

Lactosyl-sepharose binding proteins from pancreatic cancer cells show differential expression in primary and metastatic organs

Micah N Sagini¹, Karel D Klika², Agnes Hotz-Wagenblatt³, Michael Zepp¹ and Martin R Berger¹ 

¹Toxicology and Chemotherapy Unit, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany; ²Molecular Structure Analysis, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany; ³Genomics and Proteomics Core Facility, Bioinformatics-Husar Unit, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany
Corresponding author: Martin R Berger. Email: m.berger@dkfz-heidelberg.de

Impact statement

Interaction of glycan binding proteins with aberrantly expressed glycans in tumor environment is crucial for metastasis. Here, we established a work flow for investigating the presence of a subset of these proteins in PDAC cells, which bind to a lactosyl-sepharose resin. The resin had been designed to isolate proteins with lectin-like properties. The corresponding lactosyl-sepharose binding proteins (LSBPs) show affinity for galactose and other monosaccharides. A subset of the LSBPs shows also calcium dependency. The importance of these proteins is highlighted by their differential expression profiles in PDAC cells growing in primary (pancreas) and metastatic (liver and lung) organ sites. Based on their affinity for the lactosyl-resin and monosaccharides, LSBPs hold potential for PDAC diagnosis and as drug targets. This work has set the stage for further investigation of the occurrence and the role of LSBPs in patient samples using the newly established workflow.

Abstract

In normal cells, glycan binding proteins mediate various cellular processes upon recognition and binding to respective ligands. In tumor cells, these proteins have been associated with metastasis. Lactosyl-sepharose binding proteins (LSBPs) were isolated and identified in a workflow involving lactosyl affinity chromatography and label-free quantification mass spectrometry (LFQ MS). A binding study with monosaccharides was performed by micro-scale thermophoresis and nuclear magnetic resonance spectroscopy. Influence of galactose on LSBPs' binding to the lactosyl resin was investigated by competitive affinity chromatography followed by LFQ MS. An analysis of amino acids with sugar binding motifs was searched using bioinformatics tools. The expression profiles of these proteins at the mRNA level, as determined by a chip array from a pancreatic ductal adenocarcinoma (PDAC) liver metastasis model, were used for evaluating their potential role in cancer progression. Proteomics data and their respective genes were analyzed by MaxQuant and Ingenuity Pathway Analysis. In total, 1295 LSBPs were isolated and identified from Suit2-007 human pancreatic adenocarcinoma cells. Interaction studies revealed that these proteins exhibit low to moderate affinity for monosaccharide sugars. Some of these LSBPs even showed reduced affinity after calcium depletion. Among the isolated proteins were annexins and galectins in addition to other families, with no history of binding lactosyl residues. A subset of LSBPs exhibited differential profiles in the pancreas, liver, and lung environ-

ments. These modulations may be related to tumor progression. In conclusion, we show that PDAC cells contain LSBPs, a subset of which binds galactose with calcium dependency. The differential expression of these proteins in a rat model highlights their value for diagnosis and as potential drug targets for PDAC therapy. Future work will be required to validate these findings in patient samples.

Keywords: Affinity chromatography, cancer progression, glycan binding proteins, pancreatic ductal adenocarcinoma, protein expression

Experimental Biology and Medicine 2020; 245: 631–643. DOI: 10.1177/1535370220910691

Introduction

In normal mammalian cells, glycans have an important role within the post translational modification of proteins. Protein glycosylation is governed by glycosidases and

transferases that regulate the formation of complex structures.^{1,2} These structures are endowed with new functions as opposed to linear macromolecules such as proteins and lipids.³ Their complexity corresponds to a code that has the

ability to transmit information with geometrically progressing efficiency.⁴ When covalently linked to proteins, glycans mediate various functions such as cellular trafficking and localization, binding specificity, cell signaling, and thermodynamic stability.⁵ Glycan binding proteins represent a diverse group of proteins that bind simple or even complex sugar structures. They include lectins and sulfated glycosaminoglycan binding proteins.⁶ In tumor cells, the interaction of glycans with their respective receptors (glycan binding proteins) has been associated with malignant transformation and metastasis.⁷ Lectins are specific carbohydrate binding proteins, excluding antibodies, or enzymes, which use carbohydrates as substrates, and transporters of free saccharides.⁸ In lectins, sugars bind a conserved carbohydrate recognition domain (CRD) consisting of about ~130 amino acids. Binding to this module occurs either in the absence of calcium as manifested in galectins or in the presence of calcium as in C-type lectins.^{9,10}

Galectins are widely distributed in living organisms and are known to bind glycans with exposed galactose residues.¹¹ They mediate binding of proteins to various cell matrix proteins such as laminins, collagens, and fibronectins.¹² In tumor cells, galectins feature prominently in cellular processes associated with metastasis such as cell invasion, migration, angiogenesis, immune evasion, inflammation, and malignant transformation.^{13,14} Selectins are C-type lectins, which play a role in cell adhesion, leukocyte homing, and cancer metastasis. The well-known ones include L-, E-, and P-selectin, which show 50% sequence identity in their C-type lectin domains.¹⁵ Interactions with selectins are mediated by their CRDs, which recognize the sialyl Lewis^x and Lewis^a epitopes on N- or O-linked oligosaccharides.^{16,17} Annexins have recently joined the list of proteins with lectin activity, in addition to being known to bind phospholipids in the presence of calcium.¹⁸ They are significantly expressed in tumor cells and have been shown to interact with various proteins involved in tumor development.^{19,20} Some members of the annexin family, such as annexin-4, -5, and -6, have been shown to bind glycosaminoglycans and sialoglycoproteins.²¹ Although glycan binding proteins have been associated with malignant transformation, little is known about their expression profile in tumor cells given that their separation by traditional enrichment methods is challenging.⁵

Affinity chromatography is a versatile technique commonly used in the purification of lectins.²² It involves the coupling of sepharose beads to sugar moieties that bind the protein of interest. This method was successfully applied for the purification of riproximin (Rpx) using a hydrolyzed sepharose resin.²³ Rpx is a toxic type II ribosome inactivating plant lectin with affinity for glycan structures and was also shown to inhibit the growth of pancreatic adenocarcinoma cells *in vitro*.²⁴ Its purification was subsequently improved by using a lactosyl resin instead of the hydrolyzed sepharose resin.²⁵ In a carbohydrate microarray, Rpx was shown to preferentially bind to the N- and O-linked oligosaccharides bearing galactose and N-acetylgalactosamine moieties.²⁵ When rats bearing liver tumors were treated with Rpx, significant tumor regression was

observed compared to the control groups.^{24,26} This observation was presumably related to the effect of Rpx on O-linked (e.g. antigen of mucins) and N-linked glycans (e.g. carcinoembryonic antigens (CEA)) in tumor cells. These proteins have been associated with metastatic pancreatic ductal adenocarcinoma (PDAC), a lethal malignancy with poor prognosis and a survival rate of <5%. Currently, PDAC has no effective therapy, and it is predicted to be the second leading cause of cancer-related mortality by 2030.²⁷

In the present work, we isolated and identified proteins with lactosyl binding properties from human Suit2-007 PDAC cells in an established work flow. This method involved protein separation by affinity chromatography using a lactosyl resin that had been used for the purification of Rpx. With the goal of evaluating the role of the isolated proteins in PDAC progression, all proteins identified by this method were related to chip array data, which had been obtained from a PDAC liver metastasis model.²⁸ In this model, human Suit2-007 PDAC cells were orthotopically and intraportally implanted into rats for tumor growth in the primary (pancreas) as well as in the metastatic environments (liver and lung), respectively. A comparison of gene expression levels between the primary and metastatic sites was used to qualify a subgroup of the separated proteins in terms of their potential involvement in PDAC progression.

Materials and methods

Cell culture and tumor implantation

Human pancreatic Suit2-007 cells were cultured under standard conditions (37°C, 100% humidity, 5% CO₂ in air) in complete RPMI medium containing 10% fetal calf serum and 1% glutamine (Gibco, Fischer Scientific, Germany). The cells were harvested upon attaining 80% confluence. Animal (Charles River Company, Sulzfeld, Germany) experiments were performed as described elsewhere²⁸ and in line with the ethical approval made by the animals' ethics committee (Regierungspräsidium) in Karlsruhe, Germany. Tumor cells were implanted orthotopically and intraportally for their growth in the pancreas as well as in the liver and lungs, respectively. Animals were imaged once weekly for a period of four weeks. After this period, the tumor-bearing animals were sacrificed and tumor nodules re-isolated from the three organs (liver, pancreas, and lungs). The samples were used for total RNA isolation.

Purification of total RNA from tissues for chip array

The extraction of RNA as well as chip array was performed, as described elsewhere.²⁸ Total RNA was extracted from cells re-isolated from the pancreas, liver, and lungs using the Qiagen RNA isolation kit (QIAGEN GmbH, Deutschland). For chip array, hybridization of biotin-labeled cRNA samples on Illumina Human Sentrix-12 BeadChip arrays (Illumina, Inc.) was prepared according to the modified Eberwine protocol.

