

Molecular and structural basis for Lewis glycan recognition by a cancer-targeting antibody

Caroline Soliman¹, Andrew J. Guy¹, Jia Xin Chua², Mireille Vankemmelbeke², Richard S. McIntosh³, Sarah Eastwood¹, Vi Khanh Truong¹, Aaron Elbourne¹, Ian Spendlove³, Lindy G. Durrant^{2,3}, Paul A. Ramsland^{1,4,5,*}

¹School of Science, RMIT University, Melbourne, VIC, Australia.

²Scancell Limited, University of Nottingham Biodiscovery Institute, University Park, Nottingham, United Kingdom.

³Division of Cancer and Stem Cells, School of Medicine, University of Nottingham Biodiscovery Institute, University Park, Nottingham, United Kingdom.

⁴Department of Immunology, Central Clinical School, Monash University, Melbourne, Victoria, Australia.

⁵Department of Surgery Austin Health, University of Melbourne, Melbourne, Victoria, Australia.

*Corresponding author: Associate Professor Paul A. Ramsland, School of Science, RMIT University, Plenty Road, Bundoora, Victoria 3083, Australia. Phone +61 3 9925 7024, email paul.ramsland@rmit.edu.au

Keywords

Carbohydrate-binding antibody, Lewis glycans, cancer-targeting antibody, molecular docking

Abbreviations

3D, three dimensional; ADCC, antibody-dependent cellular cytotoxicity; CDC, complement-dependent cytotoxicity; CDR, complementarity-determining region; EthD-1, Ethidium homodimer; Fab, fragment antigen-binding; FUC, fucose; Fv, variable fragment; GAL, galactose; H chain, heavy chain; L chain, light chain; Le, Lewis glycan; mAb, monoclonal antibody; MD, molecular dynamics; NAG, N-acetylglucosamine; PDB, protein data bank; RMSD, root mean square deviation; SEM, scanning electron microscopy; WHO, World Health Organisation.

Abstract

Immunotherapy has been successful in treating many tumour types. The development of additional tumour-antigen binding monoclonal antibodies (mAbs) will help expand the range of immunotherapeutic targets. Lewis histo-blood group and related glycans are overexpressed on many carcinomas, including those of the colon, lung, breast, prostate and ovary, and can therefore be selectively targeted by mAbs. Here we examine the molecular and structural basis for recognition of extended Le^a and Le^x containing glycans by a chimeric mAb. Both the murine (FG88.2) IgG3 and a chimeric (ch88.2) IgG1 mAb variants showed reactivity to colorectal cancer cells leading to significantly reduced cell viability. We determined the X-ray structure of the unliganded ch88.2 fragment antigen-binding (Fab) containing two Fabs in the unit cell. A combination of molecular docking, glycan grafting and molecular dynamics simulations predicts two distinct subsites for recognition of Le^a and Le^x trisaccharides. While light chain residues were exclusively used for Le^a binding, recognition of Le^x involved both light and heavy chain residues. An extended groove is predicted to accommodate the Le^a-Le^x hexasaccharide with adjoining subsites for each trisaccharide. The molecular and structural details of the ch88.2 mAb presented here provide insight into its cross-reactivity for various Le^a and Le^x containing glycans. Furthermore, the predicted interactions with extended epitopes likely explains the selectivity of this antibody for targeting Lewis-positive tumours.

Introduction

Cancer is a leading cause of mortality worldwide, currently accounting for around 1 in 6 deaths, with the highest incidences for lung, breast and colorectal cancer. According to the World Health Organisation (WHO), lung cancer accounts for 18.4% of cancer related deaths, followed by 9.2% for colorectal cancer, with breast cancer accounting for only 6.6% due to a relatively more favourable prognosis. In addition, prostate cancer ranks second for incidence in men, although the prognosis is often favourable, and cervical cancer ranks fourth for incidence as well as mortality in women, although rates are decreasing due to early prevention methods [1]. While several different targeted therapies have recently been approved for individual or combination therapy for cancer [2], they only cover a small subset of cancers. Consequently, there is still a distinct lack of targeted treatments available for many types of cancer.

Glycosylation of malignant cells differs significantly to that of healthy cells and is considered a key contributing factor to the hallmarks of cancer [3, 4]. Aberrant glycosylation in cancer has been associated with tumour progression and metastasis, and consequently, glycans are potential targets for therapeutic antibodies [3, 5]. Lewis glycans are fucosylated determinants known as Le^a, Le^b, Le^x and Le^y and sialylated versions, sLe^x and sLe^a. Type I glycans, also known as histo-blood group antigens, include Le^a and Le^b, which are formed by a Gal β 1-3GlcNAc core. This core glycan determinant has been referred to as the Le^c antigen. While it has not been as extensively studied as other Lewis glycans, Le^c is specific to Le^a and Le^b negative red cells and has been considered a precursor to Le^a and Le^b (Figure S1) [6-10]. Related Le^x and Le^y glycans are developmental antigens, classified as type II glycans, with very restricted distribution on adult somatic cells. While Le^a and Le^x trisaccharides are structurally similar, they differ in the position of the glycosidic linkages of the outermost galactose (Gal) and fucose (Fuc) to the core N-acetyl-glucosamine (GlcNAc). Consequently, the minimal determinants are Gal β 1-3(Fuc α 1-4)GlcNAc for Le^a and Gal β 1-4(Fuc α 1-3)GlcNAc for Le^x (Figure S1) [11]. Lewis antigens are aberrantly expressed on tumours derived from tissues that are often normally negative for that specific glycan [12]. In particular, Lewis antigens are known to be over-expressed on breast, lung, colorectal, prostate and ovarian cancers [13].

Antibody recognition of Lewis glycans Le^x and Le^y has been studied in more detail as compared to other blood-group glycans. Five crystal structures of different antibodies bound to Lewis glycans have been determined in addition to three structures of unbound anti-Lewis antibodies [14]. Both BR96 and hu3S193 mAbs target Le^y and co-crystal structures of each antibody reveals carbohydrate recognition occurs in large binding pockets [15-17]. While 291-2G3-A and 54-5C10-A mAbs both target Le^x and were developed against *Schistosoma* parasites, only one has been resolved in complex with Le^x and involves a

shallow binding pocket [18, 19]. Although each of these anti-Lewis antibodies have been well characterised, with some tested in clinical trials [12], none are currently approved as therapeutic mAbs.

Recently, a mAb known as FG88.2, raised against plasma membrane extracts from colorectal cancer cells, was shown to target Lewis glycans found on a wide range of colorectal, pancreatic, gastric, non-small cell lung and ovarian tumours. When screened against glycan arrays, the FG88.2 mAb bound to Le^a-containing glycans and extended epitopes including di-Le^a, Le^a-Le^x, and Le^cLe^x. Importantly the FG88.2 mAb displayed minimal binding to many other mammalian glycans or normal tissues indicating its capacity for Lewis-positive tumour selectivity. In addition, the antibody demonstrated antibody-dependent cellular cytotoxicity (ADCC), complement-dependent cytotoxicity (CDC) and direct (caspase-independent) tumour cell killing. It was also shown to internalise, colocalise with lysosomes and deliver saporin that killed cells with subnanomolar potency. *In vivo* studies revealed potent anti-tumour efficacy in a metastatic colorectal xenograft tumour model in mice, leading to significant long-term survival. This work indicates the potential of FG88 as a therapeutic mAb for the treatment of multiple solid tumours [20].

Here we characterise the potential molecular basis for recognition of Lewis glycans (Le^a and Le^x) by the chimeric antibody ch88.2. Both mouse (FG88.2) and chimeric (ch88.2) variants showed reactivity to colorectal cancer cells, with mAb binding significantly reducing cell viability. We determined the X-ray structure of ch88.2 Fab, with both Fabs in the unit cell displaying similar structures with some differences within the CDR loops, particularly in the H chain. A mixture of molecular docking, glycan grafting and molecular dynamics simulations predicted a binding region with distinct sites for Le^a and Le^x glycan epitopes. The Le^a-Le^x hexasaccharide, which was previously identified by glycan array as the top binding motif [20], was found to interact via both its Le^a and Le^x components. Taken together, the structural and computational analysis suggest that the FG88.2/ch88.2 mAbs are capable of binding multiple extended Lewis glycans for effective targeting of Lewis-positive tumours.

Materials and Methods

Generation of ch88.2 mAb. Chimeric 88.2 (ch88.2) IgG1 vector were created by introducing the heavy and light chain variable domain genes of the FG88.2 IgG3 mAb into the pDCOrig vector. Construction of the chimeric antibody is described in detail in [21]. For this study, mAb was expressed in Expi293F™ cells using the ExpiFectamine™ 293 Transfection kit (Gibco, Life Technologies). Briefly, 100 µg plasmid DNA was combined with ExpiFectamine™ 293 Reagent in Opti-MEM Medium and added to a suspension of HEK293 cells (100 mL, 2×10^6 /mL). ExpiFectamine™ 293 Enhancers were added 20 hours later, and cells were harvested 7 days post-transfection. For harvesting, mAb-containing supernatant was filtered (0.22 µm bottle top filters, Merck Millipore), stored in sodium azide (final concentration of 0.2% w/v), and purified on protein G columns (HiTrap ProteinG HP, GE Healthcare) using an AKTA FPLC (GE Healthcare). Fractions containing mAb were eluted (100 mM glycine pH12 containing 0.05% v/v Tween 20), pooled, neutralised to pH 7.0 (with 1 M HCl) and dialyzed against PBS prior to storage.

Cell culture. The human adenocarcinoma cell line COLO205 (ATCC: CCL-222) was obtained from Cell Bank Australia, authenticated using short tandem repeat profiling and quality controlled using Biotool Mycoplasma Detection Kit-QuickTest. Cells were cultured in 25 cm² and 75 cm² tissue culture flasks under sterile conditions using complete RPMI 1640 media with L-glutamine (containing 10% Fetal Bovine Serum and 1% Penicillin-Streptomycin) at 37°C with 5% CO₂. Cells were passaged when 70% confluent using 0.25% Trypsin-EDTA and cultured to a maximum of 15 passages. Cells were harvested periodically in freeze media (FBS with 10% DMSO) for storage at -80°C and for long-term storage in liquid nitrogen.

Immunofluorescent detection of antibody binding. Glass coverslips were coated with poly-D-lysine at 1 mg/mL and cells were seeded overnight in complete RPMI media. Cells were fixed with 2% paraformaldehyde, blocked with 10% goat serum and primary antibody was then added at a concentration of 5 µg/mL in PBS for 1 hour at room temperature. Secondary antibody, being either F(ab')₂-Goat anti-Human IgG Fc PE or goat anti-mouse IgG (H+L) AF488 (Thermo Fisher Scientific), was added at 1-2 µg/ml in PBS for 1 hour at room temperature. DAPI was then used to stain the DNA and coverslips were mounted with Mowiol 4-88 (Merck). Samples were examined using a Leica DM2500 epifluorescence microscope with a DFC310 digital camera, and images were captured using LAS software (V4.1; Leica Microsystems).

Live/Dead cell staining and fluorescence microscopy. Glass coverslips were coated with poly-D-lysine at 1 mg/mL and cells were seeded overnight in complete RPMI media. Cells were treated with 30 µg/mL of antibody for 24 hours at 37°C with 5% CO₂. Samples were then incubated with a solution of 4 µM

Ethidium homodimer (EthD-1) and 2 μ M Calcein-AM in PBS for 45 minutes at room temperature. Coverslips were mounted onto slides and examined by fluorescence microscopy for red (dead) and green (live) fluorescence. Samples were examined using a Leica DM2500 epifluorescence microscope with a DFC310 digital camera, and images were captured using LAS software (V4.1; Leica Microsystems).

Flow cytometry for antibody binding and antibody-mediated killing. To assess antibody binding, cells were seeded at 1×10^5 cells/well and treated with primary antibody followed by secondary antibody (as per immunofluorescent detection method above). Monoclonal human IgG1 (F598) [22] and mouse IgG3 (49-31.1) [23] antibodies against irrelevant targets were available in house and were used as isotype controls. Samples were fixed in 2% paraformaldehyde before being analysed in the FACS Canto (10,000 cells per sample). To assess antibody killing, cells were seeded at 1×10^5 cells/well and treated with 30 μ g/mL of antibody in complete RPMI media for 24 hours at 37°C with 5% CO₂. As a control for total cell lysis (positive control), cancer cells were treated with 0.1% Triton X-100 in PBS for 15 minutes. Cells were imaged in a 1 μ g/mL propidium iodide solution on the FACS Canto (10,000 cells per sample). The resultant data was analysed by FlowJo software where cell populations were gated by cell size and complexity, following which AF488 or propidium iodide positive populations (FITC or PE emission $>10^3$) were selected. One-way ANOVA tests were used to assess statistical significance in cell death (non-viability) compared to the negative controls. Individual data points are shown along with the mean (n = 2 for ch88 antibody binding from 1 experiment or n = 4 for all other samples from 2 experiments), and standard deviation was determined for antibody killing (n = 4).

Scanning Electron Microscopy (SEM). Silicon wafer surfaces were coated with poly-D-lysine at 1 mg/mL and cells were seeded overnight in complete RPMI media before being treated with 30 μ g/mL of antibody for 24 hours at 37°C with 5% CO₂. For cellular imaging, samples were affixed using 3% glutaraldehyde, dehydrated and coated with a thin film of gold. Scanning electron micrographs were obtained using a FEI Verios 460L field-emission scanning electron microscope (FE-SEM) (FEI Company, Oregon, USA) operated with an accelerating voltage of 2-5 kV. The resultant images were analysed using the Gwyddion and Image J software suites.

Fab Production. A Pierce Fab Preparation Kit (Thermo Fisher Scientific) was used to produce Fab from IgG. Briefly, 0.5 mL of an 8 mg/mL IgG sample was digested with Papain-agarose for 6 hours at 37°C. Fc and residual intact IgG was separated from Fab using protein A affinity chromatography. Coomassie stained SDS-PAGE (precast 4-15% Bis-Tris Mini Gels and MES running buffer, BioRad) was used to examine the purity of the Fab compared to the Fc and intact IgG samples under non-reducing and reducing (β -mercaptoethanol) conditions. Fab was quantitated by absorbance at 280 nm (Nanodrop) assuming a mass extinction (E, derived concentrations in mg/mL) value of 1.0 (1.37 for intact IgG). DLS

was used to determine protein size and polydispersity using a Zetasizer Nano ZS instrument (Malvern Instruments). Cuvettes containing 100 μL of protein sample were measured (five replicates) at 25°C, and time-correlated light scattering data were analysed as size-distributions by scattered intensities (histograms of particle diameter (nm) versus % intensity). The Z-average hydrodynamic diameters (D_H , in nm) and overall polydispersity of samples were estimated by the cumulants method.

Crystallisation of ch88.2 Fab. For crystallisation, ch88.2 Fab was dialysed overnight into ultrapure water (Milli-Q) using Slide-A-Lyser MINI Dialysis devices with a 10-kDa-cutoff. Fab samples were concentrated to 10-20 mg/mL using Pall Corp Nanosep OMEGA devices with a 10 kDa cut-off membrane. Crystals of ch88.2 Fab were formed in the presence of either Le^x or Le^a trisaccharide (Carbosynth) with a 5-fold molar excess of carbohydrate over the Fab sample. Crystals were generated in sitting-drop vapor diffusion experiments at 18°C in a 96-well sitting drop plate (Corning) using the Crystal Screen High Throughput kit (Hampton Research). Reservoirs of 80 μL of each crystallisation condition were used with 1 μL of Fab: Le^x sample mixed with 1 μL of the same condition for the sitting drop. While ch88 Fab crystals with Le^a were obtained, the best Fab crystals were obtained with Le^x . Crystals suitable for X-ray diffraction were produced in the crystallisation condition composed of 20% v/v polyethylene glycol (PEG) monomethyl ether (MME) 200, 0.1 M TRIS pH 8.5 and 0.1 M Nickel (II) Chloride. For data collection, crystals were mounted in nylon loops (Hampton Research) in the drop solution and plunged into liquid nitrogen.

