Upregulation of Glycans Containing 3' Fucose in a Subset of Pancreatic Cancers Uncovered Using Fusion-Tagged Lectins

Article *in* Journal of Proteome Research · May 2015

DOI: 10.1021/acs.jproteome.5b00142

Journal of proteome research

Upregulation of Glycans Containing 3' Fucose in a Subset of Pancreatic Cancers Uncovered Using Fusion-Tagged Lectins

Sudhir Singh[1#], Kuntal Pal[1#], Jessica Yadav[1], Huiyuan Tang[1], Katie Partyka[1], Doron Kletter[2^], Peter Hsueh[1], Elliot Ensink[1], Birendra KC[3], Galen Hostetter[1], H. Eric Xu[1], Marshall Bern[2^], David F. Smith[4], Anand S. Mehta[5], Randall Brand[6], Karsten Melcher[1], and Brian B. Haab[1*]

Just Accepted Manuscript **DOI:** 10.1021/acs.jproteome.5b00142 Publication Date (Web): May 4, 2015

[1]*Van Andel Research Institute, Grand Rapids, MI*
[2]*Palo Alto Research Center, Palo Alto, CA*
[3]*Spectrum Health, Grand Rapids, MI*
[4]*Emory University, Atlanta, GA*
[5]*Drexel University, Philadelphia, PA*
[6]*University of Pittsburgh Medical Center, Pittsburgh, PA*
^*Present address: Protein Metrics, Inc., San Carlos, CA*

*Correspondence to:

Brian B. Haab, PhD
Van Andel Research Institute
333 Bostwick NE, Grand Rapids, MI 49503
616-234-5268
brian.haab@vai.org

Keywords: lectins, fucose, biomarkers, glycans, glycan arrays, antibody-lectin sandwich arrays, pancreatic cancer, sialyl-Lewis A

[#]Contributed equally

Abstract

The fucose post-translational modification is frequently increased in pancreatic cancer, thus forming the basis for promising biomarkers, but a subset of pancreatic cancer patients does not elevate the known fucose-containing biomarkers. We hypothesized that such patients elevate glycan motifs with fucose in linkages and contexts different from the known fucose-containing biomarkers. We used a database of glycan array data to identify the lectins CCL2 to detect glycan motifs with fucose in a 3' linkage; CGL2 for motifs with fucose in a 2' linkage; and RSL for fucose in all linkages. We used several practical methods to test the lectins and determine the optimal mode of detection, and we then tested whether the lectins detected glycans in pancreatic cancer patients who did not elevate the sialyl-Lewis A glycan, which is upregulated in ~75% of pancreatic adenocarcinomas. Patients who did not upregulate sialyl-Lewis A, which contains fucose in a 4' linkage, tended to upregulate fucose in a 3' linkage, as detected by CCL2, but they did not upregulate total fucose or fucose in a 2' linkage. CCL2 binding was high in cancerous epithelia from pancreatic tumors, including areas negative for sialyl-Lewis A and a related motif containing 3' fucose, sialyl-Lewis X. Thus glycans containing 3' fucose may complement sialyl-Lewis A to contribute to improved detection of pancreatic cancer. Furthermore, the use of panels of recombinant lectins may uncover details about glycosylation that could be important for characterizing and detecting cancer.

Introduction

Specific glycan motifs may serve as effective biomarkers for epithelial cancers such as pancreatic cancer. A glycan that has increased production by the majority of pancreatic adenocarcinomas is sialyl-Lewis A. The sialyl-Lewis A glycan, detected using the CA19-9 monoclonal antibody[1, 2] and related antibodies[3], is the current best biomarker for pancreatic cancer, showing elevations in the blood of about 75% of patients with pancreatic adenocarcinoma[4]. The CA19-9 test is useful as a monitor for disease responses or to confirm diagnosis in cases of extreme elevation. But because it is not elevated in a significant subset of patients and is elevated in some cases of benign biliary obstruction or other non-cancerous conditions[4], its performance is inadequate for cancer detection or diagnosis.

Previous research suggests that abnormal glycan production is a common feature of pancreatic cancers, including those without blood elevations of CA19-9. Immunohistochemistry studies show that increased secretion of Lewis and ABO blood group glycans is a near-universal feature of pancreatic adenocarcinoma[5], and truncated O-linked glycosylation resulting in the Tn and sialyl-Tn antigens occurs in almost all epithelial cancers[6] and pancreatic cancers in particular[7]. Furthermore, some of the CA19-9-low tumors elevate a glycan called sialyl-Lewis C, detected with the monoclonal antibody Dupan-2[8, 9]. Sialyl-Lewis C elevations occur mainly in the 5% of people who have homozygous germline mutations in a gene called FUT3, resulting in the inability to form sialyl-Lewis A[10]. But the above glycan motifs have not yet proven valuable in a biomarker panel for pancreatic cancer detection. If a particular glycan were elevated in the patients who are low in CA19-9, it could be valuable as part of a biomarker panel.

Several studies have linked increased fucose levels to disease[11] and have shown that the fucosylated forms of specific proteins make promising biomarker of cancer[12-15]. In

fact, the sialyl-Lewis A glycan contains fucose in the motif Siaα2-3Galβ1-3(Fucα1-4)GlcNAc (where Sia is sialic acid, Gal is galactose, Fuc is fucose, and GlcNAc is N-acetylglucosamine). But most studies have not provided information on the relative importance of particular fucose linkages or contexts. The location and surrounding context of the fucose has critical importance. For example, if it is on the 2' carbon of a galactose, it can contribute to the precursor for ABO blood group structures, whereas on the 3' or 4' of a GlcNAc in N-acetyl-lactosamine it can contribute to the formation of Lewis blood group structures that can be ligands for selectin receptors[16]. Such configurations can influence cell trafficking, immune cell migration, receptor activity, and bacterial binding to the gut, among other functions[11]. Considering the potential contributions of fucosylated glycans to cancer-promoting functions and the consequences of subtle changes in fucosylation, distinct fucose-containing motifs could characterize the tumors that do not make sialyl-Lewis A.

A powerful approach to getting more detail about fucosylated glycans in clinical samples is to use the exquisite specificity available from glycan-binding antibodies and lectins[17, 18]. Lectins can distinguish between linkages or other nuances of the overall oligosaccharide—information not easily accessible using mass spectrometry or chromatography methods. In addition, lectins function well in *in-vitro* assays that use small amounts of sample and that have the throughput and precision needed for biomarker studies, and they are amenable to recombinant production for high-content screening studies[19-21].

We previously developed a database of analyzed glycan array data[22, 23]. This tool allowed us to identify a set of lectins that could serve to probe complementary presentations of fucose. We developed and validated the recombinant lectins and used antibody-lectin sandwich arrays to test whether specific fucose presentations

characterize subsets of pancreatic cancer patients, particularly those who do not elevate sialyl-Lewis A.

Methods

Antibodies and Biological Reagents

We purchased antibodies and proteins from various sources (Table S1), and Dr. Mehta provided the recombinant *Aleuria aurantia* lectin (AAL)[24]. We obtained as kind gifts the Dupan-2 antibody from Dr. Michael Hollingsworth (Omaha, NE, USA), the DNA for RSL from Dr. Anne Imberty (Grenoble, France), and the DNA for CCL2 and CGL2 from Dr. Markus Kuenzler and Dr. Markus Aebi (Zurich, Switzerland).

We purified the antibodies to be printed onto microarray slides by dialysis (Slide-A-Lyzer, Pierce Biotechnology) to phosphate buffered saline (PBS) and by ultra-centrifugation. To link biotin on the detection antibodies and lectins, we used the EZ-Link-sulfo-NHS-LC-Biotin reagent (Pierce Biotechnology) according to the manufacturer's instructions with modifications[25]. We labeled the anti-polyHis antibody (Ab18184, Abcam) with Cyanine5 (Sulpho-Cyanine5 NHS ester, #13320, Lumiprobe) using the same protocol as for biotinylation. We purchased streptavidin with conjugated Cy5 (#SA1011, Invitrogen Life Technologies).

Recombinant Protein Production and Purification

Construction of Avitag-Lectin-H8 expression vector. The gene fragments encoding residues 1-90 of RSL, 1-151 of CGL2 and 1-143 of CCL2 were PCR amplified with the primers listed in Table S2. The resulting PCR fragments were digested with NcoI and SalI restriction endonucleases and inserted into the first T7 RNA polymerase-driven expression cassette of a modified pET-Duet1 expression vector (Novagen) to encode a fusion protein consisting of an Avitag followed by the lectin and a His8 tag. The 14 amino

acid Avitag[26] functions as a defined biotinylation site that in combination with the biotin ligase BirA expressed from the second expression unit is quantitatively biotinylated in *E. coli* in the presence of 40 µM biotin.

Construction of His8-Lectin-Avitag expression vector. To express lectins with N-terminal His8-tag and C-terminal Avitag, open reading frames encoding the fusion proteins were PCR-engineered (Table S2). The products were inserted between the Nco1 and Not1 restriction site of the pET-Duet1 variant that expressed BirA from the second expression cassette. *CGL2* has an internal NcoI site that was mutated without change of the lectin protein sequence. The primer details used for amplification and site directed mutagenesis is shown in Supplementary Table 1.

Protein expression and purification. The plasmids carrying Avitag-Lectin-His_8 and His_8-Lectin-Avitag proteins were overexpressed in E.coli BL21(DE3) (Novagen) cells. The transfected cells were grown in LB medium containing ampicillin (50 µg/mL) to an A_{600} of 0.6 at 24 °C. Protein expression was induced by 0.1 mM isopropyl β-D-1-thiogalactopyranoside. For *in-vivo* biotinylation of tagged lectin proteins, 40µM biotin was added during induction. After induction, the cells were grows at 16 °C overnight. The cells were harvested by centrifugation at 4000 rpm using the H12000 rotor (Thermo Scientific) and resuspended in lysis buffer containing 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 25 mM imidazole, and 5% glycerol. The resuspended cells were lysed using a French Press homogenizer at 10000 psi. The crude lysates were centrifuged at 18000 rpm in an SS-34 rotor (Sorvall) for 30 minutes at 4 °C, and supernatants were loaded on nickel-chelate-Sepharose HP columns (GE Healthcare). The columns were extensively washed with buffer containing 50 mM Tris-HCl (pH7.5), 400 mM NaCl, and 25 mM imidazole, and eluted with buffer composed of 50 mM Tris-HCl (pH7.5), 150 mM NaCl, 250 mM imidazole, and 5% glycerol. The eluted protein was further purified by

chromatography using a Superdex 200 or 300 gel filtration column (GE Healthcare) in buffer composed of 50 mM Tris-HCl (pH7.5) and 150 mM NaCl. In the case of the His_8-Lectin-Avitag fusion proteins, an additional anion exchange chromatography (GE Healthcare) step was performed to remove unwanted protein bands that were observed with those constructs. Selected fractions from the chromatography were analyzed by SDS-PAGE with Coomassie-blue staining.

Analytical Size-Exclusion Chromatography

We analyzed each lectin using analytical-scale size-exclusion chromatography (Shimadzu Prominence UFLC, using the Sepax Nanofilm SEC-500 column (#201500-4615, Sepax Technologies). We injected 18 µL of each sample in 50 mM Tris-HCl buffer (pH 7.5), with 150 mM NaCl as mobile phase for 12 minutes at a flow rate of 0.3 mL/min. We analyzed each lectin in pure form, and in separate runs we pre-incubated each lectin with Cy-5 conjugated anti-polyHis and Cy5 conjugated streptavidin for 1 hour at a 2:1 molar ratio of secondary reagent to lectin. We recorded UV absorbance with respect to time.