Preparation of samples for separation by affinity chromatography

Cell lysates were prepared from human Suit2-007 PDAC cells using lysis buffer (850 μ L RIPA buffer: 50 mM TRIS, 150 mM NaCl, 1.0% NP-40, 25 \times protease inhibitor 40 μ L), 10 \times Phosphostop tablet (100 μ L), and 100 mM of NaVO₃ (10 μ L) (all from Roche diagnostics, Germany). Cells were disrupted by mixing with a pipette, which was repeated after every 10 min while on ice. The lysate was centrifuged (16,400 r/min, 30 min) to obtain a clear supernatant. Re-isolated tumor nodules from in vivo samples weighing \sim 100 mg were homogenized in liquid nitrogen using a microdismembrator (Sartorius, GmbH, Germany). The resulting powder from the tumor was processed similarly to the cell lysate. The protein concentrations in lysate, and tumor homogenates were determined by Roti Nanoquant solution (Carl Roth GmbH & Co. KG, Germany).

Protein separation by affinity chromatography

For protein separation, a Tricon column (10 mm diameter, 120 mm length) was packed with lactosyl-sepharose gel (GE Life Sciences, Germany). The column binding capacity was determined using galectin-3 (kindly provided by Prof. Jürgen Kopitz, Heidelberg). Galectin-3 (1 mg/mL) was mixed with lactosyl-sepharose gel (1 mg/mL) in a falcon tube and left overnight on a rotary shaker (4°C). The tube was centrifuged (5000 r/min, 20 min) and the supernatant recovered. The bound galectin-3 was determined from the initial protein concentration and the concentration of the recovered supernatant. Clear lysate (1 mg/mL) was loaded onto the column with (20 mM) TRIS-HCL and Arg-HCL buffer at low velocity flow (0.5 mL/min). The flow rate was then increased to 1 mL/min, as the non-column binding proteins fraction started to elute from the column. When the UV signal returned to the baseline level, the column binding protein fraction was eluted with elution buffer: TRIS-HCL (20 mM), Arg-HCL (20 mM), CaCl₂ (100 mM), and galactose (100 mM). Arg-HCL was included in the buffer to minimize protein-protein interactions and non-specific binding. When the elution was complete, the column was regenerated by washing with 1 M CaCl₂ followed by equilibration with loading buffer.

Gel staining with Coomassie blue

Coomassie stain (Serva Electrophoresis GmbH, Germany) was prepared from the following stock solutions. Solution A contained 10% (w/v) ammonium sulfate, 2% (w/v) phosphoric acid in double distilled (dd) water (Merck KGaA, Darmstadt, Germany). Solution B contained 5% (w/v) Coomassie Brilliant Blue G-250 in double distilled water (dd H₂O). A washing solution containing 25% methanol in dd H₂O was prepared with 20 μ g of proteins separated by affinity chromatography and were loaded into the wells of precast PAGE gels (4–20%) (Serva Electrophoresis GmbH, Heidelberg, Germany) and separated by gel electrophoresis. The staining solution was prepared by mixing 100 mL of stock solution A with 2.5 mL stock solution B. The resulting solution was vortexed for 20 min after which

methanol (25 mL) was added and the solution again vortexed for 20–30 min. The gel was immersed in the freshly prepared colloidal Coomassie stain and left overnight with gentle shaking. The following day, the gel was destained by washing with 25% methanol in dd H₂O. The gel was again washed twice with dd H₂O till the protein bands attained the desired contrast.

Protein labeling with Alexa fluor

For thermophoresis, lactosyl-sepharose binding proteins (LSBPs) were labeled with the fluorescent dye Alexa fluor 647 (Invitrogen, Germany). Proteins (2 mg/mL) were mixed with 36 μ L of 0.1 M sodium bicarbonate buffer (pH 8.3). While stirring, 50 μ L of the reactive dye (Alexa fluor) was added drop-wise to the protein fraction. The mixture was then incubated for 1 h at room temperature while shaking at 350 r/min. A PD 10 – Sephadex G-25 (GE, Life Sciences, München, Germany) gel filtration column was equilibrated with 10 mL of phosphate buffer (1 \times , pH 8.0) at room temperature. The protein sample was applied onto the gel surface and allowed to percolate through the column. The intensely blue colored fraction, which contained the labeled proteins, was collected and excess buffer removed by centrifugation (5000g, 30 min) using 10 kDa Amicon protein filters (Merck KGaA, Darmstadt, Germany). The labeled protein fraction was sterile filtered, its concentration determined, and then stored at –20°C till further use.

Evaluation of galactose binding to LSBPs

Microscale thermophoresis

An aliquot (7 μ L) of 1 mM galactose was serially diluted to lower concentrations in Eppendorf tubes. One μ g/mL of labeled LSBPs was added to each Eppendorf tube containing galactose and briefly vortexed. The samples were transferred into capillary tubes and carefully placed in the Monolith instrument. Measurement was done using the in-built software by a computer connected to the Monolith NT.115. After 1 h, the data were retrieved for analysis. To evaluate the role of calcium in galactose binding to LSBPs, the experiment was repeated as described above and 1 μ M of ethylene glycol-bis (β -aminoethyl ether)-N,N,N',N'-tetraacetic acid (EGTA) was added to the reaction mixture. Measurement was done as before.

Nuclear magnetic resonance spectroscopy

Binding was evaluated by line broadening or pertinent chemical shift changes observed in the proton (¹H) spectrum as well as by the transfer of bulk water magnetization to the ligand, using waterLOGSY pulse sequence. Before analysis, stock solutions of five sugars (1 mM each) and the LSBPs (1 mg/mL) were prepared and stored on ice until use. These sugars included galactose (Gal), glucose (Glc), fucose (Fuc), mannose (Man), and rhamnose (Rha). A concentration of 100 μ g/mL protein and 1 mM sugar was used. The proton ¹H NMR spectra were acquired at 298 K with a Bruker Avance II NMR spectrometer. The

instrument was equipped with a 5-mm inverse-configuration probe with triple-axis-gradient capability at field strength of 14.1 T operating at 600.1 MHz for ^1H nuclei. The ^1H NMR spectra were acquired using the 1D NOESY pulse sequence. A pre-saturation of the water signal was applied using standard spectral parameters as for normal ^1H NMR spectra. This included an Aq of 2.50 s, 5 s for the PAD, a mixing time (τ_m) of 50 ms and a read-out pulse of 90° . Processing consisted of zero-filling ($\times 4$) and the application of 0.3 Hz of line broadening. WaterLOGSY spectra were also acquired using the pulse sequence as described by Dalvit with an Aq of 1.95 s, 4 s for the PAD, and a mixing time (τ_m) of 1 s.²⁹ Whereas the suppression of the water signal was effected by excitation sculpting using a 2 mssinc-shaped pulse, the selective excitation of the water signal was effected by a 5 ms Gaussian-shaped pulse.³⁰ WaterLOGSY spectra were processed with zero-filling ($\times 4$) and the application of 2 Hz of line broadening.

Identification and characterization of LSBPs

Protein fractions were concentrated using Unicorn Ultra protein concentrating tubes (0.5 mL capacity). The process of protein concentration involved thorough washing with the loading buffer to get rid of bound galactose resulting from the elution buffer. Four equal samples of protein fractions (200 $\mu\text{g}/\text{mL}$) were prepared in triplicates from the galactose-free protein stock solution. Galactose was added to the first (50 mM final concentration) and second (150 mM final concentration) set of samples, respectively. In the third set, galactose (150 mM final concentrations) and EGTA (0.5 mM final concentrations) were added. The fourth set, which served as control, did not contain galactose or EGTA. These samples were incubated for 15 min at 37°C followed by separation with affinity chromatography. For each set, the loading buffer contained the corresponding concentration of galactose. For protein fractions mixed with 50 mM of galactose, a buffer containing TRIS-HCL (20 mM), Arg-HCl (20 mM), and galactose (50 mM) was used. For protein fractions containing 150 mM galactose without and with EGTA, an equal concentration of galactose was included in the loading buffer. The samples were then applied to the column, and protein fractions, which did not bind to the column due to the presence of galactose, were collected. Proteins were precipitated from the buffer solutions by adding 400 μL of methanol to 100 μL protein solutions followed by vortexing and centrifugation (9000g for 10 s). Then, 100 μL of chloroform was added followed again by vortexing and centrifugation (9000g for 10 s). To separate the phases, 300 μL of water was added to the samples, which were vortexed vigorously followed by centrifugation (9000g, 1 min). The upper phase was removed and discarded. An aliquot of 300 μL of methanol was added to the lower chloroform phase and the interphase material containing the precipitated proteins. The mixture was vortexed and centrifuged again (9000 g, 2 min) to obtain a protein pellet. The supernatant was removed, and the dry pellets were dissolved in water and their protein concentrations determined by Roti Nanoquant solution as described previously.

Quantitative protein analysis by mass spectrometry

Proteins solutions with concentrations ranging from 0.29 to 0.152 $\mu\text{g}/\mu\text{L}$ were loaded onto an SDS-gel and allowed to run only a short distance of 0.5 cm. After Coomassie blue staining, the total sample was cut out from the gel and digested by trypsin. Samples were processed based on a slightly modified protocol³¹ carried out on the DigestProMSirobotic system (INTAVIS Bioanalytical Instruments AG). Digested samples were loaded onto a cartridge trap column packed with Acclaim PepMap300 (C18, 5 μm , 300 \AA wide pore, Thermo Scientific) and separated using a 120 min gradient from 3% to 40% ACN on a nanoEase MZ Peptide analytical column (300 \AA , 1.7 μm , 75 $\mu\text{m} \times 200 \text{ mm}$, Waters). Eluted peptides were analyzed by an online coupled Q-Exactive-HF-X mass spectrometer running in a data-dependent acquisition mode where one full scan was followed by up to 30 MS/MS scans.

