PNAS
www.pnas.org

Supplementary Information for

Mapping glycan-mediated galectin-3 interactions by live cell proximity labeling

Eugene Joeh, Timothy O'Leary, Weichao Li, Richard Hawkins, Jonathan R. Hung, Christopher G. Parker, Mia L. Huang

Correspondence to Mia L. Huang
Email: miahuang@scripps.edu

**This PDF file includes:**

Supplementary text
Figures S1 to S17
Tables S1 to S2



www.pnas.org/cgi/doi/10.1073/pnas.2009206117

**Supplementary Information Text**

**Methods**

**Generation of PX-Gal3 fusion constructs**

The full-length DNA insert for PX-Gal3 and PX-Gal3$_{\Delta 116}$ were synthesized by Genscript (New Jersey, USA). Starting from the human galectin-3 cDNA sequence, an optimized sequence that has a codon adaptation index of 0.96 was determined by Genscript using a proprietary OptimumGene algorithm and ligated into the EcoRI/NotI cloning sites of a pETDuet-1 vector. Upon transformation into BL21 (DE3) pLysS chemically competent *E. coli* and antibiotic selection using ampicillin LB agar plates (100 µg/mL), single cell colonies were picked, and the selected clones grown in LB with ampicillin (100 µg/mL). To induce protein expression, the transformed E. coli was grown in 5 mL starter cultures of LB with ampicillin overnight. The overnight culture was then added to 500 mL of fresh LB with ampicillin and shaken at 37 °C until $OD_{600}$~0.6. The cultures were then induced with IPTG (1 mM, 4 hours at 37 °C) to initiate protein expression in LB containing ampicillin (100 µg/mL). Following centrifugation (3300 xg, 15 min at 4 °C) and lysis of the cell pellet in lysis buffer on ice (5 mL; 20 mM imidazole in 2x PBS (274 mM NaCl, 5.4 mM KCl, 20 mM $Na_2HPO_4$, 3.6 mM $KH_2PO_4$) *via* sonication (1 sec ON, 3 sec OFF, 30 sec TOTAL ON, 20% amplitude), the crude protein supernatant was purified using Ni-NTA beads (1 hr at 4 °C) or an AKTA Start FPLC system for His-tag mediated purification. The column-bound proteins were washed with 25 mM imidazole and eluted using 250 mM imidazole, and further purified using α-lactose agarose beads. Upon elution with 250 mM lactose, the resulting proteins were dialyzed against PBS to remove excess lactose and stored in PBS supplemented with 2mM EDTA and 0.05% Tween-20, flash frozen and stored at -80 °C. SDS-PAGE and immunoblotting were used to verify expression and purity.

**Gels and Western blots**

Unless otherwise stated, SDS-PAGE analysis of proteins was performed using either manually cast or pre-cast protein gels (Biorad). Sample protein concentrations were normalized with 660nm Protein Assay (Pierce) and loaded using Laemmli buffer with 2-mercaptoethanol and boiled (95 °C, 5 min). Running conditions include 85-100 mV in Tris-Glycine-SDS buffer for 60-90 min. Transfers were performed using nitrocellulose or PVDF membranes (60 min, 100 mA). Blocking was performed with 5% BSA/TBST, incubations with primary antibodies were performed overnight at 4 °C, with rocking. Secondary antibody incubations were performed at RT for 1-2 hours.

**ELISAs to evaluate binding affinities**

The glycoprotein(s) of interest were immobilized at defined concentrations (specified in each context) in pH 9.6 carbonate buffer overnight at 4 °C onto high-binding plate (Nunc Maxisorp). Next day, plates were washed 3x in PBST, followed by blocking with 2% BSA/PBST (1 hr, RT). Recombinant galectin-3 (Biolegend # 599706) was then incubated either for 1-2 hrs at RT or overnight at 4 °C, with rocking. Following washing 3x with PBST, plates were incubated with anti-galectin-3 antibody, followed by an HRP-conjugated secondary antibody. Following subsequent washing steps, plates were developed with TMB substrate and quenched with 2N $H_2SO_4$. For ELISAs with asialofetuin (Sigma Aldrich # A4781), 5 mg/mL was immobilized. For ELISAs with multiple targets, equimolar (40 nM) amounts of the recombinant human proteins were immobilized onto the plate. Experiments were repeated with at least two independent biological replicates.

**Transient over-expression of PX-Gal3**

Adherent LX-2 cells seeded overnight at 50-80% confluency (1.25 x $10^5$ cells/well) were washed twice in DPBS and incubated with plasmid DNA pre-complexed (according to manufacturer's instructions) with Fugene HD and returned to the incubator. The DNA sequence encoding PX-Gal3 was inserted into a bicistronic pIRES2-AcGFP1 vector (Takarabio # 632435). GFP fluorescence was readily observed by microscopy 24 and 48 hours after incubation, and

subsequent proximity ligation steps were performed as before, while omitting the initial protein incubation step.

**Isolation of peripheral blood mononuclear cells (PBMC) from whole peripheral blood samples**
Whole peripheral blood samples were obtained from healthy deidentified donors from buffy coats supplied by OneBlood. PBMC was isolated by Histopaque gradient centrifugation. For 50 mL of whole blood, 12.5 mL of Histopaque density gradient media was added to a 50 mL Falcon Tube for each sample and 25 mL of whole blood was carefully added to each tube and centrifugation was performed at 2,000 rpm for 30 min at RT. The PBMC and the plasma was collected and combined into a 50 mL tube and centrifuged at 1500 rpm for 8 min at 4 °C. The supernatant was aspirated, and the pellet was resuspended in 10 mL of cold DBPS and centrifuged at 1500 rpm for 8 min at 4 °C. The supernatant was aspirated and another 10 mL of cold DPBS was added for a repeated wash. Once the cells were counted, they were spun down, the supernatant was removed, and the cells were resuspended to a final concentration of 50 x $10^6$ cells/mL in RPMI media (no FBS). Freshly isolated PBMC were used for all experiments.

**N-glycan glycomics MS of LX-2 cell surfaces**
Cell surface N-glycomics analysis was based on a published procedure (1), starting from LX-2 cells. The cells were cultured to ~ 90% confluency in 10 $cm^2$ plates to yield approximately 10 x $10^6$ cells per plate. Cells were harvested and washed four times by re-suspending in 4 mL ice-cold PBS followed by centrifugation (600 rpm, 10 min, 4 °C). Free glycopeptides were obtained by re-suspending cell pellets in 1.0 mL of 0.25% trypsin in PBS (Quality Biological, #118086721) and shaking at 250 rpm for 15 min at 37 °C. Samples were placed in ice for five min followed by centrifugation (8,000 xg, 10 min, 4 °C). The resultant glycopeptide-containing supernatant was heated at 100 °C for five min to deactivate trypsin. After cooling for 10 min on ice, glycans were released by treatment with 2.0 µL of PNGase F (OmpA3-PNGaseF-TEV-His6 was a gift from Martin Caffrey (Addgene plasmid # 114274 ; http://n2t.net/addgene:114274 ; RRID:Addgene_114274) for 20 hr at 37 °C (7.5 µg/µL in 25 mM Tris pH 7.2 with 50% glycerol). Following the treatment with PNGase F, peptides were removed with a C18 cartridge (Thermo Scientific HyperSep, #03251257). Column cartridges were conditioned by successive treatment with methanol, 5% aqueous acetic acid, *n*-propanol, and 5% aqueous acetic acid. Samples were first acidified to 5% acetic acid, centrifuged to remove debris (16,000 xg, 10 min, 4 °C), then loaded to the column and eluted with 5.0 mL 5% aqueous acetic acid. Column flow-through was combined with aqueous eluate and lyophilized. Experiments were repeated across two independent biological replicates.

**Preparation of PX-Gal3 enriched glycans**
Approximately 400 µL of the experimentally treated beads were suspended in 1.5 mL 0.25% trypsin in PBS (Quality Biological, #118086721) and shaken at 250 rpm for 30 min at 37 °C. This was then heated at 100 °C for five min to denature the trypsin. After cooling on ice for five min the sample was treated with 5.0 µL of PNGase F for 20 hr at 37 °C (in-house, Addgene #114274; 7.5 µg/µL in 25 mM Tris pH 7.2 with 50% glycerol) and incubated at 37 °C for 20 hr. The glycan-containing supernatant was separated from the beads, and peptides were removed with a C18 cartridge (Thermo Scientific HyperSep, #03251257) conditioned by successive treatment with methanol, 5% aqueous acetic acid, *n*-propanol, and 5% aqueous acetic acid. The sample was centrifuged to remove debris (14,000 xg, 10 min, 4 °C), loaded onto the column and eluted with 5.0 mL 5% aqueous acetic acid. The aqueous flow through and eluate were combined and lyophilized.