Data collection and structure determination. Diffraction data were collected at the Australian Synchrotron using the MX2 beamline, by the ultrafine ϕ -slicing data-collection method using an EIGER X 16M detector and the qeGUI graphical user interface (oscillation range 180°, $\Delta\phi = 0.1^\circ$). Diffraction data were auto processed on the MX2 beamline using automated indexing with *xdsme* (using *XDS* and *Pointless*) and *AIMLESS* [24, 25]. Data processing and *hkl* file conversions were implemented in the *XDS* and the CCP4 program package [26, 27]. All modelling and crystallographic refinements were performed using Phenix, COOT and REFMAC, and the CCP4 program package [28-30]. Figures were generated using Discovery Studio (Dassault Systèmes BIOVIA, USA). X-ray data was processed in the *P1* space group and the structure was determined by molecular replacement using Protein Data Bank (PDB) code 4X80. Several rounds of fitting of the atomic model to electron density and crystallographic refinement were conducted. Data collection and refinement statistics are reported in Table S1.

Molecular docking. Molecular docking was performed with Vina-Carb 1.0 [31]. Ligand structures were obtained using the Carbohydrate Builder available on Glycam-Web [<http://glycam.org>]. Structures were prepared for docking using AutoDockTools 1.5.6 [32], with residues around putative binding sites made flexible for the docking runs (H chain residues 52-58, 103-111 and L chain residues 26-32, 92-95).

Docking was performed within a search space that encompassed all of the CDR loops, centred around Arg-107H. The Vina-Carb exhaustiveness parameter was set to 16 and all other parameters were left at default values. Docking was performed on 3 different structures; Fv1, Fv2 and an equilibrated Fv structure obtained following a 100ns MD run. The top 9 docked poses were extracted for each structure.

Glycan Grafting. The Gly-Spec tool was used to assess the agreement between glycan microarray data and docked poses [33, 34]. Gly-Spec grafts complete glycan structures onto a minimal binding determinant/receptor complex and can predict whether the grafted structure will bind experimentally based on steric clashes between the grafted ligand and the protein receptor. Gly-Spec can then compare to a set of known binders from a CFG glycan microarray and report the overall agreement between the grafted structures and the array data. To assess the agreement of docked poses with CFG array data, we reduced each docked pose to its core Le^a determinant and used this structure as input to Gly-Spec. Positive binders were determined from CFG glycan array data (<http://www.functionalglycomics.org>, last accessed 01/07/20, ID = primscreen_5844) using a method based on Median Absolute Deviation (MAD) [35] with a threshold z-score of $z > 3.5$ for assigning positive binders (Figure S5). Poses that had high agreement with CFG array data (as determined with Gly-Spec) were used as starting points for MD simulations.

Molecular dynamics simulations. MD simulations were run using GROMACS 2018.2 [36, 37]. Proteins were parameterised using the AMBER99SB-ILDN forcefield [38]. Carbohydrate topologies were generated using the Glycam06j forcefield via GLYCAM-Web [<http://glycam.org>], converted to GROMACS format using ACPYPE [39, 40] and combined with the protein topology to form the complete protein-carbohydrate system. Initial coordinates for the protein-carbohydrate system were based on molecular dynamics simulations using the ch88.2 Fab structure (Fv portion only). The protein-carbohydrate system was placed in a rhombic dodecahedral box with ≥ 10 Å distance between the molecule and the edge of the box. The system was solvated using the TIP3P water model, then ionised and neutralised with Na⁺ and Cl⁻ to a concentration of 0.15 M. Energy minimization was performed using the steepest descent algorithm (5000 steps), follow by equilibration at constant volume and temperature (NVT ensemble) for 100 ps, with each replicate initialised with random velocities sampled from a Maxwell distribution at 300 K. This was followed by equilibration with constant pressure and temperature (NPT ensemble) for 300 ps. Pressure coupling was achieved via the Parrinello-Rahman barometer with a time constant of 2 ps and a reference pressure of 1 bar. Temperature coupling was achieved using the modified Berendsen thermostat, with separate temperature coupling groups for ligand/protein and water/ions respectively. A time constant of 0.1 ps was used for temperature coupling, with a reference temperature of 300 K. Equations of motion were integrated using the leap-frog integrator with a time-step of 2 fs. Hydrogen bonds were constrained during all steps using the LINCS constraint algorithm [41]. Long range

electrostatics were calculated using the Particle Mesh Ewald (PME) method with cubic interpolation. Neighbour searching was performed using the Verlet cut-off scheme with a distance cut-off of 1.0 nm for both Coulomb and van der Waals interactions. Extended simulations were run using the Spartan high-performance computer system (University of Melbourne).

Analysis of MD simulations was performed using a combination of GROMACS 2018.2 and the MDTraj Python package (v1.9.1). Ligand root-mean-square deviation (RMSD) was calculated relative to the protein backbone carbon atoms (i.e. protein backbone carbon atoms were used to remove rotational and translational movement across frames, and ligand RMSD calculated for these transformed trajectories). Hydrogen bonds were identified using MDTraj, with the Baker-Hubbard criteria used to identify hydrogen bonds ($\theta > 120^\circ$ and distance from H to acceptor atom $< 2.5 \text{ \AA}$) [42]. Hydrogen bonding matrices were generated using frames in which ligand RMSD $< 0.5 \text{ nm}$. Side-chain dihedral angles were calculated with MDTraj using frames in which ligand RMSD $< 0.5 \text{ nm}$. Plots were constructed using the Matplotlib 1.5 and Seaborn 0.9.0 Python libraries. Visualisation of docked poses was performed using USCF ChimeraX (v0.9). LigPlot+ version 2.2 was used to generate the ligand interaction diagram [43].

Results

Activity and membrane effects of mAb 88.2 against colorectal cancer cells. Previous work by Chua et al. demonstrated the ability of FG88.2 to target colorectal cancer cells [20]. In order to confirm activity of the chimeric mAb ch88.2, binding and killing experiments were conducted using colorectal cancer cell line COLO205. Both ch88.2 and FG88.2 showed strong binding to COLO205 cells by immunofluorescence staining (Figure 1A) and flow cytometry, with mean binding determined to be 99.9% for ch88.2 and 89.9% for FG88.2 and minimal binding in both the negative (0.02% and 0.05%) and isotype (0.09% and 2.3%) controls (Figure 1C). Cell viability was determined after 24 hours of antibody treatment, with mAb treatment shown to cause cell death as measured by Live/Dead staining. The average non-viability increased from 0% in the negative control to 21% after ch88.2 treatment and 44% after FG88.2 treatment, compared to 100% in the positive control (Figure 1B). Cell viability was also measured by flow cytometry through propidium iodide uptake, with non-viability significantly increasing from $4.1 \pm 1.0\%$ in the negative control (and $4.9 \pm 0.4\%$ in the isotype control) to $16.6 \pm 0.9\%$ after ch88.2 treatment and $34.6 \pm 2.9\%$ after FG88.2 treatment ($p < 0.001$), with near-complete death in the positive control ($97.1 \pm 0.5\%$) (Figure 1D).

To examine the antibody-induced membrane damage leading to cell death, scanning electron microscopy (SEM) was used. Prior to treatment, SEM images showed typical cell morphology of COLO205 with rounded regular shaped cells with no clear membrane damage (Figure 2A). Following 24 hours of ch88.2 antibody treatment, a percentage of cells appeared damaged with membrane irregularities including pores and large blebs. Some cells also retained baseline morphology, in agreement with Live/Dead staining and propidium iodide uptake data (Figure 2B). In comparison, after 24 hours of FG88.2 antibody treatment, cellular damage is more pronounced. Large pores are visible in some cells, while other cells are smaller in appearance with irregular morphology and some liquid matter visible on the surface, indicating a loss of cellular membrane integrity (Figure 2C).

Production, crystallisation, and structure determination of ch88.2. Fab was produced by papain digestion of ch88.2 IgG followed by protein A purification, and sample quality was assessed to determine suitability for structural studies. By Coomassie-stained SDS-PAGE gel the Fab was shown to be pure with a disulfide-linked dimer of 50 kDa consisting of paired heavy (H) and light (L) chains of approximately 25 kDa sizes as expected (Figure S2A). Size-distribution analysis of dynamic light scattering (DLS) data showed that the Fab sample was uniform, consisting mainly of a single population at 6.9 nm (72%) with an average hydrodynamic diameter (D_H) of 9.3 nm, compared to IgG which was mostly aggregated (91%) with an average D_H of 105.7 nm (Figure S2B).

To characterise the structural basis for recognition of Lewis glycans by ch88.2, crystallisation of the Fab:glycan complex was attempted. However, electron density was not observed for the Le^x (described here) or the Le^a trisaccharides, and crystallisation with the Le^a-Le^x hexasaccharide was not attempted. Instead, the crystal structure of ch88.2 Fab in the free form was determined to 2.3 Å ($R_{\text{work}}/R_{\text{free}} = 0.213/0.293$) from a triclinic *P1* crystal, with two Fabs in the asymmetric unit (designated here as Fab1 and Fab2). X-ray diffraction data collection and crystallographic refinement statistics are presented in Table S1. The electron density maps allowed fitting of the L and H polypeptide chains for each Fab, except for CH1 residues 141-146 (sequential numbering), which are distant from the antigen-binding site and can be disordered in Fab crystal structures [44, 45]. There is clear electron density corresponding to most binding site residues, although the maps are not as well defined around complementarity determining region 3 (CDR3) of the H chain of both Fabs. While the main-chain was generally visible, several side-chains from Ser 101H to Tyr 112H had weak or missing electron densities (Figure S2C & D).

Structure of the unbound ch88.2 Fab. The overall 3D structures of both Fabs from the asymmetric unit were very similar, with quaternary structures typical of most antibodies. The binding sites are formed by six complementarity-determining regions (CDRs), three from the L chain (identified as L1, L2, and L3) and three from the H chain (identified as H1, H2, and H3). The Fv portions, which are comprised of VL and VH (sequences shown in Figure S3), were shown to be similar with most of the minor variations occurring in the CDR loops surrounding the binding site (Figure 3C). End-on surface views of each Fab illustrate that the main difference between Fab1 (Figure 3A) and Fab2 (Figure 3B) occurs in CDR H3. This is further evident in the CDR view of each Fab. In the overlay, there is little to no variation in L1, L2, L3 and H1 CDR loops, but some variation in H2 and larger differences in H3, which can be seen in the main-chain as well as some individual side-chains (Figure 3D-F). While these differences may be due to crystal packing, with Fab1 packed against Fab2, they also demonstrate the overall plasticity of the ch88.2 Fab CDR regions.

Molecular docking of Le^a-Le^x reveals a range of potential binding poses. To elucidate the structural basis for recognition of Le^a and Le^x by the ch88.2 mAb, molecular dynamics simulations and molecular docking were used, combined with a re-analysis of previously published glycan microarray data for FG88.2 [20]. FG88.2 and ch88.2 contain the same variable chain sequences and were shown to bind to the same cell lines and tissues [20]. Thus, recognition of top glycan binders, identified with FG88.2 mAb, were explored using the structure of the chimeric variant. Microarray data revealed that FG88.2 is capable of binding a large number of Lewis-containing sugars. The top binders were a mixture of Le^a- and/or Le^x-containing glycans, with Le^a being found in most of the strongest binders, while the top binder was a Le^a-Le^x hexasaccharide (Figure 4A-C). Initial attempts to dock small Le^a or Le^x trisaccharides

proved difficult, with docked poses tending to bury the reducing end of the carbohydrate ligand. This is problematic, as the reducing end is typically attached to either another monosaccharide unit or a glycolipid/glycoprotein on the cell surface.

Therefore, to determine the likely binding mode for Le^a, we performed molecular docking on the highest binding Le^a-Le^x hexasaccharide. We used this extended structure for docking to enforce some of the structural/steric constraints that would not be captured by docking Le^a alone. Vina-Carb was used to dock the Le^a-Le^x ligand to 3 different structures. Namely, the two Fv regions from the crystallographically determined Fabs, and a third Fv structure obtained by molecular dynamics simulations to remove potential crystal packing artefacts. As the crystal structures were unliganded, docking was performed with flexible side-chains for residues around the putative binding region. The best 9 docked poses were extracted for each structure, giving a total of 27 potential docked poses. We observed very little agreement between docked poses, with a wide range of potential binding modes identified (Figure S4).

Integration of glycan microarray data highlights a likely binding configuration. To narrow down the set of likely poses, we used the Gly-Spec tool to compute the overall agreement between each pose and experimental FG88.2 glycan microarray data [33, 34]. Gly-Spec grafts whole glycan structures onto a minimal binding determinant in a glycan-protein complex, and predicts binding based on steric clashes between grafted glycans and the protein structure. This predicted binding can then be compared to experimental glycan microarray data. As Le^a is the proposed minimal binding determinant, we grafted glycans onto the Le^a portion of docked poses.

After grafting onto the Le^a portion of all 27 docked poses, only a single pose was deemed to have an acceptable level of agreement with CFG glycan array data, with an overall agreement of 87% to CFG array data (Figure 4D & E). For this pose, one glycan (glycan ID=385) was predicted to be a non-binder by Gly-Spec, despite being within the positive binding set by glycan microarray. This glycan also contains a separate Le^x moiety, suggesting that the Le^a moiety may not be the minimal binding determinant for this one glycan structure. Indeed, for this structure, the Le^a moiety has an alpha-2 linked fucose on the Gal residue, which is likely to impact binding via the Le^a portion of the glycan (Figure 4A, Table 1). The other conflicting result (glycan ID=533) was predicted to be a binder by Gly-Spec, but not by glycan microarray. However, this glycan is just below the threshold for positive binders by microarray data, suggesting that this disagreement is the result of uncertainty in the threshold for calling a positive binder, rather than a true disagreement by Gly-Spec. Overall, this suggests a strong level of agreement between this docked pose and the glycan microarray data.

While we have defined recognition based on key Lewis determinants Le^a, Le^x and Le^c, recognition also extends to sulfated glycans (glycan IDs 24, 28, 290, 291 and 492) and other branched glycans with similar core structures (glycan IDs 24, 291 and 430) (Figure 4 and Table 1). In general, the sulfated glycans appear to have lower binding strength when compared to the top binding glycans, indicating that binding is permissive of the sulfate group. For example, the sulfate group on glycan ID=28 and glycan ID=492 is positioned away from the binding site so as not to interfere with the suggested binding mode. The ch88.2 mAb also appears capable of binding to structures like glycan ID=24 and glycan ID=291, which are closely related to Le^x containing Galβ1-4(Fucα1-3)Glc instead of Galβ1-4(Fucα1-3)GlcNAc. Both of these glycans are also sulfated in one or two positions. In addition, ch88.2 is predicted to bind to glycan ID=430. This glycan contains Fucα1-3GlcNAc and Galβ1-4GlcNAc, which are both found in Le^x but does not contain a complete Lewis glycan structure. This demonstrates broad recognition of ch88.2 for Lewis and Lewis-like glycans.

