Title

Biomolecular Recognition of the Glycan Neoantigen CA19-9 by Distinct Antibodies.

Journal

Journal of molecular biology, 433(15)

ISSN

0022-2836

Authors

Borenstein-Katz, Aliza
Warszawski, Shira
Amon, Ron
et al.

Publication Date

2021-07-01

DOI

10.1016/j.jmb.2021.167099

Peer reviewed

Published in final edited form as:
J Mol Biol. 2021 July 23; 433(15): 167099. doi:10.1016/j.jmb.2021.167099.

Biomolecular Recognition of the Glycan Neoantigen CA19-9 by Distinct Antibodies

Aliza Borenstein-Katz[#1], **Shira Warszawski**[#2], **Ron Amon**[#3], **Maayan Eilon**[1], **Hadas Cohen-Dvashi**[1], **Shani Leviatan Ben-Arye**[3], **Nova Tasnima**[4], **Hai Yu**[4], **Xi Chen**[4], **Vered Padler-Karavani**[3,*], **Sarel Jacob Fleishman**[2], **Ron Diskin**[1,*]

[1]Department of Chemical and Structural Biology, Weizmann Institute of Science, 76100 Rehovot, Israel

[2]Department of Biomolecular Sciences, Weizmann Institute of Science, 76100 Rehovot, Israel

[3]Department of Cell Research and Immunology, The Shmunis School of Biomedicine and Cancer Research, The George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv 69978, Israel

[4]Department of Chemistry, University of California, Davis, CA 95616, USA

[#] These authors contributed equally to this work.

Abstract

Glycans decorate the cell surface, secreted glycoproteins and glycolipids, and altered glycans are often found in cancers. Despite their high diagnostic and therapeutic potential, however, glycans are polar and flexible molecules that are quite challenging for the development and design of high-affinity binding antibodies. To understand the mechanisms by which glycan neoantigens are specifically recognized by antibodies, we analyze the biomolecular recognition of the tumor-associated carbohydrate antigen CA19-9 by two distinct antibodies using X-ray crystallography. Despite the potential plasticity of glycans and the very different antigen-binding surfaces presented by the antibodies, both structures reveal an essentially identical extended CA19-9 conformer, suggesting that this conformer's stability selects the antibodies. Starting from the bound structure of one of the antibodies, we use the AbLIFT computational algorithm to design a variant with seven core mutations in the variable domain's light-heavy chain interface that exhibits tenfold improved affinity for CA19-9. The results reveal strategies used by antibodies to

[*]***Correspondence to Vered Padler-Karavani and Ron Diskin:*** vkaravani@tauex.tau.ac.il (*V. Padler-Karavani*), ron.diskin@weizmann.ac.il (*R. Diskin*).

CRediT authorship contribution statement

Aliza Borenstein-Katz: Investigation, Formal analysis, Writing -original draft, Visualization. **Shira Warszawski:** Software, Formal analysis, Writing -original draft. **Ron Amon:** Investigation, Resources, Writing -original draft, Visualization. **Maayan Eilon:** Investigation, Formal analysis, Visualization. **Hadas Cohen-Dvashi:** Investigation, Resources, Formal analysis. **Shani Leviatan Ben-Arye:** Resources, Formal analysis. **Nova Tasnima:** Resources. **Hai Yu:** Resources. **Xi Chen:** Resources. **Vered Padler-Karavani:** Conceptualization, Investigation, Writing -review & editing, Visualization, Formal analysis, Supervision, Funding acquisition. **Sarel Jacob Fleishman:** Conceptualization, Investigation, Writing -review & editing, Software, Formal analysis, Supervision, Resources. **Ron Diskin:** Conceptualization, Investigation, Writing -review & editing, Formal analysis, Visualization, Supervision.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: The Weizmann Institute of Science and the Tel Aviv University have filed a patent application for the AbLIFT designs.

Europe PMC Funders Author Manuscripts

specifically recognize glycan antigens and show how automated antibody-optimization methods may be used to enhance the clinical potential of existing antibodies.

Keywords

CA19-9; antibodies; structure; design; glycans

Introduction

Cancer is a leading cause of death worldwide, with epithelial carcinoma the most devastating. Changes in cell surface markers are one of the hallmarks of cancer, and antibodies that bind these markers are ideal therapeutics and/or diagnostic tools.1 Surface glycosylation is a universal feature of cells but is often altered during malignant transformation, leading to a distinct subset of antigens that are selectively and abundantly expressed on cancer cells.[2–5] This feature is intimately associated with abnormal expression of the glycosylation biosynthetic pathways, leading to variations in the basic core carbohydrate chains (glycans) conjugated to glycoproteins and glycolipids.[3,6] These aberrations particularly affect the expression of sialic acids (Sias) that cap cell surface glycans. For example, the sialyl Lewis a (SLe^a) tetrasaccharide stems from incomplete synthesis of the normal glycan Disialyl-Le^a. While both SLe^a and Disialyl-Le^a are generated via the same metabolic pathway, reduction or loss of expression of the α2 – 6-sialyltransferase (ST6GalNAc VI) during malignancy shifts the pathway towards expression of the cancer antigen SLe^a, also known as carbohydrate antigen CA19-9[7] (Figure 1). Altered glycosylation pattern often correlate with advanced cancer stage, progression and/or metastasis.[2,4,5,8] Interestingly, a recent study in mice demonstrated that CA19-9 is an active driver of pancreatitis, which leads to the development of pancreatic cancer.[9] This discovery assigns, for the first time, an active role for CA19-9 as a cancer driver. Importantly, mAbs targeting CA19-9 were able to reverse pancreatitis in this mouse model,[9] establishing CA19-9 as a prime target for cancer therapy.

SLe^a is detected on pancreatic, colorectal, stomach and liver cancers.[7,10] This cancer-associated marker is widely used in clinical practice for serological assays.[5,11,12] It is the only FDA-approved test for pancreatic cancer and is also used in assays for colorectal, gastric and biliary cancers.[5] The assay is based on a monoclonal antibody (mAb) capturing the CA19-9 antigen and is commonly used to monitor clinical response to therapy; however, it is not useful for early detection or diagnosis due to unacceptably high rates of false positive and false negative readouts.[11–13] Therefore, although this serological assay has been available for almost three decades, the interpretation of CA19-9 measurements is largely hampered by non-specific increased reads for the levels of CA19-9, either due to associated morbidity (e.g. obstruction of the biliary tree or inflammation) or due to assay-dependent variability, both in diseased and healthy subjects.[14] As a result, pancreatic cancer is often detected too late at an advanced stage resulting in a low five-year survival rate.