**Immunoprecipitation of endogenous basigin from LX-2 cells**
Confluent plates of 15 $cm^2$ plates of live LX-2 hepatic stellate cells were scraped, and cells were pelleted by centrifugation (650 xg, 3 min, 4 °C). The cells were washed with cold PBS twice and then lysed with RIPA + Protease Inhibitor Cocktail (1 mL) for 15 min on ice. The lysate was pre-clarified by centrifugation (16,000 xg, 10 min, 4 °C) and the supernatant was added to pre-

washed Protein G Sepharose Fast Flow beads (100 mL, GE) for 10 min at 4 °C on a rocker to remove any non-specific bead binding. The beads were pelleted by centrifugation (1,450 xg, 10 min, 4 °C). The immunocomplex was captured by adding the lysate and 10 µg of anti-basigin antibody (Novus Biologicals # NB500-430) concurrently with pre-washed Protein G Sepharose Fast Flow beads (100 µL, GE) overnight at 4 °C on a rocker. The beads were pelleted by centrifugation (1,450 xg, 10 min, 4 °C) and the supernatant was collected and put aside. The beads were then washed with DPBS (1 mL x 3). Basigin was eluted from the antibody on the beads by adding 100 µL of elution buffer (0.2 M glycine, pH 2.6) for 10 min at RT. The beads were pelleted and the supernatant was removed and quickly quenched with an equal volume of Tris buffer (1 M, pH 9.0). The elution and quenching were repeated two more times, for a total of three elutions and kept separate. 4x Laemmli buffer was added to the beads and heated at 95 °C for 10 min to dissociate any remaining immunocomplex off the beads.

**Harvesting N-glycans from immunoprecipitated basigin**

To prepare and purify the N-glycans captured by basigin immunoprecipitation, the eluate from the BSG-antibody beads (as well as positive and negative controls) was lyophilized, re-dissolved in 300 µL water, and proteins were denatured by heating at 80 °C for 30 minutes. After cooling on ice, the samples were digested with 2.0 µL of PNGase F for 20 hr at 37 °C (in-house, Addgene #114274; 7.5 µg/µL in 25 mM tris pH 7.2 with 50% glycerol). Free glycans were separated from protein with a 10 kDa molecular weight cutoff centrifugal filter unit (Millipore Amicon Ultra-0.5 #UFC5010). The glycan-containing filtrate was desalted with a PD minitrap G-10 column (Cytiva) and dried under vacuum.

**Preparation of released N-glycans for mass spectrometry analysis**

The released N-glycans from the cell surface, PX-Gal3, and basigin-IP samples were prepared for LC/MS analysis by reduction followed by permethylation. Reduction was carried out by dissolving the dried glycans in 200 µL of 10 mg/mL sodium borohydride (Oakwood Chemical # 042896) in 1 M $NH_4OH$ solution (Sigma-Aldrich # 09859) and heating at 60 °C for 1 hr. After cooling to RT, the samples were desalted using a PD minitrap G-10 column (Cytiva) and lyophilized. Permethylation was based on a published procedure, following the conventional milliliter scale protocol (2). Briefly, a DMSO/NaOH slurry was first prepared by dissolving 300 µL 50 wt % aqueous solution sodium hydroxide (Acros Organics # 25986025) in 600 µL methanol (Fisher Scientific # AF544). This was shaken with 8.0 mL DMSO and centrifuged (3,000 xg, 3 min, RT). The mixture was decanted, 8.0 mL fresh DMSO was added, mixed well, and centrifuged again. This was repeated three more times to yield a translucent NaOH pellet that was dissolved in 3 mL DMSO and used immediately. Prior to methylation, the dried, reduced glycans were dissolved in 200 µL DMSO and pre-incubated for 10 min at 37 °C followed by 20 min shaking at RT. Then 350 µL freshly vortexed NaOH/DMSO slurry was added, followed by an additional 5 min of mixing. To this mixture, 100 µL methyl iodide (Sigma Aldrich # 289566) was added and the sample was vigorously mixed (10 min, RT). The reaction mixture was mixed with 1 mL water and nitrogen gas was bubbled through the solution to remove excess methyl iodide. The permethylated glycans were extracted into 2 mL dichloromethane (Sigma Aldrich # D65100), washed three times with 750 µL water, and dried under a stream of nitrogen.

**LC-MS/MS instrument method and analysis for glycomics**

Permethylated glycan samples were dissolved in 20% MeCN (aq.) prior to LC-MS/MS analysis and stored at -20 °C. Chromatography was performed on an Ultimate 3,000 UHPLC (Thermo Scientific) equipped with a reversed phase Waters Acquity Peptide BEH C18 column (150 mm x 2.1 mm inner diameter, 130 Å particle size). Column temperature was maintained at 50-55 °C. Mobile phase A consisted of water (18.2 MΩ) with 0.1% formic acid (Thermo Scientific #28905), and mobile phase B was MeCN (Fisher #A994) with 0.1% formic acid. Flow rate was maintained at 100 or 120 µL/min. The elution gradient was from 20% buffer B to 60% buffer B over span of at least 90 min, increased up to 95% B over the final 15 min. The UHPLC was interfaced with an LTQ XL ETD Hybrid Ion Trap-Orbitrap ESI mass spectrometer (Thermo Scientific). The mass spectrometer was operated in positive ion mode in the mass range of 700 m/z – 2,000 m/z with a

spray voltage of 3.5 kV. $MS^2$ data was collected in data-dependent acquisition mode, with the top four most abundant ions with signal intensity over 10,000 counts selected from the full $MS^1$ scan for collision induced dissociation (CID). Calibration was performed regularly (Thermo LTQ ESI Positive Ion Calibration Solution, #88322) to ensure accuracy of 10 ppm or less. Data was processed with XCalibur 2.1 (Thermo Scientific). Relative proportions of glycans were calculated using the area of the extracted ion chromatogram (EIC) for all adducts and charge states for a particular glycan or by integrating the total ion count chromatographic peak (for basigin immunoprecipitation experiments). Glycan compositions and structures were identified using Simglycan (3) and GlycoWorkbench (4).