Search for amino acids motifs responsible for sugar binding

To identify amino acid motifs responsible for sugar binding proteins, known and novel motifs were searched in column binding proteins and controls. The known CRD sequence was searched in LSBP protein sequences against the respective PFAM-A hmm-model (PFAM (RRID:SCR_004726)),³² using the HMMER 3.0 software (<http://hmmer.org/>, default parameters). Next, new sequence patterns were designed using ricin and Rpx as a reference. The program find-patterns (Wisconsin Package) used the pattern file indicated below to search for the occurrence of patterns in the LSBPs.

```
Gal-1 1 (D,N,E)(X)[2,19]((Q,I,L)(X)[1,3](W,Y,F)(X) [1,19]
(N,Q)(N,Q) 0!
RiproxA1
D(X)[10,13]I(X)[1,2]W(X)[7,9]NQ 0!
RiproxY1
D(X)[10,13]L(X)[1,2]W(X)[7,9]NQ 0!
RicinY1
D(X)[10,13]I(X)[1,2]W(X)[7,9]NQ 0!
```

The amino acids included in the sequence search pattern were as follows: asparagine, aspartate, glutamine, leucine, isoleucine, tyrosine, and tryptophan. In addition, a de novo motif search was performed for the column binding proteins with the program MEME (v2.2) (MEME Suite – Motif-based sequence analysis tools, RRID: SCR_001783). This program, developed by Bailey and Elkan, is able to find conserved motifs in a group of unaligned sequences.³³

Statistical analysis

Proteomics data were analyzed using MaxQuant software (version 1.6.0.16) and Uniprot (RRID: SCR_004426) extracted databases.³⁴ Identification of false discovery rate cut-offs were 0.01 for both peptide and protein levels. Proteins were quantified using a label-free quantification approach based on the MaxLFQ algorithm.³⁵ A minimum of two quantified peptides per protein was required for quantification. The data were further processed using in-house compiled R-scripts and the Perseus software

package.³⁶ The mRNA levels of LSBPs were evaluated by the Ingenuity Pathway Analysis (IPA; RRID:SCR_008653) program with $p < 0.001$. A Z-score of ± 3 was used as a cut-off for selecting three functional annotations of interest. Venn diagrams for further evaluation were used (online at <http://bioinformatics.psb.ugent.be/webtools/Venn/>).

Results

Separation of proteins by affinity chromatography

Proteins derived from Suit2-007 human PDAC cells growing in vitro were separated by affinity chromatography using a lactosyl resin. The resin matrix was composed of sepharose, a linear polysaccharide of alternating D-galactose and 3,6-anhydro- α -L-galactopyranose, which was chemically linked to the lactosyl ligand. The beads had an open pore size (90 μ m) suitable for separating proteins in the range of up to 150 kDa. This resin had been used previously for the purification of Rpx from *Ximenia americana*. In this study, proteins were separated by affinity chromatography from lysates and tumor nodules of Suit2-007 cells. The former had been obtained from cells grown in vitro, the latter from the tumor nodules growing at three sites following implantation into nude rats.²⁸ Proteins, which bound to the resin, are hereafter referred to as LSBPs. Their respective chromatograms exhibited similar retention times for peaks from in vitro cell lysates and in vivo homogenates (see Figure 1(a) and (b)). Figure 1(c) represents proteins derived from cell lysate (lane A) and tumor homogenate (lane D) as well as the respective flow through

fractions (lanes B and E) and the column binding fractions (lanes C and F) that were separated by electrophoresis. The protein sizes varied between the flow through and LSBPs, respectively. The chromatogram in Figure 1(d) represents a run using the LSBP fraction with the mobile phase containing a low concentration of galactose (50 mM) as described below.

Affinity of monosaccharides to LSBPs

Microscale thermophoresis. Microscale thermophoresis (MST) was performed to evaluate the interaction of monomeric sugars (galactose, glucose, and fucose) with LSBPs. Binding was evaluated by serially diluting the sugars' concentration from 1 mM down to 4 μ M and adding 1 μ g/mL of LSBPs to each concentration. Galactose, glucose, and fucose exhibited low affinities for the labeled proteins (see Figure 2(a) for galactose binding). To test the influence of calcium on galactose binding, EGTA (1 μ M) was added to the solution containing galactose and the protein fractions. As indicated in Figure 2(b), the addition of EGTA to the reaction mixture diminished binding between galactose and LSBPs.

Nuclear magnetic resonance spectroscopy

Further evaluation of sugar binding to LSBPs was performed by nuclear magnetic resonance spectroscopy (NMR) spectroscopy. Binding was evaluated by line broadening and chemical shift changes in the 1 H spectra and the waterLOGSY pulse sequence,^{29,37} which allowed ranking of the sugars' affinity based on the respective signal sizes.

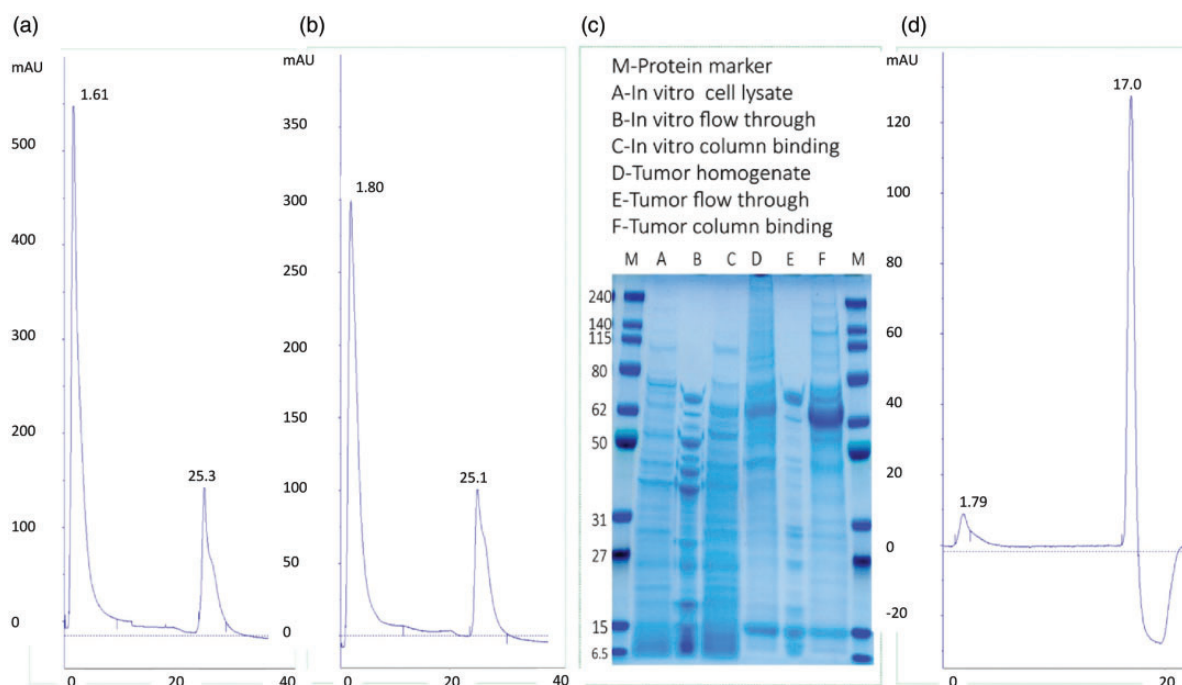


Figure 1. High-performance liquid affinity chromatography and electrophoretic separation of LSBPs. (a) and (b) are chromatograms of protein samples isolated from lysate of tumor cells grown in vitro and of tumor homogenates from in vivo samples, respectively. Protein separation was performed using a lactosyl-resin, which binds proteins (LSBPs) with an affinity for this sugar. In the chromatograms, the column binding proteins are represented by peaks with retention times between 25 and 30 min. (c) a polyacrylamide gel comparing protein separation from cell lysate (lane A) and tumor homogenate (lane D) as well as the flow through (lanes B and E) and column binding fractions (lanes C and F). The chromatogram shown in (d) represents the separation of LSBPs following the addition of galactose (50 mM) to the mobile phase. (A color version of this figure is available in the online journal.)