Stability of the proposed ch88.2 Le^a-Le^x complex. Given that ligand docking was performed on a fixed structure, with only some sidechain flexibility surrounding the CDR regions, we performed a small number of initial MD runs to identify the most stable configuration for the ch88.2 Fv:Le^a-Le^x complex. Using the single docked pose identified by Gly-Spec, we performed 3 x 100 ns MD runs. From these initial simulations, we identified the most stable configuration for the antibody-glycan complex, where Le^a-Le^x remained in position throughout the run. This structure was then run in 10 x 100 ns MD runs to assess overall stability.

The Le^a-Le^x hexasaccharide was stable in the binding site, remaining bound for 9 out of 10 runs (Figure 5A). Some fluctuations in overall ligand RMSD were observed across runs, and this was typically due to movement in the Le^a portion of the ligand (Figure 5B), with the Le^x end remaining relatively fixed (Figure 5C). Interestingly, the reverse behaviour was observed when running MD simulations of just the Le^a and Le^x portions of the ligand; Le^a was stable in 6 out of 10 runs, with minimal RMSD fluctuation (Figure 5D), whereas Le^x displayed greater RMSD variation and only remained bound in 3 out of 10 runs (Figure 5E). Overall, the apparent stability of the ligands in MD simulations is in agreement with CFG glycan microarray data, in which Le^a-Le^x is the strongest binder, followed by Le^a-containing glycans then Le^x-containing glycans (Figure 4A).

Potential interactions between ch88.2 and Le^a-Le^x. For the proposed binding pose, the Le^a-Le^x hexasaccharide occupies a shallow binding groove involving both the L and the H chains, with Le^x sitting at the interface of H and L chains, while Le^a interacts solely with L chain residues (Figure 6D-F). For the Le^a-Le^x ligand, binding is mediated by several hydrogen bonding interactions, as well as likely hydrophobic and/or CH-π interactions. In particular, CDRs 1 and 3 of the L chain are positioned around

the proposed binding site, while the H chain CDRs are too far away to form any direct hydrogen bonds with the glycan. Key residues include Tyr-62H, whose backbone is involved in two highly populated hydrogen bonds to NAG1 (Le^x), one to the nitrogen on the N-acetyl group and the other to O1. Ser-94L forms one hydrogen bond to O2 on GAL3 (Le^x), while Glu-64H forms one bond to either O2, O3 or O4 on FUC2 (Le^x). Binding also involves the N-terminal residue (Asp-1L), which forms one hydrogen bond to either O5 or O6 on NAG4 (Le^a) as well as one bond to O4 of FUC2 (Le^x) (Figure 6D, Figure 7A, Figure S6). Notably, the majority of hydrogen bonds are to the Le^x portion of the structure. However, the Le^a portion appears to be supported by CH- π interactions between the *alpha* face of GAL6 and Trp-92L. In particular, C3 of GAL6 is within 4.5 Å of the centre of the tryptophan pyrrole ring in 42.5% of simulation frames. There are also possible hydrophobic interactions occurring between GAL6 and both Trp-92L and Phe-32L. There is high degree of mobility in Trp-92L and Phe-32L sidechain dihedral angles when Le^a is not present in the binding site. With Le^a bound, the aromatic side-chains of Trp-92L and Phe-32L form a semi-rigid hydrophobic cage around GAL6 (Figure 7 & Figure 8).

Additional MD simulations were performed with just the Le^a and Le^x trisaccharides in their respective binding pockets. The interactions for each trisaccharide were largely similar to those observed for the Le^a-Le^x hexasaccharide. When examining the hydrogen bonding matrix, the frequency of hydrogen bonding interactions between each pair of residues was generally similar (Figure 6D-F). Interestingly, while there are fewer H-bonds for Le^a when compared to Le^x, the H-bonds that are seen for Le^a alone are more populated in the trisaccharide structure than in the Le^a-Le^x hexasaccharide structure. These bonds between Glu-27L and O4 of FUC5 and between Asp-1L and O1 as well as O6 on NAG4 are again further supported by likely CH- π or hydrophobic interactions involving Trp-92L and Phe-32L (Figure 7B & Figure 8). For the Le^x trisaccharide, the H-bonds are nearly identical to those observed for the Le^x portion of the hexasaccharide. The most highly populated bonds are between Tyr-62H to the nitrogen on the N-acetyl group and to O1 on NAG1, Ser-94L to O2 on GAL3, Glu-64H to either O2 or O4 on FUC2, and Asp-1L to O4 of FUC2 or O4 of GAL3 (Figure 7C & Figure 8).

Discussion

Recently, there has been an increasing trend towards targeted therapeutics for the treatment of cancer to increase specificity and reduce off-target effects. Lewis system carbohydrates (Le^a , Le^b , Le^x , Le^y) are often aberrantly expressed on tumours derived from tissues that are normally negative or have minimal expression on restricted tissues [12]. This includes carcinomas of the colon, lung, breast, prostate and ovary that can therefore be selectively targeted by anti-Lewis mAbs. A number of Lewis-binding mAbs have been structurally characterised, including five co-crystal structures and three unbound structures of anti-Lewis antibodies [14]. Although each of these antibodies have been well characterised and some have progressed through clinical trials [12], none have been approved for therapeutic use. In addition, these structures describe anti- Le^x and anti- Le^y antibodies, but no crystal structures of anti- Le^a antibodies have been determined. In this study we examine the nature of a chimeric-human IgG1 mAb that can target both Le^a - and Le^x -containing glycans.

Previous work by Chua et al. demonstrated the ability of the murine antibody FG88.2 to target colorectal cancer cells [20]. Here we also conducted further additional studies to assess the killing ability of the chimeric variant ch88.2 IgG sample, which was used in our structural studies. FG88.2 was previously shown to have direct tumour cytotoxicity via a mechanism distinct from complement-mediated or cellular-mediated effects [20]. Murine IgG3 mAbs have previously been shown to be particularly effective at agglutination [46]. Murine IgG3 are also able to self-associate via their constant regions, forming large oligomeric networks [47]. As such, we sought to determine if the chimeric ch88.2, which is built on a human IgG1 backbone, could also display direct killing activity against Lewis-expressing colorectal cancer cells. Treatment with ch88.2 IgG1 caused a reduction in cell viability, although not to the levels observed by the murine FG88.2 IgG3 (Figure 1). This was further reflected during SEM imaging, where a percentage of cells appeared damaged with membrane irregularities including pores and large blebs (Figure 2). Interestingly, by DLS, intact ch88.2 (also used for structural studies) was shown to be highly aggregated in solution (around 90%) (Figure S2). This level of aggregation may have led to increased clustering of Lewis epitopes, resulting in modest direct cell killing effects.

Several other mAbs have been developed against type II Lewis glycans, including BR96 (anti- Le^y), hu3S193 (anti- Le^y), 291-2G3-A (anti- Le^x) and 54-5C10-A (anti- Le^x). Co-crystal structures of both BR96 and hu3S193 mAbs with Le^y show very similar mechanisms for carbohydrate recognition, with almost identical binding sites and carbohydrate orientation for these two structures. In these structures, Le^y is accommodated in a large but relatively deep binding pocket, with antibody-glycan interactions dominated by the heavy chain [15-17, 48]. The mAbs 291-2G3-A and 54-5C10-A were both developed

against *Schistosoma* parasites, although only 291-2G3-A has been resolved in complex with Le^x. This structure involves a shallower binding pocket that is centrally located at the VL-VH interface [18, 19]. Notably, no crystal structures of mAbs in complex with type I Lewis glycans (Le^a or Le^b) have been determined, so comparisons can only be made to type II glycan structures. The binding mechanism proposed here for the larger Le^a-Le^x epitope involves an extended binding groove, composed of two adjoining shallow binding pockets. The binding pocket for Le^a is formed entirely by the light chain, whereas the pocket for Le^x involves significant contributions from both heavy and light chain residues. This extended binding groove with distinct sites for Le^a and Le^x may explain the high affinity for Le^a-Le^x as opposed to di-Le^a or other Le^a-containing glycans as observed by glycan microarray. Additionally, this may also explain the relatively high affinity for Le^c-Le^x, despite Le^c and Le^x being weaker binding determinants than Le^a. The fucose residue of Le^a (FUC5) is relatively unengaged in direct binding interactions, only making contact with Glu-27L in a small number of simulation frames. This suggests that FUC5 is not required for binding. As this residue is not present in the Le^c-Le^x pentasaccharide, it would likely bind in a similar way, in agreement with glycan array data.

An interesting feature that appears common to many anti-carbohydrate antibodies is the involvement of multiple aromatic residues in carbohydrate binding, in particular tyrosine (Tyr), tryptophan (Trp) and phenylalanine (Phe). Aromatic residues can participate in binding via direct hydrogen bonding with the glycan, stacking interactions between aromatic rings and the hydrophobic face of carbohydrate rings (involving CH- π interactions) or other hydrophobic interactions. These types of interactions have been noted in all previously determined co-crystal structures of anti-Lewis antibodies [15, 16, 18]. The proposed binding mode for ch88.2 is no exception, with interactions with aromatic residues playing a key role in binding to the Le^a portion of the structure. The predicted high stability of Le^a as compared to Le^x, may be explained by potential CH- π interactions with aromatic residues in the L chain via GAL6 despite Le^x being involved in a greater number of hydrogen bonding interactions. In general, CH- π interactions differ significantly between monosaccharides, with beta-D-Gal having a relatively high propensity for forming CH- π interactions via its alpha face [49]. For ch88.2, these interactions likely occur between the alpha face of GAL6 and Trp-92L. Additionally, other hydrophobic interactions may play a key role in ligand stability. It is thought that desolvation of aromatic residues and carbohydrate CH groups contributes significantly to overall antibody-glycan interactions [50]. In the binding mode predicted here, Trp-92L and Phe-32L are angled such that GAL6 is inserted end-on into a hydrophobic pocket formed by these two residues.

In summary, it is the balance of hydrogen bonding and hydrophobic interactions that engage the extended glycan structure. The combination of molecular docking, glycan grafting and molecular dynamics have provided crucial insights into the binding of Le^a- and Le^x-containing glycans to ch88.2,

highlighting the cooperative interplay between Le^a and Le^x in the extended Le^a-Le^x structure. The Le^a portion is largely surrounded by antibody residues and appears to dominate the interaction with ch88.2. However, the Le^x portion is partially exposed to solvent indicating that a larger glycan may be accommodated within the binding site. On the surface of a cancer cell the glycan targets are likely larger than the hexasaccharide reported here and presented as glycolipids or glycoproteins. Thus, the mode of recognition of Lewis glycans identified here is compatible with binding of ch88.2 to larger antigenic targets on tumour cells. Detailed characterisation of the interactions identified here could assist in structure-guided manipulation of the antibody to enhance its selectivity for tumour-related Lewis antigens and drive the development of further therapeutic antibody candidates.

Acknowledgements

This research was undertaken in part using the MX2 beamline at the Australian Synchrotron, part of ANSTO, and made use of the Australian Cancer Research Foundation (ACRF) detector. This research was also undertaken using the LIEF HPC-GPGPU Facility hosted at the University of Melbourne. This Facility was established with the assistance of LIEF Grant LE170100200. This research was performed in part at the RMIT Micro Nano Research Facility (MNRF) in the Victorian Node of the Australian National Fabrication Facility (ANFF). We acknowledge the facilities, and the scientific and technical assistance of the RMIT Microscopy & Microanalysis Facility (RMMF), a linked laboratory of Microscopy Australia. CS is supported by a research training program stipend scholarship from the Australian Government, Department of Education and Training. PAR is supported by a Vice Chancellor's senior research fellowship from RMIT University.

Conflict of Interest Statement: Caroline Soliman., Andrew J. Guy, Richard S. McIntosh, Ian Spendlove, Sarah Eastwood, Vi Khanh Truong, Aaron Elbourne and Paul A. Ramsland declare no conflict of interest. Lindy G. Durrant is director and CSO of Scancell Ltd. and has ownership interest (including patents) in Scancell Ltd. Mireille Vankemmelbeke and Jia Xin Chua are employees of Scancell Ltd.

Author Contributions: C.S., A.J.G. and P.A.R. designed research; C.S., A.J.G., S.E., V.K.T., A.E. and P.A.R. performed research; C.S., A.J.G., and P.A.R. analysed the data; J.X.C., M.V., R.S.M., I.S., L.G.D. contributed new reagents; C.S., A.J.G. and P.A.R. wrote the paper; and J.X.C., M.V., R.S.M., S.E., V.K.T., A.E., I.S. and L.G.D. reviewed and commented on the paper. All authors read and approved the final manuscript.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org [PDB ID code 6X5E].

References

1. World Health Organization. (2018) Latest global cancer data: Cancer burden rises to 18.1 million new cases and 9.6 million cancer deaths in 2018. International Agency for Research on Cancer.
2. Kaplon, H. and Reichert, J. M. (2019) Antibodies to watch in 2019. *mAbs*. **11**, 219-238.
3. Peixoto, A., Relvas-Santos, M., Azevedo, R., Santos, L. L. and Ferreira, J. A. (2019) Protein Glycosylation and Tumor Microenvironment Alterations Driving Cancer Hallmarks. *Front. Oncol.* **9**, 380-403.
4. Munkley, J. and Elliott, D. J. (2016) Hallmarks of glycosylation in cancer. *Oncotarget*. **7**, 35478-35489.
5. Durrant, L. G., Noble, P. and Spendlove, I. (2012) Immunology in the clinic review series; focus on cancer: glycolipids as targets for tumour immunotherapy. *Clin. Exp. Immunol.* **167**, 206-215.
6. Oriol, R., Pendu, J. and Mollicone, R. (1986) Genetics of ABO, H, Lewis, X and Related Antigens. *Vox Sang.* **51**, 161-171.
7. Hanfland, P., Graham, H. A., Crawford, R. J. and Schenkel-Brunner, H. (1982) Immunochemistry of the Lewis blood-group system: Investigations on the Le c antigen. *FEBS Lett.* **142**, 77-80.
8. Hanfland, P., Kordowicz, M., Peter-Katalinic, J., Pfannschmidt, G., Crawford, R. J., Graham, H. A. and Egge, H. (1986) Immunochemistry of the Lewis blood-group system: Isolation and structures of Lewis-c active and related glycosphingolipids from the plasma of blood-group O Le(a-b) nonsecretors. *Arch. Biochem. Biophys.* **246**, 655-672.
9. Le Pendu, J., Lemieux, R. U. and Oriol, R. (1982) Purification of Anti-Lec Antibodies with Specificity for β DGal(1 \rightarrow 3) β DGlcNAcO- Using a Synthetic Immunoabsorbent. *Vox Sang.* **43**, 188-195.
10. Harmening, D. (1989) *Modern Blood Banking and Transfusion Practices*, Philadelphia.
11. Stanley, P. and Cummings, R. D. (2017) Structures Common to Different Glycans. In *Essentials of Glycobiology* (Varki, A., Cummings, R. D., Esko, J. D. and al., e., eds.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
12. Dingjan, T., Spendlove, I., Durrant, L. G., Scott, A. M., Yuriev, E. and Ramsland, P. A. (2015) Structural biology of antibody recognition of carbohydrate epitopes and potential uses for targeted cancer immunotherapies. *Mol. Immunol.* **67**, 75-88.
13. Agostino, M., Farrugia, W., Sandrin, M., Scott, A., Yuriev, E. and Ramsland, P. (2012) Structural glycobiology of antibody recognition in xenotransplantation and cancer immunotherapy. In *Anticarbohydrate Antibodies* (Paul Kosma, S. M.-L., ed.), Springer (Vienna, Austria).
14. Soliman, C., Yuriev, E. and Ramsland, P. (2017) Antibody recognition of aberrant glycosylation on the surface of cancer cells. *Curr. Opin. Struct. Biol.* **44**, 1-8.
15. Jeffrey, P. D., Bajorath, J., Chang, C. Y., Yelton, D., Hellstrom, I., Hellstrom, K. E. and Sheriff, S. (1995) The x-ray structure of an anti-tumour antibody in complex with antigen. *Nat. Struct. Mol. Biol.* **2**, 466-471.