A potential obstacle to using anti-carbohydrate antibodies for theranostics is their low affinity and low specificity compared to antibodies targeting proteins.[15,16] This limitation

Europe PMC Funders Author Manuscripts

prompted development of tools to better define such antibody-antigen interactions[17] and enhance their affinity.[18] Thus, detailed structural information for the CA19-9 and its recognition by mAbs is a step towards the design of more efficient reagents in the fight against some of the most devastating cancer types. Recently, we described a new antibody-design method, called AbLIFT, that focuses design calculations on the interfaces formed between specific antibody light and heavy chain pairs as observed in crystallographic analysis.[19] This method has the potential to increase antibody binding affinity by fixing the packing between the light and heavy chains in their binding-competent forms. The AbLIFT approach starts with a user-guided definition of the positions that are targeted for design; effectively, these are restricted to positions at or around the light-heavy chain interface of the variable domain and away from positions that interact directly with the antigen. The next steps are automated: a multiple-sequence alignment of homologous antibodies is generated and informs the selection of individual mutations by a computed position specific scoring matrix (PSSM). Selected mutations are then modeled individually against the background of the parental antibody using atomistic design calculations in Rosetta, and mutations that destabilize the antibody are eliminated from consideration. In the last step, all of the combinations of individually tolerated mutations (computed in the previous steps) are atomisti-cally modeled, relaxed and ranked according to their computed energies; the lowest-energy designs are clustered to eliminate designs that are too similar to one another, and the top few dozen designs are recommended for experimental analysis. AbLIFT was used previously to optimize the computed energy of a variety of antibodies. Furthermore, in experimental screens of up to 20 designs per target antibody, several of the designs exhibited increased expressibility, thermal stability, aggregation resistance and antigen-binding affinity.[19] Recently, the same design strategy was used to substantially improve the stability and expressibility of an enzyme that exhibits anti-biofilm properties.[20] Thus, we hypothesized that AbLIFT could also be used to enhance anti-CA19-9 antibodies.

Here, we provide molecular insights into antigen recognition by two of the most well-defined anti-CA19-9 mAbs, the murine 1116NS19.9[21,22] commonly used in the CA19-9 serological test, and the human 5b1[23] that is currently investigated for cancer imaging in clinical trials.[24] We present high-resolution crystal structures of both antibodies in complex with CA19-9 antigen. These structures reveal two distinct binding solutions to a single conformer of CA19-9. We further use the state-of-the-art AbLIFT computational tool together with this structural information to design mAbs that target CA19-9 with an order of magnitude greater affinity.

Results

Antibodies 1116NS19.9 and 5b1 are specific to CA19-9

To reveal the molecular basis for the immune recognition of the tumor-associated carbohydrate antigen CA19-9, we selected two of the most widely used mAbs, 1116NS19.9 and 5b1 for structural studies. Since the presentation mode of glycans can affect their recognition,[25,26] we first evaluated antibody-glycan recognition both with glycans attached to a solid surface or in solution, as binding a soluble form of the glycan is a prerequisite for structural studies. The mAbs' variable domains were cloned and their functionality was

confirmed both as single-chain Fv (scFv) fragments, and as full-length human IgG1. First, the variable heavy and light chain fragments (VH and V_L, respectively) of 1116-NS-19–9 and 5b1 were each cloned into pETCON2 plasmid as scFv with (G4S)3 linker between the two variable domains, and were transformed to yeast cells which were induced to express cell surface scFvs.[27] Antigen recognition was evaluated in solution by flow cytometry of scFv-expressing yeast cells against SLe^a-nanoparticles with multivalent expression of antigen to resemble their presentation on cancer cells. This analysis revealed slightly higher scFv surface expression of 5b1-scFv. In addition, we observed a higher antigen-binding signal with 5b1-scFv than with 1116NS19.9-scFv (Figure 2(a), Supplementary Figure 1(a)). In both 5b1-YSD and 1116NS19.9-YSD two populations are noticed due to variability in scFv expression levels, however antigen binding of both populations seems to be similar (Figure 2(a)). Subsequently, the variable regions of the two mAbs were cloned into a human IgG1 scaffold and antigen recognition was examined. Binding of full-length antibodies to multivalent nanoparticles coated onto a solid surface showed binding to SLe^a but minimal recognition of Le^a antigen that lacks the terminal sialic acid (Figure 2(b)), implying that sialic acid recognition plays an important role for the binding of both antibodies. Additionally, binding to SLe^x antigen was minimal emphasizing the glycan linkage role in antibody recognition (Figure 2(b)), since SLe^x tetrasaccharide (Neu5Acα2 – 3Gal**β1** – **4**(Fuc**α1** – **3**)GlcNAc) is a structural isomer of SLe^a (Neu5Acα2 – 3Gal**β1** – **3**(Fuc**α1** – **4**)GlcNAc), having the same carbohydrate building blocks. These data also show that the cloned antibodies are fully functional, both against a flexible antigen in solution and with antigen fixed to a solid surface, and are specific for their target antigen SLe^a (CA19-9).

Structure determination of CA19-9 bound to the two mAbs

As a first step towards understanding the biomolecular recognition of CA19-9, we produced the two mAbs in HEK293F cells and purified them using protein-A affinity chromatography. We digested the two IgGs with papain followed by protein-A affinity and size-exclusion chromatography to obtain the antigen-binding fragments (Fab; Figure 3(a)). CA19-9 (SLe^a) antigen was chemoenzymatically synthesized as described in ref.[28], with a terminal primary amine-containing linker (SLe^aβProNH_2, Figure 1). We crystallized both the apo (without antigen) and the holo (with CA19-9 antigen) states of Fab fragments of 1116NS19.9 and 5b1 mAbs, followed by X-ray diffraction analyses at the European Synchrotron Radiation Facility. Complete data sets were collected at 1.6 A and 1.5 A resolutions for the CA19-9-bound and the apo-state of mAb 1116NS19.9, respectively (Table 1). Complete data sets were further collected at 2.4 A resolutions for the CA19-9-bound and the apo-Fab of Ab 5b1 (Table 1), and all structures were solved using molecular replacement. For the apo mAb 1116NS19.9, a human-derived Fab (PDB: 3U7W) was used as the search model and subsequently the 1116NS19.9 structure was used as a search model for solving the rest of the structure through molecular replacement. In the holo-structures of both 1116NS19.9 and 5b1, clear electron density for CA19-9 was observed, allowing accurate modeling (Figure 3(b)). In both structures, density for the propyl linker attached to the CA19-9 was either missing completely (5b1) or was only weakly visible for the first carbon atom of the linker (1116NS19.9). Hence, this linker was omitted from the models.

The two antibodies recognize a similar, low-energy conformer of CA19-9

The structures of 1116NS19.9 and 5b1 mAbs reveal that in both cases, the CA19-9 antigen binds in a groove that is formed between the V_H and V_L domains (Figure 3(c)). Typically, glycosidic bonds can freely rotate, and hence oligosaccharides inherently present an ensemble of conformations in solution.[29] Nevertheless, comparing the bound CA19-9 antigen from the structures of 1116NS19.9 and 5b1 reveals that in both cases, CA19-9 assumes a surprisingly similar conformer (Figure 4(a)) with an all-atom RMSD (root-mean-square deviation) of 0.97 Å. Of note, 1116NS19.9 and 5b1 were isolated from different species (i.e., human vs. mouse, respectively) and were elicited against CA19-9 that was displayed in two very distinct contexts (i.e., on a carcinoma cell and as a protein-conjugated antigen, respectively).

