBP ligands on breast cancer cell lines

Based on the ability of GalMBP to bind glycans characteristic of tumor cell surfaces, probing of additional breast cancer cell lines was undertaken. Direct binding assays revealed that cells of different origins differ in their capacity for binding GalMBP, with higher affinity binding observed for MCF7 and SKBR3 cells compared to ZR75-1 and T47D cells (Figure 7A and B). As a control, little binding was observed for MDA-MB-231 cells, which are not believed to be of epithelial origin. The differences in binding correlate partially with variations in the levels of Lewis[x] epitopes on glycoproteins from these cells as detected with two monoclonal anti-CD15 antibodies, with highest binding observed for MCF7 and SKBR3 cells (Figure 7C).
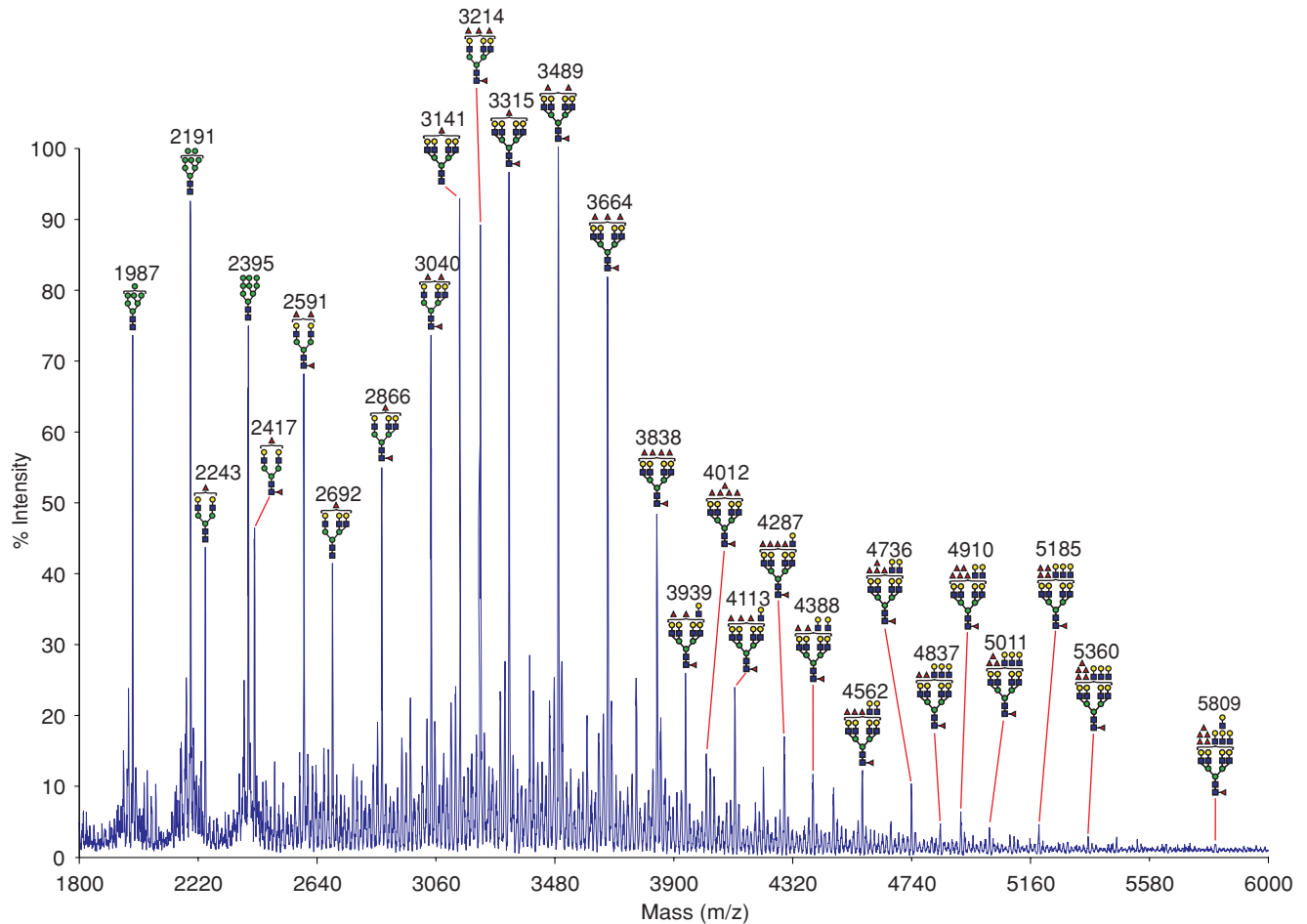
Fig. 4. Mass spectrometric analysis of *N*-glycans from CD98hc. Glycans released from purified GalMBP ligands were permethylated and subjected to MALDI MS profiling. All labeled molecular ions corresponding to complex-type glycans were subjected to MS/MS analysis. The assignment for each ion represents the most likely structure based on the precise fit between the calculated composition and the *m/z* value, taking into account the biosynthetic pathways of *N*-glycosylation in addition to the MS and MS/MS data. Values (*m/z*) listed correspond to ^{12}C throughout. Satellite peaks surrounding major signals are either artifacts of the permethylation process or are minor glycans with one or more sialylated antennae. The symbols are defined in Figure 1.