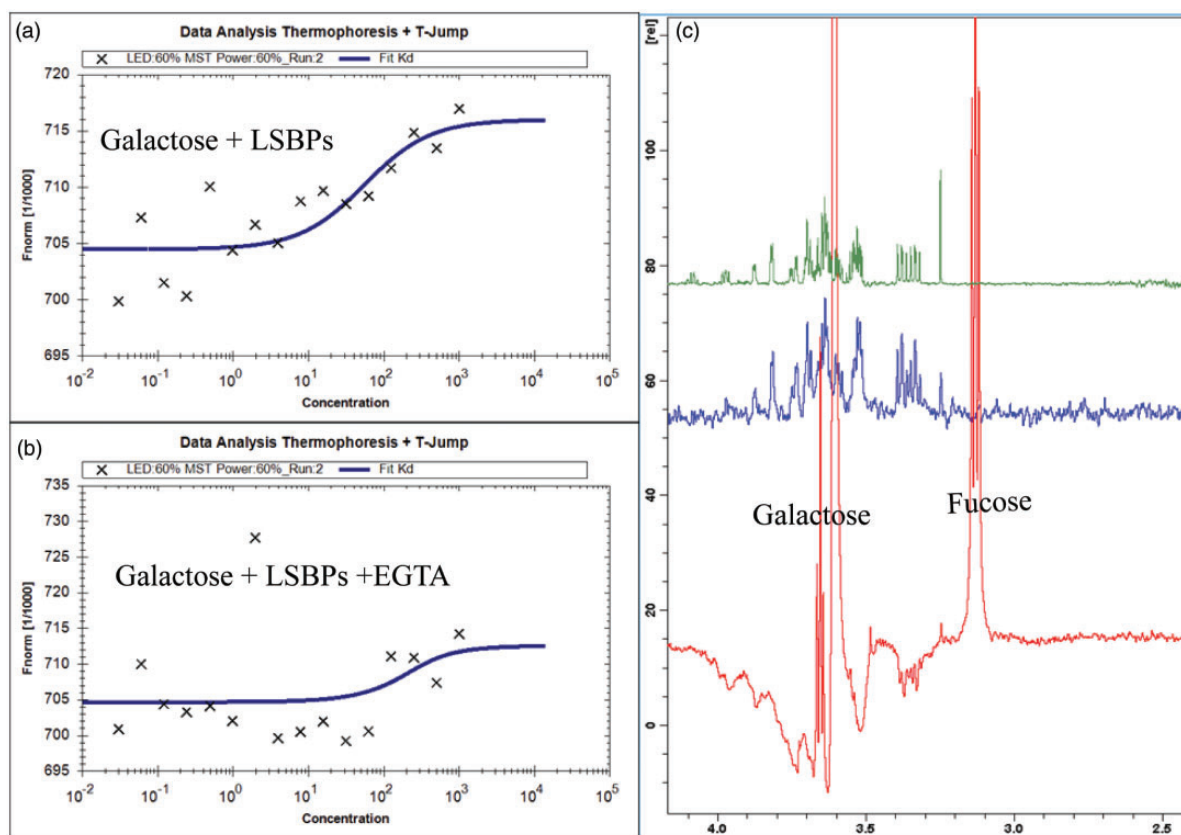


Figure 2. Affinity of monosaccharides to lactosyl-sepharose binding proteins. (a and b) Binding curves of galactose to LSBPs in the absence or presence of EGTA, respectively, as evaluated by microscale thermophoresis. (c) The NMR spectra for the binding of galactose and fucose to LSBPs as demonstrated by waterLOGSY. The waterLOGSY spectrum in blue is a control sample of sugars in the absence of LSBPs. The spectrum in red shows the negative resonances for the ring protons of galactose and fucose as well as the methyl group of fucose. The negative signals indicate the interaction of these sugars with LSBPs. For comparison, a normal ¹H spectrum (green) for galactose and fucose in the absence of LSBPs is also shown. (A color version of this figure is available in the online journal.) LSBP: lactosyl-sepharose binding protein; EGTA: ethylene glycol-bis (β -aminoethyl ether)-*N, N, N', N'*-tetraacetic acid.

In solution, the sugars examined exist predominantly as pyranose-ring forms and as α - and β -anomers. Only the pyranose-ring forms were amenable to analysis because the furanose-ring and open-chain forms of each sugar were too low in concentration. The affinity of galactose and fucose for LSBPs was moderate, based on clear line broadenings and the observed chemical shift changes in the ¹H spectra. From the comparative line broadening of the anomeric H-1s, the binding was greater for the β -anomer of galactose than for the α -anomer, and similarly in the case of fucose. The observed bindings in the ¹H spectra were confirmed by waterLOGSY spectra. In addition, notable intensities in the waterLOGSY experiment for many of the protons inferred that the binding epitope encompasses the entire molecule. A competition binding experiment was also performed to compare the affinity of β - and α -anomers of galactose and fucose. Although the difference was only slight, the β -anomer of galactose had a higher affinity compared to the β -anomer of fucose. Similarly, the α -anomer of galactose showed a slightly higher affinity in comparison to the α -anomer of fucose. However, the affinity of rhamnose to LSBPs was not clearly ascertained from the ¹H spectra. WaterLOGSY evaluation, however, confirmed the very weak binding of both α - and β -anomers of rhamnose to LSBPs. Moreover, waterLOGSY confirmed

that β -rhamnose interacted with LSBPs more strongly than α -rhamnose. In addition, mannose and glucose were also examined with similar results, i.e. for all five of these sugars, evaluation of the line broadening and the waterLOGSY spectra revealed that the two anomeric forms exhibited differential affinity with the β -anomer always showing higher affinity than the α -anomer. The order of binding for the 10 predominant sugar forms of the five sugars as evaluated by NMR is as follows; β -Gal > β -Glc > β -Fuc \gg α -Gal \approx α -Fuc \approx α -Glc \approx β -Man \gg α -Man > β -L-Rha > α -L-Rha.

Identification and characterization of LSBPs

Thoroughly washed and therefore galactose-free LSBPs were identified and characterized by label-free mass spectrometry (LFQ MS). Re-administration of these LSBPs to the lactosyl resin served as control. In parallel, other LSBP fractions were loaded to the lactosyl-resin following the addition of galactose (50 or 150 mM), without or with EGTA (0.5 mM) to the respective mobile phases. For controls, the LSBPs bound to the lactosyl-resin were eluted; for the other groups, the flow through (i.e. the fraction that was prevented from binding to the lactosyl-resin) was collected for further analysis (see Figure 1(d)). The collected fractions were precipitated from the loading buffer and also

evaluated by LFQ MS. Details of the experimental design and analysis are given in Materials and Methods section as well as in Supplementary Figures 1 and 2. In total, 1336 proteins were detected by identifying at least two related peptides. Proteins that had no measurable LFQ value ("0"; $n = 41$) in all control groups were excluded, thus decreasing the number of proteins to 1295. The distribution of these proteins regarding their localization is summarized in Table 1. From this number, only 1193 proteins were matched to the chip array data (see Supplementary Table 1). A breakdown of these (see Table 2) shows different protein groups as quantified by LFQ MS. Proteins, which were not influenced by galactose in their column-binding properties, are represented by LFQ ratios of zero ($n = 1062$). Those, which were partially influenced by addition of galactose to the mobile phase ($n = 152$), showed LFQ ratios ≥ 0 . The last group ($n = 81$) represents proteins, which showed measurable LFQ ratios > 0 in all three experimental groups. As shown in Figure 3(a) and (b), proteins were grouped according to their column binding properties in the presence of galactose, either without or with EGTA. From this analysis, LSBPs were identified showing a similar response to low (50 mM) or high (150 mM) galactose without and with EGTA concentrations (Figure 3(c) to (e)). These proteins exhibited differences in affinities for the lactosyl resin as demonstrated by their LFQ ratios. Other LSBPs showed differential binding behavior between the experimental groups (Figure 3(f) to (h)). However, most of the proteins were not influenced in their binding by the addition of galactose and/or EGTA to the mobile phase.

Search for sugar binding motifs in LSBPs

LSBPs were also characterized by searching for motifs with sugar binding amino acids. Remarkably, the CRD motif, as described in Finn et al.³² was not found in any of the LSBP proteins. Therefore, a "findpattern" search was performed using amino acids from ricin and Rpx that are known to bind galactose (see Materials and Methods section). The pattern search was performed in those proteins that showed sensitivity to galactose ($n = 81$ or 87), but did not find any hits. With the program MEME, six motifs (a-f) were identified (Figure 4(a)). These motifs had different widths and occurred at different sites of the protein sequences (see Supplementary Table 2). The six motifs were evaluated in proteins with sensitivity to galactose ($n = 81$ or 87), proteins with partial sensitivity for galactose ($n = 152$ or 157), LSBPs with no sensitivity for galactose ($n = 207$), kinases ($n = 99$), and random proteins ($n = 98$). Evaluation of these motifs in proteins showed that motifs d and/or f were associated with significantly higher scores in the proteins with sensitivity to galactose than in the other protein groups (Figure 4(b) and (c)). In total, 37 proteins were identified with motifs d and f. Of these, 14 proteins contained both motifs, 10 contained motif d only and 13 contained motif f only. All 37 proteins are shown in the Supplementary Table 3. It is worth noting that the number of galactose-sensitive and galactose partially sensitive proteins increased from 81 to 87 and 152 to 157, respectively, during the conversion of gene IDs to protein entries in the protein database (Uniprot.org).

Table 1. Overview of LSBPs identified by mass spectrometry following affinity separation.

Protein location	Identified proteins	Description of protein subgroups
Cytoplasmic proteins	650 (54.5%)	Enzymes (kinases, peptidases, phosphatases), transporters, transcription regulators, translation regulators, others
Extracellular proteins	86 (7.2%)	Enzymes (peptidases), transporters, growth factors, cytokines, others
Nucleus proteins	304 (25.5%)	Enzymes, transcription regulators, others
Plasma membrane proteins	113 (9.5%)	Enzymes (peptidases, kinases, phosphatases), transporters, transmembrane receptors, G-coupled proteins, ion channels
Protein from other locations (others)	40 (3.3%)	Transporters, enzymes (peptidases, kinases, phosphatases), others

Table 2. Breakdown of proteins of the flow through after the addition of galactose low and high concentrations to the mobile phase.^a

Proteins' distribution in the Venn diagrams	Galactose LC ^b (50 mM) – LFQ	Galactose HC ^c (150 mM)–LFQ	Galactose HC ^c (150 mM) + EGTA–LFQ
1062	0.0	0.0	0.0
81	>0.0	>0.0	>0.0
77	>0.0	0.0	0.0
22	0.0	0.0	>0.0
24	>0.0	0.0	>0.0
13	>0.0	>0.0	0.0
8	0.0	>0.0	0.0
8	0.0	>0.0	>0.0

LFQ: label-free quantification; EGTA: ethylene glycol-bis (β -aminoethyl ether)- N, N, N', N' -tetraacetic acid.