The fact that both antibodies converged to recognize a similar conformation of CA19-9 implies that the conformer observed in our crystallographic analyses is energetically preferred. To understand the possible reasons for this convergence, we analyzed previously published structures of the Lewis a core region of CA19-9 (i.e., only the Fuc, GlcNAc, and Gal).[30–33] In all of these cases, the core regions assumed a similar conformation to one that was shown computationally to represent a low-energy state.[34] Comparing the Lewis a core of CA19-9 to one of the published structures of Lewis a[33] reveals that they both assume a very similar conformation (Figure 4 (b)). Analyzing the Siaα2 – 3Gal bond in CA19-9 indicates glycosidic torsion angles of 49.5° and 119.1° for ϕ and ψ, respectively (Figure 4(c)). These torsion angles also map to an energetically preferred region as previously calculated for CA19-9 using molecular-dynamics simulations.[34] Hence, both 1116NS19.9 and 5b1 have indeed converged to recognize a low-energy conformer of CA19-9.

The unanticipated finding that the two mAbs recognize a similar conformer of CA19-9 raises the question whether they also utilize similar binding mechanisms. By superimposing the structures according to the CA19-9 antigen, it is clear that the antibodies bind CA19-9 in two distinct ways (Figure 4(d)). The two Fabs approach CA19-9 from different angles (Figure 4 (d)) and use their complementarity-determining regions (CDRs) for binding in different modes (Figure 4(e)). In the case of mAb 1116NS19.9, aside from CDRL1, all other CDRs are involved in ligand binding, and CDRL3 is the most significant contributor to molecular interactions. By contrast, mAb 5b1 does not engage CA19-9 through CDRH1 and CDRH2, while CDRH3 and CDRL1 make significant contacts with CA19-9 (Figure 4(e)). Comparing the binding of CA19-9 by 1116NS19.9 to 5b1, the former has a relatively deeper groove than the latter, which is reflected in a larger total buried surface area for the complex formation (973 and 859 $Å^2$ for 1116NS19.9 and 5b1, respectively), though both antibodies exhibit a similar binding specificity profile for CA19-9.

Molecular basis for the specificity toward CA19-9

CA19-9 differs from Disialyl-Le^a antigen by the lack of a sialic-acid (Sia) moiety that is typically connected to the carbon-6 (C6) of the GlcNAc (Figure 1). In the case of 1116NS19.9, the hydroxyl group extending from C6 directly faces the heavy-chain, leaving no room for accommodating the extra Sia in this conformation (Figure 4(f)). Moreover, the binding of CA19-9 to 1116NS19.9 is partly facilitated by a hydrogen bond that the GlcNAc

C6-hydroxyl group forms with Asn52A (Kabat numbering scheme) on CDRL2, favoring a free hydroxyl group in this position. In the case of mAb 5b1, the hydroxyl extension from C6 of the GlcNAc is not buried at the interface, as seen with 1116NS19.9 (Figure 4 (f)). Nevertheless, an additional Sia moiety cannot be accommodated unless the Sia assumes a conformation that is potentially less energetically favorable than the one observed in both antibody-bound complexes. Moreover, the free hydroxyl of C6 participates in a water-mediated interaction with the light-chain of 5b1, providing an additional selection power for CA19-9 in this conformation.

Molecular recognition of CA19-9

Considering the relatively small size of CA19-9 (819 Dalton) and its hydrophilic nature, it could be regarded as a suboptimal immunogen.[35–37] As such, the recognition of CA19-9 by the mAbs could be suboptimal and restricted to only a few contacts. Nevertheless, 1116NS19.9 and 5b1 display intricate interaction networks with CA19-9 (Figure 5(a) & 5(b)) that are likely to be critical for binding affinity and specificity to this challenging antigen. In the case of 1116NS19.9, all the hydroxyl groups of CA19-9 that face the antibody, aside from O7 of the sialic acid's hydroxyl, participate in either direct or water-mediated polar interactions with the mAb (Figure 5(a)). Many amino acid sidechains from both chains mediate polar contacts with the antigen, including heavy chain positions Trp33, Asn52A, Arg95, and Phe96, as well as the light-chain positions Tyr49, Arg50, Arg53, Tyr91, Asp92, and Arg96 (Figure 5(a)). Of these interactions, the guanidino group of the light-chain Arg50 is especially important as it forms an ideal salt-bridge with the carboxylate of the Sia (Figure 5(a)). Indeed, 1116NS19.9 with a R50A mutation in its light chain recognize the pancreatic-cancer BxPc3 cell line at a lower efficiency compared with 1116NS19.9 (Figure 6, Supplementary Figure 1(b)). Besides the polar contacts, a few additional hydrophobic interactions contribute to and complete the recognition site of CA19-9 on 1116NS19.9 (Figure 5(a)). In the case of 5b1, there is also a saturated network of polar interactions that include all the hydroxyls of CA19-9 that face the antibody, aside from O4 of the Gal (Figure 5(b)); some of these interactions are water-mediated. On the heavy-chain, Arg97, Arg98, Thr100A, Gly100B, and Ala100D, all from CDRH3, contact CA19-9 (Figure 5(b)). The lightchain residues Ser30B, Phe32, Tyr34, Arg50, Trp50, Trp91, and Asp93 complete the set of residues that form the polar binding site of CA19-9. Interestingly, the carboxyl group of the Sia in the case of 5b1 is exposed to solvent and does not contribute to the binding of CA19-9. In addition, and similarly to 1116NS19.9, several hydrophobic interactions contribute to the binding of the CA19-9 antigen (Figure 5(b)).

Binding-induced conformational changes

In order to bind CA19-9, 1116NS19.9 and 5b1 may need to undergo conformational changes. To evaluate this possibility, we compared the apo and the holo-structures of 1116NS19.9 (Figure 7(a)) and of 5b1 (Figure 7(b)). Indeed, the superimposition of the apo and holo 1116NS19.9 reveals significant conformational changes between the structures (Figure 7(a)). Trp33 and Arg95 from CDRH1 and CDRH3, respectively, form a tight cation-π interaction in the apo-state of 1116NS19.9. For binding CA19-9 in the holostructure, this cation-π interaction breaks, and both residues assume different rotamers that allow them to accommodate and to engage with CA19-9 through hydrogen bonds (Figure 7(a)). Also,

Asn53 of CDRH2 assumes a different rotamer that allows the entire loop to move closer to CA19-9, enabling Asn52A to form a hydrogen bond with the hydroxyl of the GlcNAc-C6 (Figures 7(a) & 4(f)). In the light chain of 1116NS19.9, Arg50 of CDRL2 forms cation-π interaction with a nearby tyrosine, and upon binding to CA19-9, this interaction breaks, and Arg50 assumes a different rotamer that allows it to form a salt-bridge with the carboxyl group of the Sia moiety (Figure 7(a)). In contrast to these substantial structural rearrangements of 1116NS19.9, the superimposition of the apo and holo-structures of 5b1 reveals that the binding of CA19-9 does not involve any notablen conformational changes (Figure 7(b)). Hence, not only do these antibodies recognize CA19-9 at different angles (Figure 4(d)), they also evolved to use different binding mechanisms. The binding site of 5b1 is pre-configured for binding, whereas the conformational changes of 1116NS19.9 that are coupled to the abolishment of cation-π interactions suggest an induced-fit mechanism.