Analysis of the *N*-linked glycans purified on GalMBP from the additional cell lines confirmed that SKBR3 cells display multiple glycans with outer arm fucosylation, including tri- and tetra-antennary structures, although the degree of fucosylation is reduced compared to the MCF7 cells and glycans from ZR75-1 and T47D cells bear even fewer such structures (supplementary Figure 1).

The relatively paucity of CD15 antigens and related epitopes on the glycoproteins from ZR75-1 and T47D cells indicated that there might be other GalMBP ligands on these cells that account for some of the binding observed. The ability of GalMBP to interact with the T antigen suggested that *O*-linked glycans displayed on cell surface mucins such as MUC1 could be additional targets (Burchell et al. 2001). The presence of such mucins would not easily be detected by mass spectrometry because few tryptic peptides can be released from the core polypeptide of the mucin, so the presence of MUC1 was demonstrated by using a monoclonal antibody to probe gels of the GalMBP ligands (Figure 7D). The blot reveals the presence of MUC1 in the material purified from MCF7, SKBR3, and ZR75-1 cells.

Mass spectrometry of the *O*-linked glycans on glycoproteins purified from the cells lines confirmed the presence of T antigen as the predominant *O*-linked glycan from all but the T47D cells, in which the T antigen is mostly sialylated (supplementary Figure 2).

These results demonstrate that the relative contributions of fucosylated *N*-linked structures and *O*-linked T antigen vary between different breast cancer cell lines. In addition to differences in the relative contributions of these two classes of glycans, the different molecular weights of glycoproteins bearing CD15 antigens on the different cell lines suggested that these are carried on different protein scaffolds. This conclusion was confirmed by proteomic analysis, which demonstrated extensive variability in the glycoproteins purified on GalMBP (supplementary Table III). Although CD98hc, the predominant carrier on MCF7 cells, is also detected in the extracts from SKBR3 and T47D cells, it is not the predominant protein present in these cases and it is not detected in the case of the ZR75-1 cells. These results suggest that the extent of tumor-specific glycosylation in different cell lines may depend on the abundance of different
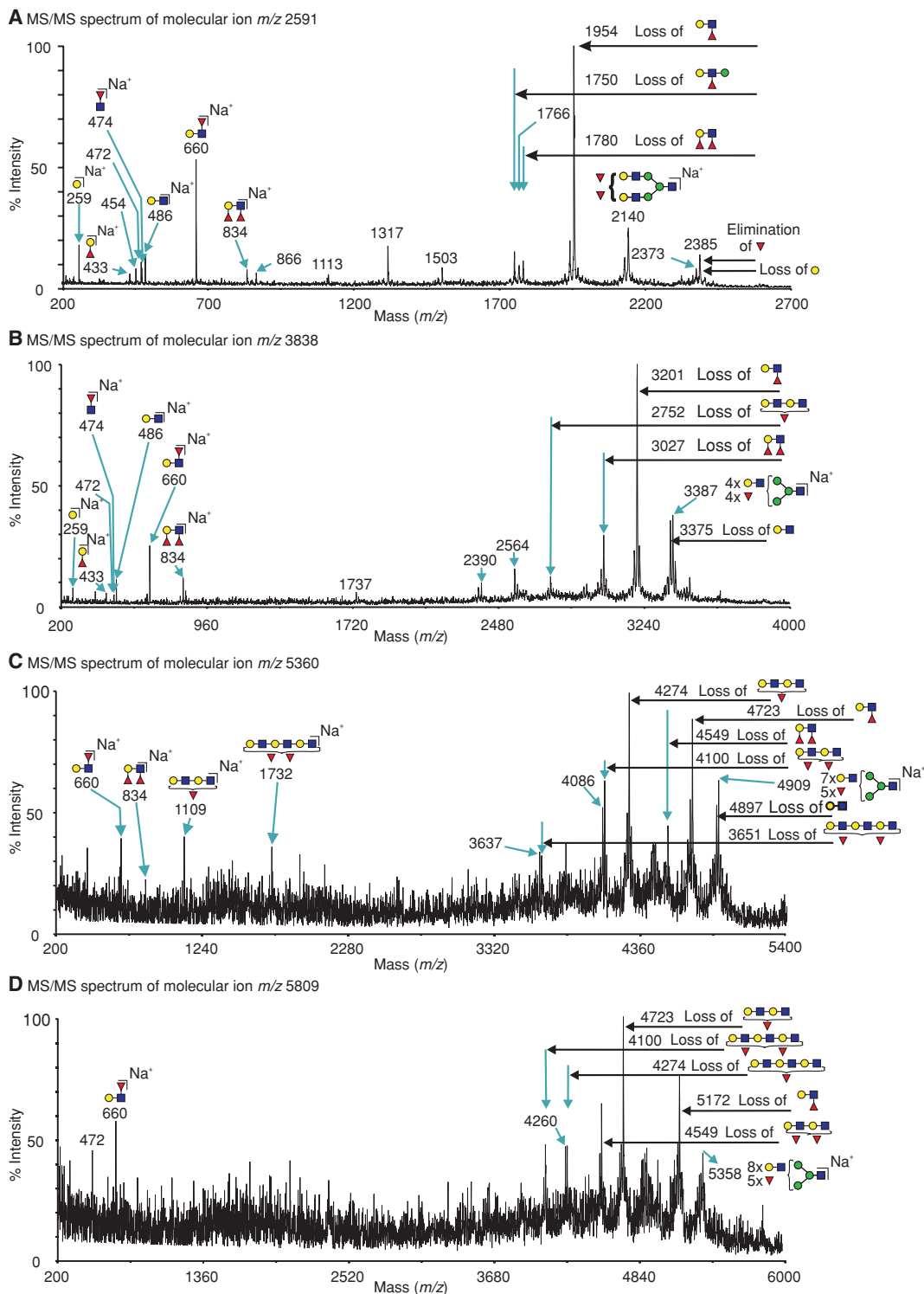
Fig. 5. MALDI-TOF/TOF spectra of permethylated *N*-glycans from CD98hc. Assignments of the fragment ions are shown either as loss of glycan moieties from the molecular ion (see annotations on horizontal arrows) or as antenna fragments which are B ions for cleavage at HexNAc and C ions for cleavage at hexoses (Domon and Costello 1988). Assigned masses correspond to the ^{12}C isotope. The annotated ions not assigned in the schematics correspond to double or triple cleavages. Symbols are defined in Figure 1. (**A**) Fragmentation of the *m/z* 2591 molecular ion (Fuc$_3$Hex$_5$HexNAc$_4$) produced ions at *m/z* 660 and 1954, corresponding to cleaving off of terminal Fuc-(Hex-HexNAc). The presence of terminal Lewis^x is demonstrated by the diagnostic ion at *m/z* 2385, resulting from elimination of fucose attached at the 3 position of GlcNAc. (**B**) Fragments from the molecular ion at *m/z* 3838 (Fuc$_5$Hex$_7$HexNAc$_6$) include ions at *m/z* 660, corresponding to Lewis$^{x/a}$, *m/z* 834, corresponding to Lewis$^{y/b}$, and *m/z* 474 corresponding to core fucosylation. (**C**) Dominant fragments from the ion at *m/z* 5360 (Fuc$_6$Hex$_{10}$HexNAc$_9$), at *m/z* 1109, and at *m/z* 4274 indicate the presence of a two-repeat poly *N*-acetyllactosamine bearing one fucose residue. (**D**) Fragmentation of the ion at *m/z* 5809 (Fuc$_6$Hex$_{11}$HexNAc$_{10}$) reveals the presence of three-repeat structures bearing one or two fucose residues at *m/z* 4100 and 4274.
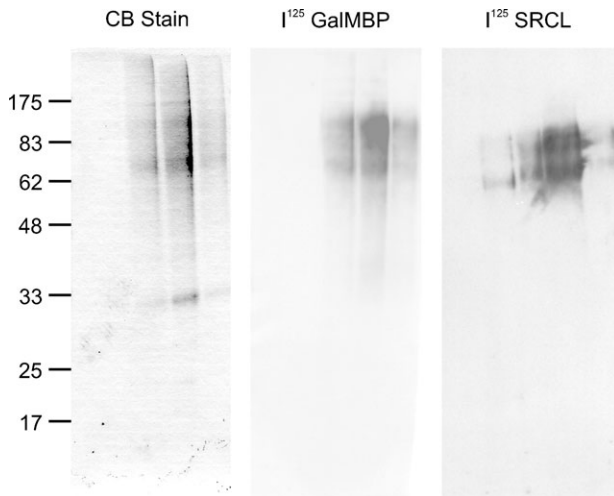
Fig. 6. Demonstration of GalMBP ligands from MCF7 cells binding to the scavenger receptor C-type lectin. Purified MCF7 ligands were analyzed by SDS–polyacrylamide gel electrophoresis. Parallel lanes were stained with Coomassie blue or blotted onto nitrocellulose membranes and probed with ^{125}I-GalMBP or ^{125}I-scavenger receptor C-type lectin (SRCL).