16. Ramsland, P. A., Farrugia, W., Bradford, T. M., Mark Hogarth, P. and Scott, A. M. (2004) Structural Convergence of Antibody Binding of Carbohydrate Determinants in Lewis Y Tumor Antigens. *J. Mol. Biol.* **340**, 809-818.
17. Farrugia, W., Scott, A. M. and Ramsland, P. A. (2009) A possible role for metallic ions in the carbohydrate cluster recognition displayed by a Lewis Y specific antibody. *PLoS One.* **4**, e7777.
18. van Roon, A. M., Pannu, N. S., de Vrind, J. P., van der Marel, G. A., van Boom, J. H., Hokke, C. H., Deelder, A. M. and Abrahams, J. P. (2004) Structure of an anti-Lewis X Fab fragment in complex with its Lewis X antigen. *Structure.* **12**, 1227-1236.
19. de Geus, D. C., van Roon, A. M., Thomassen, E. A., Hokke, C. H., Deelder, A. M. and Abrahams, J. P. (2009) Characterization of a diagnostic Fab fragment binding trimeric Lewis X. *Proteins.* **76**, 439-447.
20. Chua, J. X., Vankemmelbeke, M., McIntosh, R. S., Clarke, P. A., Moss, R., Parsons, T., Spendlove, I., Zaitoun, A. M., Madhusudan, S. and Durrant, L. G. (2015) Monoclonal Antibodies Targeting LecLex-Related Glycans with Potent Antitumor Activity. *Clin. Cancer. Res.* **21(13)**, 2963-2974.
21. Vankemmelbeke, M., McIntosh, R. S., Chua, J. X., Kirk, T., Daniels, I., Patsalidou, M., Moss, R., Parsons, T., Scott, D., Harris, G., Ramage, J. M., Spendlove, I. and Durrant, L. G. (2020) Engineering the human Fc-region enables direct cell killing by cancer glycan-targeting antibodies without the need for immune effector cells or complement. *Cancer Res.*, canres.3599.2019.
22. Cywes-Bentley, C., Skurnik, D., Zaidi, T., Roux, D., Deoliveira, R. B., Garrett, W. S., Lu, X., O'Malley, J., Kinzel, K., Zaidi, T., Rey, A., Perrin, C., Fichorova, R. N., Kayatani, A. K. K., Maira-Litràn, T., Gening, M. L., Tsvetkov, Y. E., Nifantiev, N. E., Bakaletz, L. O., Pelton, S. I., Golenbock, D. T. and Pier, G. B. (2013) Antibody to a conserved antigenic target is protective against diverse prokaryotic and eukaryotic pathogens. *Proc. Natl. Acad. Sci. U. S. A.* **110**, E2209.
23. Hogarth, P. M., Edwards, J., McKenzie, I. F., Goding, J. W. and Liew, F. Y. (1982) Monoclonal antibodies to the murine Ly-2.1 cell surface antigen. *Immunology.* **46**, 135-144.
24. Casanas, A., Warshamanage, R., Finke, A. D., Panepucci, E., Olieric, V., Noll, A., Tampe, R., Brandstetter, S., Forster, A., Mueller, M., Schulze-Briese, C., Bunk, O. and Wang, M. (2016) EIGER detector: application in macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **72(Pt 9)**, 1036–1048.
25. Aragao, D., Aishima, J., Cherukuvada, H., Clarcken, R., Clift, M., Cowieson, N., Ericsson, D. J., Gee, C., Macedo, S., Mudie, N., Panjkar, S., Price, J. R., Riboldi-Tunncliffe, A., Rostan, R., Williamson, R. and Caradoc-Davies, T. T. (2018) MX2: a high-flux undulator microfocus beamline serving both the chemical and macromolecular crystallography communities at the Australian Synchrotron. *J Synchrotron Radiat.* **1**, 885-891.
26. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S.,

- Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235-242.
27. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125-132.
28. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213-221.
29. Emsley, P., Lohkamp, B., Scott, W. G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486-501.
30. Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240-255.
31. Nivedha, A. K., Thieker, D. F., Makeneni, S., Hu, H. and Woods, R. J. (2016) Vina-Carb: Improving Glycosidic Angles during Carbohydrate Docking. *J. Chem. Theory Comput.* **12**, 892-901.
32. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. and Olson, A. J. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785-2791.
33. Grant, O. C., Tessier, M. B., Meche, L., Mahal, L. K., Foley, B. L. and Woods, R. J. (2016) Combining 3D structure with glycan array data provides insight into the origin of glycan specificity. *Glycobiology.* **26**, 772-783.
34. Grant, O. C., Xue, X., Ra, D., Khatamian, A., Foley, B. L. and Woods, R. J. (2016) Gly-Spec: a webtool for predicting glycan specificity by integrating glycan array screening data and 3D structure. *Glycobiology.* **26**, 1027-1028.
35. Coff, L., Chan, J., Ramsland, P. A. and Guy, A. J. (2020) Identifying glycan motifs using a novel subtree mining approach. *BMC Bioinformatics.* **21**, 42-59.
36. Berendsen, H. J. C., van der Spoel, D. and van Drunen, R. (1995) GROMACS: A message-passing parallel molecular dynamics implementation. *Comput Phys Commun.* **91**, 43-56.
37. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E. and Berendsen, H. J. (2005) GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701-1718.
38. Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O. and Shaw, D. E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins.* **78**, 1950-1958.
39. Sousa da Silva, A. W. and Vranken, W. F. (2012) ACPYPE - AnteChamber PYthon Parser interface. *BMC Res. Notes.* **5**, 367.
40. Bernardi, A., Faller, R., Reith, D. and Kirschner, K. N. (2019) ACPYPE update for nonuniform 1–4 scale factors: Conversion of the GLYCAM06 force field from AMBER to GROMACS. *SoftwareX.* **10**, 100241.

41. Hess, B. (2008) P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J. Chem. Theory Comput.* **4**, 116-122.
42. Baker, E. N. and Hubbard, R. E. (1984) Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* **44**, 97-179.
43. Laskowski, R. A. and Swindells, M. B. (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **51**, 2778-2786.
44. Fan, Z. C., Goldstein, B. Z., Guddat, L. W., Thakur, A., Landolfi, N. F., Co, M. S., Vasquez, M., Queen, C., Ramsland, P. A. and Edmundson, A. B. (1999) Comparison of the three-dimensional structures of a humanized and a chimeric Fab of an anti-gamma-interferon antibody. *J. Mol. Recognit.* **12**, 19-32.
45. Sela-Culang, I., Alon, S. and Ofran, Y. (2012) A systematic comparison of free and bound antibodies reveals binding-related conformational changes. *Journal of Immunology (Baltimore, Md. : 1950).* **189**, 4890.
46. Klaus, T. and Bereta, J. (2018) CH2 Domain of Mouse IgG3 Governs Antibody Oligomerization, Increases Functional Affinity to Multivalent Antigens and Enhances Hemagglutination. *Front. Immunol.* **9**, 1096.
47. Haji-Ghassemi, O., Blackler, R. J., Martin Young, N. and Evans, S. V. (2015) Antibody recognition of carbohydrate epitopes. *Glycobiology.* **25**, 920-952.
48. Sheriff, S., Chang, C. Y., Jeffrey, P. D. and Bajorath, J. (1996) X-ray structure of the uncomplexed anti-tumor antibody BR96 and comparison with its antigen-bound form. *J. Mol. Biol.* **259**, 938-946.
49. Hudson, K. L., Bartlett, G. J., Diehl, R. C., Agirre, J., Gallagher, T., Kiessling, L. L. and Woolfson, D. N. (2015) Carbohydrate-Aromatic Interactions in Proteins. *J. Am. Chem. Soc.* **137**, 15152-15160.
50. Asensio, J. L., Arda, A., Canada, F. J. and Jimenez-Barbero, J. (2013) Carbohydrate-aromatic interactions. *Acc. Chem. Res.* **46**, 946-954.

Figures Legends

Figure 1. Activity of mAb 88.2 against colorectal cancer cells. A) Binding of ch88.2 (red) and FG88.2 (green) to COLO205 cells (DNA in blue) by fluorescence microscopy. Negative controls (no primary) are shown for each mAb. **B)** Death of colorectal cancer cells following 24 hours of treatment with ch88.2 or FG88.2. Cells were stained with Calcein-AM live stain (green) and EthD-1 dead stain (red). **C)** Binding of ch88.2 and FG88.2 to COLO205 cells by flow cytometry with respective negative (no-primary) and isotype controls for each mAb. Mean binding is shown as a horizontal bar (n = 2 for ch88.2, n = 4 for FG88.2). **D)** Death of colorectal cancer cells following 24 hours of treatment with ch88.2 or FG88.2. Non-viability was determined by propidium iodide uptake measured by flow cytometry, with mean percentage uptake shown as a horizontal bar (n = 4; *** indicates p < 0.001). Scale bars in panels A & C represent 100 μ m.

Figure 2. SEM images of antibody-induced cellular damage. Images of colorectal cancer cells showing differences in cellular morphology **A)** before treatment (negative control), and after antibody treatment with **B)** ch88.2 and **C)** FG88.2. COLO205 cells were grown on poly-D-lysine coated silicon chips overnight and treated with antibody for 24 hours. Yellow arrows point to examples of cellular and membrane damage. Scale bars are as indicated in each image.

Figure 3. Structure of ch88.2 Fab. A & B) End-on solvent-accessible view of each Fab of ch88.2. **C)** Overlay of the two Fabs of ch88.2, each depicted as ribbons. The Fv region of each antibody was overlaid by sequence alignment. In each panel, the VL domain is shown in blue and the VH domain is shown in orange. CDRs are shown in pale blue (L chain) and pale orange (H chain). **D)** CDR residues from ch88.2 Fab1 (chains L and H) with L chain CDRs depicted as pale grey sticks and H chain CDRs depicted as green sticks. **E)** CDR residues from ch88.2 Fab2 (chain A and B) with L chain CDRs depicted as pale blue sticks and H chain CDRs depicted as orange sticks. **F)** Overlay of CDR residues from ch88.2 Fab1 (shown in pale grey and green) and Fab2 (shown in pale blue and orange).

Figure 4. Identification of likely binding poses for the Le^a-Le^x hexasaccharide bound to ch88.2 using molecular docking and glycan microarray data. A) Glycan microarray data for ch88.2 showing the presence of Le^a, Le^x or Le^c motifs in top binders. Terminal and non-terminal motifs are distinguished for each Lewis motif; any motif with additional carbohydrate linkages or modifications on any sugar within the motif is considered to be non-terminal. There were no non-terminal Le^c motifs present. Each glycan

is identified by Glycan array ID (CFG array version 5.1). Antibody was tested at 50 $\mu\text{g}/\text{mL}$. Two Lewis motifs within the same glycan are indicated by darker shades of blue. RFU is shown on a linear scale. A full list of all binding glycans is provided in Table 1. **B & C)** The top binder in the ch88.2 glycan microarray is a $\text{Le}^{\text{a}}\text{-Le}^{\text{x}}$ hexasaccharide shown as a cartoon (in SNFG nomenclature) and depicted as sticks (coloured by atom type). **D & E)** Agreement between docked poses and glycan microarray data for all 27 docked poses examined. Binder agreement is defined as the percentage of experimental binders which are also predicted to bind following glycan grafting, and conversely for non-binder agreement. Each pose is represented in grey, with the top pose shown in red.

Figure 5. Ligand stability during molecular dynamics simulations. **A)** RMSD values for the entire ligand across 10 independent MD runs of a $\text{Le}^{\text{a}}\text{-Le}^{\text{x}}$ hexasaccharide. **B)** RMSD values for the Le^{a} portion of the ligand across 10 independent MD runs of a $\text{Le}^{\text{a}}\text{-Le}^{\text{x}}$ hexasaccharide. **C)** RMSD values for the Le^{x} portion of the ligand across 10 independent MD runs of a $\text{Le}^{\text{a}}\text{-Le}^{\text{x}}$ hexasaccharide. **D)** RMSD values for the entire ligand across 10 independent MD runs of a Le^{a} trisaccharide. **E)** RMSD values for the entire ligand across 10 independent MD runs of a Le^{x} trisaccharide.

Figure 6. Le^{a} - and Le^{x} -containing ligands are predicted to bind to ch88.2 in an extended binding groove predominantly involving L chain CDR1 and CDR3 residues. **A)** Representative pose for a $\text{Le}^{\text{a}}\text{-Le}^{\text{x}}$ hexasaccharide bound to ch88.2. **B)** Representative pose for a Le^{a} trisaccharide bound to ch88.2. **C)** Representative pose for a Le^{x} trisaccharide bound to ch88.2. Each antibody is depicted a transparent surface (L chain in purple and H chain in pink) with key binding residues depicted as sticks. Ligands are depicted as sticks coloured by atom type (C, green; O, red; N, blue; H, white). **D)** Hydrogen bonding matrix for a $\text{Le}^{\text{a}}\text{-Le}^{\text{x}}$ hexasaccharide bound to ch88.2 across 10 x 100ns independent MD runs. **E)** Hydrogen bonding matrix for a Le^{a} trisaccharide bound to ch88.2 across 10 x 100ns independent MD runs. **F)** Hydrogen bonding matrix for a Le^{x} trisaccharide bound to ch88.2 across 10 x 100ns independent MD runs.