Enhancing 1116NS19.9 affinity using computational antibody design

Increasing sensitivity of CA19-9 recognition is of great interest for advancing diagnosis and management of CA19-9-positive malignancies in general, and of pancreatic cancer in particular. Focusing on the clinically used 1116NS19.9, we sought to enhance its affinity toward CA19-9. Conventional approaches to design or engineer higher-affinity antibody variants typically direct mutations to the antigen-binding surface.[27] The intricate and highly polar antigen-binding surface revealed by the crystallographic analyses of the two antibody-bound complexes suggested, however, that mutations at these surfaces might destabilize the complex. By contrast, the recently published AbLIFT computational antibody-design method focuses design calculations on the interfaces formed between the antibody light and heavy chains, away from the antigen-binding surface (available as a web server for academic users http://AbLIFT.weizmann.ac.il).[19] While these interfaces are often tightly packed and therefore challenging for computational or experimental strategies, optimizing them mitigates the risk of altering the critical antibody-antigen interactions; these interactions, as revealed by our crystallographic analysis, are remarkably intricate in this particular case. Furthermore, based on the induced-fit mechanism revealed for 1116NS19.9 recognition of CA19-9, we hypothesized that optimizing the interactions observed in its holo-state structure would improve CA19-9 binding affinity.

We applied AbLIFT to the coordinates of 1116NS19.9 bound to CA19-9 and selected 17 variants (Supplementary Table 1) that were calculated to have the most significant favorable change in Rosetta free energy (ΔΔG) of the variable domain and were different by at least three mutations from the parental antibody and from other designs. Based on visual inspection, we chose heavy-chain positions Asp35, Thr93, Thr94, and Tyr98 (Figure 8(a)) and light chain residues Ser43, Asp56, Tyr87, and Phe98 (Figure 8(b)) for design. In a preliminary screen, we experimentally expressed all 17 designs and tested their binding using surface plasmon resonance (SPR) to a monomeric CA19-9 as an analyte in steady-state binding experiments with five-point concentration series ranging from 0.488 μM to 125 μM. These assays reflect the true microscopic affinities in the absence of avidity effects. Out of 17 variants, 15 either did not bind CA19-9 at all, or bound but displayed affinities that were weaker or similar to 1116NS19.9. Two designs (i.e., AbLIFT-2 & AbLIFT-15) displayed higher binding affinities to CA19-9, of which we selected AbLIFT-15 for further

analysis. Using steady-state analyses and an extended 14-points concentration series ranging from 500 μM to 0 μM of CA19-9, we measured for AbLIFT-15 a K_D value of 1.7 μM, compared to a K_D value of 14.7 μM for 1116NS19.9 (Figure 8(c)), which represents a nearly tenfold improvement in affinity. We further measured the affinity of 5b1 to CA19-9 as a reference in this avidity-free system and determined the K_D value to be 12.8 μM. Hence, AbLIFT-15 exhibits tenfold higher affinity compared to both 1116NS19.9 and 5b1 anti-CA19-9 antibodies.

AbLIFT-15 has a total of seven mutations compared to 1116NS19.9, a large number of core mutations relative to mutants obtained in conventional antibody engineering and design efforts (T93A, T94V, and Y98F in the heavy chain and S43P, D56P, Y87W, and F98W in the light chain). To verify that the molecular recognition of CA19-9 by the designed antibody did not differ substantially from the parental antibody, we determined the bound complex structure at 1.4 Å resolution (Table 1). Remarkably, despite seven core mutations, the structure reveals an almost identical main-chain structure (RMSD of 0.22 Å for 220 Cα atoms of the variable regions) (Figure 8(a) & (b)). Crucially, binding to CA19-9 was very similar to the conformation observed in the parental antibody 1116NS19.9 as predicted by the design model.

We also noted that one of the designed spacefilling mutations, Y87W (light chain) (Figure 8(b)) impacted a position that is conserved to phenylalanine or tyrosine in the vast majority of antibodies. In fact, the IMGT antibody resource[38] shows no occurrence of tryptophan at this position in any mouse or human antibody germline. The evolutionary-likelihood scores (PSSM) of phenylalanine and tyrosine at position 87 are both relatively high; and yet, because of the physicochemical similarity between the aromatic amino acids, tryptophan exhibits a PSSM score of 0 and passes the threshold of tolerated mutations. Thus, despite the position's high evolutionary conservation, the fact that in this case, the mutation also fills a cavity favored it in the atomistic design steps. Remarkably, considering the highly unusual mutation, the crystallographic structure showed very high accuracy in this as in all the other designed sidechains (Figure 8(b)). Thus, the AbLIFT strategy, though it is constrained by evolutionary calculations, yields mutations that are rarely observed in either natural or engineered antibodies.

AbLIFT-15 retains desired biophysical and biochemical properties

Introducing multiple mutations at the VH/VL interface of 1116NS19.9 may have undesired effects on the stability, specificity, or reactivity of the antibody. To evaluate that, we performed several different assays. First, we determined the temperature denaturation midpoints (T_m) for the Fab portions of 1116NS19.9 and AbLIFT-15 using nano-deferential scanning fluorimetry (Figure 8 (d)). In the presence of CA19-9 and also in its absence, the Fabs of both 1116NS19.9 and AbLIFT-15 exhibits high thermal stability with Tm values at above 80 °C (Figure 8(d)). Interestingly, AbLIFT-15 was somewhat less stable than 1116NS19.9 with a ΔT_m = 2.3 °C in the absence of CA19-9 and ΔT_m = 1.0 °C in the presence of CA19-9. These differences in Tm indicate some undesired effect of the mutations, perhaps by destabilizing V_H or V_L. Alternatively, this small reduction in T_m may reflect an unavoidable consequence of reducing the configurational entropy of the antibody,

which results in a higher free energy. Nonetheless, the reduction in thermostability is marginal, and both 1116NS19.9 and AbLIFT-15 could be regarded as thermostable.