target proteins as well changes in the expression of components of the glycosylation machinery of the cell.

Discussion

The results described here demonstrate that breast cancer cells vary widely in the nature of their glycosylation in spite of their similar origins and phenotypes, with two classes of tumor-associated glycosylation present: branched, outer arm-fucosylated *N*-linked glycans, and *O*-linked glycans bearing the T antigen. However, the most abundant protein carriers of the *N*-linked glycans vary dramatically between different cell lines, with MCF7 cells being unusual in that a single glycoprotein is the predominant carrier of the heavily fucosylated structures. These results emphasize the importance of looking at more than one breast cancer cell line when examining glycosylation differences. They also provide a molecular explanation of variability in the response of breast cancers to treatments with antibodies directed against carbohydrate epitopes (Mordoh et al. 1995) and they suggest that screening may help to identify tumors more likely to respond to treatment with anti-CD15 antibodies. In addition to providing an important tool for such analysis, GalMBP can be used as a surrogate for endogenous receptors, such as the endothelial scavenger receptor C-type lectin, that may interact with tumor cell glycans during metastasis but which are much more difficult to obtain in quantities sufficient for use as affinity reagents.

The fact that GalMBP binds two subsets of galactose-terminating glycans that are commonly present on tumor cells makes it a powerful tool for identifying and analyzing glycoproteins present on tumor cell lines. Purification of glycoproteins by affinity chromatography on GalMBP requires minimal sample preparation prior to affinity chromatography and the high yields, coupled with the sensitivity of mass spectrometry, allowed characterization of both protein and glycan components without requiring radiolabeling or immunoprecipitation. The bacterial expression system for the C-terminal fragment of GalMBP allows easy production of large quantities of trimeric protein with

clusters of binding sites and will facilitate creation of further variants with modified selectivity for different sets of glycans.