^aTable 2 gives the breakdown of proteins shown in Figure 3, which were quantified by MS analysis and their respective LFQ ratios. A value below the detection limit (0.0001) indicates that the binding of these proteins to lactosyl-sepharose resin was not influenced by galactose (50 and 150 mM) concentrations.

^bGalactose low concentration.

^cGalactose high concentration.

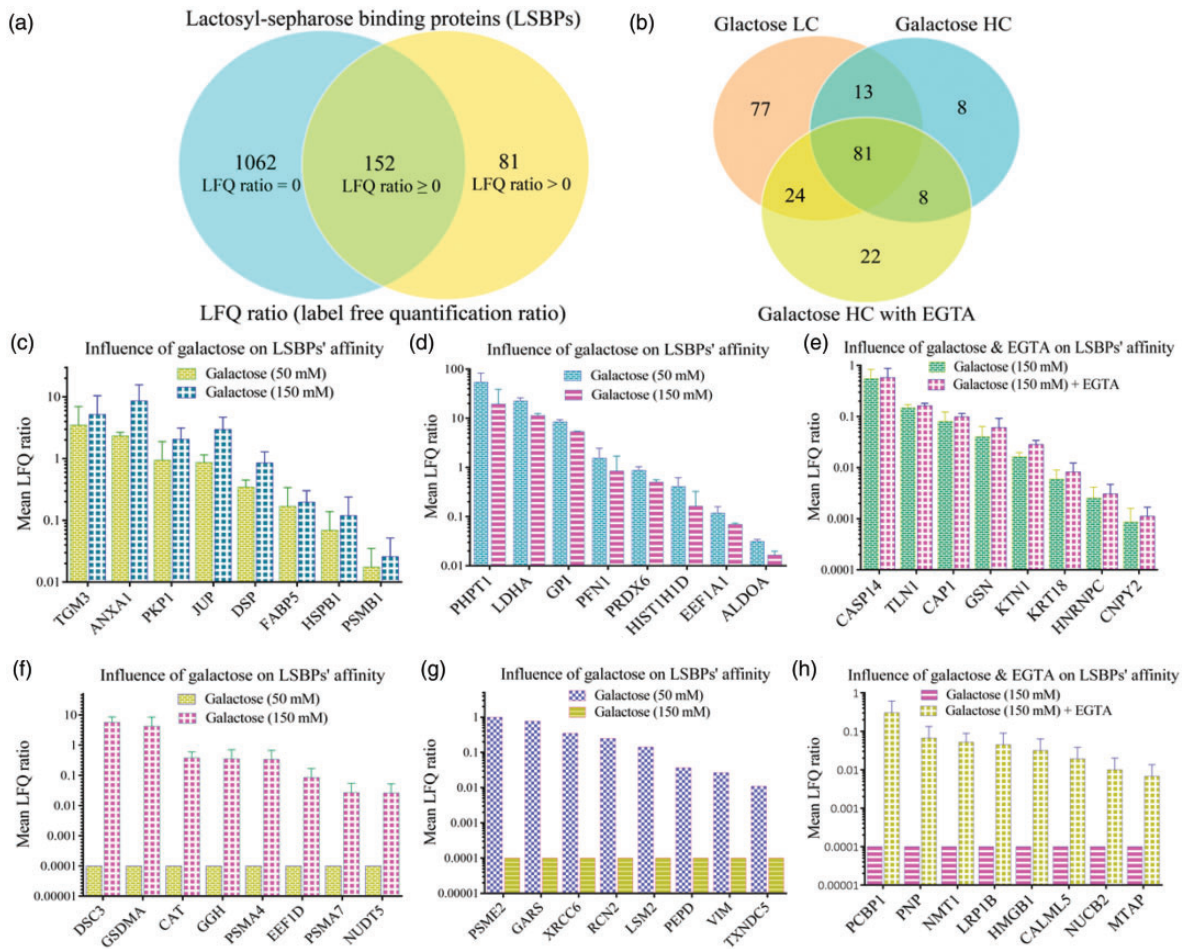


Figure 3. Label-free quantification of lactosyl-sepharose binding proteins by MS analysis. (a) The Venn diagram of protein groups as characterized by LQ MS. The experiment was performed to identify LSBPs, which had been isolated from PDAC cell lysates. Re-administration of these LSBPs to the lactosyl-resin served as control. For competitive elution, mobile phases of the other experimental groups contained LSBPs and galactose at low (galactose LC, 50 mM) or high (galactose HC, 150 mM) concentrations. In a third group, EGTA (0.5 mM) was added to the high galactose concentration. For controls, the LSBPs bound to the lactosyl-resin were eluted; for the other groups, the flow through (i.e. the fraction that was prevented from binding to the lactosyl-resin by galactose) was collected. Data on LQ ratios were obtained by comparing the experimental groups with controls resulting in three groupings: the galactose non-sensitive group ($n = 1,062$) with LFQ values = 0, the galactose partially sensitive group ($n = 152$) with values LFQ values ≥ 0 and the galactose-sensitive group ($n = 81$) with LFQ values > 0 . (b) The distribution of all galactose-sensitive proteins ($n = 233$) with respect to their LQ ratios. Altogether, 81 proteins showed LFQ values > 0 in all three experimental groups, 45 proteins showed LFQ values > 0 in two experimental groups, and 107 proteins showed LFQ values > 0 in only one experimental group. (c to e) proteins, which exhibited similar intensities of binding to the lactosyl resin, in the presence of low and high galactose concentrations, without and with EGTA. Bargraphs F to G exemplify distinct differences in column binding between high and low galactose concentrations without and with EGTA. For non-detectable LFQ values, a constant value of 0.0001 was used which corresponds to the detection limit. (A color version of this figure is available in the online journal.)

Expression of LSBPs in a PDAC liver metastasis model

After isolating and identifying LSBPs from Suit2-007 PDAC cells growing in vitro, we related these proteins to their mRNA expression levels, which had been measured before from Suit2-007 cells growing in vivo. As detailed earlier,²⁸ a chip array had been performed comparing the expression profiles of genes in Suit2-007 PDAC cells growing in the pancreas, liver, and lung environments of an established rat model. These tumor cells had been re-isolated from the respective organs following orthotopic as well as intraportal tumor cell implantation. The corresponding gene profiles showed the expression of individual genes (fold change) in one environment compared to the others. The aim of this comparison was to assess the potential importance of LSBPs for the metastatic process.

When designing the experiment for the separation of cellular LSBPs, we had anticipated to isolate proteins

with lectin-like properties, comparable to that of Rpx, as well as others known to bind sugars, like members of the annexin and galectin families. Based on the results of our chip array experiment, we had knowledge on the differential expression status of the genes in question. Comparison of the mRNA expression fold change of some of these genes attracted our interest for investigating the presence of glycan binding proteins in tumor cells (see Figure 5(a) and (b) for the expression profiles of some members of the annexin and galectin families). When relating the pancreatic expression to the expression in vitro, only annexin-2 and -3 deviated from the basal expression (fold change = 1), showing fold changes of 4.3 and 0.36, respectively. A comprehensive analysis of all known annexins is shown in Supplementary Table 4. When relating their expression in liver or lung to pancreas, however, all shown family members were significantly modulated (see Figure 5(b)).