We next evaluated the specificity of AbLIFT-15 using a printed sialoglycan micro-array (Supplementary Table 2). This array analysis indicated that AbLIFT-15 is highly specific to CA19-9 (AcSLe^a-glycan #83), GcSLe^a (glycan #86) and their corresponding 9-O-acetylated derivatives (glycan #87 and #88, respectively) (Figure 9(a)). The importance of the sialic acid and fucose residues for antibody recognition was also demonstrated, since Le^a (glycan #84) and Neu5Acα2 – 3Galβ1 – 3GlcNAcβProNH$_2$ (non-fucosylated-SLe^a, glycan #13) showed no binding at all. The specificity of AbLIFT-15 to sialylated glycans is also evident by a reduction in binding to the colorectal adenocarcinoma WiDr cell line following a sialidase treatment of the cells (Figure 9(b), Supplementary Figure 1(c)). Also, AbLIFT-15 exhibits a similar cell-recognition pattern to 1116NS19.9 as it robustly stains the BxPc3 cancer cell line but only marginally stains HEK293T control cells (Figure 9(c), Supplementary Figure 1(b)). Interestingly, AbLIFT-15 recognizes BxPc3 at higher efficiency compared with 1116NS19.9 (Figure 9(c)). Along these lines, a systematic staining of WiDr using a concentration series of both 1116NS19.9 and AbLIFT-15 clearly demonstrates superiority of AbLIFT-15 in detection efficiency (Figure 9(d), Supplementary Figure 1(c)). Overall, AbLIFT-15 exhibits augmented activity that retains the selectivity and specificity of 1116NS19.9.

Discussion

By contrast to proteins, glycans are more challenging target immunogens for the humoral immune system. The low immunogenicity of aberrant glycans is due in part to their potential flexibility and their highly hydrophilic nature.[39] For this reason, understanding the molecular details of how antibodies recognize glycans may provide important biophysical insights for designing nextgeneration diagnostics and therapeutics. Here, we provide structural information for the recognition of CA19-9 by two different mAbs. Both 1116NS19.9 and 5b1 bind CA19-9 using extensive polarinteraction networks which allow them to bind this small antigen with a K_D of 14.7 lM and 12.8 lM, respectively. Despite being isolated from different hosts (*i.e.*, mouse and human, respectively) and against different targets, 1116NS19.9[21,22] and 5b1[23] recognize an almost identical conformation of CA19-9. This observation strongly implies that the configuration of CA19-9, as determined here, represents a preferred conformation of this carbohydrate antigen. Comparison to a calculated energy profile of CA19-9[34] corroborated our notion that this conformation represents a low-energy state. These findings further suggest that the binding of antibodies to CA19-9 may be partially restricted in cases where the antigen cannot freely rotate with respect to the protein-surface that it modifies due to specific molecular interactions that it makes. In such scenarios, and since each antibody recognizes CA19-9 at a different angle, using a combination of antibodies like 1116NS19.9 and 5b1 may provide a more complete detection of CA19-9 than using each of the mAbs alone. This could be advantageous for either therapy or diagnostics.

Being relatively small and hydrophilic makes CA19-9 a challenging target for molecular binding and for conventional antibody-optimization strategies that target the

complementaritydetermining region for mutation. We previously used library screening to identify variants of 1116NS19.9 that can bind to CA19-9 more tightly.[18] In this current study, we extended our efforts and used structural data together with innovative computational design to produce AbLIFT-15. All the seven mutations that we have introduced to AbLIFT-15 are at the interface between the V_H and V_L and do not directly contact CA19-9. Despite the successful bottom line of the design procedure, we noted that most of the designs failed to bind the antigen, whereas in previous AbLIFT applications most of the designs were functional.[19] The lower success rate here may be ascribed to the following reasons: (i) The parental antibody's affinity is low and may be at the detection limit, increasing the likelihood that even antibodies that exhibit somewhat lower affinity are not detected; (ii) The glycan binding surface is geometrically very intricate and slight deviations in the packing at the light-heavy chain interface (induced by the designed mutations) may propagate to affect the orientation of critical determinants of molecular recognition. In this connection, the extreme accuracy of the AbLIFT-15 design as revealed by its crystallographic analysis is notable and suggests that the stringency may be quite high in this particular antigen; and (iii) We implemented a large number of designed mutations in this case (3–8 mutations per design) relative to previous implementations of this and similar algorithms (typically, up to 6),[20,40,41] and the large number of mutations may raise the likelihood of design inaccuracy. Nevertheless, given the unusual complexity of this design problem, the design of two substantially improved variants (of 17 tested) validates the use of AbLIFT to automatically, effectively and through a modest experimental effort enhance antibodies that have been mostly recalcitrant to conventional optimization approaches.

The shortcomings of CA19-9 as a marker for early diagnosis and screening for pancreatic cancer include false-positive results for patients with a benign pancreatic disease[11–13,42] and prompted the search of other biomarkers.[43] However, recent research demonstrates CA19-9-induced pancreatitis as a driving force for pancreatic cancer in a mouse model, justifying pancreatic-cancer monitoring for individuals with pancreatitis.[9] False-negative results are reported for 5–10% of the population.[42] In addition to the false positive and negative subpopulations, the sensitivity and specificity of the CA19-9 screening are not sufficient.[11–13,44] While false positive and negative results due to benign illness and genotype cannot be eliminated through the application of enhanced antibodies towards CA19-9, tighter binding may improve the positive predictive score for screening tests. Future research will be directed to studying whether the enhanced anti-CA19-9 antibodies provide a benefit in earlier and more accurate diagnosis for pancreatic cancer, either by detecting CA19-9 alone or in combination with other biomarkers.[44] In addition, stronger binding to cells that are decorated with CA19-9, as we demonstrated here for AbLIFT-15, could facilitate the use of anti-CA19-9 antibodies as immunotherapeutic agents.

Materials and methods

Cloning of antibodies into yeast surface display (YSD) system and functional assay

Sequences of the variable domains (V_H and V_L) of the anti-SLe[a] mouse antibody 1116NS19.921 and the human antibody 5b1[23] were used to design scFv of (N'-Vh) – (GGGGSGGGGSGGGGS linker)–(C'-V_L), and DNA fragments synthesized by Integrated

DNA Technologies Inc. (IDT, Israel). The scFv DNA sequence was optimized for codon usage compatible with expression in human cells, without altering the amino acid sequence. In addition, the scFv sequence was flanked by plasmid homology regions at the 5' and 3' ends (36 and 45 nucleotides, respectively). The flanking regions contained 5'-NdeI and 3'-BamHI restriction enzyme cloning site in-frame with the scFv. Then, EBY100 yeast cells were transformed with each synthesized scFv and *NdeI* BamHI digested plasmid for *in vivo* ligation, as described.[18] The resulting scFv contained N' HA and C' c-Myc tags (encoded in the plasmid) that allowed to monitor surface expression.