Thermal Shift Assays

A 10 µL reaction volume was used in 96 well plates. The reaction consisted of 2 µg of protein, 1 µL of 50X diluted stock of SYPRO Orange dye (Invitrogen Life Technologies), and TBS buffer with or without ligand oligosaccharide. We purchased the oligosaccharides Blood Group B type 2 tetrasaccharide (#G422, V-Labs, Covington, LA), Glycyl-2'- fucosyllactose (#NH302), Lewis X trisaccharide (#LN303), and Glycyl-lacto-N-tetraose (#NH403). We incubated each oligosaccharide with the lectin at various molar ratios on ice for 30 minutes. We used a real-time PCR system (Step One Plus, Applied Biosystems) to run a melt curve program at a ramp rate of 1˚ C and a temperature range of 30˚ C to 100˚ C. We recorded fluorescence at each temperature.

Glycan Array Analysis

The glycan synthesis and array core facility of the Consortium for Functional Glycomics (CFG) performed the glycan array experiments and the primary analysis according to the methods presented previously[27]. We downloaded data from www.functionalglycomics.org that previously had been obtained using lectins and glycan-binding antibodies supplied by various investigators. In addition, we sent the recombinant version of CCL2 with biotinylation at the C-terminus to the CFG core facility for processing on their glycan array version 5.2. For detailed analyses of the datasets, we used the GlycoSearch analysis program[28], and for mining glycan array data to find particular lectins, we used the GlycanBinder database[23], which derives information from the CFG website.

Immunohistochemistry Analysis

We used automated staining (Ventana Discovery Ultra) to perform immunohistochemistry (IHC) on sections cut from formalin-fixed, paraffin-embedded blocks. We used the same protocol for the anti-sLeA (9L426), anti-sLeX (CSLEX1) primary antibodies and CCL2 lectin: antigen retrieval using the Ventana CC1 buffer for 36 minutes at 95 °C; primary antibody incubation at a 1:200 dilution and CCL2 incubation at 15 µg/mL for 32 minutes at 37 °C; and secondary antibody (Ventana Umap HRP-conjugated anti-mouse) for 12 minutes at 37 °C. The development step used the diaminobenzadine chromagen according to preset parameters in the Ventana platform.

Antibody-Lectin Sandwich Array Analysis of Plasma Samples

Plasma sample collection. All collections took place at the University of Pittsburgh Medical Center following informed consent of the participants and prior to any surgeries or procedures. All blood samples were collected according to the standard operating procedure from the Early Detection Research Network and were frozen at -70 °C or

colder within 4 hours of time of collection. Aliquots were shipped on dry ice and thawed no more than three times prior to analysis.

Antibody array fabrication and use. The antibody array methods followed those presented earlier[29-31] with slight modifications. We printed forty-eight identical arrays containing various antibodies (Table S1) onto glass microscope slides coated with ultra-thin nitrocellulose (PATH Slides, Grace BioLabs) using a contact printer (Aushon 2470, Aushon BioSystems). We printed six replicates of each antibody in randomized positions within each array. After printing, hydrophobic borders were imprinted onto the slides (SlideImprinter, The Gel Company, San Francisco, CA) to segregate the arrays and allow for individual sample incubations on each array.

The arrays were blocked using 1% bovine serum albumin (BSA) in phosphate buffered saline (PBS) plus 0.5% Tween-20 for one hour at room temperature. The slides were rinsed in 0.5X PBS plus 0.1% Tween-20, washed in the same buffer for 15 minutes, and dried by brief centrifugation at 160 x g, with printed arrays facing outside.

To prevent lectin binding to the spotted capture antibodies, we treated the spotted antibodies (in the arrays to be detected with RSL, CGL2, or CCL2) with an enzyme to remove their glycans. We incubated PNGaseF (P0704S, New England Biolabs) diluted 1:100 in enzyme reaction buffer on each array at 37 °C for 12 hours, and then washed and dried the arrays.

The plasma samples were diluted two-fold into PBS containing 0.1% Tween-20, 0.1% Brij-35, an IgG blocking cocktail (200 µg/mL mouse and rabbit IgG and 100 µg/mL goat and sheep IgG (Jackson ImmunoResearch)) and protease inhibitor (Complete Mini EDTA-free Tablet, Roche Applied Science). After briefly washing the arrays in PBS/0.1% Tween-20, the samples were incubated on the arrays overnight at 4 °C. The arrays were washed in three changes of PBS/0.1% Tween-20 for three minutes each and dried by

centrifugation (Eppendorf 5810R, rotor A-4-62, 1500 x g for three minutes), and a biotinylated lectin or antibody (3 µg/mL in PBS with 0.1% BSA and 0.1% Tween-20) was incubated on the arrays for one hour at room temperature. After washing and drying the arrays as above, Cy5-conjugated streptavidin (Roche Applied Science) prepared at 2 µg/mL in PBS with 0.1% BSA and 0.1% Tween-20 was incubated for one hour at room temperature, followed by a final wash and dry. The fluorescence images were acquired as above.

Analysis and Data Processing

We used the software GenePix Pro 5.1 (Molecular Devices, Sunnyvale, CA) to quantify and analyze the images using both automatic and manual spot finding features. The intensity of each spot was calculated by subtracting the local background from the median intensity of each spot. The quantified results for each image were further processed using the Grubb's test to remove outlier spots among the six replicates for each spotted antibody or protein. The remaining replicate spots were used to calculate the geometric mean for each captured protein or antibody.

The results were processed using Microsoft Office Excel, and the figures were prepared using GraphPad Prism (GraphPad Software) and Canvas XIV (ACD Systems).

Results

Identifying Lectins Binding Diverse Fucose Presentations

A variety of common motifs contain fucose in the α1-2, α1-3, α1-4, or α1-6 linkages (Fig. 1A). We queried a database called GlycanBinder, which contains raw data, analyzed data, and metadata from publicly-available glycan array experiments[23], to identify lectins that have high probabilities of binding motifs containing each linkage (Fig. 1B). The queries returned lectins that could be used to detect fucose in all linkages or in particular

linkages (Fig. 1C and Tables S3-S6). For example, the lectin from *Ralstonia solanum* (RSL)[32] binds fucose in all linkages; *Coprinopsis cinerea* lectin 2 (CCL2)[33] binds motifs containing fucose in the α1-3 linkage but not in the 2', 4', or 6' linkages; and *Coprinopsis cinerea* galectin 2 (CGL2)[34] binds motifs containing fucose primarily in the α1-2 linkage (Fig. 1C). We were not able to find a lectin that exclusively bound fucose in the α1-4 or α1-6 configurations. The glycan array data did show that the lectins from *Pisum sativum* and *Lens culinaris* may be specific for 6'-linked fucose (Table S6), as reported earlier[35], but a closer examination of the glycan array data suggested a requirement for certain mannose linkages in each cases, rendering them not exclusive binders to 6'-linked fucose. We did not have glycan array data for the lectin from *Pholiota squarrosa*, described by Kobayashi and coworkers as specific for α1-6 fucose[36].

We more fully analyzed the glycan array data for CCL2, CGL2, and RSL to better characterize their binding specificities. We previously developed a method called Motif Segregation[37, 38] and software called GlycoSearch[23, 28] to analyze glycan array data. RSL bound fucose regardless of the presentation: all glycans containing fucose had higher levels than those that did not (glycan array data not shown). CCL2 mainly bound fucose in a 3' linkage, but with some limitations: If a fucose was present on an adjacent monosaccharide in a 2' linkage (e.g. Fucα1-2Galβ1-4(Fucα1-3)GlcNAcβ, the Lewis Y motif), no binding occurred, and if the motif was close to the spacer (and the array surface), no binding occurred. CGL2 mainly bound motifs similar to blood group A and blood group B (e.g. Galβ1-3(Fucα1-2)Galβ1-4), as well as blood group H (Fucα1-2Galβ1-4). CGL2 also bound glycans terminating in 3SO4Galβ in the absence of fucosylation, but nearly all the glycans bound by CGL2 contain fucose in a 2' linkage but not in any other linkage (glycan array data not shown). Therefore the motifs bound by

CCL2 and CGL2 are mutually exclusive and contain fucose in a 3’ and 2’ linkage, respectively.

Ulex europeaus agglutinin (UEA-1), frequently used to detect fucose in a 2’ linkage, did not appear in our search. A closer examination of the UEA glycan array data showed that it is mainly a blood group H binder, but that in many cases it does not bind where Fucα1-2Galβ is present (analysis not shown). These cases mainly are where the 3’ carbon of the galactose is occupied (which is not blood group H) and where the galactose is linked to the 3’ carbon of the next saccharide (as in Type 1 blood group H). Therefore UEA-1 is not as general as CGL2 for binding motifs containing fucose in a 2’ linkage.

Effects of Fusion Tags On Native Multimer Formation and Glycan Binding

We selected RSL, CCL2, and CGL2 for recombinant production and testing. Because we did not know which terminus would function best for tagging and detection, or which tag would work best, we produced two different versions of each lectin. In one version, we placed a poly-histidine on the N-terminus and a biotin—via the Avitag construct[26]—on the C-terminus, and in the other version, the tags were switched.

The lectins form multimers in their natural states (Fig. 2A), so we examined the relative amounts of monomer and multimer using size-exclusion chromatography followed by polyacrylamide gel electrophoresis. Both constructs for each lectin produced the correct size of monomeric lectin (Fig. 2B and Fig. S1), but the formation of multimers was different between the constructs. For RSL, the version with biotin at the C-terminus properly formed a trimer, but the version with biotin at the N-terminus tended to dimerize (Fig. 2B). The N-terminus of RSL is more centrally located in the trimeric structure than the C-terminus (Fig. 2A), so the larger Avitag on the N-terminus could be more disruptive than poly-histidine. Similarly for CCL2, the version with Avitag at the C-terminus formed

trimers (Fig. S1), but the N-Avitag version just formed the monomer. The C-terminus could be farther removed from the subunit interfaces, based on the crystal structure of the monomeric form of CCL2 (Fig. 2A). (A crystal structure for multimeric CCL2 is not available.) Recent research similarly found that N-terminal tagging of CCL2 disrupted dimer formation (M. Kuenzler, personal communication). For CGL2, the N-terminus Avitag version formed trimers and the C-terminus Avitag version formed dimers (Fig. S1). The crystal structure of CGL2 shows that its native form is as a tetramer (Fig. 2A), so potentially the presence of the tags destabilizes the full tetrameric form but allows the formation of trimers. Thus we obtained proper production of each version of each lectin, but for RSL and CCL2, the biotin at the N-terminus was more disruptive to multimer formation than biotin at the C-terminus.

We used thermal shift assays, analytical size-exclusion chromatography, and binding to protein arrays to test the glycan binding capabilities of the lectins (Fig. S2). The assays confirmed proper glycan binding for each version of each protein, and they confirmed that both the poly-histidine and biotin tags were accessible to the secondary detection reagents at each terminus.

Differences Between the Lectins in Cancer-Specific Binding

At this point we were able to test the hypothesis that subsets of pancreatic cancer patients elevate fucose-containing motifs besides sLeA. We used antibody-lectin sandwich arrays[29] to analyze plasma samples from pancreatic cancer and control subjects. For each lectin we used the version that was biotinylated at the C–terminus, based on the better multimerization and generally higher signals when incubated on the glycoprotein arrays, as compared to the version with biotin at the N-terminus (data not shown).

Probing fucose linkages

We selected plasma samples from pancreatic cancer patients who were either elevated or not elevated in sLeA (as determined in previous research[31]), along with samples from patients with pancreatitis and from healthy individuals. We incubated each sample on multiple antibody arrays, and we probed each array with one of the lectins or an antibody (Fig. 3A). The arrays contained 9 different antibodies (Table S1) targeting sialyl-Lewis A or other cancer-associated glycans.