In a recent study, isolation of ligands for wild-type MBP from SW1116 human colon tumor cells by affinity chromatography revealed that unmodified rabbit MBP binds to heavily fucosylated glycans, including those present on CD26 (Kawasaki et al. 2009). The recognition mechanism of this wild-type protein is different from GalMBP because fucose would be the principal sugar interacting with the binding site, but the best ligands bear multiple fucosylated *N*-acetyllactosamine repeats and thus would overlap with those bound by GalMBP. It is interesting that this approach leads to identification of a pattern of glycoprotein ligands different to any of the breast cancer cell lines, with CD26 and CD98hc identified as the major carriers of glycans that react with wild-type MBP.

The glycan array profile of GalMBP reveals a preference for galactose-terminating glycans that either bear terminal fucose residues or terminate in galactose-linked β1-3 to GalNAc. GalMBP makes primary interactions with terminal galactose residues (Iobst and Drickamer 1994; Kolatkar and Weis 1996). The glycan array results indicate that this binding is enhanced through secondary interactions with an adjacent terminal fucose residue. The preference for terminal fucose shows some similarity to that of the macrophage galactose receptor (Coombs et al. 2006), but the selectivity of GalMBP for Galβ1-3GalNAc has not been observed in other galactose-binding receptors. The glycan array results demonstrate that, in the context of glycans, GalMBP binds preferentially to galactose-terminated ligands, which is distinct from the preferential binding to GalNAc observed for the asialoglycoprotein receptor and the roughly equal binding of galactose- and GalNAc-containing ligands to the macrophage galactose receptor (Coombs et al. 2006). These results indicate that further targeting to novel sets of glycans may be achieved by additional engineering of the binding site, employing the large body of information on the structural basis for ligand recognition by various C-type lectins (Drickamer 1999; Feinberg et al. 2007) in conjunction with rapid characterization of the specificities developed using the glycan array.

In the presence of complement, full-length GalMBP binds to and lyses erythrocytes pretreated with neuraminidase to expose terminal galactose in the form of the Thomsen–Friedenreich antigen (T antigen) on red blood cell glycans (Bray et al. 1981) (unpublished observations). The ability of GalMBP to lyse neuraminidase-treated red blood cells highlights a potential therapeutic role for the protein either through complement-dependent lysis, by complement-mediated phagocytosis by leukocytes, or by a complement-independent mechanism that has been referred to as MBP-dependent cell cytotoxicity (Ma et al. 1999).

Material and methods

Preparation of proteins

The C-terminal fragment of GalMBP was expressed by inducing bacteria in the presence of Ca^{2+} as described for the scavenger receptor C-type lectin (Coombs et al. 2005). Protein released by sonication was purified on galactose-Sepharose by the method used previously (Iobst and Drickamer 1994). Full-length wild-type MBP-A was prepared following the published procedure
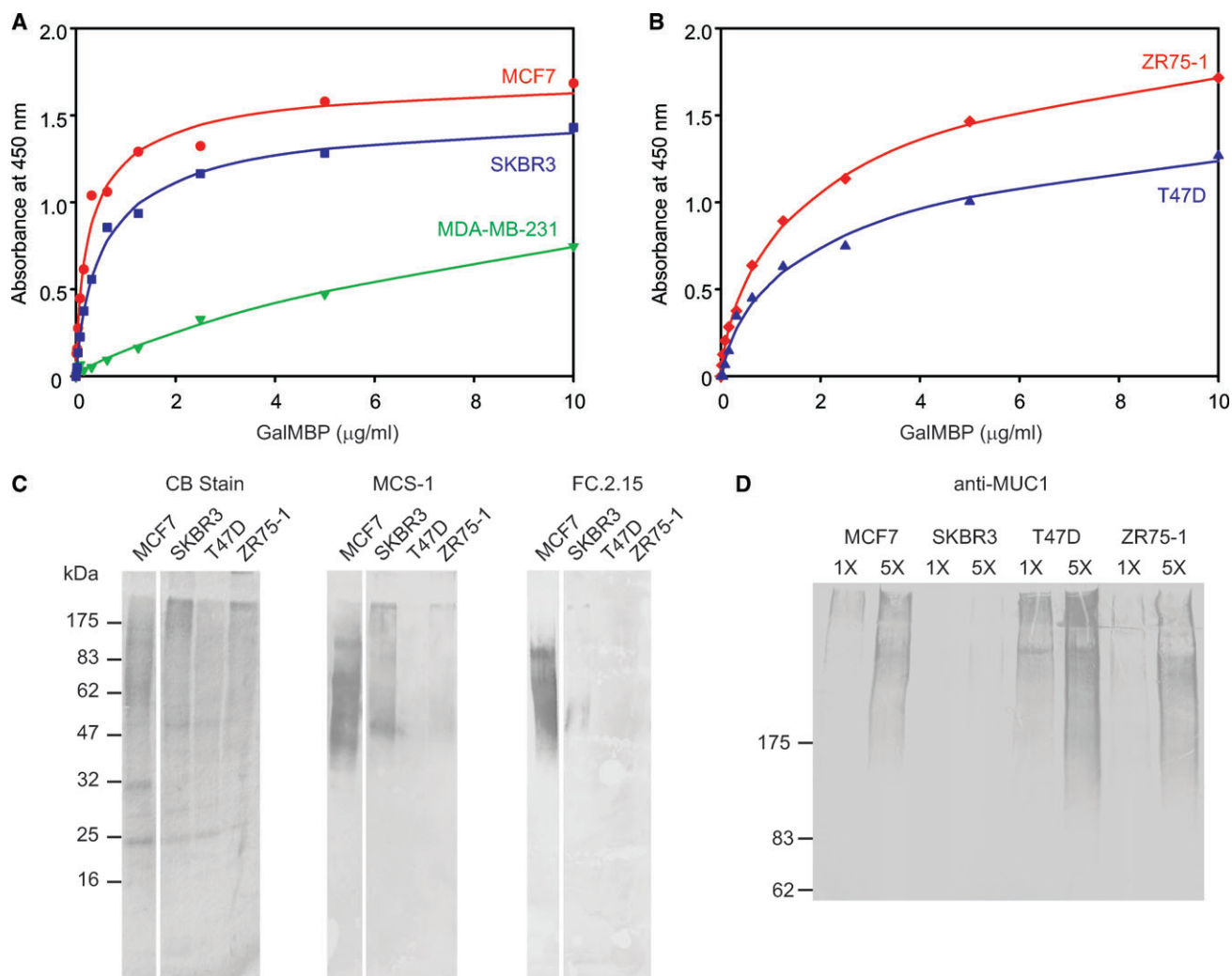
Fig. 7. Binding of GalMBP to a panel of breast cancer cell lines and purification of glycoprotein ligands. (**A** and **B**) Binding of GalMBP was assayed as described in Figure 2A. (**C** and **D**) Gel electrophoresis of GalMBP ligands from multiple tumor cell lines purified by affinity chromatography on immobilized GalMBP. (**C**) Gels (17.5%) were stained with Coomassie blue or blotted onto nitrocellulose and probed with monoclonal anti-Lewis[x] antibodies followed by the peroxidase-conjugated goat anti-mouse secondary antibody. (**D**) Gels (5%) were blotted on nitrocellulose and probed with the monoclonal anti-MUC1 antibody followed by the peroxidase-conjugated goat anti-mouse secondary antibody.