Induction of scFv expression on YSD system

To obtain scFv surface expression on yeast cells, 1116NS19.9-scFv-pETCON2 or 5b1-scFv-pETCON2 transfected yeast cells were cultured in SD-Trp a synthetic defined media (SD) lacking Tryptophan (Trp) [2% glucose (Sigma), 0.67% yeast nitrogen base w/o amino acids (BD), 0.54% Na_2HPO_4 (Sigma), 0.86% NaH2PO4 (Sigma) and 0.192% yeast synthetic drop-out medium supplements without Trp (Sigma)] at 30 °C, passaged 1:10 each day for three days, then scFv was expressed by changing the media to SG-Trp a synthetic galactose (SG) based media [2% galactose (Sigma), 0.2% glucose, 0.67% yeast nitrogen base w/o amino acids, 0.54% Na_2HPO_4, 0.86% NaH_2PO_4, and 0.192% yeast synthetic drop-out medium supplements without Trp] and the temperature to 20 °C, and cells were grown overnight to obtain scFv-YSD cells.

Assessment of scFv functional reactivity by Fluorescence-Activated Cell Sorting (FACS)

Induced scFv-YSD cells were functionally analyzed for antigen binding by FACS, as described,[18] with some modifications. Briefly, we used target antigens in a nanoparticle expression mode with multivalent expression on ~ 30 kDa polyacrylamide polymers carrying biotin tags (PAA-Bio; biotin ~ every 5th amide group with 7-9 glycans per particle). Thus, we used polyvalent SLe^a-PAA-Bio glycans nanoparticles. The non-specific target antigen Le^a-PAA-Bio was used as a negative control. 5×10^6 1116NS19.9-scFv-expressing yeast cells or 5b1-scFv-expressing yeast cells were washed with 1 ml assay buffer (PBS, 0.5% ovalbumin) then incubated with 1 μM SLe^a-PAA-biotin antigen and 1:50 diluted mouse-anti-c-Myc (4 μg/ml), both in 50 μl assay buffer, and for negative control, cells were incubated with 50 μl assay buffer, then all incubated for 1 h at room temperature (RT) with rotation. Cells were washed with 1 ml ice cold assay buffer, then incubated for 40 min on ice with APC-streptavidin and Alexa-Fluor-488-goat-anti-mouse IgG1 diluted 1:50 (10 μg/ml) and 1:200 (10 μg/ml) respectively in 50 μl assay buffer. Cells were washed with 1 ml ice cold PBS, then resuspended in 500 μl PBS. Cell fluorescence was measured by FACSort flow cytometry (Becton Dickinson) and analyzed with Kaluza analysis software. Double positive ($APC\text{-}Ag^+AF488\text{-}Ab^+$) yeast cells exemplified functional cloned scFv constructs (1116NS19.9-scFv and 5b1-scFv).

Cloning and expression of antibodies as IgGs

Cloning was done by Gibson assembly as described,[18] with some modifications. Variable heavy and light fragments of 1116NS19.9 or 5b1 were amplified by PCR. Reaction was made in Q5 reaction buffer, with 1 μl of plasmid DNA template (65–98 ng), 200 μM each dNTP, 1 U Q5 hot start high fidelity DNA polymerase (New England Biolabs), 500 nM each

primer (Supplementary Table 3 primers #1–4 for 1116NS19.9 or primers #5–8 for 5b1) complete volume to 50 µl with PCR grade water. PCR conditions were 95 °C for 2 min followed by 30 cycles of 95 °C for 30 s, 61 °C for 60 s, 72 °C for 60 s, and final incubation of 72 °C for 5 min. To remove template segments, the PCR product was supplemented with 6 il of 10 x CutSmart Buffer, 20 U DpnI (New England Biolabs), and completed the volume to 60 µl with PCR grade water, then incubated at 37 °C for 1 h. PCR digested fragments were purified from agarose gel by Zymoclean Gel DNA Recovery Kit (Zymo Research). Heavy and light chain full IgG p3BNC expression plasmids were divided to three parts for PCR amplification, variable region, left and right arms. Left and right arms of heavy and light p3BNC plasmids were amplified, digested and purified as described for variable regions using relevant primers (Supplementary Table 3; primers #9–14). Of each fragment, variable region, right and left arms, 25 ng were taken for Gibson assembly. Reaction was made in isothermal reaction buffer containing 5% PEG 8000, 100 mM Tris-HCl pH 7.5,10 mM $MgCl_2$, 10 mM DTT, 0.2 mM of each dNTP and 10 mM NAD. To this buffer we added 0.04 U T5 exonuclease (NEB), 0.25U Phusion polymerase (NEB) and 40 U Taq DNA ligase (NEB), and ligation was made at 50 °C for 1 h. Plasmids were electroporated into XL1 Escherichia. coli, to validate the sequence and producing high amount of p3BNC expression plasmids. Human embryonic kidney 293a cells were then used to produce full length whole Ab 1116NS19.9 or Ab 5b1 from their respective p3BNC expression plasmids that were transfected with polyethylenimine reagent (PEI; Polysciences).

Functionality of cloned secreted 1116NS19.9-hIgG1 and 5b1-hIgG1 antibodies was tested by ELISA against antigen-coated plates.

Antibodies specificity by ELISA

Specificity was examined by binding of full-length 1116NS19.9-hIgG1 and 5b1-hIgG1 antibodies to various glycans by ELISA assay. 96-wells plate was coated with SLe^a-PAA-Bio, Le^a-PAA-Bio, SLe^x-PAA-Bio or PAA (GlycoTech) in duplicates at 0.25 µg/well overnight at 4 °C. Wells were blocked with blocking buffer (PBS pH7.4, 1% ovalbumin) for 1 h at RT. Blocking buffer was removed and 1116NS19.9-hIgG1 or 5b1-hIgG1 antibodies at 10 µg/ml in blocking buffer was added at 100 µl/well, then incubated for two hours at RT. Plates were washed three times with PBST (PBS pH 7.4, 0.1% Tween), then incubated for 1 h at RT with HRP-goat-anti-human IgG 0.11 µg/ml in 100 µl PBS. After washing three times with PBST, wells were developed with 140 µl ofO-phenylenediamine in 100 mM citrate-PO_4 buffer, pH 5.5, and the reaction stopped with 40 µl of H_2SO_4 (4 M). Absorbance was measured at 490 nm on SpectraMax M3 (Molecular Devices). Specific binding was defined by subtracting the background readings obtained with the secondary antibody only.

Synthesis of $SLe^a\beta ProNH_2$

The tumor-associated carbohydrate antigen SLe^a in the form of $SLe^a\beta ProN_3$ [Neu5Acα2 – 3Galβ1 – 3(Fucα1 – 4)GlcNAcβO$(CH_2)3N_3$] was synthesized as previously described.[28] It was used to synthesize $SLe^a\beta ProNH_2$ by catalytic hydrogenation as described below. To a stirred solution of SLe^a-βProN3 (5 mg) in water-methanol solution (2.1 ml, 1:2 by volume) in a round bottom flask (50 ml), 10% pallidum on charcoal Pd/C (2 mg) was added. The mixture was stirred under a hydrogen environment for 2 h. The solution was then passed

through a filter to remove the catalyst. The solvent was removed under vacuum and the residue was dissolved in 0.5 ml of deionized water, frozen, and lyophilized to produce SLe^aProNH_2 as a white powder.