Sialyl-Lewis A levels, measured using anti-sLeA detection at two different anti-sLeA capture antibodies, were elevated in three of the cancer patient samples relative to the control samples (Fig. 3B). Probing the same capture antibodies with CCL2 showed overlapping but divergent results (Fig. 3C). Using two different sialyl-Lewis A capture antibodies, some of the cancer samples that did not have elevations in sLeA showed elevations in CCL2 detection, along with moderate elevation in two control subjects on one of the two captures. In contrast, both CGL2 and RSL showed no cancer-associated elevations on the sialyl-Lewis A antibodies (Fig. 3D and 3E) or on any other antibody (not shown). Thus total fucose levels (detected with RSL), and fucose in 2' linkages (detected with CGL2), were not elevated in the cancer patients relative to the controls, but certain cancer patients who did not produce fucose in the 4' motif (detected with sialyl-Lewis A antibodies) produced fucose in the 3' motif instead (detected with CCL2).

Determining the Specificity of Recombinant CCL2 by Glycan Array

Because the recombinant CCL2 showed the most promise for cancer-specific detection, we characterized its binding specificity in more detail. We incubated CCL2 that was tagged with C-terminus biotin on the version 5.2 glycan array from the CFG at three different CCL2 concentrations, and we detected CCL2 binding with Cy5-labeled streptavidin. Similar to the purified, native CCL2 (Fig. 4A), the recombinant CCL2 exclusively bound glycans containing 3' fucose (Fig. 4B). The newer glycan array

included glycans with fucosylated LAc-di-NAc, which is GalNAcβ1,4(Fucα1,3)GlcNAc. CCL2 bound glycans containing this fucosylated LAc-di-NAc, indicating CCL2 does not simply bind Lewis X, which is Galβ1,4(Fucα1,3)GlcNAc. As observed with native CCL2, recombinant CCL2 only binds 3'-linked fucose in the absence of neighboring 2' fucose (Fig. 4C). These data demonstrate that neither the fusion tag nor alterations to the quaternary state of CCL2 affected its binding specificity.

Probing the CCL2 Target Glycans As a Candidate Biomarker

We further explored the preliminary findings in a larger set of samples, consisting of 27 samples from pancreatic cancer patients and 17 samples from patients with acute or chronic pancreatitis. Using a cutoff chosen to optimize accuracy of discriminating cases from controls, 14 of the cancer cases and 3 of the controls were elevated in sLeA (Fig. 5A). Detection with CCL2 of the same samples and the same capture antibody showed elevations in some of the samples that were not elevated in the sLeA sandwich assay, using a cutoff that gave only 1 false-positive elevation (Fig. 5B).

We further analyzed CCL2 detection in a set of 200 plasma samples, collected from 109 patients with pancreatic cancer and 91 patients with benign pancreatic disease, including chronic/acute pancreatitis and benign biliary obstruction. Capture with either anti-sLeA or anti-sialyl-Lewis X (sLeX) followed by CCL2 detection resulted in statistically significant discrimination of the cases from controls (Fig. 5B). CCL2 detection also functioned well for detecting a glycoform of captured MUC5AC that was significantly elevated in pancreatic cancer.

The sLeX glycan, which contains fucose in a 3' linkage, previously was found elevated in the tissue of about 30% of pancreatic cancers[39], so it potentially was the main glycan bound by CCL2 in pancreatic cancer. To determine whether CCL2 detects a different group of patients than either sLeA or sLeX, we directly compared the binding levels

between detection with CCL2 and detection with either anti-sLeA or anti-sLeX. Clearly divergent groups of cancer patients were elevated in the various assays; some patients were exclusively elevated with CCL2 detection. This result indicates that CCL2 is not simply binding sLeX—as also shown by the glycan array data (Fig. 3)—and that CCL2 provides complementary detection to both sLeA and sLeX.

CCL2-Bound Glycans in Pancreatic Tumors

We next asked whether CCL2-detected glycans were produced in the tumor tissue of patients with elevated plasma levels and whether CCL2-bound glycans have similar expression to those detected by sLeA and sLeX antibodies. We obtained matched tumor tissue and plasma from 14 patients who had a surgical resection for pancreatic cancer. A comparison of the plasma levels of sLeA and CCL2 detection (using the assays shown in Fig. 5A) agreed with the previous results, in that selected cancer samples showed elevations in one or the other assay (Fig. 6A).

The CCL2 staining was strong in the neoplastic epithelia of most of the tumors (Fig. 6B and S3). The high staining in the secretory epithelial cells is consistent with the concept that the plasma elevations are due to secretions from the cancer cells, although the levels did not necessarily match between plasma and tissue. The CCL2 staining was higher than sLeA in some cases, showing the potential of CCL2 as a complementary marker to sLeA. Furthermore, the staining pattern of CCL2 was different also than that of sLeX, indicating that the cancer-associated glycan bound by CCL2 is not simply sLeX.

Discussion

In this study, we used a novel experimental approach to uncover information about cancer glycosylation that could have relevance to cancer detection. The approach involved the use of a new database to search for lectins with particular characteristics;

recombinant production of two variants each of lectins with complementary specificities for fucose; and the testing of the lectins for cancer detection using a microarray format. We found that one lectin in particular, CCL2, bound glycans that are overexpressed in a subset of pancreatic cancer patients that does not show elevations in the current best biomarker for pancreatic cancer, sialyl-Lewis A. If this result is validated in larger sets, assays using CCL2 or reagents with a similar specificity could be used in combination with assays for sLeA to provide more accurate detection of cancer than currently possible. Furthermore, this finding supports the general hypotheses that subsets of patients are characterized by distinct glycans, and that additional glycans to further improve cancer detection are yet to be found.

An advantage of using lectins is that they already exist in nature; we can tap into the glycan-binding capabilities already discovered, which cover an increasingly broad range of motifs as researchers find more and more lectins. But identifying a lectin with an unusual specificity, or even obtaining a clear characterization of a lectin's specificity**,** could be difficult. The available information about lectins is often minimal and found in obscure journals. The GlycanBinder database addresses this limitation by incorporating analyzed data, raw data, and metadata from thousands of glycan array experiments. The database was valuable for this research by enabling identification of lectins that bind outer-arm fucose in particular linkages, as opposed to the well-known fucose-binding lectins that bind most linkages. We expect that the database will be useful for a broad range of studies. A web-accessible version is in development, as well as a version that incorporates glycan array data from other platforms besides the CFG's.

Recombinant production of the lectins was preferable to purification from natural sources because we could achieve higher purity and have control of the tagging process. But the recombinant production of lectins also has the possibility of altering glycan binding. We

therefore used a series of practical assays to test the glycan binding activity of the lectins and to compare between four modes of detection for each lectin. The methods we used—size-exclusion chromatography, thermal shift assays, competition assays, and binding to immobilized glycoproteins—combined with an analysis of features of the crystal structures, enabled us to find the optimal mode of detection for each lectin. In each case, the best mode was the biotin tag on the C-terminal. This mode appears to be best because the C-terminus is more accessible than the N-terminus and because the comparatively bulky Avitag is less disruptive to native multimerization when on the C-terminus. Whether this system is best for all recombinant lectins is not known; likely one would need to test and optimize each lectin. The approaches used here provide a practical strategy for such optimization.

Of the three lectins, only CCL2 detection showed elevations selectively in the cancer patients. Since CCL2 binds motifs containing 3’ fucose, mainly in the Lewis X and sialyl-Lewis X glycans, its elevation in patients low in sialyl-Lewis A suggests alternative paths of glycan synthesis, involving production of Lewis antigens with 3’ fucose in the absence of production of Lewis antigens with 4’ fucose. But because CCL2 detection did not correlate with sLeX detection, the elevations in CCL2 detection are not simply due to sLeX; other glycans are elevated that contain 3’ fucose. Further research could obtain more information about such glycans. In addition, further studies should address whether a biomarker panel combining CCL2 detection with sLeA detection can provide improved discrimination of pancreatic cancer from non-cancer conditions. If the improvement is validated, these and related glycans could form tests with performance good enough to be used for the detection and diagnosis of pancreatic cancer.

Previous studies also found over-expression of sLeX in some pancreatic tumors. In an immunohistochemistry study of 30 patients with pancreatic cancer, sLeX staining was

present in 30% of patients, relative to sLeA staining in 87% of patients[39]. This relative rate of sLeA and sLeX staining generally agrees with our findings, but the study did not examine plasma or the use of alternate glycans to detect cancers that are low in sLeA. A recent study using high-throughput glycan analysis using chromatography and glycosidase digestions found that the plasma protein ceruloplasmin carries elevated levels of sLeX in patients with pancreatic cancer[40].

Other studies investigated the idea that other glycans besides sLeA could be useful for the detection of pancreatic cancer. Many studies used monoclonal antibodies that, like the CA19-9 antibody, were raised by immunization with cancer cells. Combinations of CA19-9 with various antibodies such as CA 50[41], Span-1[41], CA 242[42] and others generally produced minimal added value over CA19-9 alone. Perhaps the most promising antibody for this purpose was DUPAN2[8], which binds a non-fucosylated relative of sLeA called sialyl Lewis C[43] (Siaα2,3Galβ1,3GlcNAc). Separate studies demonstrated that some of the CA19-9-low patients show elevations in DUPAN-2[9, 44], but a biomarker panel based on these assays never became established. Miyamoto and coworkers found increased levels of a novel glycan, 6-sialylated, type-1 H antigen, in 3 out of 6 pancreatic tumors that were negative for CA19-9[45], but the researchers did not demonstrate this structure as a blood biomarker.

An interesting feature of the result with CCL2 was that the higher 3-linked fucose was not widespread over all capture antibodies; it was on particular capture antibodies such as anti-sLeA (Ab1) (Fig. 5C), representing the co-expression of 3’ fucose and sialyl-Lewis A. This observation may indicate that the co-expression of certain motifs—on the same molecule or on physically-linked molecules—is as important as the elevated expression of individual motifs. The elevation of two linked motifs may reflect biological effects. For example, specific groupings of glycans could cluster cell-surface receptors to

induce signaling[46-48]. The use of a “hybrid” sandwich assay—the capture and detection of different glycans—may be a useful way to examine the co-expression of particular glycans, and to evaluate the use of such molecules as biomarkers.

We are only beginning to understand some of the biological information contained in glycosylation[49]. An effective way to probe the information in biological samples is to use recombinantly-produced, fusion-tagged lectins. The approach demonstrated here adds to the available options for finding and using lectins in biomedical research. Using this approach, we found that glycans detected by CCL2—those with fucose in a 3’ linkage—are overexpressed in a subset of pancreatic cancer patients that does not elevate sLeA. We intend to pursue the possibility that the detection of such glycans with complementary patterns of elevation in subsets of pancreatic cancer patients could result in improved detection of pancreatic cancer.

Acknowledgements

We gratefully acknowledge the Early Detection Research Network (U01CA152653), the Alliance of Glycobiologists for the Detection of Cancer (U01CA168896), and the Van Andel Research Institute for support of this work. We also thank the supporters of the Purple Community philanthropic foundation for their generous donations. We thank Drs. Markus Aebi and Markus Kuenzler for supplying the DNA for CCL2 and CGL2; Dr. Anne Imberty for supplying the DNA for RSL; Dr. Michael A. Hollingsworth for supplying the Dupan-2 antibody; and the Consortium for Functional Glycomics for providing the glycan array datasets and experiments.

Supporting Information Available:

Supplementary Tables

Table S1: Antibodies and Proteins Used
Table S2: Primer Information
Table S3: Binders to alpha 1,2 fucose
Table S4: Binders to alpha 1,3 fucose

Supplementary Figures

This material is available free of charge via the Internet at http://pubs.acs.org.