(Wallis and Drickamer 1999). A version of GalMBP corresponding to full-length MBP-A was expressed in Chinese hamster ovary cells following the same strategy used for wild-type MBP-A (Wallis and Drickamer 1997), except that the protein was purified on galactose-Sepharose. The truncated form of GalMBP was labeled with fluorescein isothiocyanate in the bicine buffer (Powlesland et al. 2006) or with [^{125}I]Bolton-Hunter reagent (Coombs et al. 2005). The glycan array developed by the Consortium for Functional Glycomics was screened using the standard buffer conditions (www.functionalglycomics.org). Scavenger receptor C-type lectin was prepared and labeled as previously described (Coombs et al. 2005).

Binding of GalMBP to cells

MCF7 cells were grown as previously described (Elola et al. 2007). Cell lines MDA-MB-231, T47D, and ZR-75-1 were obtained from the European Collection of Cell Cultures, and SKBR3 cells were purchased from the American Type Culture Collection. Cells were grown in the media recommended by the suppliers. Confluent cells in 96-well plates were washed twice with 250 µL of medium without phenol red, incubated with wild-type MBP-A or GalMBP in medium for 30 min at room temperature, washed four times with medium, incubated for 60 min at room temperature with a 1:800 dilution of anti-MBP-A antibody in medium containing 5% bovine serum albumin, washed again four times with medium, incubated for 60 min at room temperature with a 1:800 dilution of goat anti-rabbit antibody conjugated to horseradish peroxidase (30 µg/mL) (MP Biomedicals, Illkirch, France), and washed four times with medium and incubated with 100 µL of peroxidase substrate (Pharmingen, BD Biosciences, San Jose, CA) for 10 min at room temperature in the dark. Reactions were terminated with 50 µL of 2 M HCl and were centrifuged at $250 \times g$ for 5 min. Supernatants were transferred to a 96-well ELISA plate and the absorbance at 450 nm was read using a Victor^3 plate reader from Perkin Elmer (Waltham, MA).

Cell surface biotinylation

MCF7 cells grown to 80% and 40% confluence in 100 mm plates were washed twice with PBS supplemented with 0.1 mM $CaCl_2$ and 1 mM $MgCl_2$ and incubated with gentle shaking for 20 min at 4°C in supplemented PBS containing 1 mg/mL sulfo-NHS-biotin (Perbio Science, Cramlington, UK). Cells were washed twice and incubated with gentle shaking for a further 40 min at 4°C with supplemented PBS containing 100 mM glycine, and washed twice with supplemented PBS and harvested by scraping into supplemented PBS and centrifugation at $450 \times g$ for 2 min. Each pellet was dissolved in the biotinylation lysis buffer (150 mM NaCl, 100 mM Tris-Cl, pH 7.4, 1 mM EDTA, 1% Triton X-100) containing protease inhibitors, sonicated for 5 s, and incubated at 4°C for 10 min with end-over-end mixing. Solubilized pellets were centrifuged at $14,000 \times g$ for 5 min and portions of the supernatants were reserved for analysis of total protein content. The remaining supernatants were incubated with a 50% suspension of avidin-conjugated agarose beads (Perbio Science) for 60 min at 4°C with end-over-end mixing. Beads were pelleted by centrifugation at $14,000 \times g$ for 5 min. The pellets were washed four times with the biotinylation lysis buffer, and bound proteins were eluted by boiling at 100°C for 5 min in the SDS–polyacrylamide gel sample buffer containing 2-mercaptoethanol.