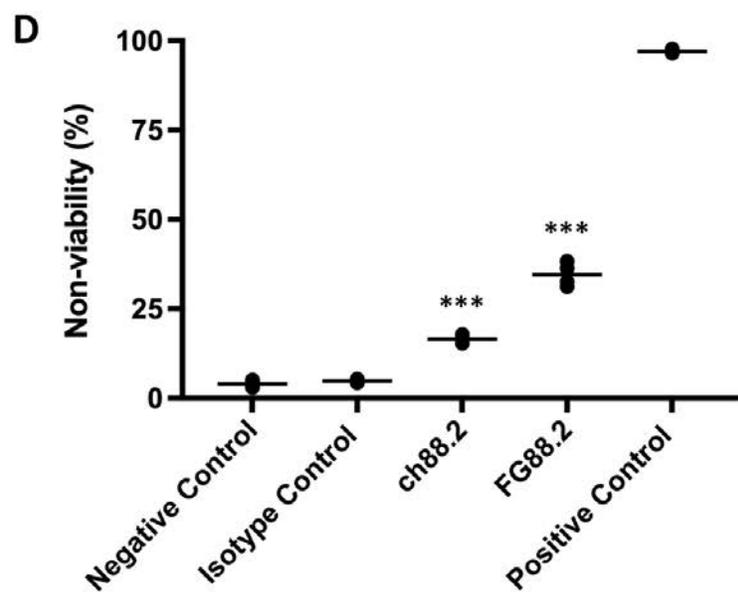
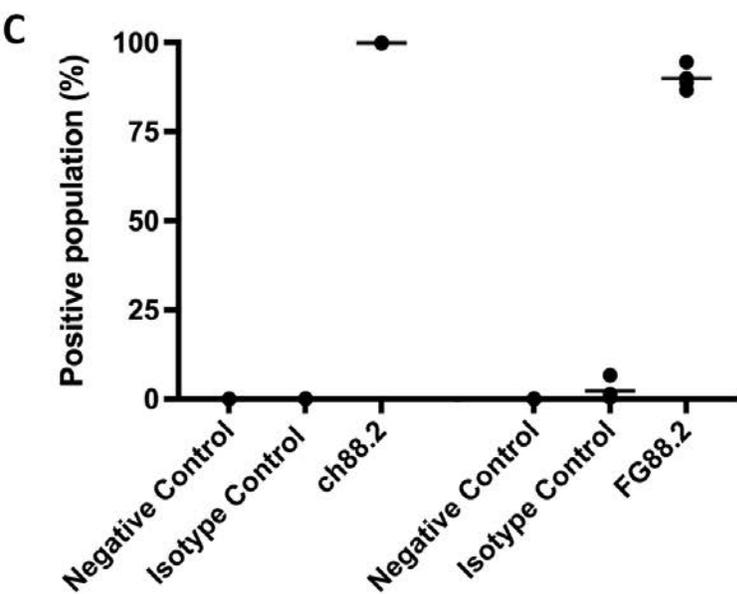
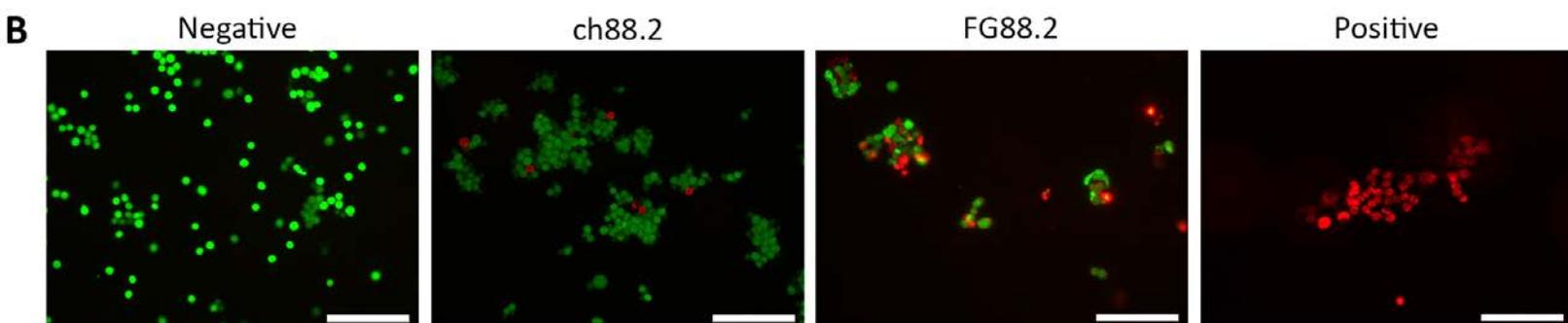
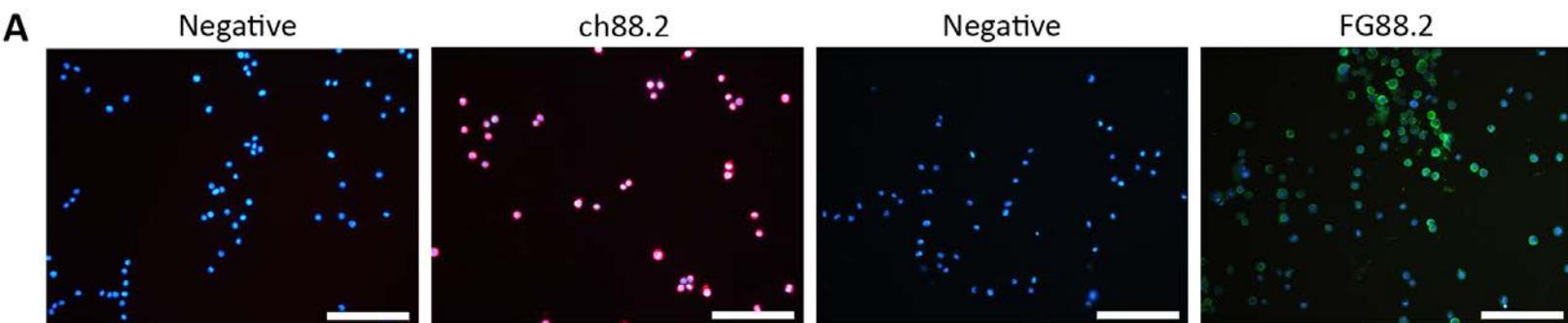
Figure 7. Potential binding of Le^{a} - and Le^{x} - containing ligands to ch88.2 involves a number of hydrogen bonds. Hydrogen bonding interactions are shown between ch88.2 and **A)** a $\text{Le}^{\text{a}}\text{-Le}^{\text{x}}$ hexasaccharide, **B)** a Le^{a} trisaccharide and **C)** a Le^{x} trisaccharide. Ligand and antibody residues are represented as sticks coloured by atom type (O, red; N, blue; H, white), with backbone carbons coloured blue for L chain, orange for H chain and green for ligands. Glycan residues are labelled using the same numbering scheme introduced in Figure 4C. Hydrogen bonds that are present in at least 20% of the MD simulations for each ligand are shown as dotted grey lines. In the case of chemically identical atoms only a single representative hydrogen bond is shown.

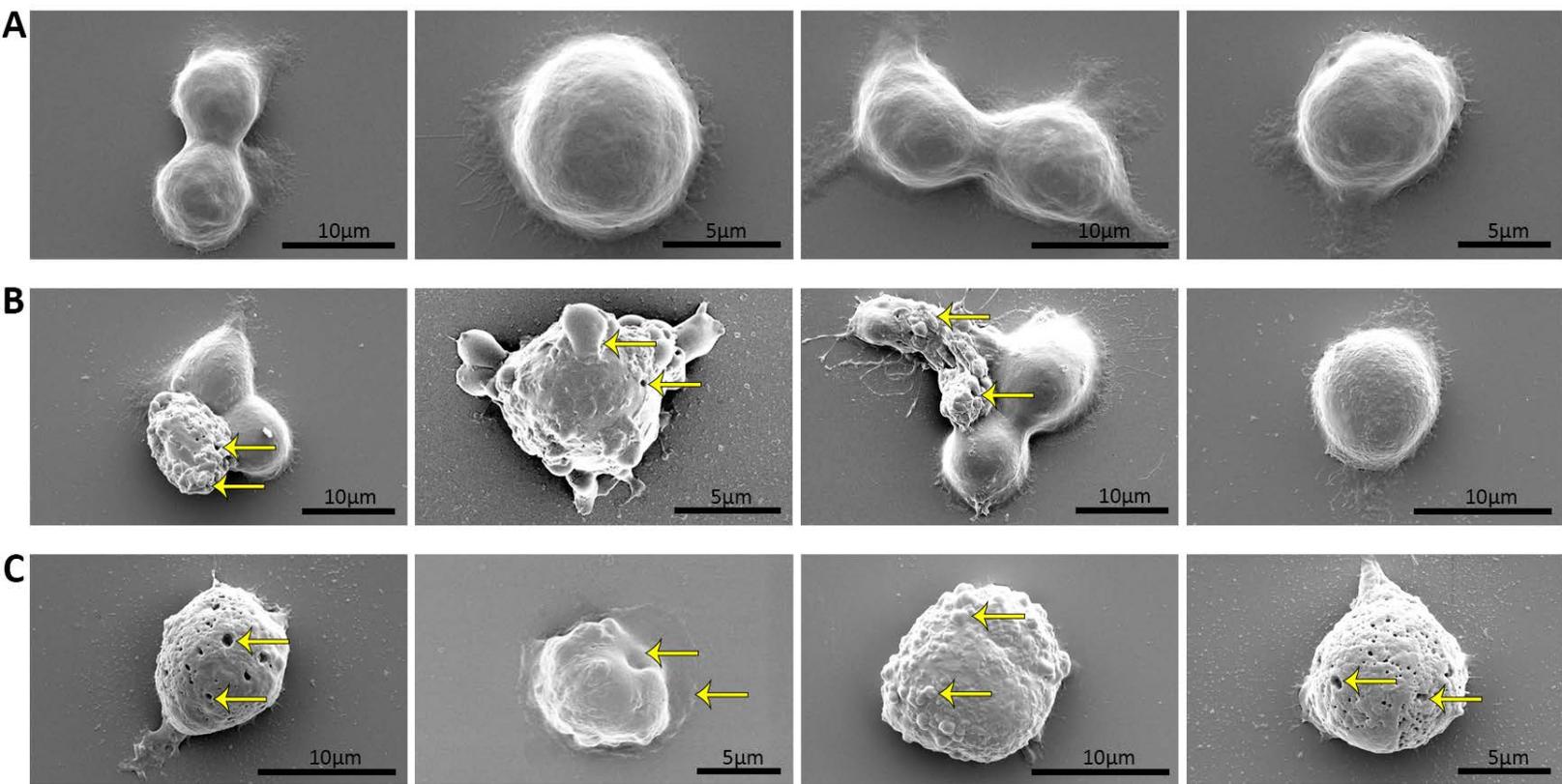
Figure 8. Flexibility in Trp-92L and Phe-32L aromatic residues in the presence of Le^a-Le^x, Le^a and Le^x. Both Trp-92L and Phe-32L residues are in the binding pocket that engages with the Le^a portion of the ligand (Le^a and Le^a-Le^x panels), and these residues are relatively stable compared to simulations without Le^a in which these residues do not make contact with the carbohydrate ligand (Le^x panel). χ_1 and χ_2 dihedral angles (displayed from 0-360° for ease of representation) for both Trp-92L and Phe-32L were determined for all simulation frames in which ligand RMSD < 0.5 nm.

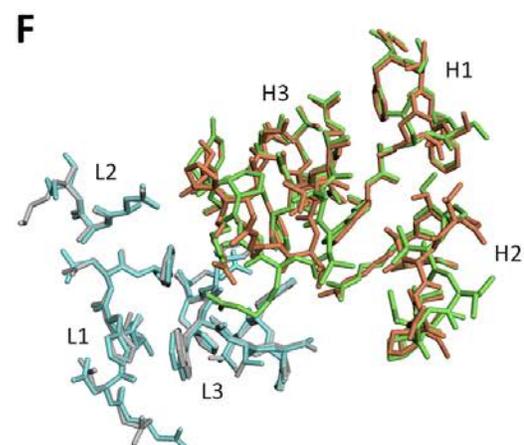
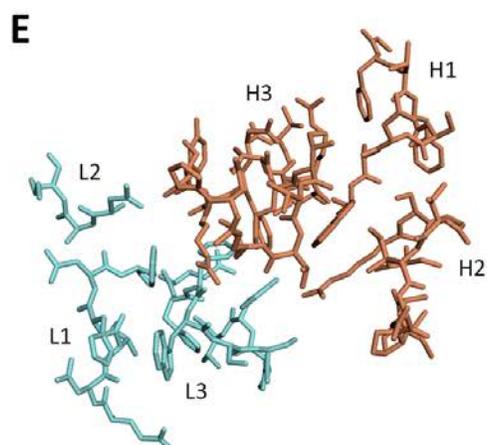
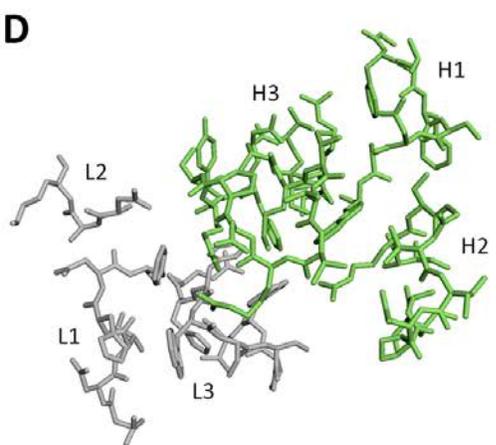
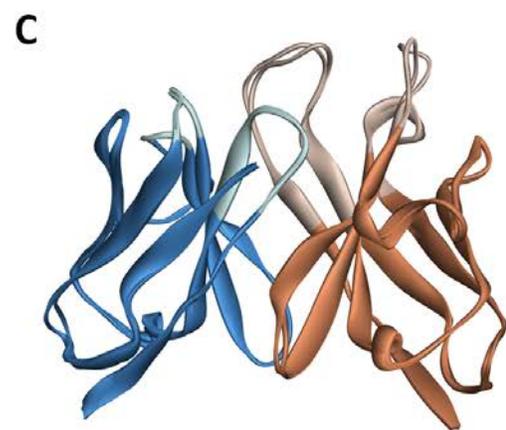
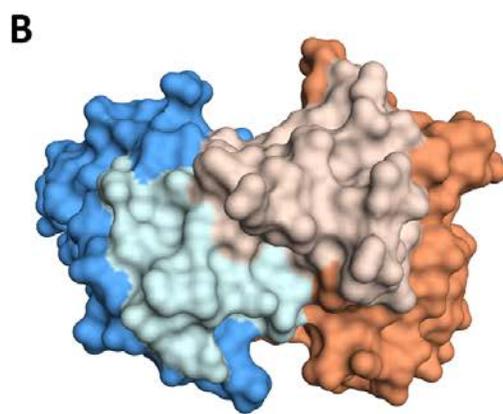
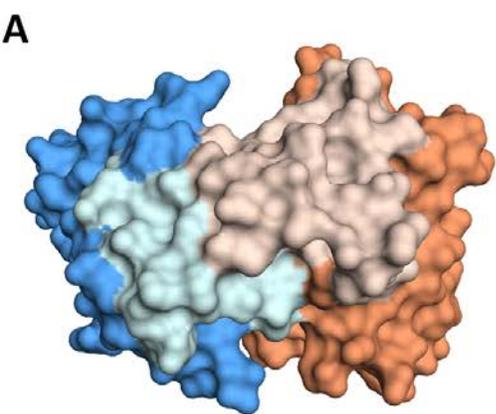
Table 1. Sequences of top binding glycans from FG88.2 microarray.