References

1. Herlyn, M.; Steplewski, Z.; Herlyn, D.; Koprowski, H., Colorectal carcinoma-specific antigen: detection by means of monoclonal antibodies. *Proc Natl Acad Sci U S A* **1979,** 76, (3), 1438-42.
2. Brockhaus, M.; Magnani, J. L.; Herlyn, M.; Blaszczyk, M.; Steplewski, Z.; Koprowski, H.; Ginsburg, V., Monoclonal antibodies directed against the sugar sequence of lacto-N-fucopentaose III are obtained from mice immunized with human tumors. *Arch Biochem Biophys* **1982,** 217, (2), 647-51.
3. Partyka, K.; Maupin, K. A.; Brand, R. E.; Haab, B. B., Diverse monoclonal antibodies against the CA 19-9 antigen show variation in binding specificity with consequences for clinical interpretation. *Proteomics* **2012,** 12, 2212-20.
4. Goonetilleke, K. S.; Siriwardena, A. K., Systematic review of carbohydrate antigen (CA 19-9) as a biochemical marker in the diagnosis of pancreatic cancer. *Eur J Surg Oncol* **2007,** 33, (3), 266-70.
5. Itzkowitz, S. H.; Yuan, M.; Ferrell, L. D.; Ratcliffe, R. M.; Chung, Y. S.; Satake, K.; Umeyama, K.; Jones, R. T.; Kim, Y. S., Cancer-associated alterations of blood group antigen expression in the human pancreas. *J Natl Cancer Inst* **1987,** 79, (3), 425-34.
6. Ju, T.; Wang, Y.; Aryal, R. P.; Lehoux, S. D.; Ding, X.; Kudelka, M. R.; Cutler, C.; Zeng, J.; Wang, J.; Sun, X.; Heimburg-Molinaro, J.; Smith, D. F.; Cummings, R. D., Tn and sialyl-Tn antigens, aberrant O-glycomics as human disease markers. *Proteomics Clinical Applications* **2013**.
7. Remmers, N.; Anderson, J. M.; Linde, E.; Dimaio, D. J.; Lazenby, A.; Hans, W.; Mandell, U.; Clausen, H.; Yu, F.; Hollingsworth, M. T., Aberrant expression of mucin core proteins and O-linked glycans associated with progression of pancreatic cancer. *Clinical Cancer Research* **2013,** 19, 1981-93.
8. Metzgar, R. S.; Gaillard, M. T.; Levine, S. J.; Tuck, F. L.; Bossen, E. H.; Borowitz, M. J., Antigens of human pancreatic adenocarcinoma cells defined by murine monoclonal antibodies. *Cancer Res* **1982,** 42, (2), 601-8.
9. Takasaki, H.; Uchida, E.; Tempero, M. A.; Burnett, D. A.; Metzgar, R. S.; Pour, P. M., Correlative study on expression of CA 19-9 and DU-PAN-2 in tumor tissue and in serum of pancreatic cancer patients. *Cancer Res* **1988,** 48, (6), 1435-8.
10. Nishihara, S.; Yazawa, S.; Iwasaki, H.; Nakazato, M.; Kudo, T.; Ando, T.; Narimatsu, H., Alpha (1,3/1,4)fucosyltransferase (FucT-III) gene is inactivated by a single amino acid substitution in Lewis histo-blood type negative individuals. *Biochem Biophys Res Commun* **1993,** 196, (2), 624-31.
11. Becker, D. J.; Lowe, J. B., Fucose: biosynthesis and biological function in mammals. *Glycobiology* **2003,** 13, (7), 41R-53R.
12. Thompson, S.; Cantwell, B. M.; Cornell, C.; Turner, G. A., Abnormally-fucosylated haptoglobin: a cancer marker for tumour burden but not gross liver metastasis. *Br J Cancer* **1991,** 64, (2), 386-90.
13. Okuyama, N.; Ide, Y.; Nakano, M.; Nakagawa, T.; Yamanaka, K.; Moriwaki, K.; Murata, K.; Ohigashi, H.; Yokoyama, S.; Eguchi, H.; Ishikawa, O.; Ito, T.; Kato, M.; Kasahara, A.; Kawano, S.; Gu, J.; Taniguchi, N.; Miyoshi, E., Fucosylated haptoglobin is a novel marker for pancreatic cancer: A detailed analysis of the oligosaccharide

structure and a possible mechanism for fucosylation. *Int J Cancer* **2006,** 118, (11), 2803-2808.
14. Aoyagi, Y.; Mita, Y.; Suda, T.; Kawai, K.; Kuroiwa, T.; Igarashi, M.; Kobayashi, M.; Waguri, N.; Asakura, H., The fucosylation index of serum alpha-fetoprotein as useful prognostic factor in patients with hepatocellular carcinoma in special reference to chronological changes. *Hepatol Res* **2002,** 23, (4), 287.
15. Wang, M.; Long, R. E.; Comunale, M. A.; Junaidi, O.; Marrero, J.; Di Bisceglie, A. M.; Block, T. M.; Mehta, A. S., Novel fucosylated biomarkers for the early detection of hepatocellular carcinoma. *Cancer Epidemiology, Biomarkers & Prevention* **2009,** 18, (6), 1914-21.
16. McEver, R. P., Selectin-carbohydrate interactions during inflammation and metastasis. *Glycoconj J* **1997,** 14, (5), 585-91.
17. Sharon, N.; Lis, H., History of lectins: from hemagglutinins to biological recognition molecules. *Glycobiology* **2004,** 14, (11), 53R-62R.
18. Arnaud, J.; Audfray, A.; Imberty, A., Binding sugars: from natural lectins to synthetic receptors and engineered neolectins. *Chemical Society Reviews* **2013,** 42, (11), 4798-813.
19. Yim, M.; Ono, T.; Irimura, T., Mutated plant lectin library useful to identify different cells. *Proc Natl Acad Sci U S A* **2001,** 98, (5), 2222-5.
20. Hsu, K. L.; Gildersleeve, J. C.; Mahal, L. K., A simple strategy for the creation of a recombinant lectin microarray. *Mol Biosyst* **2008,** 4, (6), 654-62.
21. Maenuma, K.; Yim, M.; Komatsu, K.; Hoshino, M.; Tachiki-Fujioka, A.; Takahashi, K.; Hiki, Y.; Bovin, N.; Irimura, T., A library of mutated Maackia amurensis hemagglutinin distinguishes putative glycoforms of immunoglobulin A1 from IgA nephropathy patients. *J Proteome Res* **2009,** 8, (7), 3617-24.
22. Kletter, D.; Bern, M.; Haab, B. B., Mining and using glycan array data with the GlycoSearch analysis program and GlycanBinder database. In *Glycoscience: Biology and Medicine*, Naoyuki Taniguchi, T. E., Gerald Hart, Peter Seeberger, Chi-Huey Wong, Ed. Springer: Japan, 2014.
23. Kletter, D.; Singh, S.; Bern, M.; Haab, B. B., Global comparisons of lectin-glycan interactions using a database of analyzed glycan array data. *Molecular & Cellular Proteomics* **2013,** 12, (4), 1026-35.
24. Romano, P. R.; Mackay, A.; Vong, M.; DeSa, J.; Lamontagne, A.; Comunale, M. A.; Hafner, J.; Block, T.; Lec, R.; Mehta, A., Development of recombinant Aleuria aurantia lectins with altered binding specificities to fucosylated glycans. *Biochem Biophys Res Commun* **2011,** 414, (1), 84-9.
25. Haab, B. B.; Partyka, K.; Cao, Z., Using antibody arrays to measure protein abundance and glycosylation: considerations for optimal performance. *Current Protocols in Protein Science* **2013,** 73, unit 27.6.
26. Smith, P. A.; Tripp, B. C.; DiBlasio-Smith, E. A.; Lu, Z.; LaVallie, E. R.; McCoy, J. M., A plasmid expression system for quantitative in vivo biotinylation of thioredoxin fusion proteins in Escherichia coli. *Nucleic Acids Research* **1998,** 26, (6), 1414-20.
27. Blixt, O.; Head, S.; Mondala, T.; Scanlan, C.; Huflejt, M. E.; Alvarez, R.; Bryan, M. C.; Fazio, F.; Calarese, D.; Stevens, J.; Razi, N.; Stevens, D. J.; Skehel, J. J.; van Die, I.; Burton, D. R.; Wilson, I. A.; Cummings, R.; Bovin, N.; Wong, C. H.; Paulson, J. C.,

Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc Natl Acad Sci U S A* **2004,** 101, (49), 17033-17038.
28. Kletter, D.; Cao, Z.; Bern, M.; Haab, B., Determining Lectin Specificity from Glycan Array Data using Motif Segregation and GlycoSearch Software. *Current Protocols in Chemical Biology* **2013,** 5, 1-13.
29. Chen, S.; LaRoche, T.; Hamelinck, D.; Bergsma, D.; Brenner, D.; Simeone, D.; Brand, R. E.; Haab, B. B., Multiplexed analysis of glycan variation on native proteins captured by antibody microarrays. *Nat Methods* **2007,** 4, (5), 437-44.
30. Yue, T.; Goldstein, I. J.; Hollingsworth, M. A.; Kaul, K.; Brand, R. E.; Haab, B. B., The prevalence and nature of glycan alterations on specific proteins in pancreatic cancer patients revealed using antibody-lectin sandwich arrays. *Mol Cell Proteomics* **2009,** 8, (7), 1697-707.
31. Yue, T.; Maupin, K. A.; Fallon, B.; Li, L.; Partyka, K.; Anderson, M. A.; Brenner, D. E.; Kaul, K.; Zeh, H.; Moser, A. J.; Simeone, D. M.; Feng, Z.; Brand, R. E.; Haab, B. B., Enhanced discrimination of malignant from benign pancreatic disease by measuring the CA 19-9 antigen on specific protein carriers. *PLoS ONE* **2011,** 6, (12), e29180.
32. Kostlanova, N.; Mitchell, E. P.; Lortat-Jacob, H.; Oscarson, S.; Lahmann, M.; Gilboa-Garber, N.; Chambat, G.; Wimmerova, M.; Imberty, A., The fucose-binding lectin from Ralstonia solanacearum. A new type of beta-propeller architecture formed by oligomerization and interacting with fucoside, fucosyllactose, and plant xyloglucan. *J Biol Chem* **2005,** 280, (30), 27839-49.
33. Schubert, M.; Bleuler-Martinez, S.; Butschi, A.; Walti, M. A.; Egloff, P.; Stutz, K.; Yan, S.; Collot, M.; Mallet, J. M.; Wilson, I. B.; Hengartner, M. O.; Aebi, M.; Allain, F. H.; Kunzler, M., Plasticity of the beta-trefoil protein fold in the recognition and control of invertebrate predators and parasites by a fungal defence system. *PLoS Pathog* **2012,** 8, (5), e1002706.
34. Walser, P. J.; Haebel, P. W.; Kunzler, M.; Sargent, D.; Kues, U.; Aebi, M.; Ban, N., Structure and functional analysis of the fungal galectin CGL2. *Structure* **2004,** 12, (4), 689-702.
35. Tateno, H.; Nakamura-Tsuruta, S.; Hirabayashi, J., Comparative analysis of core-fucose-binding lectins from Lens culinaris and Pisum sativum using frontal affinity chromatography. *Glycobiology* **2009,** 19, (5), 527-36.
36. Kobayashi, Y.; Tateno, H.; Dohra, H.; Moriwaki, K.; Miyoshi, E.; Hirabayashi, J.; Kawagishi, H., A novel core fucose-specific lectin from the mushroom Pholiota squarrosa. *J Biol Chem* **2012,** 287, (41), 33973-82.
37. Porter, A.; Yue, T.; Heeringa, L.; Day, S.; Suh, E.; Haab, B. B., A motif-based analysis of glycan array data to determine the specificities of glycan-binding proteins. *Glycobiology* **2010,** 20, (3), 369-380.
38. Maupin, K. A.; Liden, D.; Haab, B. B., The fine specificity of mannose-binding and galactose-binding lectins revealed using outlier-motif analysis of glycan array data. *Glycobiology* **2011,** 22, (1), 160-9.
39. Pour, P. M.; Tempero, M. M.; Takasaki, H.; Uchida, E.; Takiyama, Y.; Burnett, D. A.; Steplewski, Z., Expression of blood group-related antigens ABH, Lewis A, Lewis B, Lewis X, Lewis Y, and CA 19-9 in pancreatic cancer cells in comparison with the patient's blood group type. *Cancer Res* **1988,** 48, (19), 5422-6.