Affinity purification of ligands on immobilized GalMBP

The C-terminal fragment of GalMBP, 5 mg in 2 mL of 150 mM NaCl, 100 mM HEPES, pH 8.0, 50 mM $CaCl_2$, was coupled to 2 mL of Affigel 10 (BioRad Laboratories, Hercules, CA) at 4°C for 4 h following the manufacturer's directions and equilibrated in the loading buffer (150 mM NaCl, 25 mM Tris-Cl, pH 7.8, 25 mM $CaCl_2$). Cells (2×225 cm^2 flasks) were grown to confluence, washed in PBS, harvested by scraping into PBS followed by centrifugation at $450 \times g$ for 2 min, resuspended in the cell lysis buffer (150 mM NaCl, 25 mM Tris-Cl, pH 7.8, 2 mM $CaCl_2$) containing 1% Triton X-100 and protease inhibitors, sonicated for 10 s and incubated on ice for 30 min. Following centrifugation at $3500 \times g$ for 5 min, lysate was passed over the GalMBP affinity column, which was washed with the cell lysis buffer containing 0.1% Triton X-100. Bound proteins were eluted with the elution buffer (150 mM NaCl, 25 mM Tris-Cl, pH 7.8, 2.5 mM EDTA) containing 0.1% Triton X-100. For repurification, elution fractions were pooled, $CaCl_2$ was added to 25 mM, and the protein was again passed over the GalMBP affinity column. The column was washed with the high salt loading buffer, in which the NaCl concentration was increased to 1.25 M, and with the loading buffer, both containing 0.1% Triton X-100, and ligands were eluted in the elution buffer containing 0.1% Triton X-100.

Western blot analysis of GalMBP ligands

Samples dissolved in the sample buffer containing 1% 2-mercaptoethanol and boiled at 100°C for 5 min were analyzed by SDS–polyacrylamide gel electrophoresis and electroblotted onto Protran nitrocellulose membranes (Whatman, Maidstone, UK) (Burnette 1981). For probing with radiolabeled receptors, blots were blocked overnight at 4°C with 2% hemoglobin in the loading buffer, incubated with ^{125}I-GalMBP or ^{125}I-scavenger receptor C-type lectin for 2 h at room temperature with gentle shaking and washed four times with Tris-

buffered saline. Radioactivity was detected using a Phosphorimager SI from Molecular Dynamics (Sunnyvale, CA). Probing with anti-Lewis^x antibodies was performed as previously described (Elola et al. 2007). Blots were probed for MUC1 by incubating overnight with hybridoma cell culture supernatant containing monoclonal antibody HMFG2, provided by Joy Burchell and Joyce Taylor-Papadimitriou (Kings College London), followed by alkaline phosphatase-conjugated goat anti-mouse IgG from Jackson ImmunoResearch (West Grove, PA).

Analysis of purified ligands by mass spectrometry

Solvents, acetic acid, and trifluoroacetic acid were purchased from Romil (Cambridge, UK). Proteins present in elution fractions from the GalMBP-agarose column were precipitated in 10% trichloroacetic acid and analyzed by SDS–polyacrylamide gels. For proteomic analysis (Abu-Qarn et al. 2007), bands were excised, stain was removed by incubation in acetonitrile, and gel pieces dried in a vacuum centrifuge. Proteins present were reduced with 10 mM dithiothreitol and alkylated with 55 mM iodoacetic acid, prior to digestion with sequencing grade-modified trypsin (E.C.3.4.21.4 Promega, Madison, WI). Peptides were extracted from gel pieces by incubation with 0.1% trifluoroacetic acid for 10 min, followed by the addition of acetonitrile for 15 min. The extraction process was repeated once. Peptides were reduced in volume using a vacuum centrifuge and spotted onto a MALDI stainless steel plate with α-cyano-4 hydroxycinnamic acid used as a matrix.

Peptides were subjected to MALDI MS profiling, complemented with MS/MS sequencing of the 10 most abundant ions in each sample, on an Applied Biosystems 4800 MALDI-TOF/TOF mass spectrometer. For peptide mass fingerprint data, peak picking was conducted using GPS Explorer software version 3.6 (Applied Biosystems, Foster City, CA). A signal-to-noise threshold of 10 was used, sequazyme peptide mass standards were used as external standards for calibration and no contaminant ions were excluded. For MS/MS experiments, peak list generation and database searching were conducted using GPS Explorer software version 3.6 (Applied Biosystems) with the default parameters. Both MS and MS/MS data were used to search 283,454 entries in release 54.2 of the SwissProt database with version 2.2 of the Mascot database search algorithm (www.matrixscience.com) with the following parameters: peptide masses were fixed as monoisotopic, partial oxidation of methionine residues was considered, partial carboxymethylation of cysteine residues was considered, the mass tolerance was set at 75 ppm for precursor ions and 0.1 Da for fragment ions, and tryptic digests were assumed to have no more than one missed cleavage. Peptide matches from both MS and MS/MS data were used to generate probability-based Mowse protein scores. Scores greater than 55 were considered significant ($P < 0.05$) (Perkins et al. 1999). In order to broaden the search for other possible proteins in the GalMBP-binding fractions from MCF7 cells, the mass tolerance for precursor ions was increased to 100 ppm, taxonomy was broadened to include all species, and the possibility of two missed cleavages was allowed.