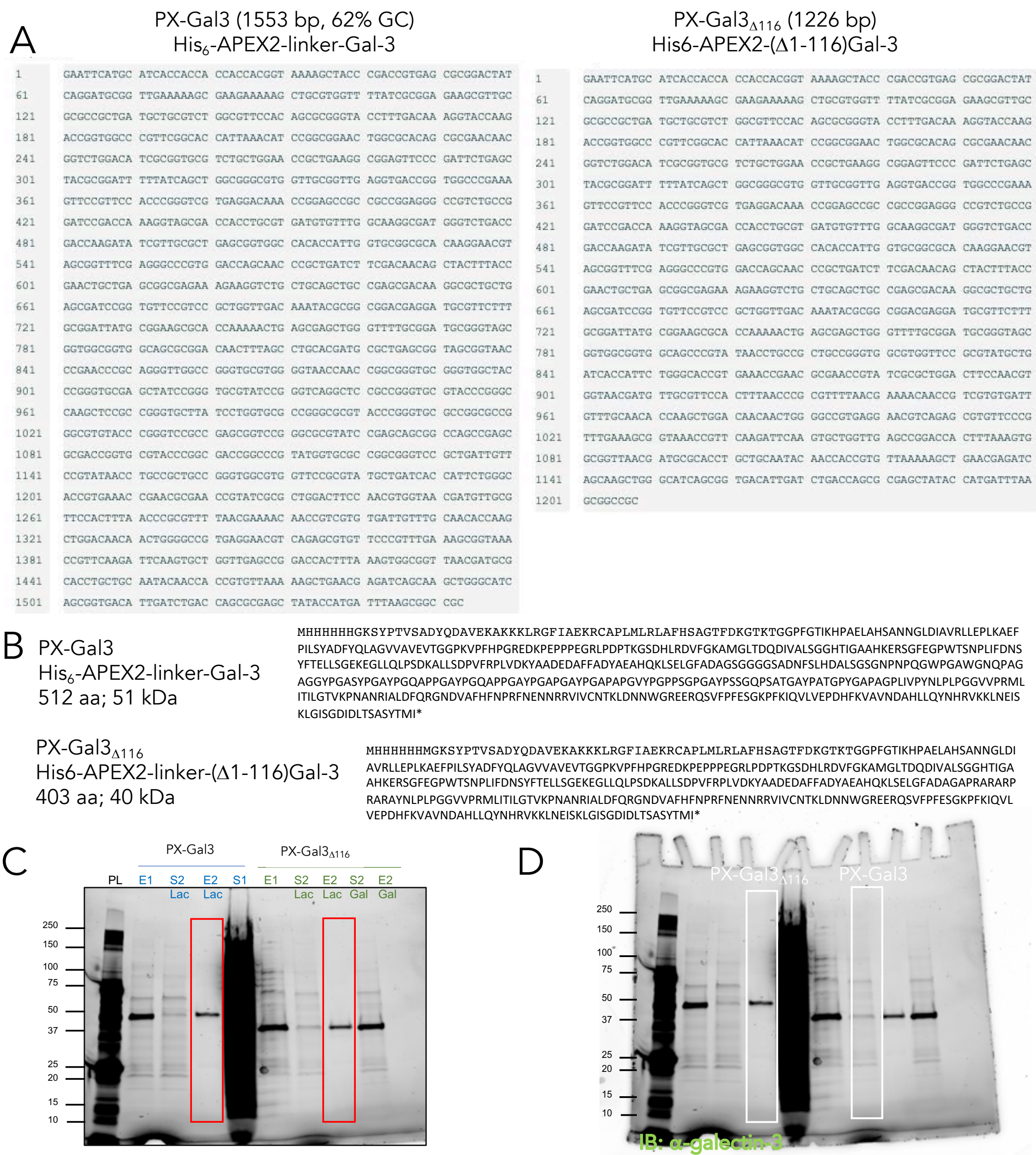



**Fig. S1. Expression of APEX2-galectin-3 fusion constructs. (A)** DNA sequence of inserted PX-Gal3 and PX-Gal3$_{\Delta 116}$ constructs, including nucleotide sequences used for cloning. **(B)** Predicted translated protein sequences. **(C)** SDS-PAGE analysis of proteins eluted from nickel beads or columns (E1) and further purification with lactose beads (E2). **(D)** Western blotting for galectin-3 shows that the resulting constructs are immuno-positive.

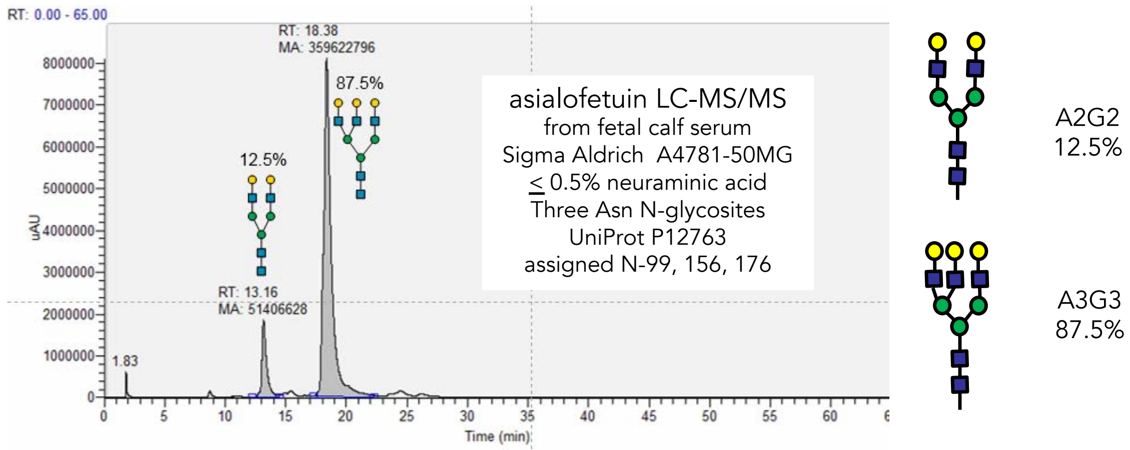



**Fig. S2. LCMS analysis of N-glycans present in commercial asialofetuin used in ELISA experiments (Fig. 2B).** Terminal galactose residues (yellow circle), which are high affinity ligands for galectin-3, are abundant.

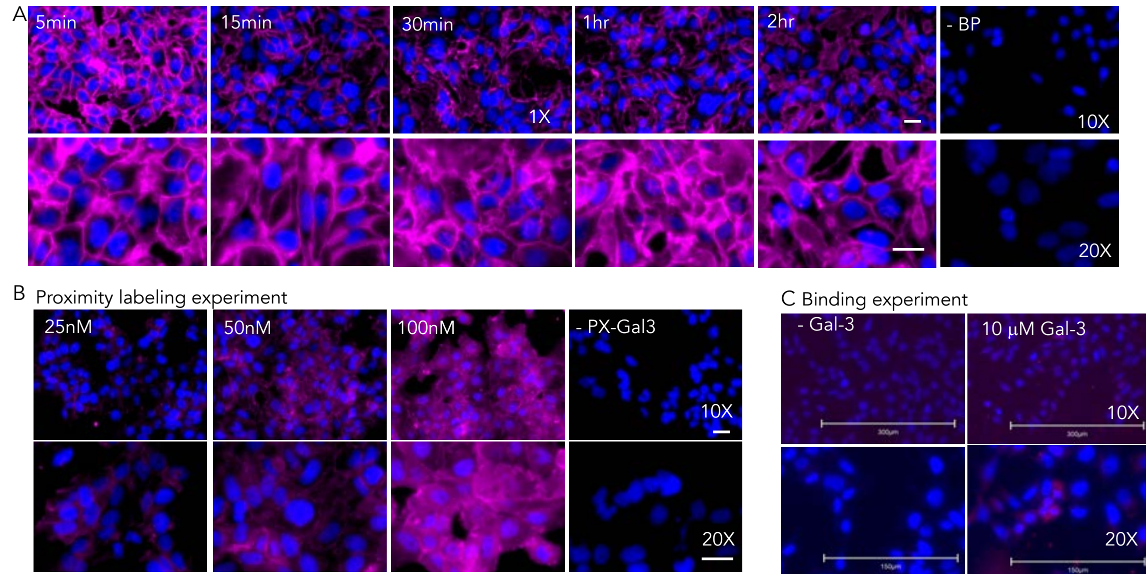



**Figure S3. Fluorescence micrographs to show the dependence of incubation time and protein concentration towards in situ proximity labeling.** Cells were stained for biotinylated interactors with Cy5-Streptavidin (purple) and nuclei are stained with Hoescht 33342 (blue). Top and bottom panels show images at 10x and 20x magnification, respectively. Scale bars: 30 μm **(A)** Five-minute incubation of LX-2 cells with PX-Gal3 (100 nM) is sufficient to achieve robust labeling. **(B)** There is a dose-dependent increase in labeling of interactors PX-Gal3 following 30 mins of protein incubation (37 °C). **(C)** Detection of the binding of galectin-3 (10 μM) to LX-2s requires significantly higher protein concentrations and incubation periods (overnight, 4 °C).

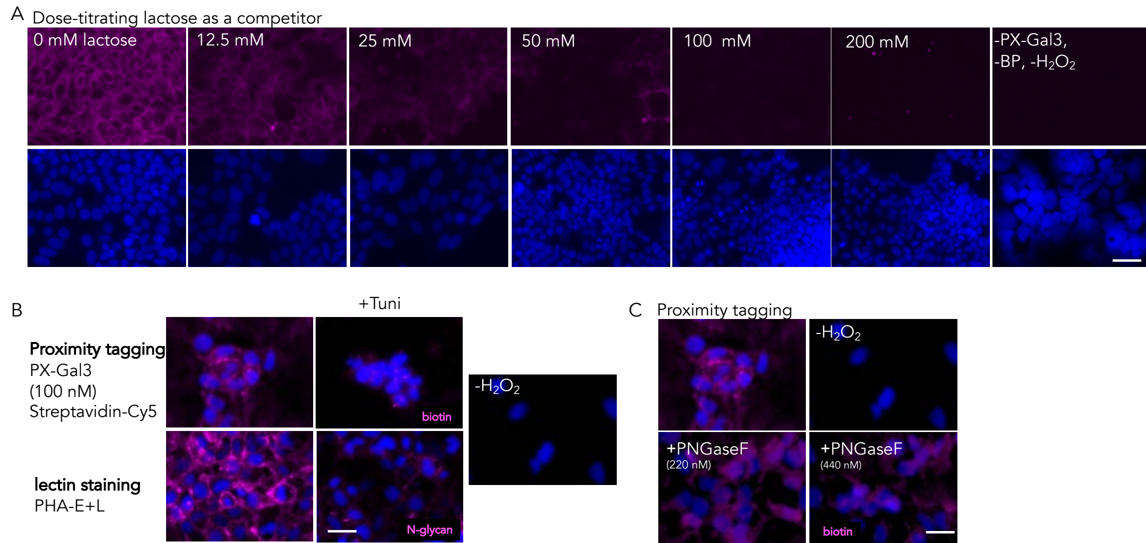
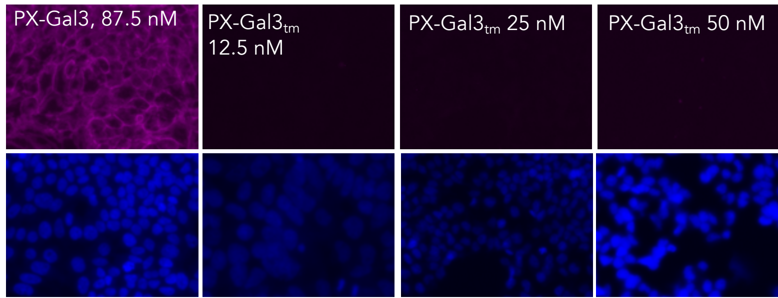