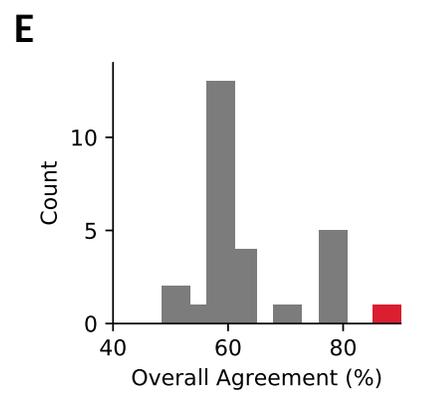
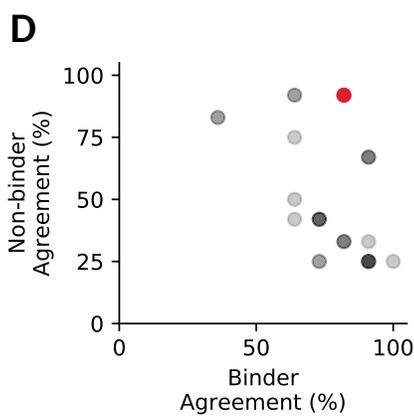
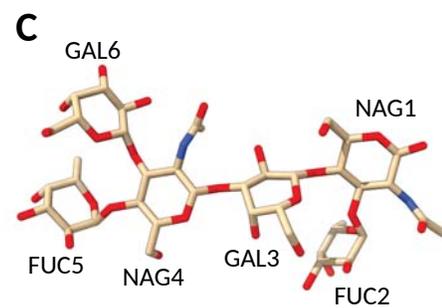
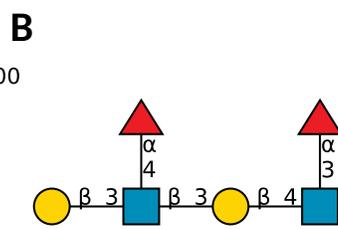
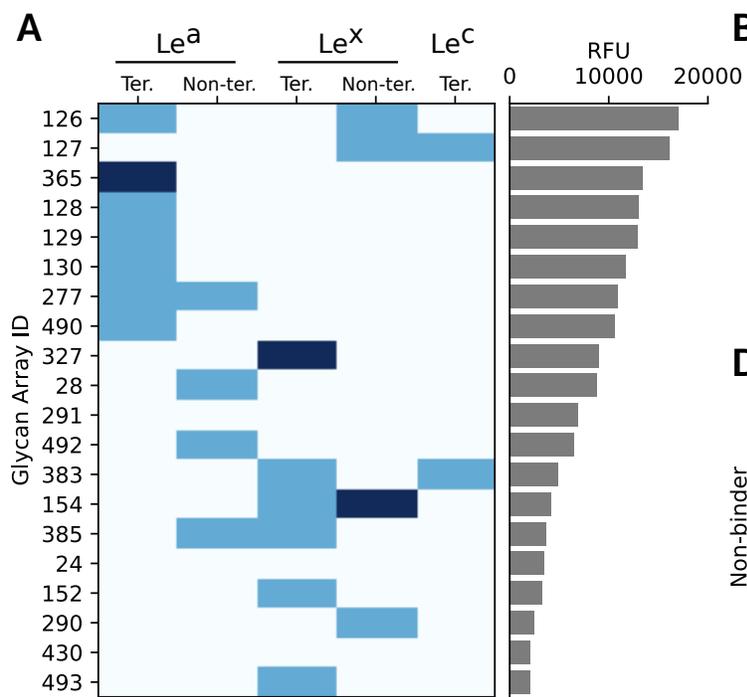
Glycan Array ID	Glycan Sequence
126	Gal β 1-3(Fuca1-4)GlcNAc β 1-3Gal β 1-4(Fuca1-3)GlcNAc β -Sp0
127	Gal β 1-3GlcNAc β 1-3Gal β 1-4(Fuca1-3)GlcNAc β -Sp0
365	Fuca1-4(Gal β 1-3)GlcNAc β 1-2Man α 1-6(Fuca1-4(Gal β 1-3)GlcNAc β 1-2Man α 1-3)Man β 1-4GlcNAc β 1-4(Fuca1-6)GlcNAc β -Sp22
128	Gal β 1-3(Fuca1-4)GlcNAc-Sp0
129	Gal β 1-3(Fuca1-4)GlcNAc-Sp8
130	Fuca1-4(Gal β 1-3)GlcNAc β -Sp8
277	Gal β 1-3(Fuca1-4)GlcNAc β 1-3Galb1-3(Fuca1-4)GlcNAc β -Sp0
490	Gal β 1-3(Fuca1-4)GlcNAc β 1-6GalNAc α -Sp14
327	Gal β 1-4(Fuca1-3)GlcNAc β 1-2Man α 1-6(Galb1-4(Fuca1-3)GlcNAc β 1-2Man α 1-3)Man β 1-4GlcNAc β 1-4GlcNAc β -Sp20
28	(3S)Gal β 1-3(Fuca1-4)GlcNAc β -Sp8
291	Gal β 1-4(Fuca1-3)(6S)Glc β -Sp0
492	(3S)Gal β 1-3(Fuca1-4)GlcNAc β -Sp0
383	Gal β 1-4(Fuca1-3)GlcNAc β 1-6(Gal β 1-3GlcNAc β 1-3)Gal β 1-4Glc-Sp21
154	Gal β 1-4(Fuca1-3)GlcNAc β 1-3Galb1-4(Fuca1-3)GlcNAc β 1-3Gal β 1-4(Fuca1-3)GlcNAc β -Sp0
385	Gal β 1-4(Fuca1-3)GlcNAc β 1-6(Fuca1-4(Fuca1-2Gal β 1-3)GlcNAc β 1-3)Gal β 1-4Glc-Sp21
24	(3S)Gal β 1-4(Fuca1-3)(6S)Glc-Sp0
152	Gal β 1-4(Fuca1-3)GlcNAc β -Sp8
290	Gal β 1-4(Fuca1-3)(6S)GlcNAc β -Sp0
430	Fuca1-3GlcNAc β 1-6(Gal β 1-4GlcNAc β 1-3)Gal β 1-4Glc-Sp21
493	Gal β 1-4(Fuca1-3)GlcNAc β 1-6(Neu5Ac α 2-6(Neu5Ac α 2-3Gal β 1-3)GlcNAc β 1-3)Gal β 1-4Glc-Sp21

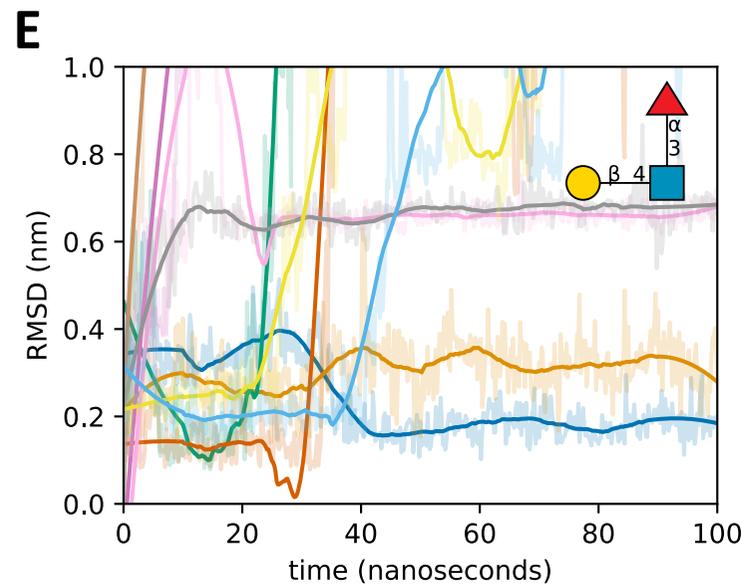
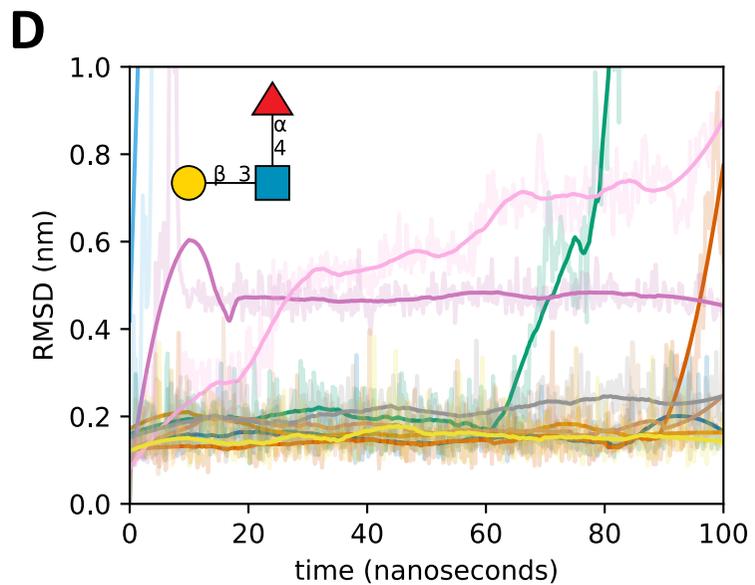
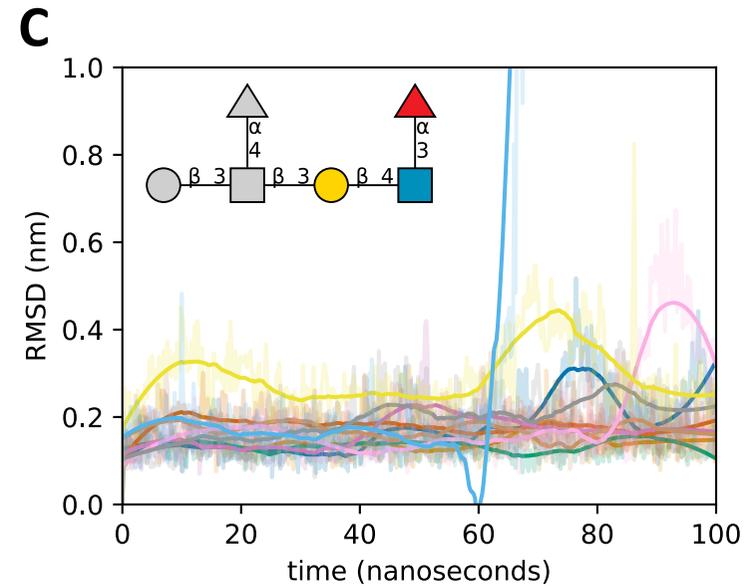
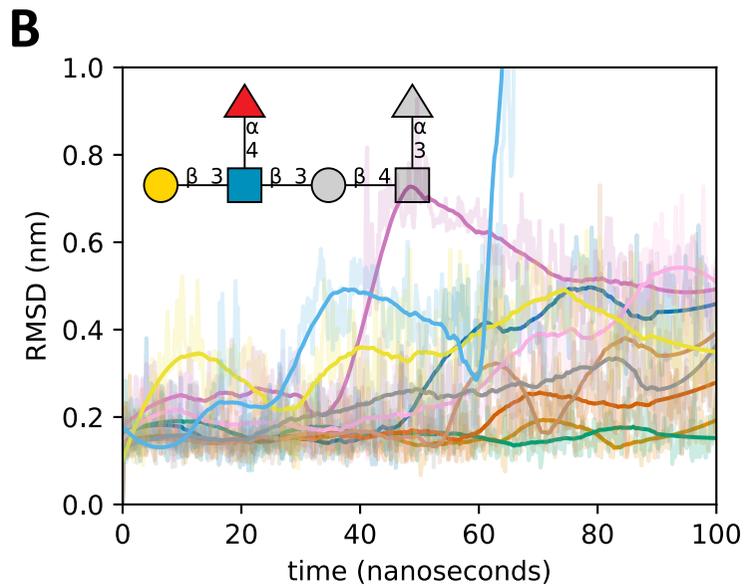
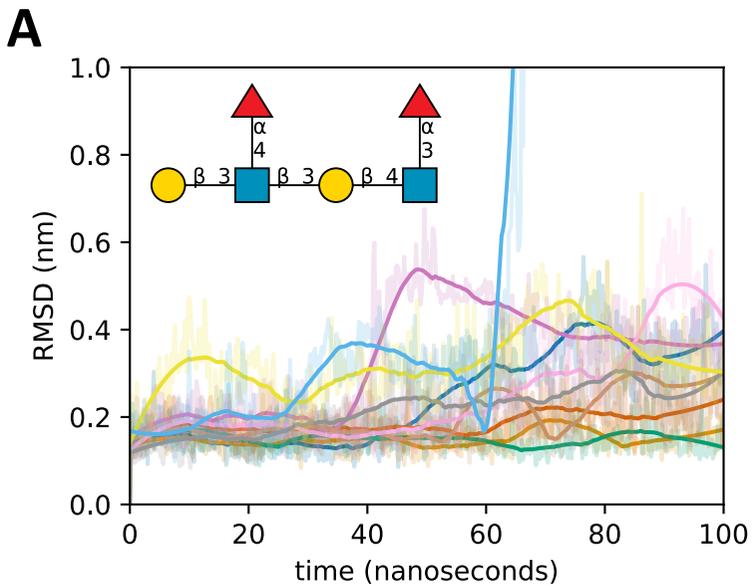
Glycans that are sulfated are listed as 3S or 6S, and Sp refers to a spacer group.

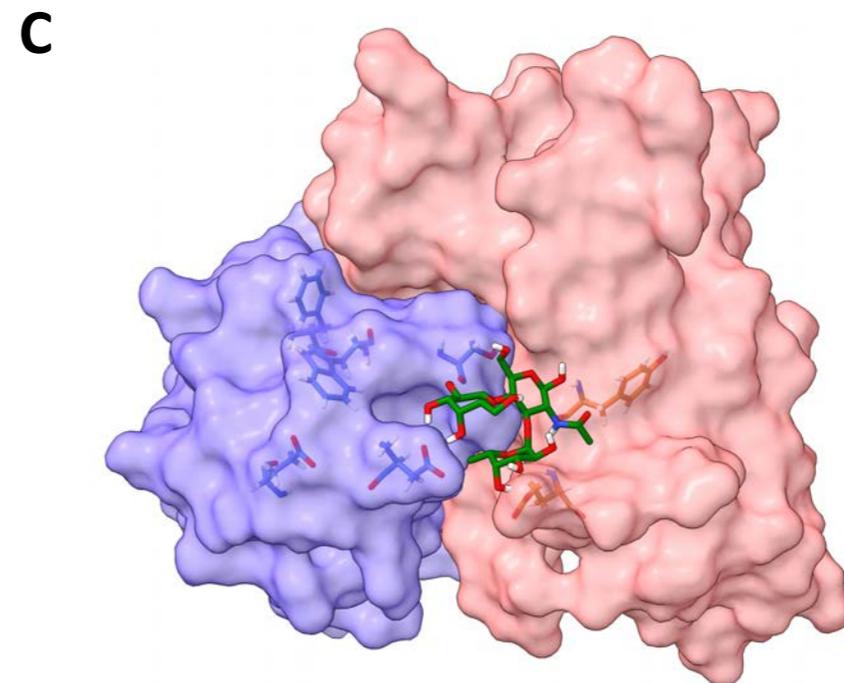
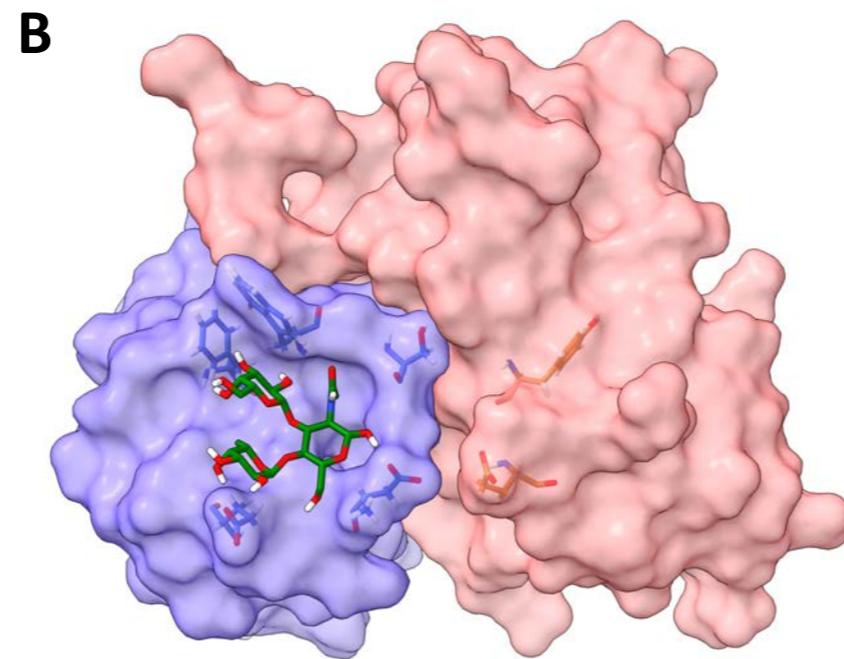
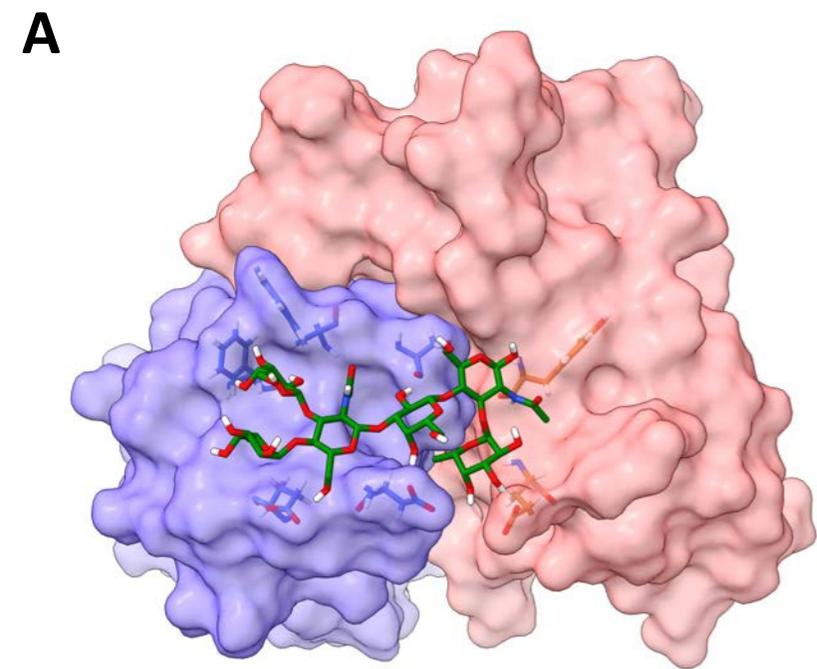