For glycan analysis, proteins eluted from GalMBP were precipitated trichloroacetic acid and subjected to glycan release and permethylation as previously described (Sutton-Smith

and Dell 2005). Samples were redissolved in 10 µL of methanol:water (4:1), mixed with an equal volume of 10 mg/mL 2,5-dihydroxybenzoic acid matrix and subjected to MALDI MS profiling, with selected ions further analyzed by MS/MS. The MALDI profile of glycans in the major elution fraction from the reverse phase cartridge is shown. Both MALDI MS and MS/MS data were acquired on an Applied Biosystems 4800 MALDI-TOF-TOF mass spectrometer.

Analysis of GalMBP ligands by liquid chromatography and tandem mass spectrometry

GalMBP ligands were precipitated with trichloroacetic acid, lyophilized, dissolved in 25 µL of 8 M urea containing 10 mM HEPES, pH 8.0, reduced with dithiothreitol, alkylated with iodoacetic acid, and digested in solution with sequencing grade modified trypsin (Promega) (Foster and Mann 2005). Tryptic peptides were dried on a vacuum centrifuge, resuspended in 80 µL 0.1% trifluoroacetic acid, desalted, and purified by nano-liquid chromatography using an Ultimate 3000 (LC Packings, Dionex) fitted with a Pepmap analytical C-18 nanocapillary. After loading in 2% acetonitrile containing 0.1% trifluoroacetic acid, the column was eluted with a gradient from 8–45% acetonitrile over 60 min at a flow rate of 0.3 µL/min. Samples were spotted directly onto a MALDI steel target plate using a Probot (LC Packings, Dionex, Sunnyvale, CA) with α-cyano-4 hydroxycinnamic acid as a matrix. Samples were subjected to MALDI-MS profiling with major peaks selected for collision-induced dissociation and sequencing by tandem mass spectrometry.

Fragment ions detected in the MS/MS spectra was used to search the SwissProt database with the Mascot database search algorithm (www.matrixscience.com) for peptide sequences consistent with the ion fragments observed. For each search, the parameters used were: peptide masses fixed as monoisotopic, partial oxidation of methionine residues considered, partial carboxymethylation of cysteine residues considered, mass tolerance 100 ppm, fragment ion tolerance 0.3 Da, tryptic digests assumed to have no more than one missed cleavage, and taxonomy group searched *Homo sapiens*. Probability-based Mowse scores were generated for each peptide sequence identified and only values of greater than 95% significance were considered (Perkins et al. 1999). Positive identification of a protein required identification of three or more unique peptide sequences identified with ions scores of greater than 95% significance. CD98hc was screened for potential *O*-glycosylation sites using the NetOGlyc 3.1 server (www.cbs.dtu.dk/services/NetOGlyc).

Supplementary data

Supplementary data for this article are available online at http://glycob.oxfordjournals.org/.

Funding

Acknowledgements

Conflict of interest statement

None declared.

Abbreviations

CRD, carbohydrate recognition domain; GalMBP, galactose-binding mutant of MBP-A; MBP, mannose-binding protein; MALDI, matrix-assisted laser desorption-ionization; MS, mass spectrometry; PBS, phosphate-buffered saline; TOF, time-of-flight.

References

Abu-Qarn M, Yurist-Doutsch S, Giordano A, Trauner A, Morris HR, Hitchen P, Medalia O, Dell A, Eichler J. 2007. *Haloferax volcanii* AglB and AglD are involved in *N*-glycosylation of the *S*-layer glycoprotein and proper assembly of the surface layer. *J Mol Biol.* 374:1224–1236.

Bray J, Lemieux RU, McPherson TA. 1981. Use of a synthetic hapten in the demonstration of the Thomsen-Friedenreich (T) antigen on neuraminidase-treated human red blood cells and lymphocytes. *J Immunol.* 126:1966–1969.

Burchell JM, Mungul A, Taylor-Papadimitiriou J. 2001. *O*-Linked glycosylation in the mammary gland: Changes that occur during malignancy. *J Mammary Gland Biol Neoplasia.* 6:355–364.

Burnette WN. 1981. "Western blotting": Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal Biochem.* 112:195–203.

Coombs PJ, Graham SA, Drickamer K, Taylor ME. 2005. Selective binding of the scavenger receptor C-type lectin to Lewis[x] trisaccharide and related glycan ligands. *J Biol Chem.* 280:22993–22999.

Coombs PJ, Taylor ME, Drickamer K. 2006. Two categories of mammalian galactose-binding receptors distinguished by glycan array profiling. *Glycobiology.* 16:1C–7C.

Dell A. 1994. Mass spectrometry of carbohydrate-containing biopolymers. *Methods Enzymol.* 230:108–132.

Deves R, Boyd CA. 2000. Surface antigen CD98(4F2): Not a single membrane protein, but a family of proteins with multiple functions. *J Membr Biol.* 173:165–177.

Domon B, Costello CE. 1988. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate J.* 5:397–409.

Dorfman RF, Gatter KC, Pulford KA, Mason DY. 1986. An evaluation of the utility of anti-granulocyte and anti-leukocyte monoclonal antibodies in the diagnosis of Hodgkin's disease. *Am J Pathol.* 123:508–519.

Drickamer K. 1999. C-Type lectin-like domains. *Cur Opin Struct Biol.* 9:585–590.

Dube DH, Bertozzi CR. 2005. Glycans in cancer and inflammation: Potential for therapeutics and diagnostics. *Nat Rev Drug Discov.* 4:477–488.