Large-scale protein expression and purification

To produce large amount of recombinant Abs for crystallization, thermostability assays, and FACS analyses we transfected HEK293F cells maintained in FreeSyle medium (Gibco) with the p3BNC plasmids encoding the heavy and the light chains of Abs. As a transfection reagent, we used 40 kDa polyethyleneimine (PEI) (Polysciences) with a DNA / polyethyleneimine ratio of 1 μg/3 μl with a total of 1 mg DNA per 1 L of cells at 1 M cells/ml. Cells were maintained for 5–7 days in suspension before harvesting the supernatants. After clarifying the supernatants by centrifugation, Abs were captured using protein-A affinity chromatography (GE Lifesciences). Abs were eluted using 0.1 M citric acid pH 3.0 buffer, which was subsequently adjusted to pH 8.0 using Tris-HCl. For obtaining the Fab portions, papain enzyme (Sigma-Aldrich) was used to digest Abs in enzyme and protein ratio being ~ 1:80. Cutting buffer contained 20 mM Cysteine-HCl (Sigma-Aldrich) and 10 mM EDTA tittered to pH 7.0 with Tris. Cutting was performed for 90 minutes in 37 °C. Negative protein-A was performed to remove Fc fragments followed by SEC on a Superdex200 10/300 column (GE Lifesciences) in TBS buffer.

Crystallization

For protein crystallization, we used a mosquito crystallization robot (TTP Labtech) to set vapor diffusion in sitting drop experiments using 96-well iQ plates (TTP Labtech). At each well, we tested three ratios of protein (80, 120, and 160 nl) to reservoir (120 nl). PEGrx-HT screen (Hampton Research) was used to identify initial hits for apo-Ab 1116NS19.9 for the condition containing 0.10% w/v n-octyl-β-D-glucoside, 0.1 M sodium citrate tribasic dihydrate pH 5.5, and 22% w/v polyethylene glycol 3,350. Further optimization was done by growing the crystals in 7.5% ethylene glycol for cryopreservation. Protein with ligand CA19–9 was mixed in a ratio of 1:1.2 protein to CA19–9, protein samples gave crystals when grown in 24-well sitting plates with a 1:1 ratio of protein to reservoir. Ab 5b1 apo- and holo-protein crystals were crystallized using the same reservoir conditions containing 0.1 M NaCl, 0.1 M bis-tris propane pH 9.0, 18% polyethylene glycol 1,500 and 5% glycerol. The protein to CA19–9 ratio was 1:1, and the protein to reservoir ratio was 1.75:1. All crystals were grown at 20 °C.

Data collection, structure solution and refinement

X-ray diffraction data were collected at the European Synchrotron Radiation Facility (ESRF) using a Pilatus 6 M detector at 100 K. Data up to 1.5 A at beamline ID23–1 was collected for the apo and holo Fab 1116NS19.9 belonging to the tetragonal and orthorhombic space groups, respectively. Data were indexed, integrated, and scaled using XDS.[45] We used Phaser[46] to obtain a molecular replacement solution with the structure of NIH45–46 (PDB: 3u7w) and used the solved structure for molecular replacement of Fab 5b1 holo- and apo-proteins. Data for holo- and apoAb 5b1 were collected at beamline ID23–2 using a Pilatus 2 M detector at resolutions of 2.4 Å for the apo and holo structures, which belong to the hexagonal space group with 3 and 6 molecules in the asymmetric unit,

respectively. All models were manually traced into electron density maps using Coot[47] and refined using Phenix Refine[48] in an iterative fashion.

Surface Plasmon Resonance (SPR) measurements

All measurements were performed using a Biacore T200 (GE Healthcare) at 25 °C. Abs were immobilized to a protein-A chip from a stock of 20 μg/ml at similar surface densities. CA19–9 glycan in TBS buffer with azide 0.02% was used as an analyte in a series of increasing concentrations ranging from 0.488 μM to 125 μM using five steps for preliminary screening and ranging from 0 to 500 μM with 14 steps total for in-depth analyses. Affinity constants were calculated by measuring binding levels at steady states and fitting binding curves to the data using the Biacore T200 evaluation software. We used washing with binding buffer without a regeneration step for achieving baseline signal after each injection. The protein-A sensor chip was eventually regenerated using 10 mM glycine-HCl, pH 1.5.

Nano Differential Scanning Fluorimetry (nDSF)

All experiments were preformed Using the Promethus NT.48 device (TEMPER technologies) in duplicate samples of 0.2 mg/ml Fab 1116NS19.9 and Fab AbLIFT-15 with or without 100 μM CA19–9 were measured in increments of 1 °C/minute starting from 35 °C. Data analysis was performed with the PR. ThermControl v2.1.1 software.

Structure analysis and representation

Analyses and structural figures were generated using PyMol. Buried surface area calculations were performed using the AreaMol tool in the CCP4 program suite.[49]

AbLIFT design

AbLIFT was applied essentially as described.[19] Briefly, starting from the structure of 1116NS19.9, we manually selected eight positions at the interface between the variable light and heavy chains for design calculations. A multiple sequence alignment of all homologs was obtained using default parameters and a position-specific scoring matrix was computed using PSI-BLAST.[50] At each position, mutations that exhibited a PSSM score < -1 were eliminated from the design sequence choices. Next, each remaining mutation was modeled using the Rosetta atomistic modelling and design pack-age51 and relaxed using the talaris14 energy func-tion,[52] which is dominated by van der Waals contacts, hydrogen bonding and solvation. Mutations that exhibited a total energy > +1 Rosetta energy units higher than the relaxed structure of the parental antibody were further eliminated. With the remaining identities, 130,111 combinations of mutations at 8 positions that were at least 3 mutations different from the parental antibody were modeled and relaxed in Rosetta. The mutants were then ranked according to their all-atom energy, clustered by requiring that each multipoint mutant exhibit at least 3 mutations relative to any other, and the top 17 designs were selected for experimental analysis.

Mutagenesis

Specific point mutations were introduced using the Quikchange II protocol (Agilent).