40. Balmana, M.; Sarrats, A.; Llop, E.; Barrabes, S.; Saldova, R.; Ferri, M. J.; Figueras, J.; Fort, E.; de Llorens, R.; Rudd, P. M.; Peracaula, R., Identification of potential pancreatic cancer serum markers: Increased sialyl-Lewis X on ceruloplasmin. *Clin Chim Acta* **2015,** 442C, 56-62.
41. Kobayashi, T.; Kawa, S.; Tokoo, M.; Oguchi, H.; Kiyosawa, K.; Furuta, S.; Kanai, M.; Homma, T., Comparative study of CA-50 (time-resolved fluoroimmunoassay), Span-1, and CA19-9 in the diagnosis of pancreatic cancer. *Scand J Gastroenterol* **1991,** 26, (7), 787-97.
42. Kawa, S.; Tokoo, M.; Hasebe, O.; Hayashi, K.; Imai, H.; Oguchi, H.; Kiyosawa, K.; Furuta, S.; Homma, T., Comparative study of CA242 and CA19-9 for the diagnosis of pancreatic cancer. *Br J Cancer* **1994,** 70, (3), 481-6.
43. Kawa, S.; Tokoo, M.; Oguchi, H.; Furuta, S.; Homma, T.; Hasegawa, Y.; Ogata, H.; Sakata, K., Epitope analysis of SPan-1 and DUPAN-2 using synthesized glycoconjugates sialyllact-N-fucopentaose II and sialyllact-N-tetraose. *Pancreas* **1994,** 9, (6), 692-7.
44. Kawa, S.; Oguchi, H.; Kobayashi, T.; Tokoo, M.; Furuta, S.; Kanai, M.; Homma, T., Elevated serum levels of Dupan-2 in pancreatic cancer patients negative for Lewis blood group phenotype. *Br J Cancer* **1991,** 64, (5), 899-902.
45. Shida, K.; Korekane, H.; Misonou, Y.; Noura, S.; Ohue, M.; Takahashi, H.; Ohigashi, H.; Ishikawa, O.; Miyamoto, Y., Novel ganglioside found in adenocarcinoma cells of Lewis-negative patients. *Glycobiology* **2010,** 20, (12), 1594-606.
46. Gestwicki, J. E.; Cairo, C. W.; Strong, L. E.; Oetjen, K. A.; Kiessling, L. L., Influencing receptor-ligand binding mechanisms with multivalent ligand architecture. *J Am Chem Soc* **2002,** 124, (50), 14922-33.
47. Dam, T. K.; Brewer, C. F., Multivalent lectin-carbohydrate interactions energetics and mechanisms of binding. *Advances in Carbohydrate Chemistry and Biochemistry* **2010,** 63, 139-64.
48. Brewer, C. F.; Miceli, M. C.; Baum, L. G., Clusters, bundles, arrays and lattices: novel mechanisms for lectin-saccharide-mediated cellular interactions. *Curr Opin Struct Biol* **2002,** 12, (5), 616-23.
49. Varki, A. P.; Baum, L. G.; Bellis, S. L.; Cummings, R. D.; Esko, J. D.; Hart, G. W.; Linhardt, R. J.; Lowe, J. B.; McEver, R. P.; Srivastava, A.; Sarkar, R., Working group report: the roles of glycans in hemostasis, inflammation and vascular biology. *Glycobiology* **2008,** 18, (10), 747-9.

Figure Captions

Figure 1. Identification of lectins recognizing specific fucose linkages. A) Each fucose linkage appears in a variety of basic motifs, with several variants beyond those shown. B) We queried the GlycanBinder database to identify the lectins with the highest probability of binding α1-2, α1-3, α1-4, or α1-6 fucose. We examined the output to identify those that were selective for one of the motifs and further analyzed the glycan array data of particular lectins to characterize specificity. C) Certain lectins bind motifs containing fucose in all linkages, but other lectins bind motifs containing fucose only in specific linkages. The numbers are the Motif Scores, which indicate the probability that a lectin binds a given motif. The Motif Scores ≥ 3 are shown, which are those with high statistical significance ($p < 0.001$).

Figure 2. Purification and assessment of multimeric lectins. A) The crystal structures obtained from the PDB show the tetrameric conformation of CGL2 and the trimeric form of RSL. Only the monomeric form was available for CCL2. For each lectin, the C-terminus is more extended from the protein than the N-terminus. B) We used size-exclusion chromatography (top) to purify and assess the size and state of both versions of the trimeric RSL (data for other two lectins shown in Fig. S1), and we analyzed selected fractions by polyacrylamide gel electrophoresis (bottom). For RSL, both versions produced the right size of monomer, but the version with Avitag on the C-terminus formed more trimer. The assessments of multimerization are based on the elution time in size-exclusion chromatography (comparing to a size standard) and the sizes of the bands in the PAGE analysis of the unheated samples. We selected fractions for further purification by ion-exchange chromatography (not shown) and selected fractions with high purity for use in the experiments.

Figure 3. Testing for unique presentations of fucose in subsets of cancer patients. A) We printed arrays containing diverse antibodies, each targeting its own distinct glycan. We incubated a plasma sample on each antibody array and probed each array with a glycan-binding antibody or lectin. Each capture:detection combination measures the co-expression of the glycan motifs bound by the capture and detection reagents. B-E) The columns represent the fluorescence from the indicated capture:detection pairs for each sample. The dashed lines represent thresholds defining biomarker elevation, and the asterisks indicate the samples that are elevated in one of the markers.

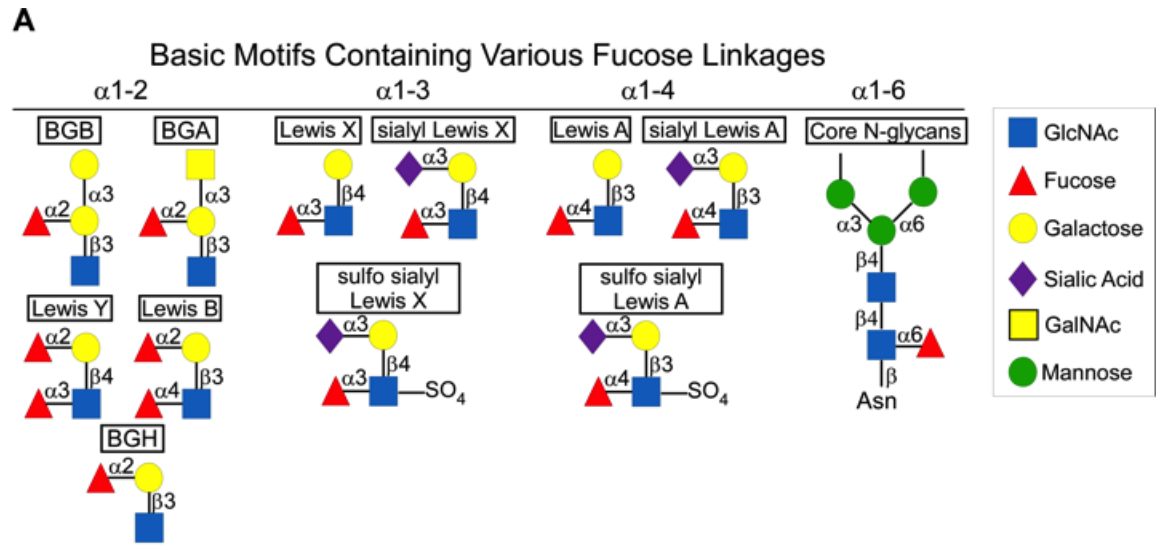
Figure 4. Analysis of the specificities of purified CCL2 and recombinant CCL2. A) We obtained data from the CFG for native, purified CCL2 incubated on array version 3.1. The columns indicate the signal at selected glycans. Glycans containing the Fucα1-3 motif and not containing the Lewis Y motif are at left, followed by glycans containing both Fucα1-3 and Lewis Y, followed by glycans containing neither. None of the 433 glycans not included in the graph showed signal above those in the graph. B) We also obtained new glycan array data using the recombinant version of CCL2 with biotinylation at the C-terminus incubated on CFG array version 5.2. The x-axis presents measurements from the various glycans, ordered by glycan ID. The structures of the top glycans are given, showing the Fucα1-3 motif in each case. C) The data from panel B are presented according to the ordering of panel A, showing the good agreement between the recombinant lectin and the purified lectin.

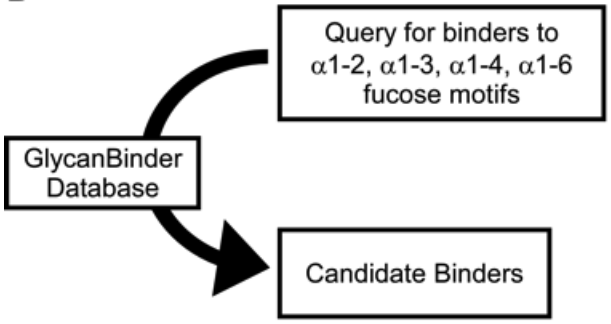
Figure 5. CCL2 detection as a biomarker to complement to sLeA. A) The measurements from the anti-sLeA sandwich assay (upper panel) are ordered from high to low for the samples from cancer patients and control subjects. The dashed line represents a cutoff chosen to give optimal discrimination between the cancer and control patient samples. The measurements from anti-sLeA capture and CCL2 detection (lower

panel) are ordered the same as in panel A. The dashed line represents a cutoff chosen to give no false-positive elevations in addition to those identified using the anti-sLeA assay in panel A. The asterisks indicate samples that showed elevations using CCL2 detection but not using sLeA detection. B) Each point represents a patient sample, detected using the indicated capture and detection reagents. C) The graphs show a comparison of the levels of the cancer samples and control sample for the indicated assays. In each graph, the same capture antibodies were used, but the detection reagents are different. The measurements obtained from detection with CCL2 are presented in the x-axes, and the measurements obtained using sLeA (left) and sLeX (right) are in the y-axex. The dashed line represents cutoffs chosen to give no false positive elevations. Some of the cancer patient samples are elevated only using CCL2 detection.

Figure 6. Complementary glycan expression in pancreatic tumors. A) The measurements using sLeA detection (left) and CCL2 detection (right) are ordered by decreasing levels of the sLeA sandwich pair. B) The three panels left, center and right shows staining with the anti-sLeA antibody, the anti-sLeX antibody, and CCL2, respectively.