Elola MT, Capurro MI, Barrio MM, Coombs PJ, Taylor ME, Drickamer K, Mordoh J. 2007. Lewis x antigen mediates adhesion of human breast carcinoma cells to activated endothelium: Possible involvement of the endothelial scavenger receptor C-type lectin. *Breast Cancer Res Treat.* 101:161–174.

Feinberg H, Taylor ME, Weis WI. 2007. Scavenger receptor C-type lectin binds to the leukocyte cell surface glycan Lewis x by a novel mechanism. *J Biol Chem.* 282:17250–17258.

Foster LJ, Mann M. 2005. Protein identification and sequencing by mass spectrometry. In: Celis JE, editor. *Cell Biology: A Laboratory Handbook.* San Diego: Academic Press. p. 415–425.

Glinsky VV, Huflejt ME, Glinsky GV, Deutscher SL, Quinn TP. 2000. Effects of Thomsen-Friedenreich antigen-specific peptide P-30 on β-galactoside-mediated homotypic aggregation and adhesion to the endothelium of MDA-MB-435 human breast carcinoma cells. *Cancer Res.* 60:2584–2588.

Hakomori S. 1996. Tumor malignancy defined by aberrant glycosylation and sphingo(glyco)lipid metabolism. *Cancer Res.* 56:5309–5318.

Hara K, Kudoh H, Enomoto T, Hashimoto Y, Masuko T. 1999. Malignant transformation of NIH3T3 cells by overexpression of early lymphocyte activation antigen CD98. *Biochem Biophys Res Commun.* 262:720–725.

Iobst ST, Drickamer K. 1994. Binding of sugar ligands to Ca^{2+}-dependent animal lectins: II. Generation of high affinity galactose binding by site-directed mutagenesis. *J Biol Chem.* 269:15512–15519.

Kawasaki N, Lin C-W, Inoue R, Khoo K-H, Kawasaki N, Ma BY, Oka S, Ishiguro M, Sawada T, Ishida H et al. 2009. Highly fucosylated *N*-glycan ligands for mannan-binding protein expressed specifically on CD26 (DPPIV) isolated from a human colorectal carcinoma cell line, SW1116. *Glycobiology* 19:437–450.

Kim YJ, Borsig L, Varki NM, Varki A. 1998. P-Selectin deficiency attenuates tumor growth and metastasis. *Proc Natl Acad Sci USA.* 95:9325–9330.

Kim YJ, Varki A. 1997. Perspectives on the significance of altered glycosylation of glycoproteins in cancer. *Glyconjugate J.* 14:569–576.

Kobata A, Amano J. 2005. Altered glycosylation of proteins produced by malignant cells, and application for the diagnosis and immunotherapy of tumours. *Immunol Cell Biol.* 83:429–439.

Kolatkar A, Weis WI. 1996. Structural basis of galactose recognition by C-type animal lectins. *J Biol Chem.* 271:6679–6685.

Ma Y, Uemura K, Oka S, Kozutsumi Y, Kawasaki N, Kawasaki T. 1999. Anti-tumor activity of mann-binding protein *in vivo* as revealed by a virus expression system: Mannan-binding protein-dependent cell-mediated cytotoxicity. *Proc Natl Acad Sci USA.* 96:371–375.

Macdonald JS. 1999. Carcinoembryonic antigen screening: Pros and cons. *Semin Oncol.* 26:556–560.

Mordoh J, Silva C, Albarellos M, Bravo AI, Kairiyama C. 1995. Phase I clinical trial in cancer patients of a new monoclonal antibody FC-2.15 reacting with tumor proliferating cells. 17:161–160.

Nakamura E, Sato M, Yang H, Miyagawa F, Harasaki M, Tomita K, Matsuoka S, Noma A, Iwai K, Minato N. 1999. 4F2 (CD98) heavy chain is associated covalently with an amino acid transporter and controls intracellular trafficking and membrane topology of 4F2 heterodimer. *J Biol Chem.* 274:3009–3016.

Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis.* 20:3551–3567.

Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis.* 20:3551–3567.

Powlesland AS, Ward EM, Sadhu SK, Guo Y, Taylor ME, Drickamer K. 2006. Novel mouse homologs of human DC-SIGN: Widely divergent biochemical properties of the complete set of mouse DC-SIGN-related proteins. *J Biol Chem.* 281:20440–20449.

Sutton-Smith M, Dell A. 2005. Analysis of carbohydrates/glycoproteins by mass spectrometry. In: Celis JE, editor. *Cell Biology: A Laboratory Handbook.* San Diego: Academic Press. p. 415–425.

Takenaka Y, Fukumori T, Raz A. 2004. Galectin-3 and metastasis. *Glycoconjugate J.* 19:543–549.

Wallis J, Drickamer K. 1999. Molecular determinants of oligomer formation and complement fixation in mannose-binding proteins. *J Biol Chem.* 274:3580–3589.

Wallis R, Drickamer K. 1997. Asymmetry adjacent to the collagen-like domain in rat liver mannose-binding protein. *Biochem J.* 325:391–400.

Weis WI, Taylor ME, Drickamer K. 1998. The C-type lectin superfamily in the immune system. *Immunol Rev.* 163:19–34.

Yagita H, Masuko T, Hashimoto Y. 1986. Inhibition of tumor cell growth in vitro by murine monoclonal antibodies that recognize a proliferation-associated cell surface antigen system in rats and humans. *Cancer Res.* 46:1478–1484.