D

	TYR62/H	GLU64/H	ASP1/L	SER94/L	GLU27/L	ASN28/L
NAG1	0.88	0.00	0.00	0.03	0.00	0.00
FUC2	0.00	0.64	0.32	0.00	0.00	0.00
GAL3	0.00	0.00	0.23	0.86	0.00	0.00
NAG4	0.00	0.00	0.54	0.00	0.12	0.00
FUC5	0.00	0.00	0.00	0.00	0.13	0.02
GAL6	0.00	0.00	0.00	0.00	0.06	0.10

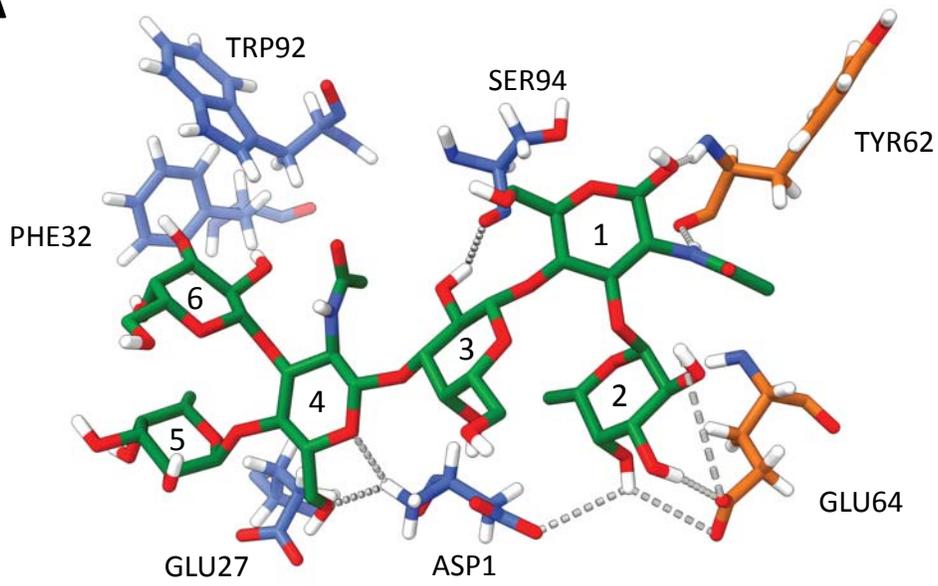
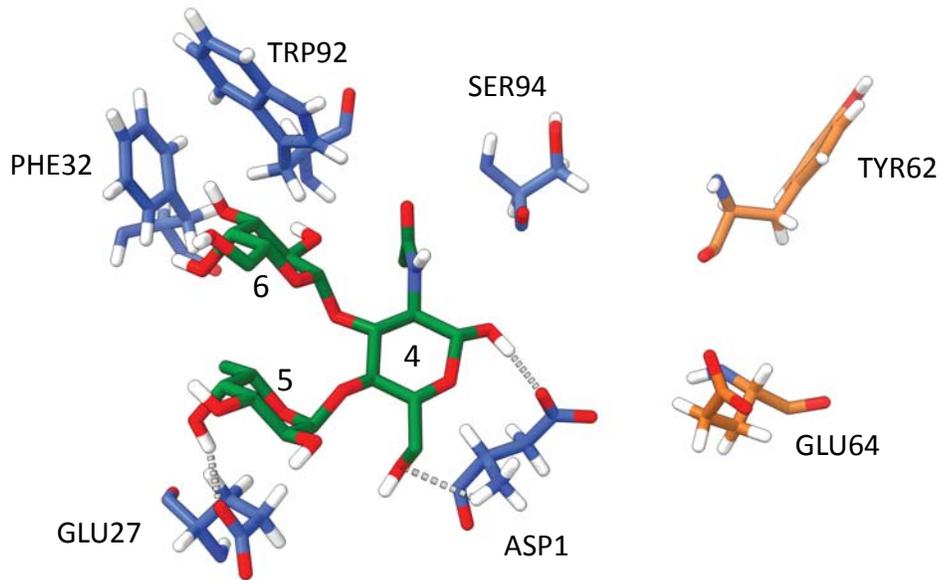
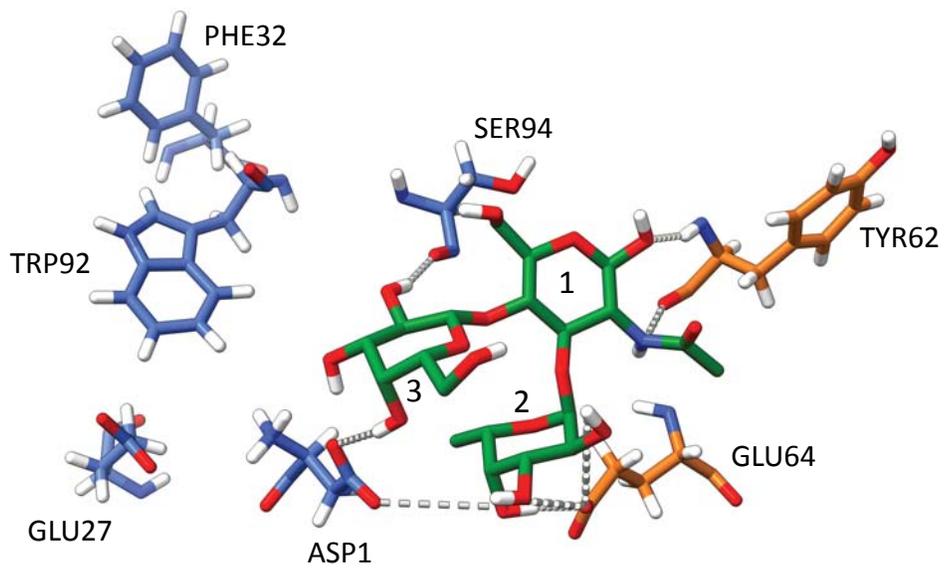
E

	TYR62/H	GLU64/H	ASP1/L	SER94/L	GLU27/L	ASN28/L
NAG4	0.00	0.00	0.77	0.09	0.02	0.00
FUC5	0.00	0.00	0.00	0.00	0.45	0.10
GAL6	0.00	0.00	0.00	0.00	0.00	0.00

F

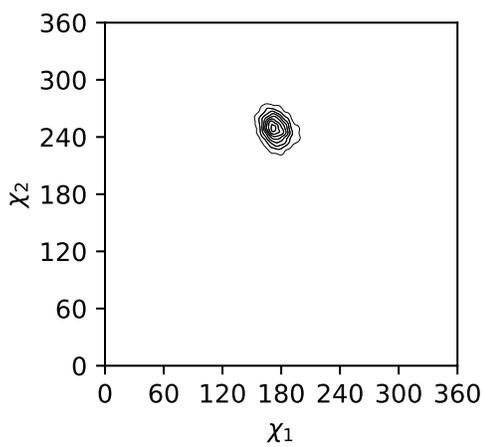
	TYR62/H	GLU64/H	ASP1/L	SER94/L	GLU27/L	ASN28/L
NAG1	0.91	0.02	0.00	0.02	0.00	0.00
FUC2	0.00	0.50	0.38	0.00	0.00	0.00
GAL3	0.00	0.00	0.30	0.87	0.00	0.00

H-bond occupancy (%)

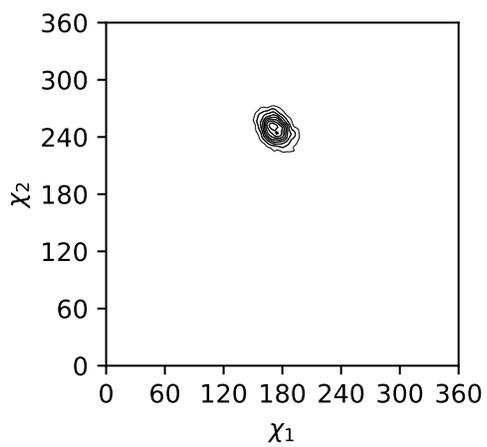
A**B****C**

Trp92

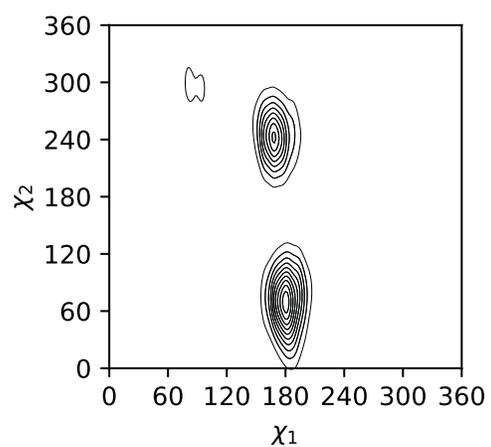
Le^aLe^x



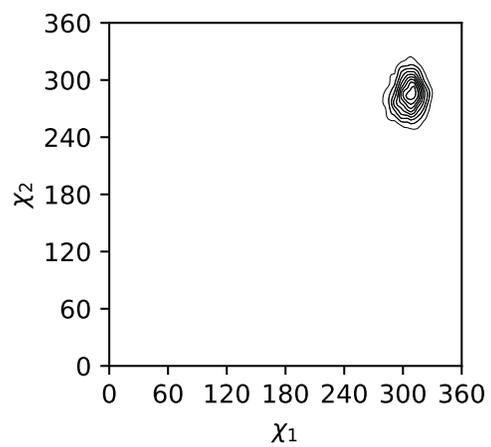
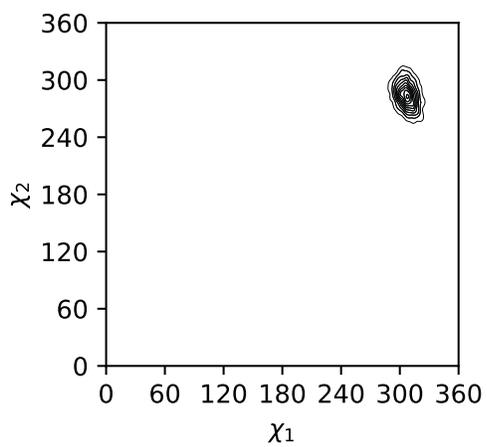
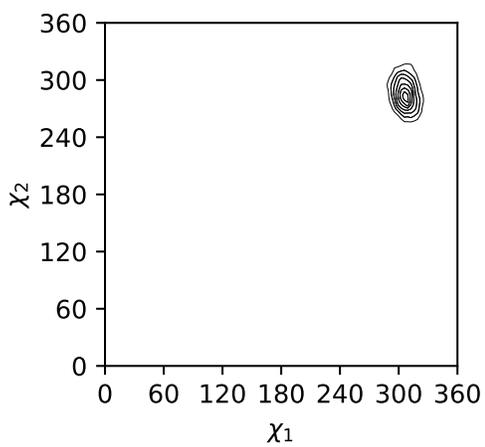
Le^a



Le^x



Phe32



Supporting Information

Molecular and structural basis for Lewis glycan recognition by a cancer-targeting antibody

Caroline Soliman¹, Andrew J. Guy¹, Jia Xin Chua², Mireille Vankemmelbeke², Richard S. McIntosh³, Sarah Eastwood¹, Vi Khanh Truong¹, Aaron Elbourne¹, Ian Spendlove³, Lindy G. Durrant^{2,3}, Paul A. Ramsland^{1,4,5,*}

¹School of Science, RMIT University, Melbourne, VIC, Australia.

²Scancell Limited, University of Nottingham Biodiscovery Institute, University Park, Nottingham, United Kingdom.

³Division of Cancer and Stem Cells, School of Medicine, University of Nottingham Biodiscovery Institute, University Park, Nottingham, United Kingdom.

⁴Department of Immunology, Central Clinical School, Monash University, Melbourne, Victoria, Australia.

⁵Department of Surgery Austin Health, University of Melbourne, Melbourne, Victoria, Australia.

Material Included

Figure S1. Schematic diagrams of Lewis histo-blood group and related antigens.

Figure S2. Production and crystallization of ch88.2 Fab.

Table S1. Data collection and crystallographic refinement statistics.

Figure S3. Light (VL) and heavy (VH) variable domain sequences of ch88.2.

Figure S4. Docked poses for Le^a-Le^x hexasaccharide.

Figure S5. Glycan array RFU data.

Figure S6. Interactions between ch88.2 Fab and Le^a-Le^x.

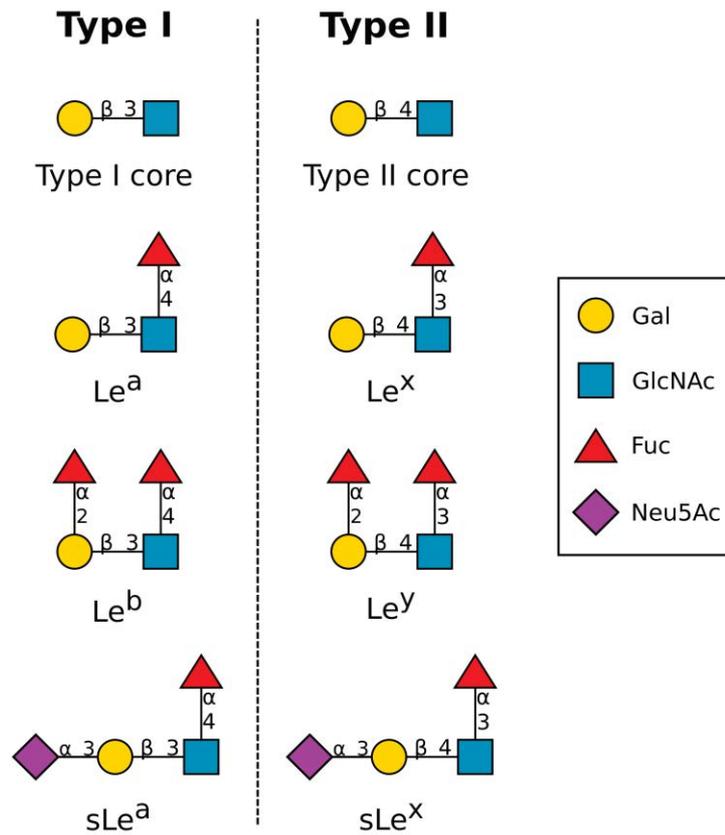


Figure S1. Schematic diagrams of Lewis histo-blood group and related antigens. Type I glycans include a core structure of Galβ1-3GlcNAc (known as Le^c), with terminal epitopes shown for Le^a trisaccharide, Le^b tetrasaccharide and sialylated version sLe^a. Type II glycans include a core structure of Galβ1-4GlcNAc, with terminal epitopes shown for Le^x trisaccharide, Le^y tetrasaccharide and sialylated version sLe^x. Each glycan is depicted as a cartoon in SNFG nomenclature.