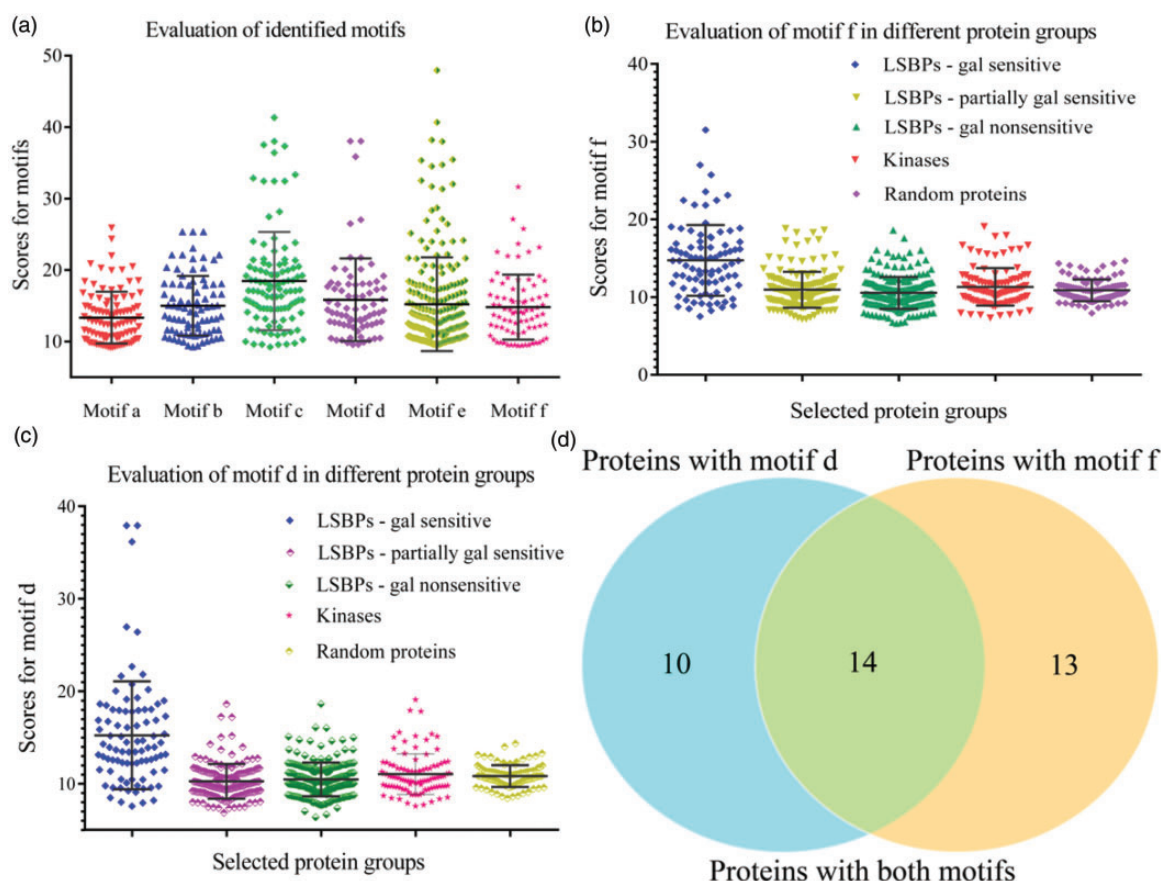


Figure 4. Sugar binding motifs for lactosyl-sepharose binding proteins. (a) A statistical comparison of six motifs (a–f) identified by a pattern search in the galactose sensitive proteins (LFQ ratio >0). (b) and (c) A comparison of scores obtained from fitting motifs d and f to five protein groups. The five protein groups consisted of: galactose-sensitive group ($n = 81$, LSBPs gal sensitive), galactose partially sensitive, ($n = 152$; LSBPs partially sensitive), galactose non-sensitive group ($n = 207$; LSBPs gal nonsensitive), kinases ($n = 99$; kinases), and random artificial proteins ($n = 98$; random proteins). (d) A Venn diagram for those proteins containing motifs d and f. (A color version of this figure is available in the online journal.)
LSBP: lactosyl-sepharose binding protein.

Clearly, PDAC cells growing in secondary organs expressed substantially higher levels of respective mRNA than those growing in the primary tissue (pancreas). In addition to concentrating on protein families known to have lectin-like properties, we also used the pool of all LSBPs for extracting the mRNA modulation of the respective genes ($n = 1193$) from the chip array data. All 1193 genes, with their respective expression ratios and p values were analyzed by the IPA. From the IPA results, three functional annotations were selected based on a significant Z-score (>3), p value, and their relevance to tumor growth and progression. These functional annotations included genes grouped under the terms “cell movement,” “cell signaling and interaction,” and “cell development.” To identify genes with single and shared functional annotations, a Venn diagram was constructed (Figure 5(c)). The Venn diagram shows the distribution of genes for LSBPs ($n = 93$) for the three functional annotations selected from IPA. Further selection ($n = 43$) was performed for those genes showing significant expression in at least one or more organ environment(s) (see Supplementary Table 5). Figure 5(d) and (e) represents bar graphs for selected genes, which were significantly up-regulated in the pancreas as well as in the liver and lungs.

Discussion

In this study, we established a workflow involving affinity chromatography and LFQ MS for the isolation and identification of lactosyl-sepharose binding from Suit2-007 human PDAC cells. By relating the respective mRNA expression of these proteins to chip array data obtained from the same cell line in a PDAC liver metastasis model, we found a subset of genes, which have the potential to contribute to PDAC progression and metastasis. The separation of these proteins by affinity chromatography was based on a lactosyl resin, which is not commonly used for protein purification. However, this resin had been used for purification of the plant lectin Rpx, which binds glycan structures and also inhibits proliferation of PDAC cells. The sugar composition of this resin was the key determinant in its affinity for Rpx. The resin is composed of sepharose beads chemically linked to a lactosyl ligand. In addition to a stable ligand, the resin had a wide fractionation range (up to 150 kDa) and can withstand high pH values. With this method, we anticipated isolating proteins with glycan binding properties similar to Rpx.

By applying tumor cell lysate to the lactosyl resin, we isolated and identified 1295 proteins with appropriate affinities. Results from LFQ MS and binding experiments

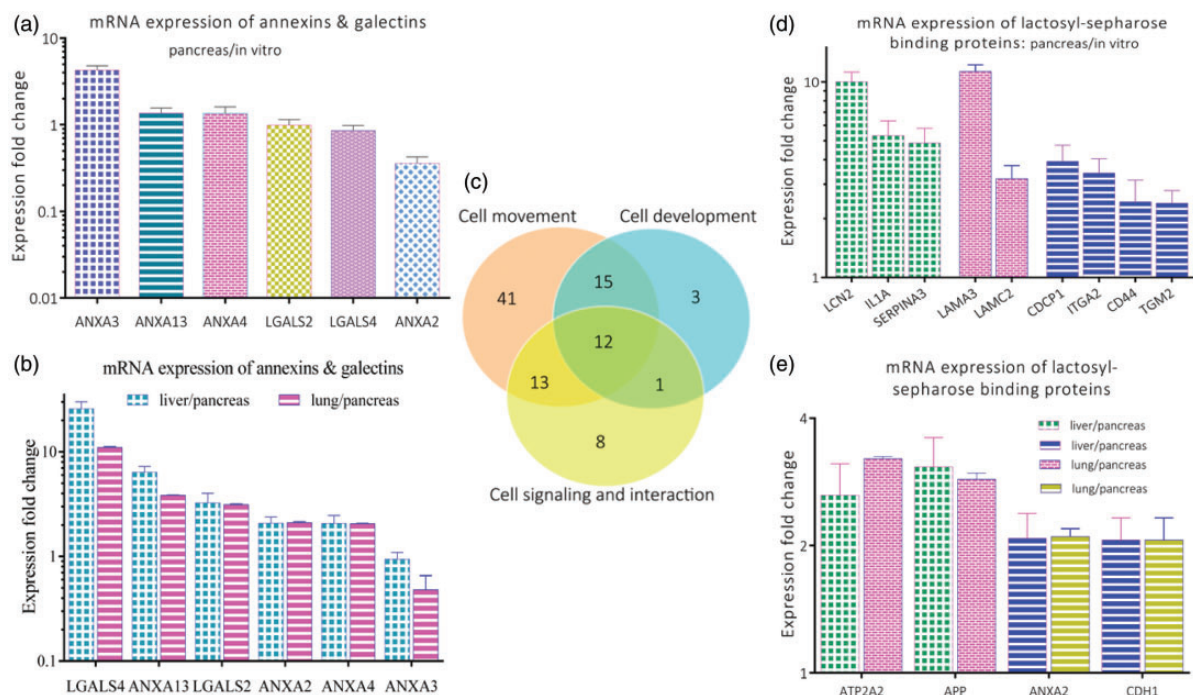


Figure 5. mRNA expression of lactosyl-sepharose binding proteins. (a and b) The gene expression profiles of selected galectin and annexin family members as determined by a chip array using tumor cells re-isolated from the pancreas, liver, and lung environments of a PDAC liver metastasis rat model. In (a), the fold change of annexins and galectins in the pancreas was compared to their expression in cells grown in vitro. In (b), the fold change of annexins and galectins in the liver and lung was compared to their expression in the pancreas. (c) A Venn diagram consisting of 93 genes as evaluated by IPA. These genes were part of the 1193 genes from IPA resulting in various functional annotations. Three functional annotations were selected based on their predicted increased activation state and a significant Z-score of >3.918 . They were grouped under the terms “cell movement,” “cell signaling and interaction” as well as and “cell development.” From these genes ($n = 93$), 43 were selected based on their significant modulation in at least one or more environments (see Supplementary Table 5). (d) and (e) show bargraph for 9 and 4 genes, respectively, which were selected based on their significant mRNA expression modulation in the pancreas. These genes were assigned to green (cell movement), blue (cell movement and cell signaling), pink (cell movement, and cell development), and yellow (cell movement, cell signaling, and cell development). (A color version of this figure is available in the online journal.)

revealed that this method was robust and reproducible. The resin enriched proteins known to bind sugars including the members of the galectin and annexin families, as well as other protein families with previously unknown properties with regard to glycan binding. One of the challenges associated with MS-based proteomics technology is the limitation in detection of low abundance proteins from any given sample. The current study was not exempted from this limitation, since we were not able to detect all the proteins we anticipated to isolate. For instance, only a few members of the annexin and galectin protein families were detected by LFQ MS following their separation by affinity chromatography. To further characterize the galactose binding proteins, we added galactose to the mobile phase at two concentrations, to compete with the resin for the galactose binding proteins. This resulted in a number of proteins (18%), which exhibited decreased binding to the resin under these conditions of competitive elution. These findings suggest that galactose may have induced conformational changes in the binding sites of these proteins, thus reducing their affinity for the resin. In addition, the size of a sugar binding pocket has also been shown to influence the affinity of proteins for sugars. For instance, enzymes, which are involved in sugar transport, have to provide a deep cleft for tight binding, as opposed to lectins, which have shallow clefts and interact with sugars transiently.³⁸ Although lectins are known to bind small carbohydrates with low

affinity, their binding is enhanced during spatial clustering (multivalency effect).³⁹