Figures





C

Motif Score	Coprinopsis cinerea lectin (CCL2)	Coprinopsis cinerea galectin 2 (CGL2)	Ralstonia solanacearum lectin (RSL)	Aleuria aurantia lectin (AAL)	Aspergillus fumigatus lectin (AFL)	Shiga-like toxin (SLT)
	Lectins					
Fuc ß	-	-	-	-	-	-
Fuc α 1-2	-	6.0	11.8	7.4	6.8	-
Fuc α 1-3	9.3	-	8.6	8.9	7.5	6.4
Fuc α 1-4	-	-	5.6	5.4	5.7	-
Fuc α 1-6	-	-	3.8	-	-	3.2
Fuc α anywhere	5.5	3.6	16.4	13.2	11.7	5.7

Figure 1

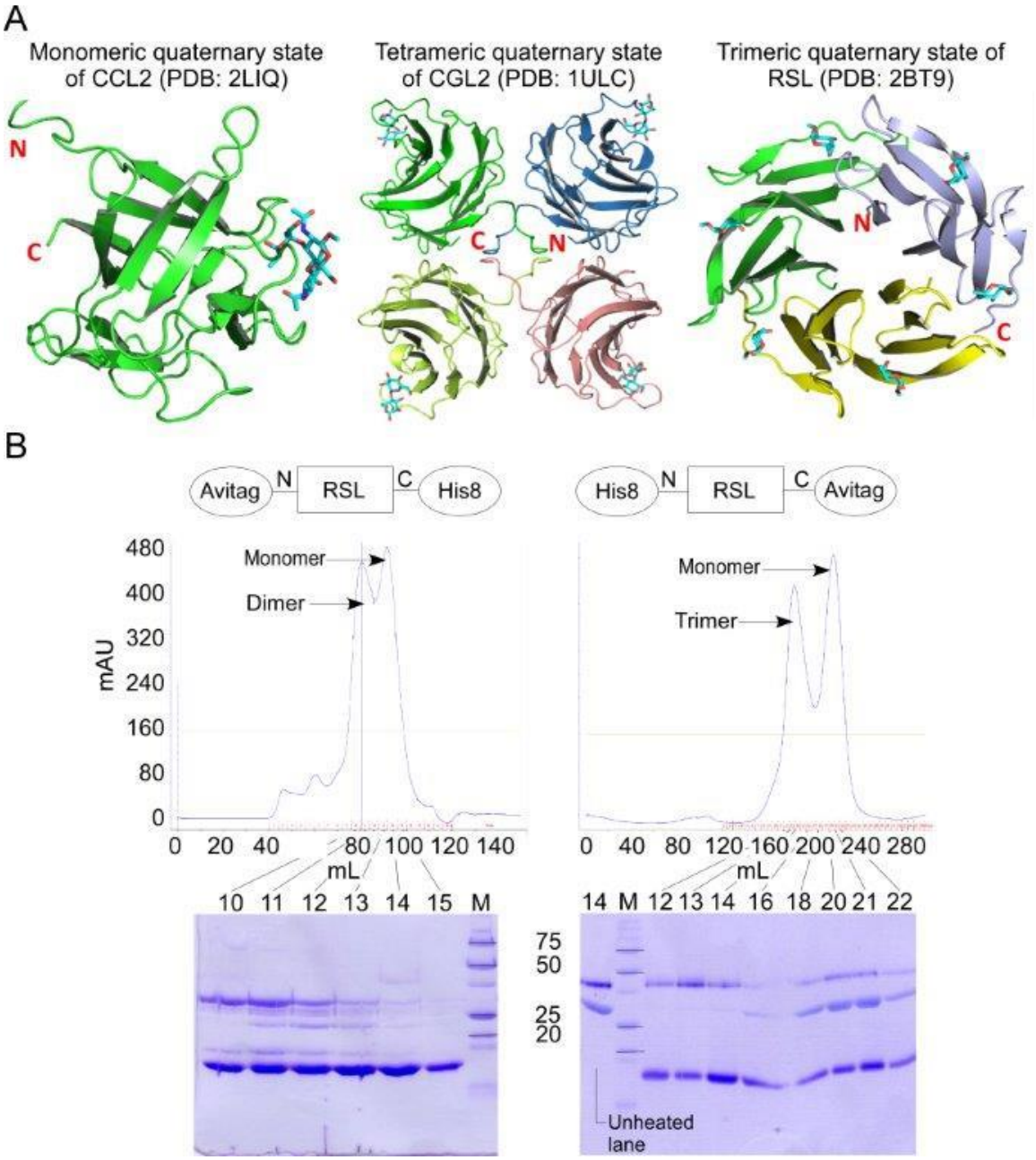

A
Monomeric quaternary state of CCL2 (PDB: 2LIQ)
Tetrameric quaternary state of CGL2 (PDB: 1ULC)
Trimeric quaternary state of RSL (PDB: 2BT9)
N
C
B
Avitag
RSL
His8
Monomer
Dimer
Trimer
mAU
mL
10 11 12 13 14 15 M
14 M 12 13 14 16 18 20 21 22
75
50
25
20
Unheated lane