**Figure S4. Glycan-dependent interactions are highly represented in the labeling of HSCs with PX-Gal3 (A)** There is significant loss of fluorescence following co-incubation of PX-Gal-3 (full, 100 nM, 30 min) with lactose, even at 12.5 mM. Full competition was observed at 100 mM lactose concentrations. **(B**) LX-2s pre-treated with tunicamycin (2 μg/mL; 24 hr), an N-glycosylation inhibitor, showed reduced staining of cell surface glycans by lectin staining (biotinylated PHA-E+L 20 μg/mL, followed by streptavidin-Cy5), and also reduced fluorescence by proximity tagging (PX-Gal3, 100 nM, 30 min). **(C)** LX-2s pre-treated with PNGAse F (24 hr) at the indicated concentrations also showed reduced proximity tagging.

Proximity labeling with triple mutant PX-Gal3$_{tm}$
(200 nM; R144S, R186S, G182A)



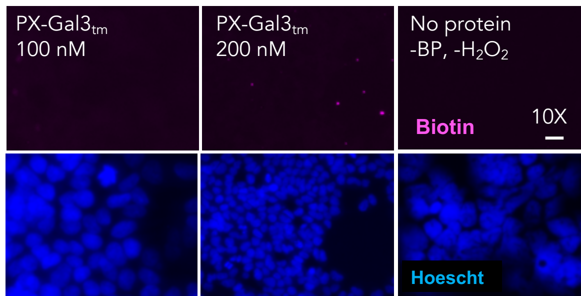




**Figure S5.** A mutant of PX-Gal3 (PX-Gal-3$_{tm}$), where three of the most important residues required for binding glycans in galectin-3 are mutated: R144S, R186S, G182A, show poor proximity labeling, indicating that glycan-mediated interactions are required for binding to LX-2s.

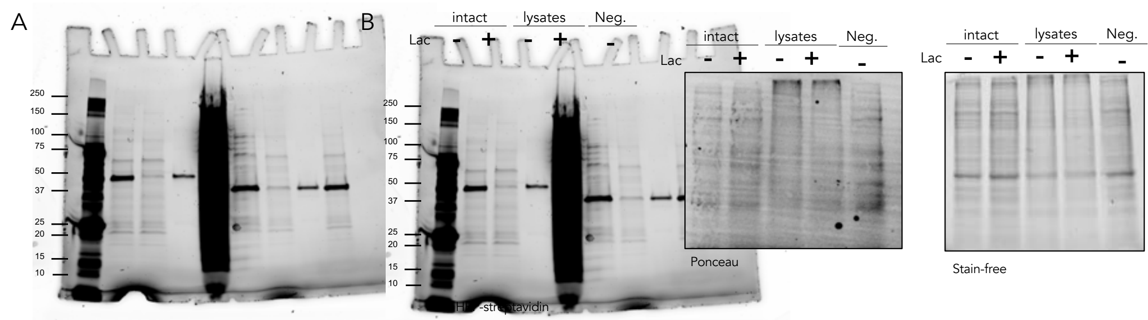



**Figure S6. Comparison of proximity labeling method in intact cells versus cell lysates.** 10 μg of total protein lysate (determined by the Lowry method) from each condition was loaded onto the gel. Only a subset of proteins is labeled in intact cells compared to lysates. **(A)** Full-length Western blot to complement Fig. 2E. **(B)** Repeat of the same experiment, showing relatively similar protein loading among different conditions by Ponceau staining and imaging by Stain-Free imaging (photoactivation followed by UV imaging).

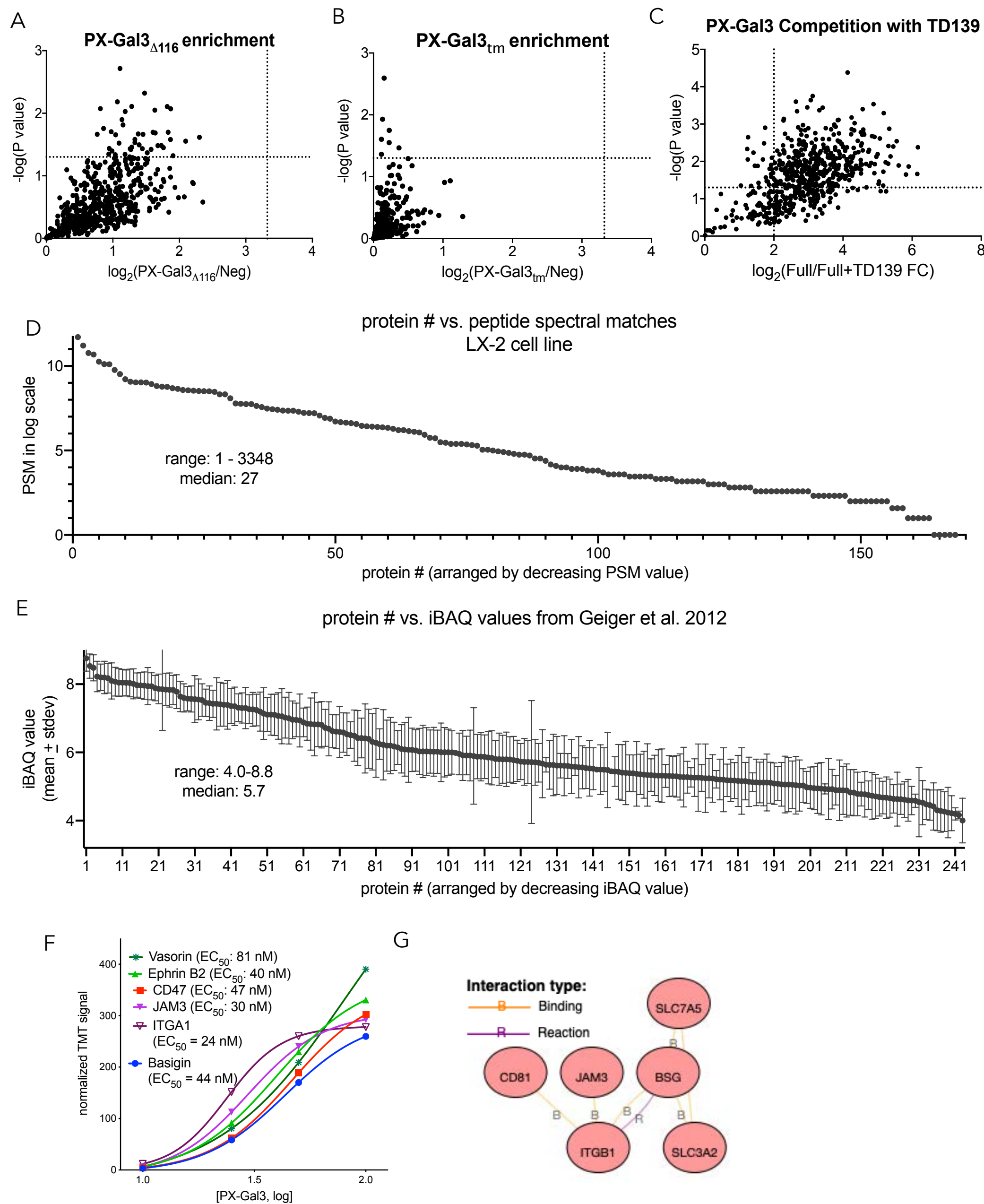



**Figure S7. Additional data from quantitative proteomics analyses collected from exogenous incubations. (A)** PX-Gal3$_{\Delta116}$ (100 nM) does not significantly enrich proteins over a negative control (Neg). **(B)** PX-Gal3$_{tm}$ does not significantly enrich proteins over Neg. **(C)** Interactions of PX-Gal3 are competed by co-incubation with TD139 (15.4 μM). **(D)** Peptide spectral matches (PSMs) for proteins enriched by PX-Gal3 (100 nM) show that they represent a wide range of abundance in LX-2 cells. **(E)** Mean iBAQ (intensity-based absolute quantification) values of proteins enriched by PX-Gal3 also show that they comprise a range of protein abundances. Mean iBAQ values were calculated as the average iBAQ value across eleven cell lines, as determined by Geiger et al. (5) **(F)** Apparent binding affinity values ($EC_{50}$) determined for selected proteins using quantitative MS-based proteomics. **(G)** Integrated Pathway Analysis of binding relationships based on PX-Gal3 enriched proteins.