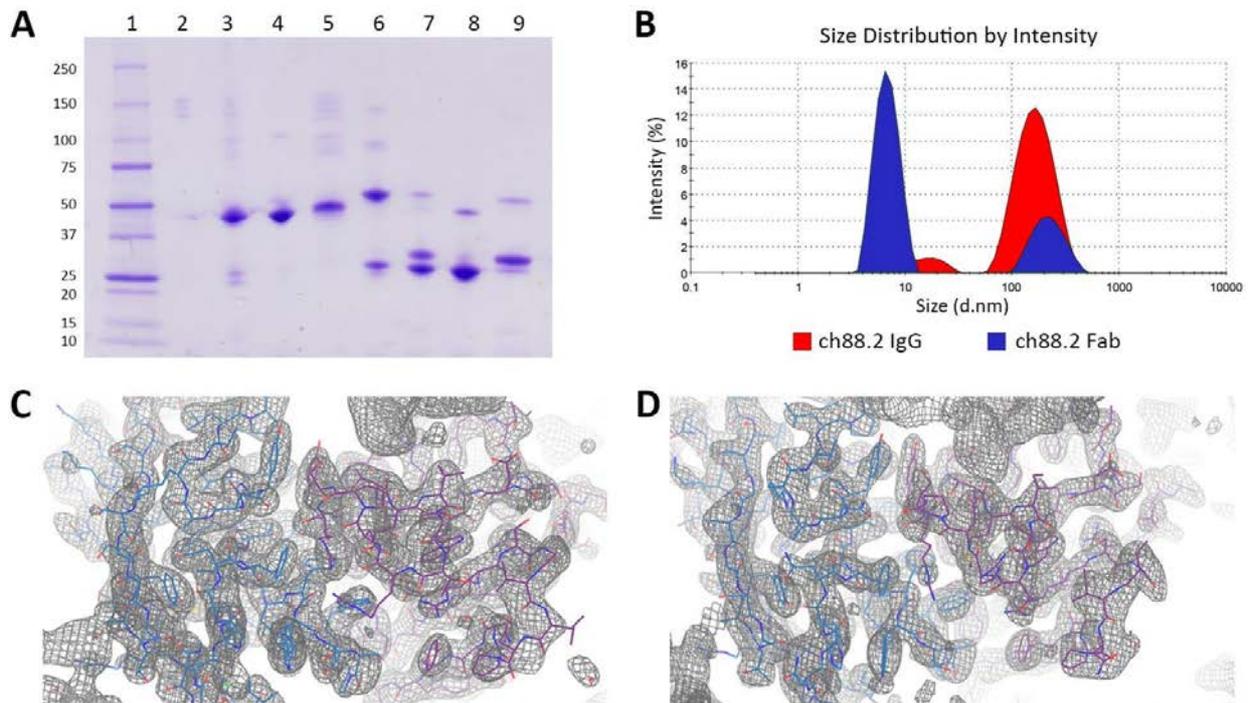


Figure S2. Production and crystallization of ch88.2 Fab. **A)** Coomassie stained SDS-PAGE gel. Lanes correspond to: 1, protein marker; 2 and 6, ch88.2 IgG; 3, and 7, antibody digest; 4 and 8, purified Fab; 5 and 9, undigested IgG and Fc fraction. Samples in lanes 6-9 were reduced using β -mercaptoethanol. **B)** Size distribution analysis of DLS data for ch88.2 IgG and ch88.2 Fab. **C & D)** Composite omit $2F_o - F_c$ map (displayed at 1.5σ level) for crystal structure of ch88.2 showing Fab1 binding site & Fab2 binding site. Residues are depicted as lines with L chain in blue and H chain in purple.

Table S1. Data collection and crystallographic refinement statistics. Values in parentheses refer to the highest resolution shell for each data set. Collection statistics were compiled from XDS, and refinement statistics from Phenix and CCP4, using MolProbity for Ramachandran statistics.

Parameter	ch88.2 Fab
Data collection	
Space group	<i>P1</i>
Unit cell dimensions (Å)	a = 38.3, b = 68.2, c = 91.3
Unit cell angles (°)	$\alpha = 110.5$, $\beta = 99.2$, $\gamma = 90.1$
Resolution range (Å)	50 – 2.3 (2.4 – 2.3)
Number of unique reflections	44428 (6787)
Data completeness (%)	92.6 (93.5)
Average multiplicity	1.8 (1.7)
<i>R</i> -factor	0.05 (0.62)
<i>R</i> _{meas}	0.07 (0.88)
Mean <i>I</i> / σ (<i>I</i>)	7.9 (1.2)
CC ½ (%)	99.8 (71.9)
Crystallographic refinement	
<i>R</i> _{work}	0.230 (0.349)
<i>R</i> _{free}	0.284 (0.381)
Average B-factor from Wilson plot (Å ²)	59.5
rmsd from ideal values	
Bond lengths (Å)	0.007
Bond angles (°)	1.0
Ramachandran plot values (%)	
Favored regions	92.5
Allowed regions	6.7
Outliers	0.8
Average B-factor (Å²)	
Protein atoms	68.4
Water	63.0

VL (108)

1 10 20 30 40 50

1 10 20 30 40 50 60

DIQMTQSPTSLASVGETVTITCRTS**ENI**-----**HNFL**TWYQQKQKSPQVLVY**NA**-----**KTLPD**

 60 70 80 90 100 (Seq)

70 80 90 100 110 120 (IMGT)

GVP-SRFGSG--SGTQYSLKINSLQPEDFGTYYC**QHFWS**SPWTFGGGTKLEIKR

VH (126)

1 10 20 30 40 50 60

1 10 20 30 40 50 60

EVKLEESGG-GLVQPGGSMKLSCAAS**GFTF**----**SDAW**MNWVRQSPEKGLEWVAE**IRSKVINPAI**YYAE

 70 80 90 100 110 120 (Seq)

70 80 90 100 110 120 (IMGT)

SVK-ERFTILRDDSKSSVYLQMNLSLRAEDTGIYYC**SRSTMITTRDPSRYFDV**WGAGTTVTVSS

Figure S3. Light (VL) and heavy (VH) variable domain sequences of ch88.2. Sequences are shown with both the sequential numbering used in the PDB file (black) and IMGT numbering scheme (green) for comparison. CDR sequences are indicated in bold (orange).

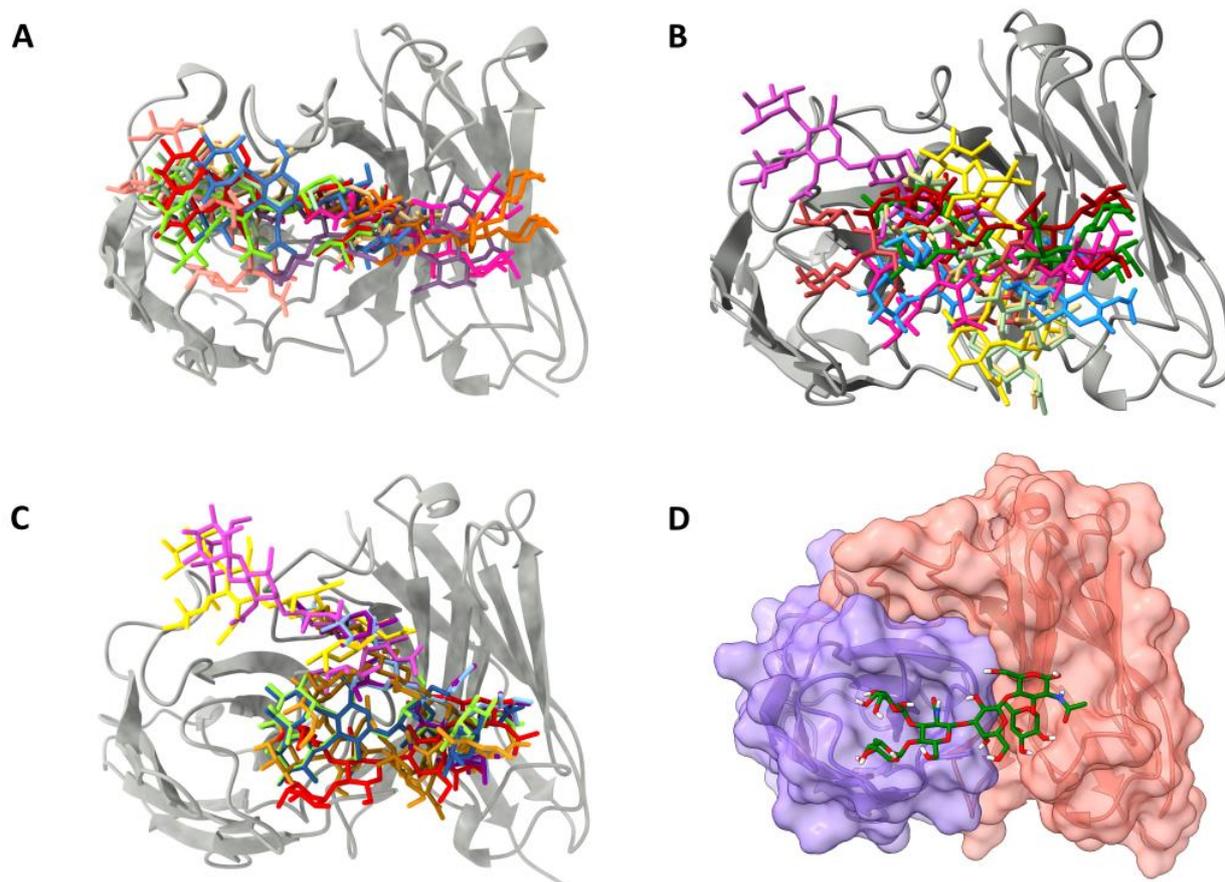
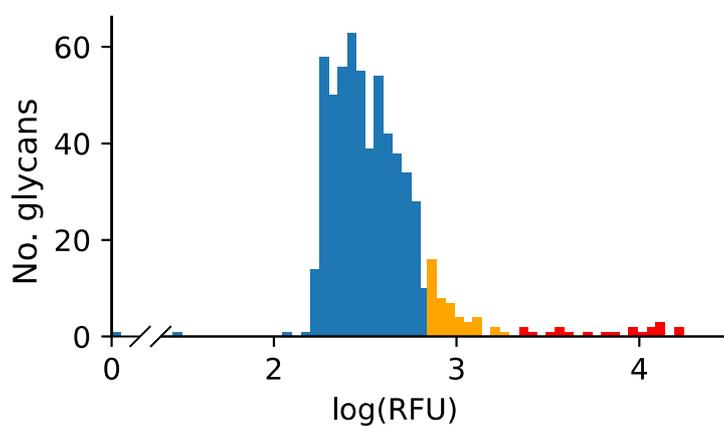


Figure S4. Docked poses for Le^a-Le^x hexasaccharide. Docking of Le^a-Le^x was performed using Vina-Carb for Fv1, Fv2 and an Fv structure obtained following molecular dynamics equilibration (Fv-MD). **A)** An overlay of the 9 docked poses for ch88.2 Fv1: Le^a-Le^x. **B)** An overlay of the 9 docked poses for ch88.2 Fv2: Le^a-Le^x. **C)** An overlay of the 9 docked poses for ch88.2 Fv-MD: Le^a-Le^x. Each antibody (Fv) is shown as grey ribbons with each pose depicted as coloured sticks. **D)** Final docked pose selected following screening and analysis of all results. The antibody is depicted as a transparent surface (L chain in purple and H chain in pink) overlaid on a ribbons representation. The ligand is depicted as sticks coloured by atom type (C, green; O, red; N, blue; H, white).



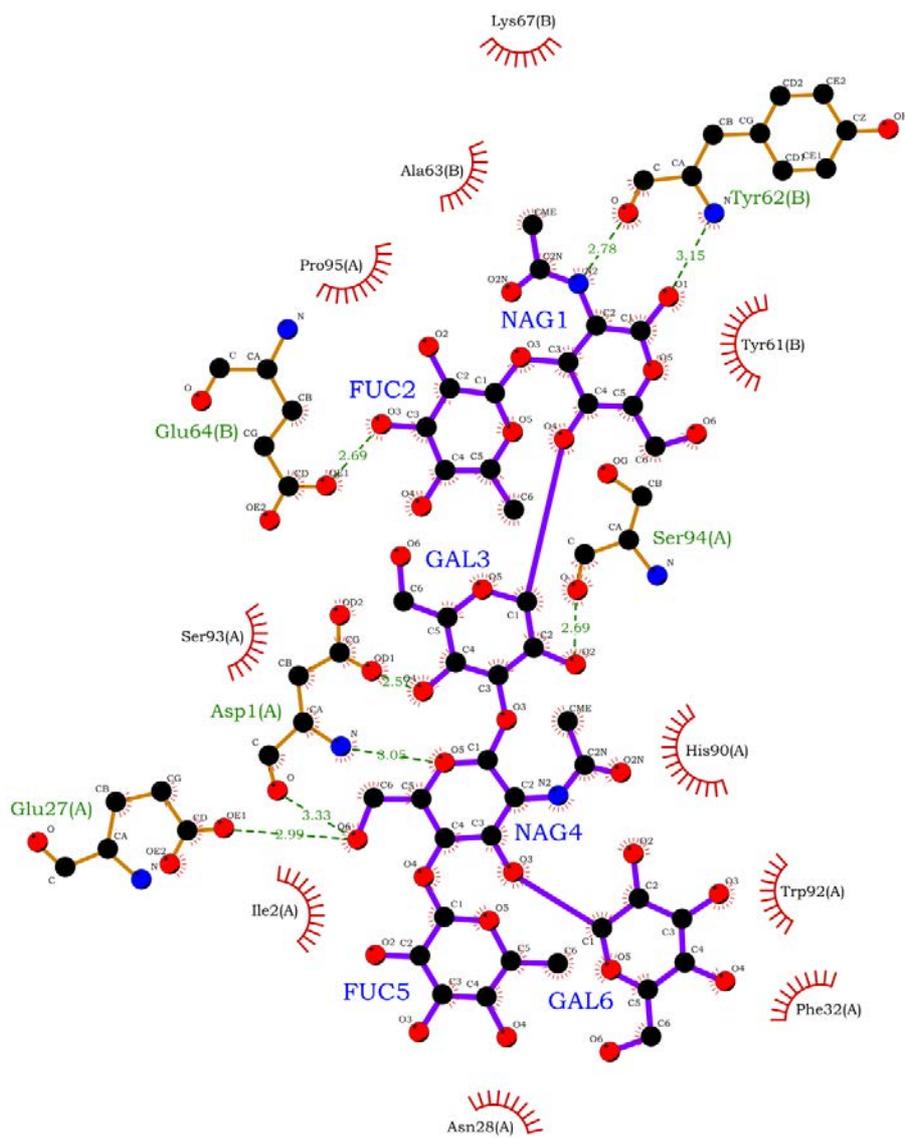


Figure S6. Interactions between ch88.2 Fab and Le^a-Le^x. Residues from the light (A) and heavy (B) chains of ch88.2 Fab are forming hydrogen bonding interactions (indicated by dotted green lines with distances) and van der Waals interactions (indicated by red spokes) with the Le^a-Le^x hexasaccharide (indicated as NAG1, FUC2, GAL3, NAG4, FUC5 and GAL6).