Figure 2

Probing fucose linkages

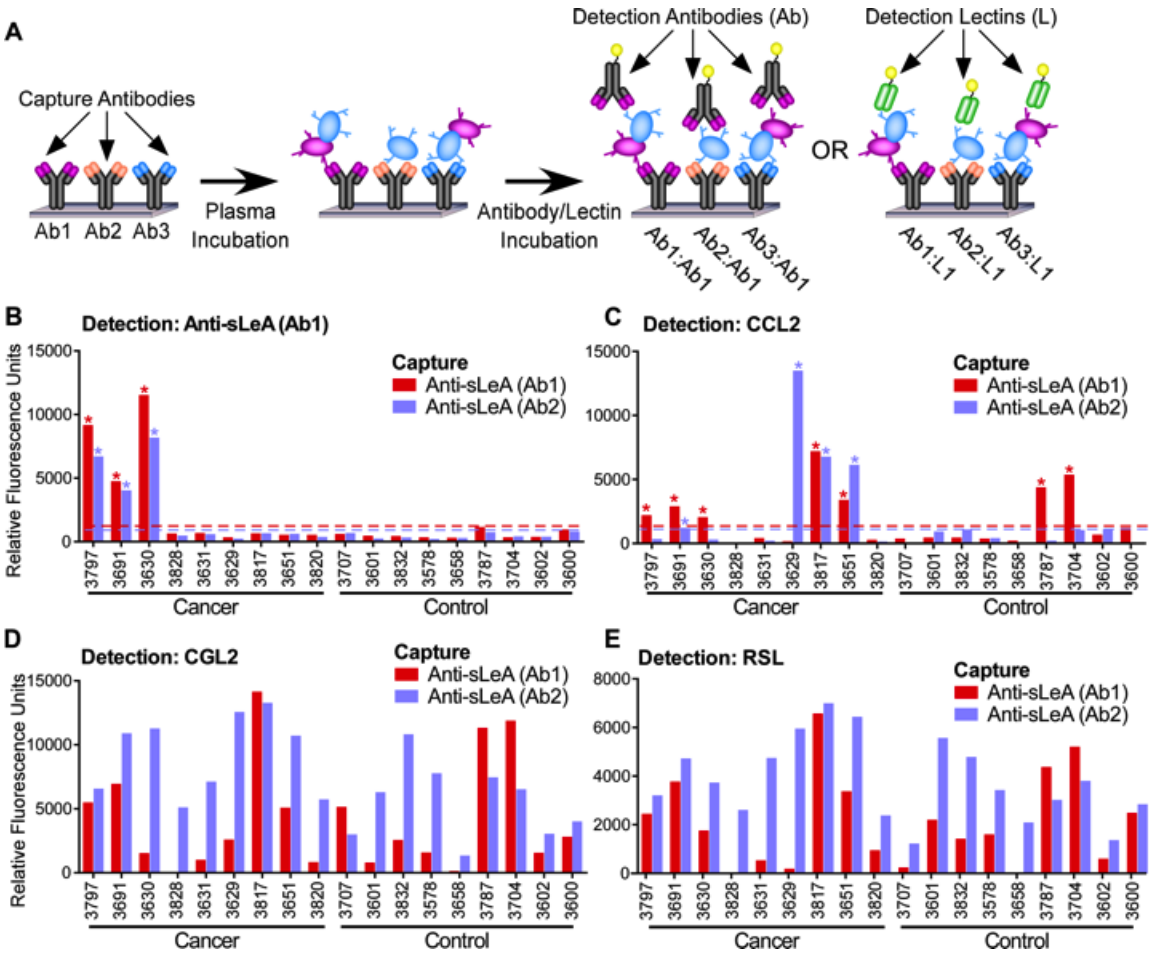


Figure 3

Probing fucose linkages

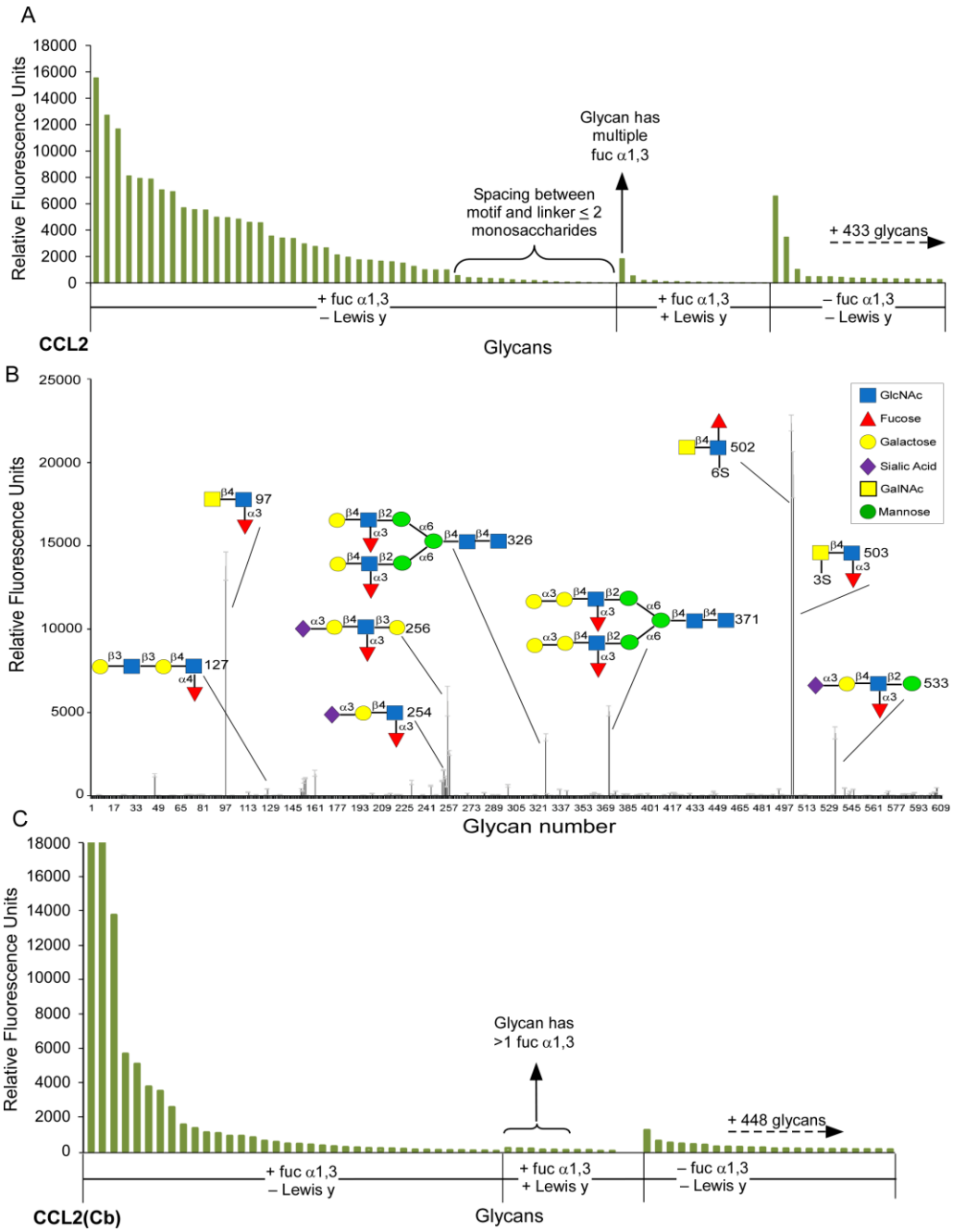


Figure 4

Probing fucose linkages

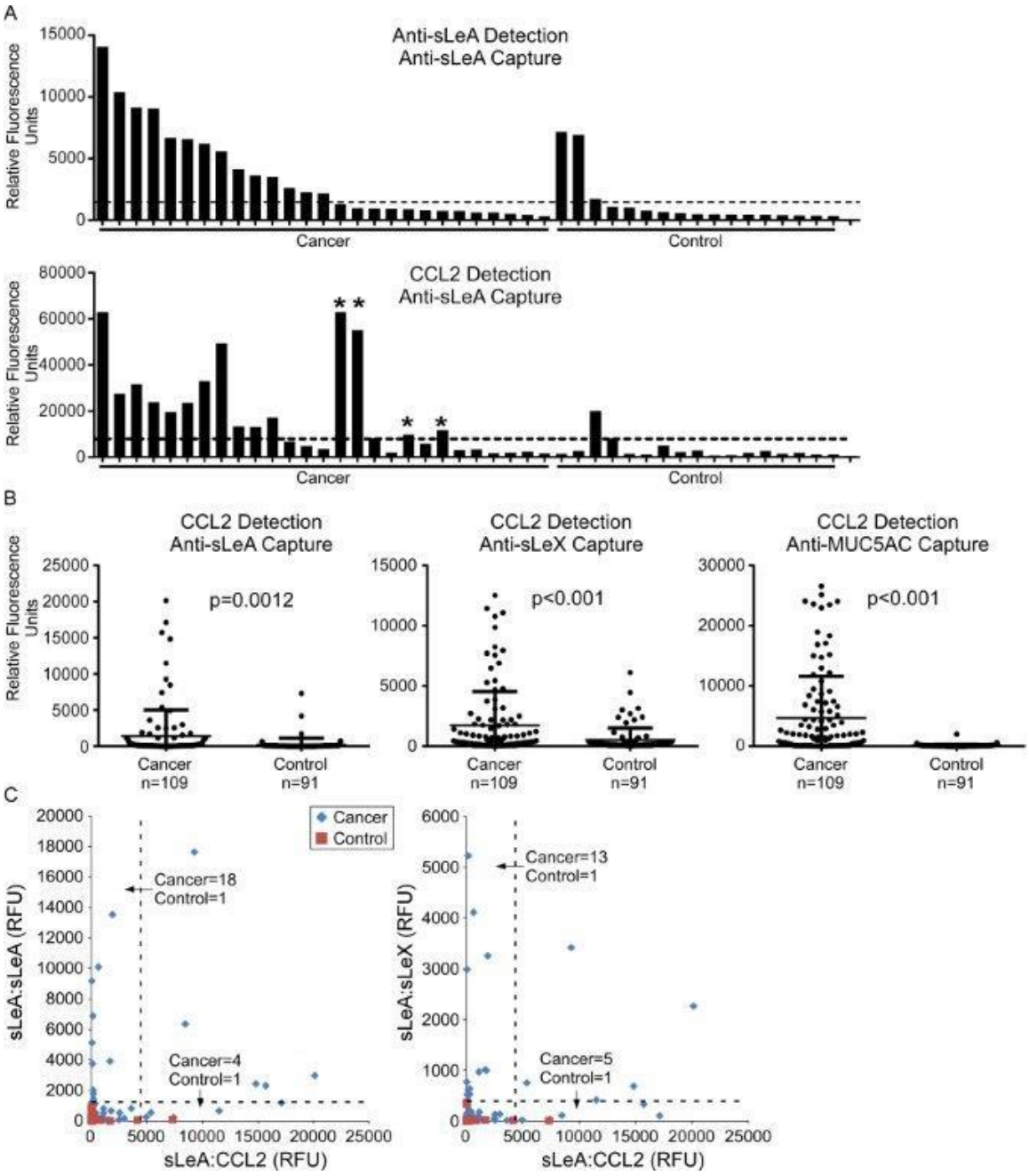


Figure 5

Probing fucose linkages

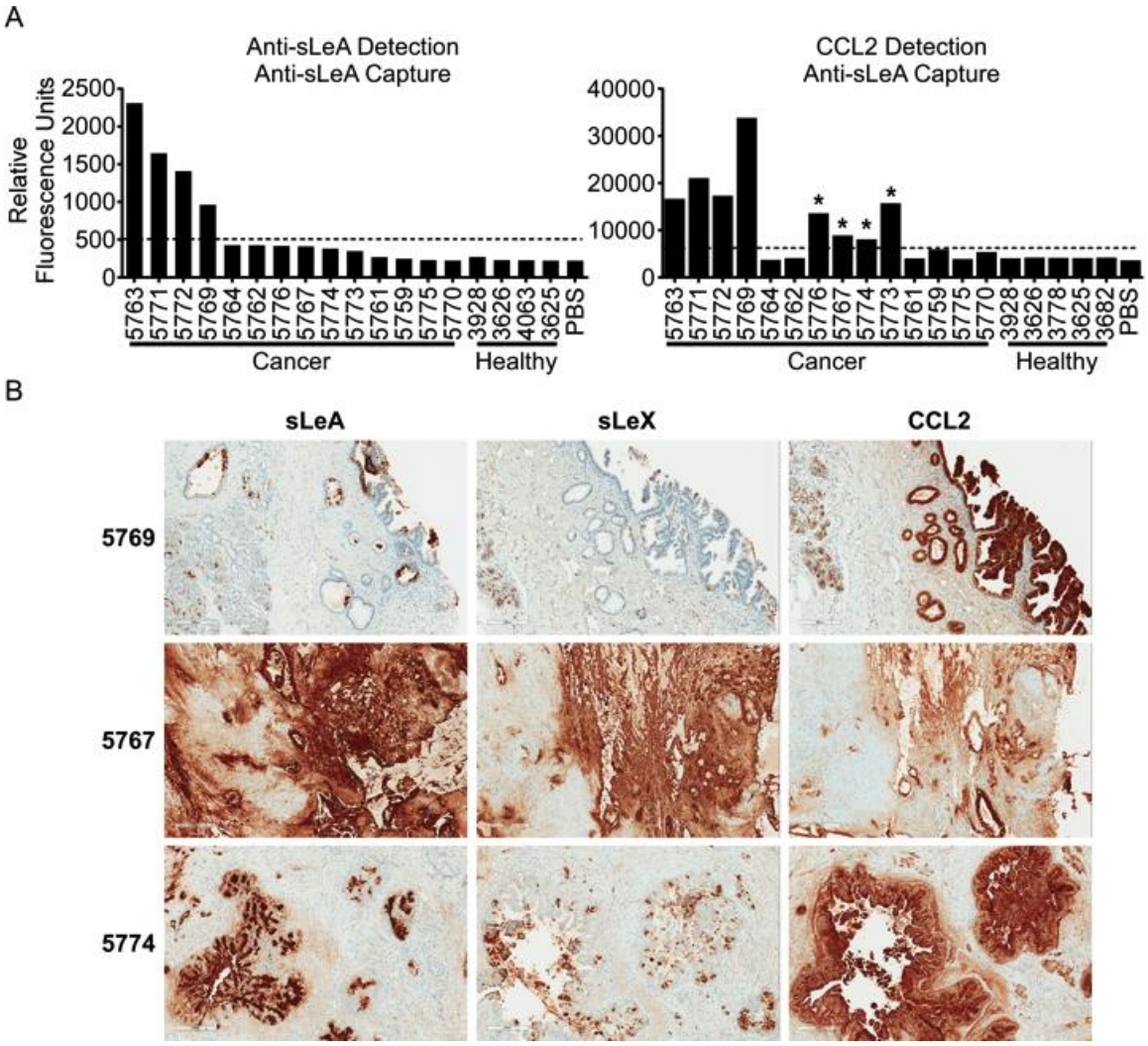
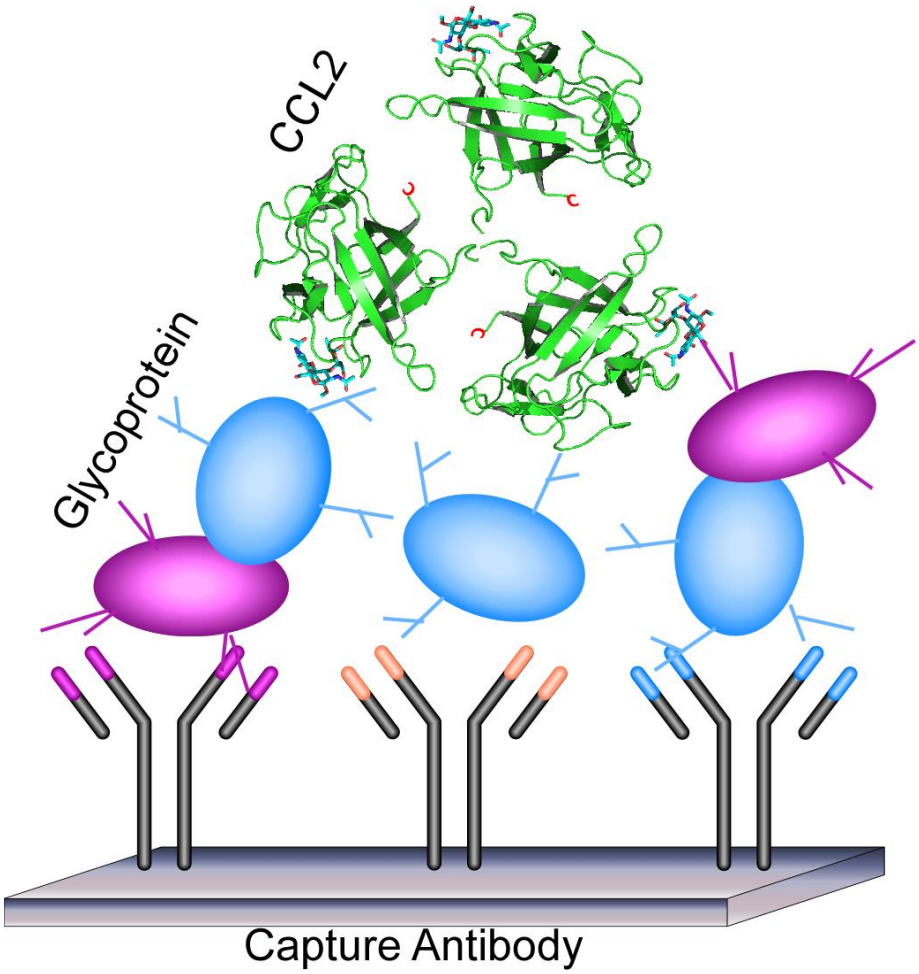


Figure 6



TOC Graphic

Supplementary Information, Singh et al., Probing fucose linkages

Contents:

Supplementary Figures

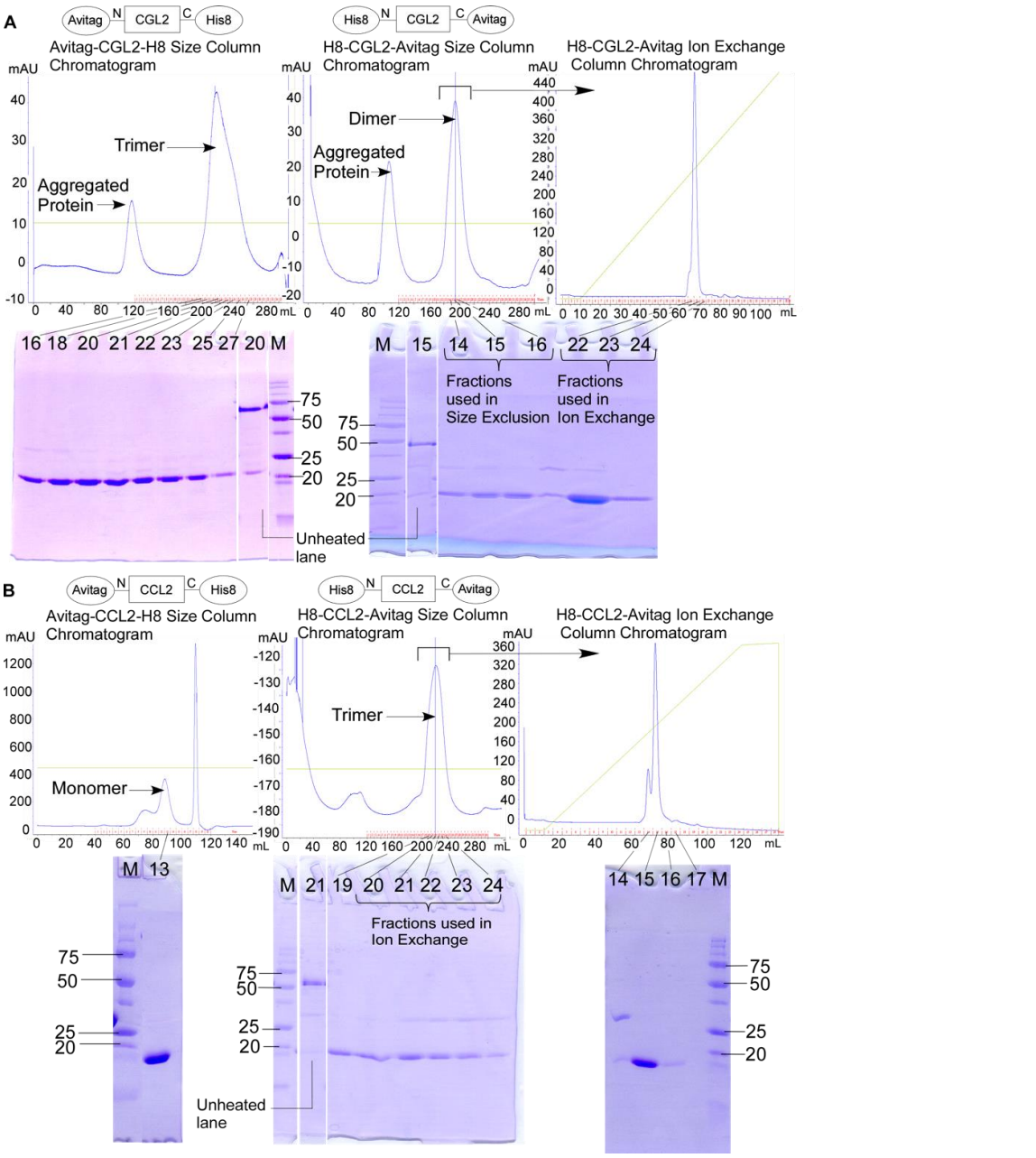


Figure S1: Purification and evaluation of CGL2 and CCL2. We used size-exclusion chromatography to assess the sizes and states of both versions of CGL2 (A) and CCL2 (B). The indicated fractions were run by polyacrylamide gel electrophoresis (bottom images in A and B). Certain fractions from both lectins were used for further purification by ion-exchange chromatography (right). CGL2 formed a trimer with the N-Avitag version and a dimer with the C-Avitag version. CCL2 was a monomer with the N-Avitag version and a trimer with the C-Avitag version. The assignments of quaternary state are based on the elution time in size-exclusion chromatography (comparing to a size standard) and the sizes of the bands in the PAGE analysis of the unheated samples.

Probing fucose linkages

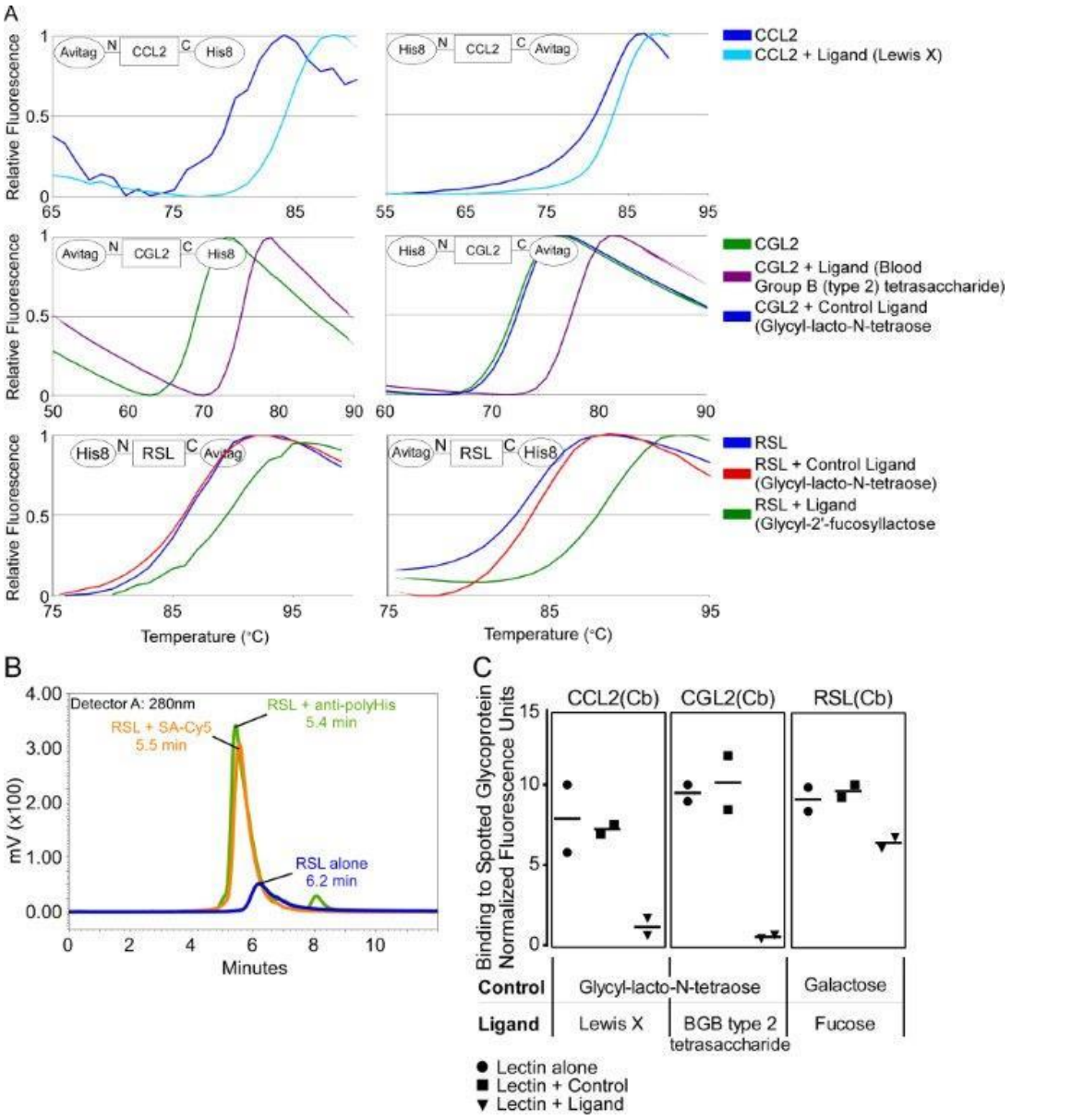


Figure S2: Confirmation of glycan binding. A) Thermal shift assays. Thermal shift assays were useful as an initial test for glycan binding because they are homogenous (entirely in solution) and do not require a secondary detection reagent. A solution containing the lectin, a glycan ligand, and a fluorophore is gradually heated, and fluorescence is measured at each temperature. Upon denaturation of the lectin, the lectin exposes hydrophobic regions to which the fluorophore adheres, stabilizing the fluorophore and increasing fluorescence. Thus the temperature at which fluorescence increases provides a measure of the stability of the protein. A lectin that has specifically bound a glycan ligand should be more stable than a free lectin. For both versions of both lectins, we observed a shift in stability when the lectins were incubated with their target glycans. This result is consistent with proper glycan binding for each protein. B) Confirmation of tag accessibility using analytical size-exclusion chromatography. To test the

accessibility of the fusion tag to a secondary detection reagent, we incubated each lectin with either streptavidin or an anti-poly-histidine antibody and separated the mixture by analytical size-exclusion chromatography. In each case, the protein elution was entirely shifted to a higher molecular weight, indicating that both the biotin and poly-histidine tags were accessible at either terminus. The example here shows the version of RSL with Avitag on the N-terminus. C) Competition assays to assess binding to the glycans of immobilized glycoproteins. We incubated the lectin alone, the lectin plus a control glycan, and the lectin plus a specific glycan ligand on microarrays containing glycoproteins, and we then measured the binding of the lectin to the glycoproteins. The glycans were incubated at 300-fold, 50-fold, and 250-fold molar excess relative to CCL2, CGL2, and RSL, respectively. The representative data presented here shows the version of each lectin with Avitag on the C-terminus, detected with the SA-Cy5 secondary reagent. In each case, we saw reduced binding to the glycoproteins upon pre-incubation with a competing ligand but not upon pre-incubation with a non-specific glycan.

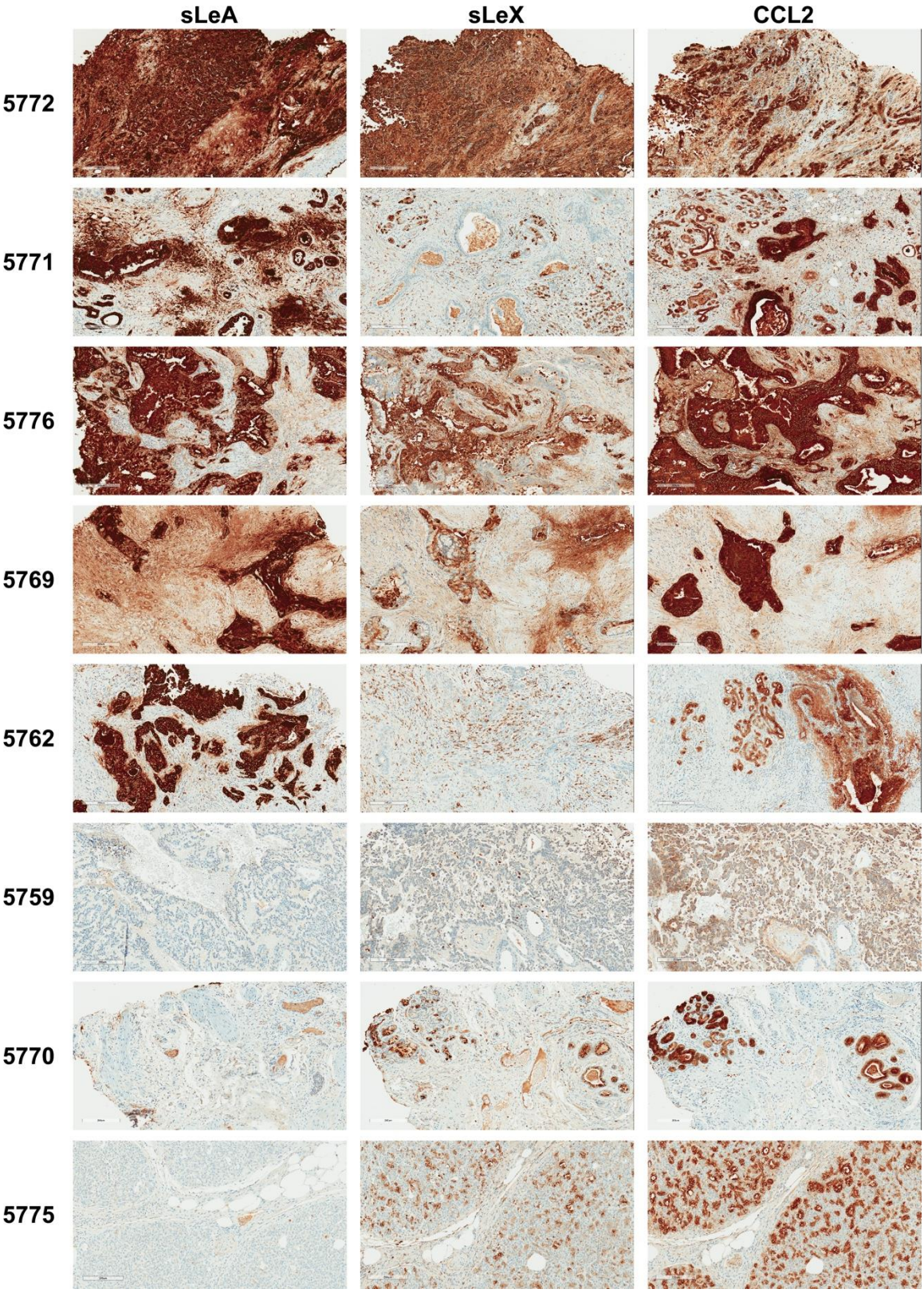


Figure S3. Complementary staining patterns of sLeA, sLeX, and CCL2.