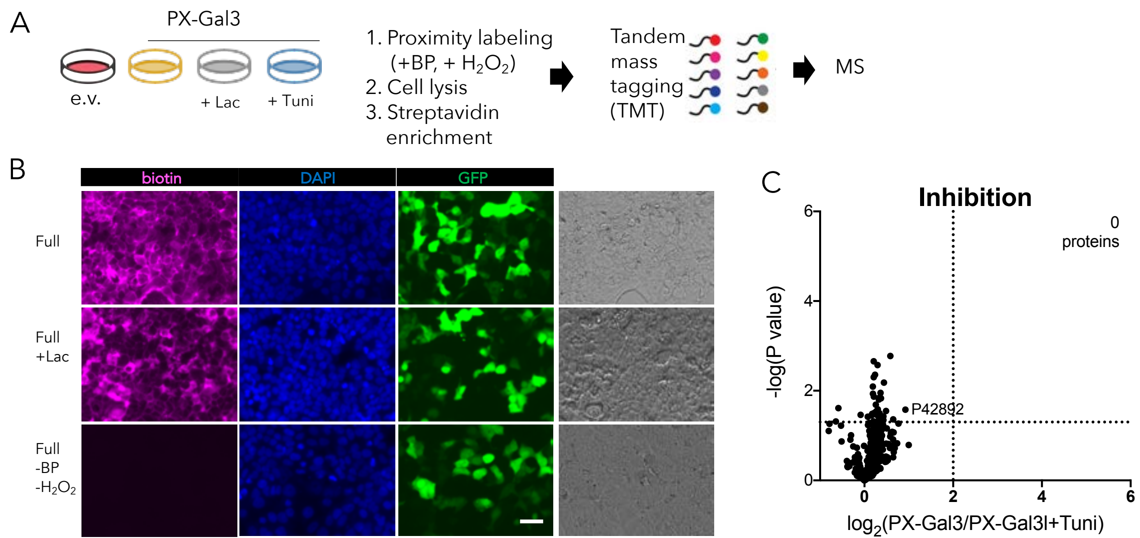



**Figure S8. Additional data collected from transfection experiments**. **(A)** LX-2 cells were transiently transfected with either empty vector (e.v.) or vector carrying PX-Gal3. Following transfection, cells were additionally treated with excess lactose (100 mM) or tunicamycin (10 μg/mL) for 24 hours, prior to performing proximity labeling. **(B)** In contrast to the exogenous treatments of the fusion proteins, addition of exogenous lactose to cells failed to significantly compete for interactions, most likely due to its intracellular impermeability. Scale bar: 20 μm **(C)** Treatment with tunicamycin, an inhibitor of N-linked glycosylation, did not compete for PX-Gal3 interactions.

| | Full-length protein | Recombinant protein | | | |
|---|---|---|---|---|---|
| | Assigned N-glycosites of full-length protein (UniProt) | Accession # | Cat. # | Predicted MW /Observed MW | Sequence of Recombinant Protein |
| **basigin-2-Fc** | N-44, N-160 of basigin-2 (equivalent to N-160 and 268 of basigin-1) | NP_940991.1 | 10186-H02H | 46.8 kDa/ 58-65 kDa | extracellular domain (Met 1-His 205) of human CD147 basigin-2 precursor was expressed with the fused Fc region of human IgG1 at the C-terminus. *** We confirmed the presence of N-glycans on residues Asn 44 and 152 of basigin-2 and Asn-73 and 299 of the Fc domain*** |
| **CD81** | none | NP_004347.1 | 14244-HNCH | 9.9 kDa/ 10 kDa | (Phe113-Lys201) was expressed with two additional amino acids (Gly&Pro) at the N-terminus. |
| **vasorin** | N-101, 117, 273, 500, 528 | Q6EMK4 | 13854-H08H | 61.1 kDa / 66-76 kDa | (Met1-Pro575) was expressed with a polyhistidine tag at the C-terminus. |
| **ephrin-B1** | N-139 | NP_004420.1 | 10894-H08H | 24.5 kDa/ 38 kDa | extracellular domain (Met 1-Lys 237) was fused with a polyhistidine tag at the C-terminus. |
| **neuroplastin** | N-171, 197, 229, 284, 296, 317 | NP_059429.1 | 15881-H08H | 23.2 kDa | (Met1-Leu221) was expressed with a polyhistidine tag at the C-terminus. *** missing 4 glycosylation sites*** |
| **CD9** | N-52, 53 | NP_001760.1 | 11029-H08H | 10 kDa / 11 kDa | second extracellular domain (Ser 112-Ile 195) of human CD9 was fused with a polyhistidine tag at the C-terminus and a signal peptide at the N-terminus. *** missing both glycosylation sites *** |
| **CD47** | N-23, 34, 50, 73, 111, 206 | NP_942088.1 | 12283-HCCH | 14.5 kDa | (Met1-Pro139) was expressed with six amino acids (LEVLFQ) at the C-terminus. *** missing last Asn-206 glycosylation site *** |

**Table S1. Commercially purchased recombinant proteins used for ELISAs.** Note that these proteins may not reflect endogenous protein sequences and glycosylation states found in LX-2 cells.

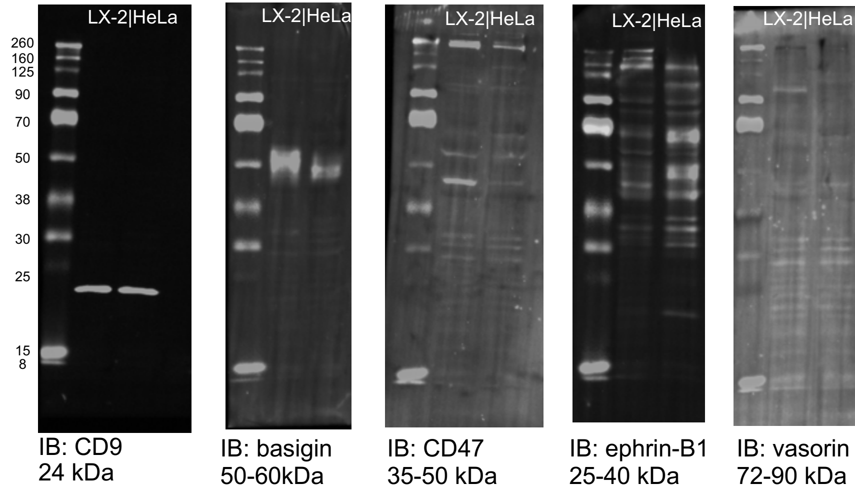



**Figure S9. Western blots showing endogenous expression of various proteins in LX-2 cells.** 10 μg each of cell lysates from LX-2 HSCs or from HeLa (human cervical cancer cells).

| | rhBasigin | rhCD81 | rhvasorin | rhEphrin-B1 | rh-Neuroplastin | rh-CD9 | rh-CD47 |
|---|---|---|---|---|---|---|---|
| log(agonist) vs. response (three parameters) | | | | | | | |
| Best-fit values | | | | | | | |
| Bottom | 0.03046 | 0.007451 | 0.007233 | -0.007259 | 0.03977 | -0.02879 | 0.02969 |
| Top | 1.131 | 1.097 | 1.115 | 1.134 | 1.191 | 1.208 | 1.192 |
| LogEC50 | 3.377 | 3.048 | 3.217 | 3.114 | 3.619 | 2.851 | 3.613 |
| EC50 | 2385 | 1117 | 1649 | 1302 | 4156 | 709.7 | 4100 |
| Span | 1.101 | 1.089 | 1.107 | 1.141 | 1.151 | 1.237 | 1.162 |
| 95% CI (profile likelihood) | | | | | | | |
| Bottom | 0.004798 to 0.05585 | -0.03674 to 0.05078 | -0.01808 to 0.03228 | -0.06184 to 0.04618 | -0.006713 to 0.08528 | -0.1294 to 0.06803 | 0.01379 to 0.04547 |
| Top | 1.070 to 1.197 | 1.026 to 1.172 | 1.065 to 1.167 | 1.041 to 1.233 | 1.042 to 1.385 | 1.078 to 1.348 | 1.137 to 1.251 |
| LogEC50 | 3.284 to 3.471 | 2.919 to 3.177 | 3.137 to 3.298 | 2.961 to 3.267 | 3.415 to 3.831 | 2.630 to 3.067 | 3.544 to 3.683 |
| EC50 | 1925 to 2959 | 829.6 to 1503 | 1370 to 1986 | 915.0 to 1849 | 2602 to 6775 | 426.7 to 1166 | 3496 to 4818 |
| Goodness of Fit | | | | | | | |
| Degrees of Freedom | 18 | 18 | 18 | 18 | 18 | 18 | 18 |
| R squared | 0.9915 | 0.9821 | 0.9931 | 0.9731 | 0.9673 | 0.9415 | 0.9961 |
| Sum of Squares | 0.02416 | 0.06055 | 0.02197 | 0.09772 | 0.08471 | 0.2838 | 0.01006 |
| Sy.x | 0.03663 | 0.05800 | 0.03493 | 0.07368 | 0.06860 | 0.1256 | 0.02364 |
| | | | | | | | |
| Number of points | | | | | | | |
| # of X values | 21 | 21 | 21 | 21 | 21 | 21 | 21 |
| # Y values analyzed | 21 | 21 | 21 | 21 | 21 | 21 | 21 |