To investigate the binding of monosaccharides to LSBPs, microscale thermophoresis and NMR spectroscopy were performed. For NMR, we used mixtures of monosaccharides to directly compare the relative binding properties through competitive binding and thereby eliminate inconsistencies in samples, which can occur when evaluating samples containing only individual monosaccharides. By line broadening and chemical shift changes in the ^1H spectra, and waterLOGSY experiments, galactose, glucose, and fucose only showed a weak affinity for LSBPs. Two other sugars (mannose and rhamnose) that were also included in the analysis exhibited even weaker affinities. From these five sugars, galactose exhibited the highest affinity for LSBPs. Their affinity could be graded as follows: galactose $>$ glucose $>$ fucose $>$ mannose $>$ rhamnose. The binding intensities of the β - and α -anomeric forms of these sugars were also distinguished with the β -anomers showing higher affinity than their corresponding α -anomers in all cases. These studies revealed that PDAC cells contained galactose binding proteins, which bind strongly to the resin, thus requiring a high salt buffer containing galactose for elution. The difference in the binding intensities of monosaccharides to glycan binding proteins has been explained previously.⁴⁰ Galactose was shown to interact strongly with glycans compared to glucose due to the

difference in configuration at the C₄ position, for which the hydroxyl group is axially oriented in galactose but equatorial in glucose for the preferred ring conformation of the pyranose-ring form. Thus, the mode of interaction of these sugars with the aromatic amino acids (tryptophan, tyrosine, or phenylalanine) at the binding site differs, whereby galactose interacts with the aromatic residues at C₃-H, C₄-H, C₅-H, and C₆-H, while glucose interacts with the aromatic residues either at C₅-H and C₆-H or at C₁-H, C₅-H and/or C₄-OH.⁴⁰ The differences in the orientation of the C-OH of the sugar, with respect to the aromatic amino acids at the binding site, may explain the observed differences in the affinity of galactose and glucose for LSBPs as demonstrated by NMR. In this study, we delineated the affinity of selected monosaccharides for LSBPs. In their simplest forms, these sugars are linked by glycosidic bonds for forming disaccharides, such as lactose (glucose and galactose), sucrose (glucose and fructose), and maltose (glucose and glucose). As both, galactose and glucose exhibited moderate affinity for LSBPs, future studies will be required to evaluate the affinity of lactose for LSBPs in comparison to individual monosaccharide components.

The binding studies also revealed that the interaction of galactose with LSBPs was calcium dependent. This was evident as the addition of EGTA to a reaction mixture containing galactose and LSBPs diminished the observed binding strength. These results proved that LSBPs were composed of Ca²⁺-dependent galactose binding proteins akin to the C-type lectins and others known to bind glycans in a calcium-dependent manner. Some members of the annexin family, for instance, have been shown to bind glycosaminoglycans and sialoglycoproteins in the presence of calcium. The C-type lectins are known to bind sugars in a calcium-dependent manner through their CRD.^{41,42} Calcium is crucial in maintaining the configuration of binding modules of these lectins.⁴³

Focusing on proteins that exhibited sensitivity to galactose, we performed a search to identify motifs associated with sugar binding. Initially, we focused on the known CRD motif, but there was no match in any of the LSBPs. Then we concentrated on a motif, which was found in plant lectins that had been isolated by the lactosyl-resin used in this study. The search involved a flexible approach with amino acids previously known to bind galactose. Although the galactose binding sites do not have a unique recognition template for sugar binding, a number of amino acids have been identified, which are pivotal in sugar-protein interactions. In addition to tryptophan, tyrosine, and/or phenylalanine, which characterize the galactose binding site, we also considered arginine, aspartate, and glutamate, which can form a bidentate interaction with the neighboring hydroxyls of the sugar.³⁸ From our search, we identified six motifs, of which only two (d and f) appeared relevant because of their amino acid composition. Motif d was considered the most relevant for having two sugar binding amino acids (tyrosine and phenylalanine) as well as others (leucine and aspartate), which are known from the galactose binding domains of ricin and Rpx. The significance of these two motifs in the context of galactose binding requires experimental validation with

assays involving the deletion or substitution of specific amino acids predicted to bind this sugar. However, this aspect remains open for future investigations, since it is beyond the scope of this work.

Metastatic disease is the underlying cause of cancer mortality worldwide accounting for 90% deaths.⁴⁴ One of the defining features of metastasis is a complexity characterized by a sequence of cellular events cascading from the primary site to the colonization of distant organs.⁴⁵ Investigations on the role of genes in metastatic progression are limited by the quality of animal models mimicking the true nature of the disease as in humans. In the present study, we employed a previously established rat model for mimicking either the growth of the primary PDAC or of its liver metastasis. Clearly, a model giving access to both conditions in the same rat would be more helpful for determining the identity of LSBPs and their contribution to metastatic progression.

By relating the mRNA expression of LSBPs in different organ environments, we identified genes by the IPA, which were associated with tumor growth and progression. These genes were grouped under three functional terms of "cell movement," "cell signaling and interaction" and "cell development," which are intertwined with the cascade of metastasis. As reported in our previous work, the analysis of filtered genes using IPA program resulted in the identification of protein-glutamine γ -glutamyl transferase 2 (TG2) among other genes, which showed modulation of expression that was associated with metastasis.²⁸ TG2 was then characterized in vitro to confirm its role in tumor progression. Future studies will focus not only on the role of this gene in vivo but also on that of others for characterizing and validating their respective roles in tumor progression. In this respect, we identified other interesting genes in the pancreas tissue, which are associated with various functional annotations. Those related to "cell movement" include lipocalin2 (LNC2), interleukin 1 alpha (IL1A), and serpin family A member 3 (SERPINA3).

Laminins (LAMA3 and LAMC2) were associated with "cell signaling," while CUB domain-containing protein 1 (DCPI), integrin alpha-2 (ITGA2) CD44 antigen (CD44), and protein-glutamine γ -glutamyl transferase 2 (TGM2) featured in all three functional annotations, i.e. in "cell movement," "cell signaling," and "cell development." In the liver, SERCA-type calcium pumps (ATP2A2) and amyloid- β A 4 protein (APP) featured in cell signaling but were involved in cell movement in the lung environment. In addition, annexinA2 (ANXA2) and E-cadherin (CDH1) were related to all three functional annotations in both liver and lung environments.

Literature information supports these genes having a role in various cancers, including PDAC. LCN2, for instance, was confirmed as one of the most over-expressed genes in PDAC. Its expression in pancreatic intraepithelial neoplasia lesions correlated with disease progression.⁴⁶ Similarly, CD44 is significantly expressed in human pancreatic tumors as well as in their metastases and is associated with patient survival.⁴⁷ In the extracellular matrix, CD44 antigen interacts with numerous ligands like hyaluronic acid, osteopontin, chondroitin, collagen,

fibronectin, and sulfated proteoglycans.⁴⁸ ITGA2 is pivotal in tumor progression and metastasis through the enhancement of processes such as cell migration, invasion, proliferation, and survival.⁴⁹ ANXA2 was also shown to be involved in the development of pancreatic cancer.²⁰ CDH1 is associated with tumor progression, particularly in regard to epithelial mesenchymal transition (EMT) processes. When silenced in PDAC cells, CDH1 induced EMT resulting in metastatic progression *in vivo*.⁵⁰

Conclusion

In conclusion, we report the isolation and identification of LSBPs by a newly established workflow involving lactosyl affinity chromatography and LFQ MS. Some of these proteins include the well-known family of galectin and annexin proteins. Unexpectedly, the resin enriched other protein families, which have not been described before with respect to carbohydrate binding. These proteins contained a fraction that showed some degree of calcium dependency for their sugar binding. Although their interaction with monomeric sugars was moderate, their affinity to the lactosyl resin was remarkably strong, warranting a buffer containing high salt conditions. At the mRNA level, some of these proteins were differentially expressed in the primary and secondary organ environments of a PDAC liver metastasis model, a feature that indicates their possible role in metastasis. The affinity of glycan binding proteins makes them attractive in diagnostic evaluation and as targets for PDAC therapy.

Authors' contributions: MNS and MRB designed the experiments. MNS, KDK, AH-W and MZ performed experiments. MNS and MRB wrote the article.

ACKNOWLEDGEMENTS

We acknowledge Prof. Dr. Dieter Kübler and Dr. Martina Schnölzer for their expert opinion regarding this work, the genomics and proteomic core facility team for chip array and mass spectrometry for sample analysis, Dr. Marineta Kovacheva for advice on protein labeling, Nobert König for tips and training on the working of the ÄKTA. We further thank the Mass Spectrometry based Protein Analysis Unit of the Genomics and Proteomics Core Facility at the German Cancer Research Center (DKFZ), Heidelberg, Germany, for performing the LFQ MS analyses.


DECLARATION OF CONFLICTING INTERESTS

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was funded by the MOHEST Kenya/NACOSTI/DAAD scholarship award (MSN).

ORCID iD

Martin R Berger  <https://orcid.org/0000-0002-3783-5365>

SUPPLEMENTAL MATERIAL

Supplemental material for this article is available online.