**Table S2. Curve fitting analyses of ELISA data in Figure 4B.**

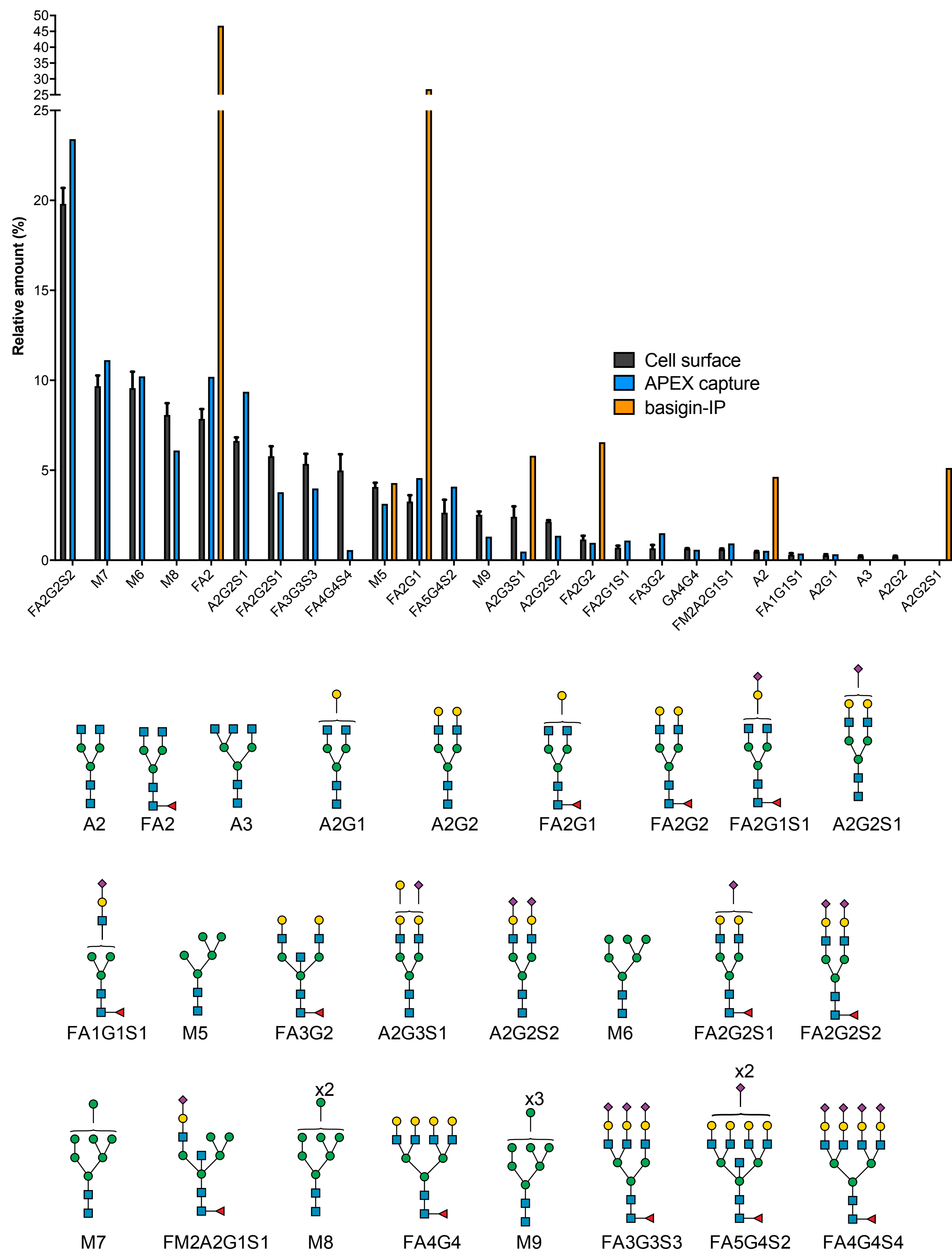



**Figure S10. Full composition analysis of the most abundant N-glycans found on LX-2 cell surfaces, captured by PX-Gal3, or from immunoprecipitated (IP) basigin.** Top panel represents the relative amounts of each N-glycan found in LX-2 cell surfaces, PX-Gal3 enriched samples, or basigin-IP. Bottom panel shows key to N-glycan nomenclature used.

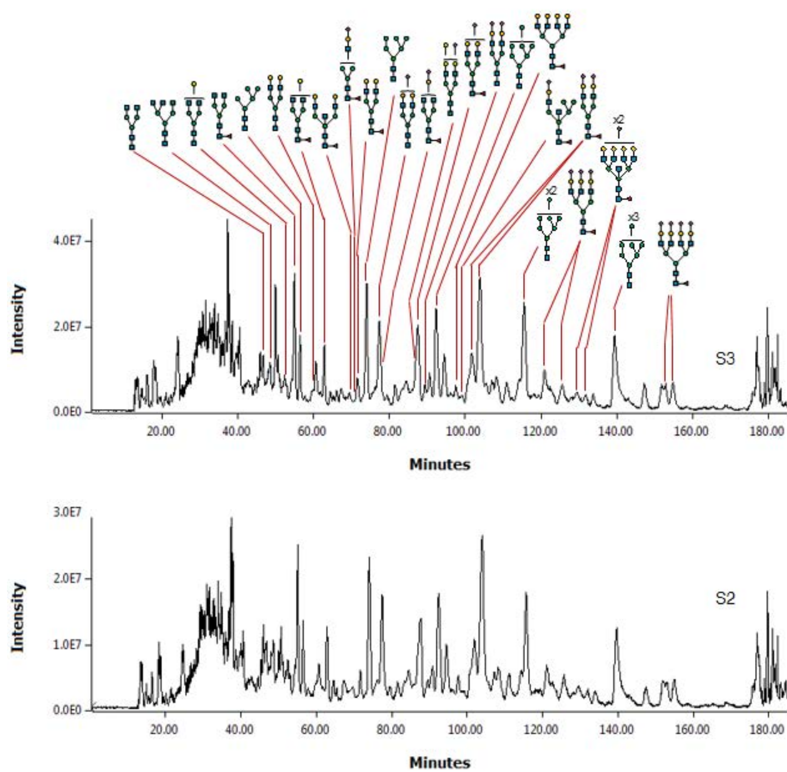



**Figure S11. Glycomics MS analysis of LX-2 cell surfaces**. Total ion chromatograms of permethylated N-glycans from two independent biological replicates (S3, S2).

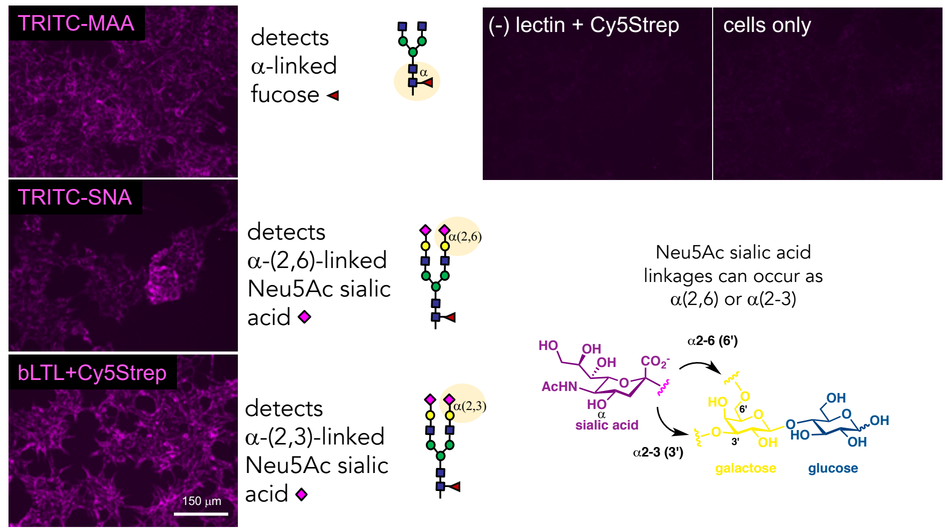



**Figure S12. Lectin staining of LX-2 hepatic stellate cells**. HSCs stain positively for the presence of α-linked fucoses and both α(2,6) and α(2,-3)-linked sialic acid linkages. LX-2 cells were seeded overnight in DMEM+ 10% FBS onto pre-coated (>10 min, RT, extensive washing with DPBS after) poly-L-lysine coated glass slide chambers. Next days, cells were fixed in 4% PFA/PBS (10 min, RT) and blocked (2% BSA/PBS + 0.1% TX, 1 hr, RT). Lectins (either TRITC-MAA, TRITC-SNA or bLTL; 1:200 in same buffer) were incubated for 1 hr at RT. The bLTL condition was washed and probed with Cy5-streptavidin (1:500 in same buffer). After washing and removal of the chambers, antifade was added and coverslips were attached. Cells were imaged with the appropriate filters.

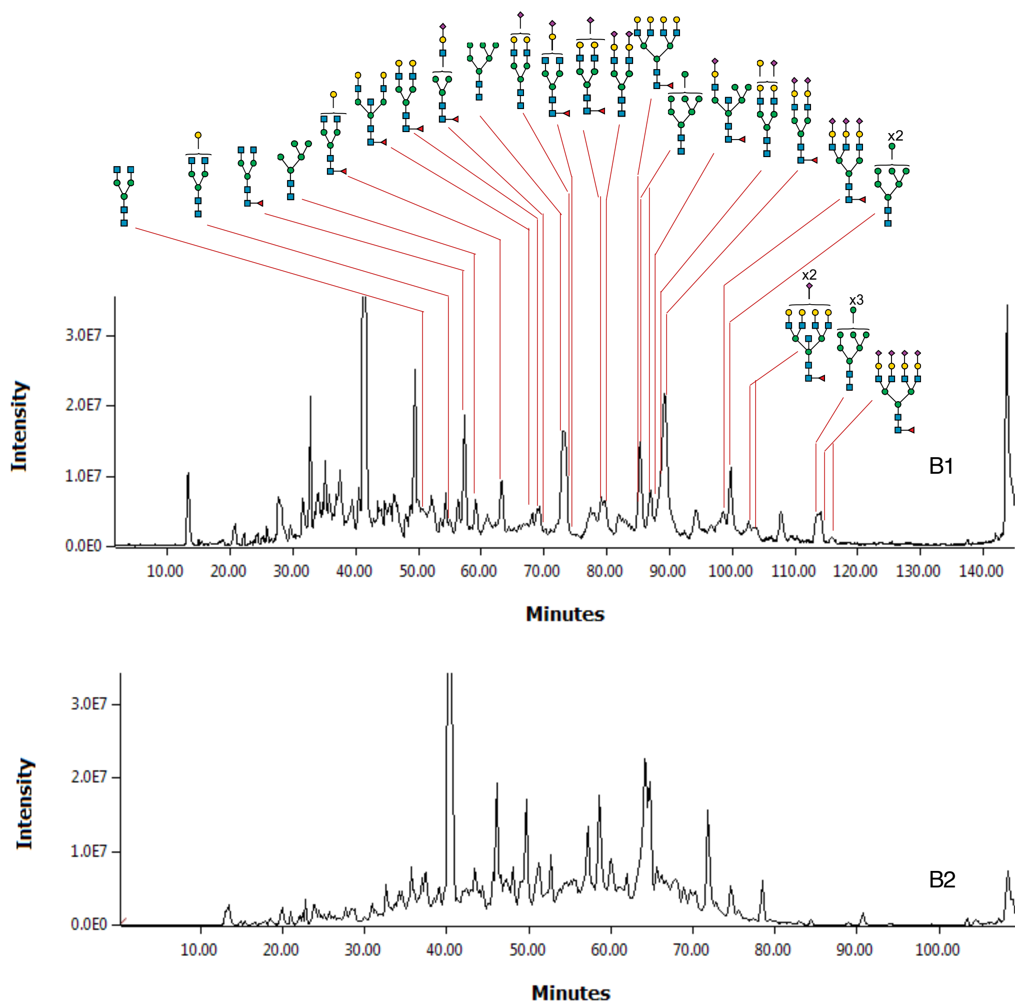



**Figure S13. Additional data for PX-Gal3 enriched glycans.** Total ion chromatograms of permethylated N-glycans from two independent biological replicates (B1, B2). Note that a shorter run time was used in B2.

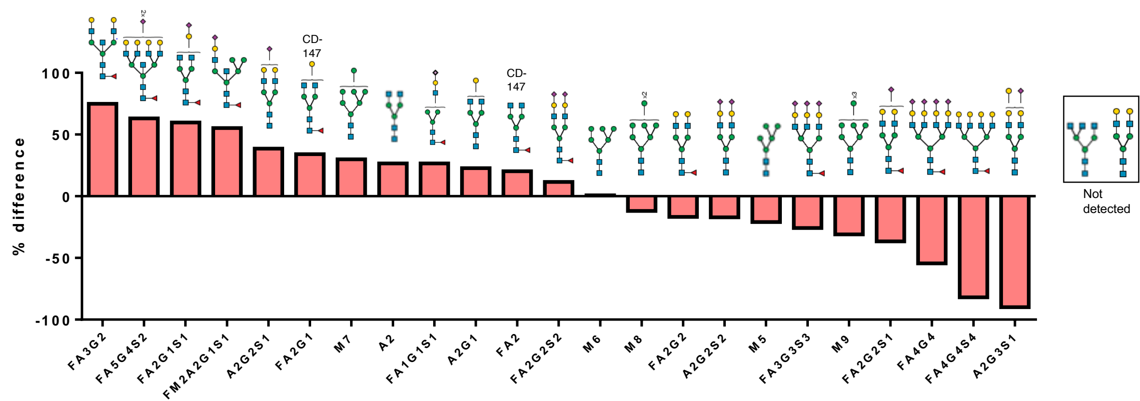



**Figure S14. Full Comparison of PX-Gal3 enriched and cell surface N-glycans.** % difference was calculated as the difference between N-glycan abundance between PX-Gal3 and cell surface divided by the abundance of the cell surface N-glycan.