REFERENCES

- Ohtsubo K, Marth JD. Glycosylation in cellular mechanisms of health and disease. *Cell* 2006;**126**:855–67
- Zhao L, Shah JA, Cai Y, Jin J. O-GlcNAc code' mediated biological functions of downstream proteins. *Molecules* 2018;**23**(8):1967
- Kronewitter SR, Marginean I, Cox JT, Zhao R, Hagler CD, Shukla AK, Carlson TS, Adkins JN, Camp DG 2nd, Moore RJ, Rodland KD, Smith RD. Polysialylated N-glycans identified in human serum through combined developments in sample preparation, separations, and electrospray ionization-mass spectrometry. *Anal Chem* 2014;**86**:8700–10
- Kaltner H, Gabius HJ. Sensing glycans as biochemical messages by tissue lectins: the sugar code at work in vascular biology. *Thromb Haemost* 2019;**119**:517–33
- Palaniappan KK, Bertozzi CR. Chemical glycoproteomics. *Chem Rev* 2016;**116**:14277–306
- Taylor ME, Drickamer K, Schnaar RL, Etzler ME, Varki A. Discovery and Classification of Glycan-Binding Proteins. In: rd, Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, Seeberger PH (eds) *Essentials of glycobiology*. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 2015, pp.361–72
- Hauselmann I, Borsig L. Altered tumor-cell glycosylation promotes metastasis. *Front Oncol* 2014;**4**:eCollection 2014
- Sharon N, Lis H. Lectins: cell-agglutinating and sugar-specific proteins. *Science* 1972;**177**:949–59
- Johannes L, Jacob R, Leffler H. Galectins at a glance. *J Cell Sci* 2018;**131**:pii:jcs208884
- Khan KA, McMurray JL, Mohammed F, Bicknell R. C-type lectin domain group 14 proteins in vascular biology, cancer and inflammation. *FEBS j* 2019;**286**:3299–332
- Ludwig AK, Kaltner H, Kopitz J, Gabius HJ. Lectinology 4.0: altering modular (ga)lectin display for functional analysis and biomedical applications. *Biochim Biophys Acta Gen Subj* 2019;**1863**:935–40
- Ghazarian H, Itoni B, Oppenheimer SB. A glycobiology review: carbohydrates, lectins and implications in cancer therapeutics. *Acta Histochem* 2011;**113**:236–47
- Miller MC, Zheng Y, Zhou Y, Tai G, Mayo KH. Galectin-3 binds selectively to the terminal, non-reducing end of beta(1->4)-galactans, with overall affinity increasing with chain length. *Glycobiology* 2019;**29**:74–84
- Belo AI, van der Sar AM, Tefsen B, van Die I. Galectin-4 reduces migration and metastasis formation of pancreatic cancer cells. *PLoS One* 2013;**8**:e65957
- Korniluk A, Kaminska J, Kiszlo P, Kemona H, Dymicka-Piekarska V. Lectin adhesion proteins (P-, L- and E-selectins) as biomarkers in colorectal cancer. *Biomarkers* 2017;**22**:629–34
- Phillips ML, Nudelman E, Gaeta FC, Perez M, Singhal AK, Hakomori S, Paulson JC. ELAM-1 mediates cell adhesion by recognition of a carbohydrate ligand, sialyl-Lex. *Science* 1990;**250**:1130–2
- Lowe JB, Stoolman LM, Nair RP, Larsen RD, Berhend TL, Marks RM. ELAM-1-dependent cell adhesion to vascular endothelium determined by a transfected human fucosyltransferase cDNA. *Cell* 1990;**63**:475–84
- Varki A, Kornfeld S. P-Type lectins. In: rd, Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, Seeberger PH (eds) *Essentials of glycobiology*. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 2015, pp.423–33
- Nakahara S, Raz A. Biological modulation by lectins and their ligands in tumor progression and metastasis. *Anticancer Agents Med Chem* 2008;**8**:22–36

20. Yoneura N, Takano S, Yoshitomi H, Nakata Y, Shimazaki R, Kagawa S, Furukawa K, Takayashiki T, Kuboki S, Miyazaki M, Ohtsuka M. Expression of annexin II and stromal tenascin C promotes epithelial to mesenchymal transition and correlates with distant metastasis in pancreatic cancer. *Int J Mol Med* 2018;**42**:821–30
21. Ishitsuka R, Kojima K, Utsumi H, Ogawa H, Matsumoto I. Glycosaminoglycan binding properties of annexin IV, V, and VI. *J Biol Chem* 1998;**273**:9935–41
22. Hage DS, Anguizola JA, Bi C, Li R, Matsuda R, Papastavros E, Pfaunmiller E, Vargas J, Zheng X. Pharmaceutical and biomedical applications of affinity chromatography: recent trends and developments. *J Pharm Biomed Anal* 2012;**69**:93–105
23. Voss C, Eyol E, Frank M, von der Lieth CW, Berger MR. Identification and characterization of riproximin, a new type II ribosome-inactivating protein with antineoplastic activity from *Ximenia americana*. *FASEB J* 2006;**20**:1194–6
24. Adwan H, Murtaja A, Kadhim Al-Tae K, Pervaiz A, Hielscher T, Berger MR. Riproximin's activity depends on gene expression and sensitizes PDAC cells to TRAIL. *Cancer Biol Ther* 2014;**15**:1185–97
25. Bayer H, Ey N, Wattenberg A, Voss C, Berger MR. Purification and characterization of riproximin from *Ximenia americana* fruit kernels. *Protein Expr Purif* 2012;**82**:97–105
26. Murtaja A, Eyol E, Xiaoqi J, Berger MR, Adwan H. The ribosome inhibiting protein riproximin shows antineoplastic activity in experimental pancreatic cancer liver metastasis. *Oncol Lett* 2018;**15**:1441–8
27. Kenner BJ, Chari ST, Maitra A, Srivastava S, Cleeter DF, Go VL, Rothschild LJ, Goldberg AE. Early detection of pancreatic cancer—a defined future using lessons from other cancers: a white paper. *Pancreas* 2016;**45**:1073–9
28. Sagini MN, Zepp M, Bergmann F, Bozza M, Harbottle R, Berger MR. The expression of genes contributing to pancreatic adenocarcinoma progression is influenced by the respective environment. *Genes Cancer* 2018;**9**:114–29
29. Dalvit C, Pevarello P, Tato M, Veronesi M, Vulpetti A, Sundstrom M. Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water. *J Biomol NMR* 2000;**18**:65–8
30. Hwang TL, Shaka AJ. Water suppression that works. Excitation sculpting using arbitrary wave-forms and pulsed-field gradients. *J Magn Reson Series A* 1995;**112**:275–9
31. Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* 2006;**1**:2856–60
32. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M. Pfam: the protein families database. *Nucleic Acids Res* 2013;**42**:D222–D30
33. Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* 1994;**2**:28–36
34. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* 2016;**11**:2301–19
35. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* 2014;**13**:2513–26
36. Tyanova S, Cox J. Perseus: a bioinformatics platform for integrative analysis of proteomics data in cancer research. *Methods Mol Biol* 2018;**1711**:133–48
37. Dalvit C. Homonuclear 1D and 2D NMR experiments for the observation of Solvent-Solute interactions. *J Magn Reson B* 1996;**112**:282–8
38. Taroni C, Jones S, Thornton JM. Analysis and prediction of carbohydrate binding sites. *Protein Eng* 2000;**13**:89–98
39. Pawar NJ, Diederichsen U, Dhavale DD. Multivalent presentation of carbohydrates by 3(14)-helical peptide templates: synthesis, conformational analysis using CD spectroscopy and saccharide recognition. *Org Biomol Chem* 2015;**13**:11278–85
40. Sujatha MS, Balaji PV. Identification of common structural features of binding sites in galactose-specific proteins. *Proteins* 2004;**55**:44–65
41. Gorelik EU, Raz A. On the role of cell surface carbohydrates and their binding proteins (lectins) in tumor metastasis. *Cancer Metastasis Rev* 2001;**20**:245–77
42. Sugawara HK, Kurisu G, Fujimoto T, Aoyagi H, Hatakeyama T. Characteristic recognition of N-acetylgalactosamine by an invertebrate C-type lectin, CEL-I, revealed by X-ray crystallographic analysis. *J Biol Chem* 2004;**279**:45219–25
43. Drickamer K. Engineering galactose-binding activity into a C-type mannose-binding protein. *Nature* 1992;**360**:183–6
44. Dillekas H, Rogers MS, Straume O. Are 90% of deaths from cancer caused by metastases? *Cancer Med* 2019;**8**:5574–6
45. Guan X. Cancer metastases: challenges and opportunities. *Acta Pharm Sin B* 2015;**5**:402–18
46. Gomez-Chou SB, Swidnicka-Siergiejko AK, Badi N, Chavez-Tomar M, Lesinski GB, Bekaii-Saab T, Farren MR, Mace TA, Schmidt C, Liu Y, Deng D, Hwang RF, Zhou L, Moore T, Chatterjee D, Wang H, Leng X, Arlinghaus RB, Logsdon CD, Cruz-Monserrate Z. Lipocalin-2 promotes pancreatic ductal adenocarcinoma by regulating inflammation in the tumor microenvironment. *Cancer Res* 2017;**77**:2647–60
47. Li L, Hao X, Qin J, Tang W, He F, Smith A, Zhang M, Simeone DM, Qiao XT, Chen ZN, Lawrence TS, Xu L. Antibody against CD44s inhibits pancreatic tumor initiation and postradiation recurrence in mice. *Gastroenterology* 2014;**146**:1108–18
48. Zhao S, Chen C, Chang K, Karnad A, Jagirdar J, Kumar AP, Freeman JW. CD44 expression level and isoform contributes to pancreatic cancer cell plasticity, invasiveness, and response to therapy. *Clin Cancer Res* 2016;**22**:5592–604
49. Desgrosellier JS, Cheresch DA. Integrins in cancer: biological implications and therapeutic opportunities. *Nat Rev Cancer* 2010;**10**:9–22
50. von Burstin J, Eser S, Paul MC, Seidler B, Brandl M, Messer M, von Werder A, Schmidt A, Mages J, Pagel P, Schnieke A, Schmid RM, Schneider G, Saur D. E-cadherin regulates metastasis of pancreatic cancer in vivo and is suppressed by a SNAIL/HDAC1/HDAC2 repressor complex. *Gastroenterology* 2009;**137**:361–71

(Received December 6, 2019, Accepted February 12, 2020)