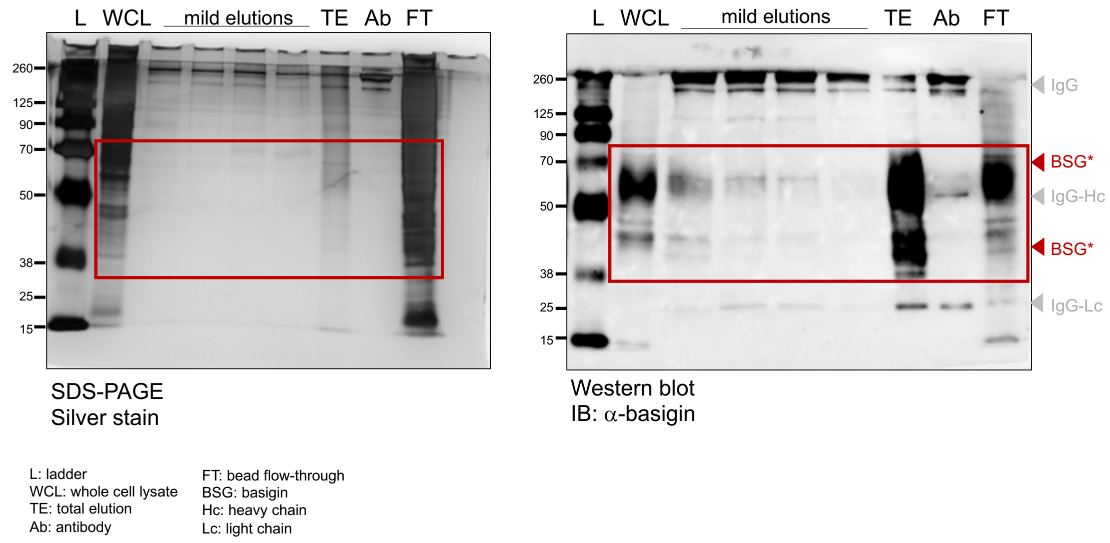



**Figure S15.** SDS-PAGE and Western blot analysis of basigin-IP (immunoprecipitated) from LX-2 cells.

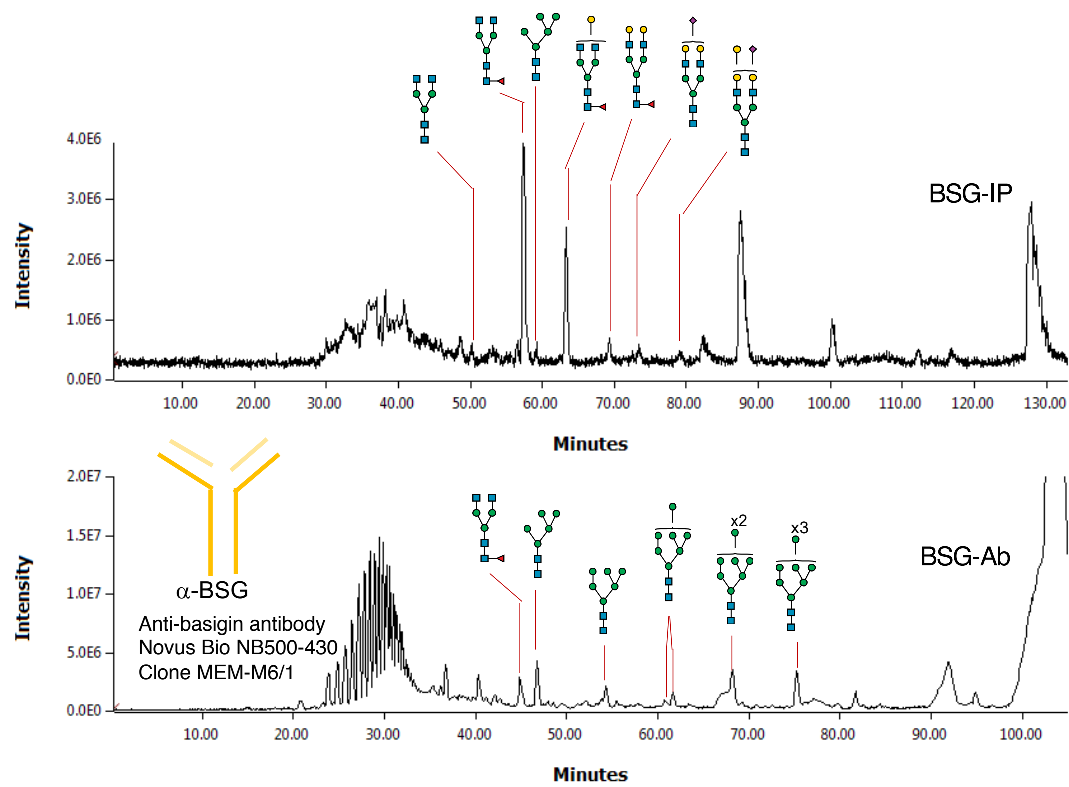



**Figure S16. Additional glycomics data for BSG-IP.** Total ion chromatograms of permethylated N-glycans from a basigin-IP sample. Bottom chromatogram represents the N-glycans from the immunoprecipitating antibody alone. Note that a shorter run time was used in BSG-Ab.

## Basigin-2-Fc

Recombinant protein
Sinobio 10186-H02H

Basigin-1 vs. Basigin-2 isoform comparison

- Found in recombinant protein
- Missing in recombinant protein

N-glycosite identification:

Asn-160 (basigin-1)
Asn-44 (basigin-2)
ILLTCSLJDSATEVTGHR

Asn-268 (basigin-1)
Asn-151 (basigin-2)
ITDSEDKALMJGSESR

Asn-73 (Fc)
JDSKNTLYLNMNSLR

Asn-299 (Fc)
TKPREEQYJSTYR

N-glycosite mapping:

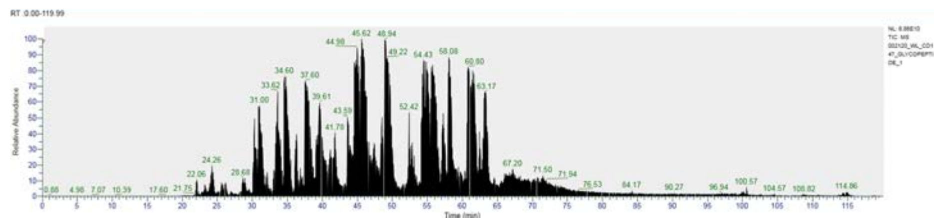



**Figure S17. Analysis of N-linked glycosites in recombinant basigin-Fc.** 20 µg of basigin-Fc (SinoBio # 10186-H02H) was dissolved in 50 µL of 50 mM ammonium bicarbonate buffer. To this, 10 µL of 50 mM dithiothreitol (DTT) was added and incubated at 55 °C for 30 min. Then 10 µL of fresh prepared iodoacetamide (IAA, 400 mM) solution was added and the mixture was incubated in the dark at room temperature for 30 min. The sample was transferred to a 10kDa centrifuge filter, then 300 µL of 50 mM ammonium bicarbonate buffer was added to the sample and centrifuged at 14,000 xg for 10 min, repeated twice to remove excess salt. The samples were transferred to a new tube and digested with 1 µg of sequencing grade trypsin (Promega V5111) and CaCl2 (1 mM) at 37 °C overnight. Following, the digested sample was heated at 95 °C for 5 min and subsequently filtered through a 0.2 µm filter and vacuum-centrifuged to near dryness via speed-vacuum centrifugation. The samples were stored in -80 °C until ready for injection. The samples were re-dissolved in 20 µL Buffer A (99.9% $H_2O$, 0.1% formic acid) and loaded onto an Orbitrap Fusion mass-spectrometer. 3 µL of the sample was loaded onto a Thermo Acclaim PepMap100 precolumn (75 µM × 2 mm) and eluted on a Thermo Acclaim PepMap RSLC analytical column (75 µM × 15 cm). Buffer A (0.1% formic acid in $H_2O$) and Buffer B (0.1% formic acid in acetonitrile) were used to establish the 120 min gradient comprised of 90 min of 2−35% B, 8 min of 35−60% B, and 2 min of 60−95% B, 8 min of 95% B, followed by re-equilibrating at 2% B for 2 min. The flow rate was 0.3 µL/min. Peptides were then analyzed on a Thermo Orbitrap Fusion Lumos proteomic mass spectrometer in a data-dependent manner, with automatic switching between MS and MS/MS scans using a cycle time 3 s. MS spectra were acquired at a resolution of 240,000 with an AGC target value of $5\times10^5$ ions or a maximum integration time of 100 ms. The scan range was limited from 375 to 2,000 m/z. Peptide fragmentation was first performed via a step higher-energy collisional dissociation (HCD) with the energy set at 15%, 30% and 45%, respectively, with an AGC target value of $5\times10^4$ and a maximum integration time of 48 ms. Additionally, collision induced dissociation (CID) fragmentation was triggered by detection of a 204.0865 ion signal, with the collision energy set at 35% and an activation time of 10 ms. MS spectra were acquired at a resolution of 30,000 with an AGC target value of $5\times10^4$ ions and a maximum integration time of 54 ms. The raw data was analyzed using pGlyco 2.2.2 software, following the manual's instruction. In general, the protein database was from UniProt (ID P35613), the mass tolerance for precursors and fragment ions were set as ± 5 ppm. and ± 20 ppm., respectively. Maximal missed cleavage was set as carbamidomethylation on cysteine residues (C +57.022 Da) as a fixed modification, oxidation on methionine (M +15.995 Da) as a variable modification. The N-glycosylation sequon (N-X-S/T, X

≠ P) was modified by changing ‘N’ to ‘J’. The N-glycosylation database was set as default in the software.

### SI References


1. H. Hamouda *et al.*, Rapid analysis of cell surface N-glycosylation from living cells using mass spectrometry. *Journal of proteome research* **13**, 6144-6151 (2014).
2. A. Shajahan, N. Supekar, C. Heiss, P. Azadi, High-Throughput Automated Micro-permethylation for Glycan Structure Analysis. *Anal Chem* **91**, 1237-1240 (2019).
3. A. Apte, N. S. Meitei, Bioinformatics in glycomics: glycan characterization with mass spectrometric data using SimGlycan. *Methods Mol Biol* **600**, 269-281 (2010).
4. A. Ceroni *et al.*, GlycoWorkbench: A Tool for the Computer-Assisted Annotation of Mass Spectra of Glycans. *Journal of proteome research* **7**, 1650-1659 (2008).
5. T. Geiger, A. Wehner, C. Schaab, J. Cox, M. Mann, Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins. *Mol Cell Proteomics* **11**, M111 014050 (2012).