Europe PMC Funders Author Manuscripts

Sialoglycan microarray nanoprinting

Arrays were fabricated with NanoPrint LM-60 Microarray Printer (Arrayit) on epoxide-derivatized slides (Corning 40044) with 16 sub-array blocks on each slide. Glycoconjugates were distributed into one 384-well source plates using 4 replicate wells per sample and 8 μl per well (Versions 13.1). Each glycoconjugate (Supplementary Table 2) was prepared at 100 μM in an optimized print buffer (300 mM phosphate buffer, pH 8.4). To monitor printing quality, replicate-wells of human IgG (80, 40, 20, 10, 5, 0.25 ng/μl in PBS + 10% glycerol) and AlexaFlour-555-Hydraside (Invitrogen A20501MP, at 1 ng/μl in 178 mM phosphate buffer, pH 5.5) were used for each printing run. The arrays were printed with four 946MP3 pins (5 μm tip, 0.25 μl sample channel, ~100 μm spot diameter; Arrayit). Each block (sub-array) has 20 spots/row, 20 columns with spot to spot spacing of 275 μm. The humidity level in the arraying chamber was maintained at about 70% during printing. Printed slides were left on arrayer deck over-night, allowing humidity to drop to ambient levels (40–45%). Next, slides were packed, vacuum-sealed and stored at RT until used.

Sialoglycan microarray binding assay

Slides were developed and analyzed as previously described,[53] with some modifications. Slides were rehydrated with dH_2O and incubated for 30 min in a staining dish with 50 °C prewarmed ethanolamine (0.05 M) in Tris-HCl (0.1 M, pH 9.0) to block the remaining reactive epoxy groups on the slide surface, then washed with 50 °C pre-warmed dH_2O. Slides were centrifuged at 200 x g for 5 min then fitted with ProPlate™ Multi-Array 16-well slide module (Invitrogen) to divide into the sub-arrays (blocks). Slides were washed with PBST (0.1% Tween 20), aspirated, and blocked with 200 μL/sub-array of blocking buffer (PBS/OVA, 1% w/v ovalbumin, in PBS, pH 7.3) for 1 h at RT with gentle shaking. Next, the blocking solution was aspirated and 100 μL/block of purified antibodies in 1.92×10^{-2} – 4.81×10^{-3} μg/mL diluted in PBS/OVA were incubated with gentle shaking for 2 h at RT. Slides were washed four times with PBST, then with PBS for 2 min. Bound antibodies were detected by incubating with secondary detection diluted in PBS, 200 μL/block at RT for 1 h, Cy3-anti-human IgG 0.4 μg/mL (Jackson ImmunoResearch). Slides were washed four times with PBST then with PBS for 10 min followed by removal from ProPlate™ Multi-Array slide module and immediately dipping in a staining dish with dH2O for 10 min with shaking, then centrifuged at 200 x g for 5 min. Dry slides immediately scanned.

Array slide processing

Processed slides were scanned and analyzed as described[53] at 10 μm resolution with a GenePix 4000B microarray scanner (Molecular Devices) using 350 gain. Image analysis was carried out with GenePix Pro 6.0 analysis software (Molecular Devices). Spots were defined as circular features with a variable radius as determined by the GenePix scanning software. Local background subtraction was performed.

Flow cytometry analyses with cancer cells

WiDr cells were obtained from American Type Culture collection (ATCC), cells were grown in Dulbecco's Modified Eagle Medium (DMEM; biological industries) supplemented with 10% heat inactivated fetal bovine serum (FBS), 2 mM L-glutamine, 100 units/ml penicillin

and 0.1 mg/ml streptomycin. For binding assays, cells were collected from plates using 10 mM EDTA. 4×10^5 cells were incubated with 10–0.156 μg/ml of 1116NS19.9 or AbLIFT-15 IgGs diluted in FACS buffer (PBS with 0.5% fish gelatin) for 1 h on ice, followed by incubation with Cy3-AffiniPure goat-anti-human IgG diluted 1:100 (15 μg/ml) in FACS buffer for 40 min on ice. Fluorescence was measured by CytoFLEX flow cytometry. To confirm cancer cells binding specificity, sialidase FACS assay was performed, in which WiDr cells were collected from plates using 10 mM EDTA. 2×10^5 cells were divided into Eppendorf tubes and incubated for four hours at 37 °C with either PBS, 100 mU active *Arthrobacter ureafaciens S al dase* (AUS) (EY Laboratories, San Mateo, CA, USA) or 100 mU inactive AUS (pre-incubated in 90 °C for 30 min) in PBS. Then, cells were washed with FACS buffer, stained with 1 μg/ml AbLIFT-15 full-length IgG antibodies, followed by washing, secondary antibody labeling and fluorescence measurement, as described above.

BxPc3 cells were grown in RPMI medium (biological industries) and HEK293T cells in DMEM (Gibco). HEK293T and Bxpc3 cells were seeded on 10-cm plates and detached 4 d later using trypsin. The cells were washed by centrifugation at 400xg for 5 min and resuspension in PBS supplemented with 0.5 % BSA (Sigma). Cells were aliquoted and incubated for 1 h with different concentrations of the antibodies diluted in PBS with 0.5 % BSA, washed, and incubated with a 1:500 dilution of goat anti-human Cy3 conjugated secondary antibody (Jackson) for 30 min. Secondary antibody stained cells were used as a negative control. Analyses were performed using an LSR II flow cytometer (BD Biosciences) and FlowJo cell analysis software (FlowJo, LLC, Ashland, Ore.).

Accession numbers

Coordinates and structure functions were deposited in the protein data bank under the accession codes that are listed in Table 1.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We are grateful to Royant Antoine and Zubieta Chloe at the ESRF, Grenoble, France, for assisting in using beamlines ID23–1 and ID23–2. We thank Richard Bryce and Irfan Alibay from University of Manchester, U.K., for assisting in the interpretation of the energy profile of CA19–9. We thank Irit Sagi from the Weizmann Institute of Science, Israel, for providing the BxPc3 cell line. The Diskin lab is supported by research grants from the Ernst I Ascher foundation, Ben B. and Joyce E. Eisenberg Foundation, Estate of Emile Mimran, Jeanne and Joseph Nissim Center for Life Sciences Research, Dov and Ziva Rabinovich Endowed Fund for Structural Biology, Donald Rivin, Stanley and Tanya Rossby Endowment Fund, Natan Sharansky, Dr. Barry Sherman Institute for Medicinal Chemistry, as well as from the Israel Science Foundation (grants No. 3147/19, 209/20, & 682/16). Research in the Fleishman lab was supported by a Consolidator Award from the European Research Council (815379) and by a charitable donation in memory of Sam Switzer. Research in the Padler-Karavani lab was funded by the European Union H2020 Program grants (ERC-2016-STG-716220), UC Davis – Israel in Research Proposal award UC Davis (to V.P-K and X.C.).

References

1. Kintzing JR, Filsinger Interrante MV, Cochran JR. Emerging strategies for developing next-generation protein therapeutics for cancer treatment. Trends Pharmacol Sci. 2016; doi: 10.1016/j.tips.2016.10.005
2. Stowell SR, Ju T, Cummings RD. Protein glycosylation in cancer. Annu Rev Pathol Mech Dis. 2015; doi: 10.1146/annurev-pathol-012414-040438
3. Varki, A; Kannagi, R; Toole, B; Stanley, P. Glycosylation Changes in Cancer. 2015.
4. Padler-Karavani V. Aiming at the sweet side of cancer: Aberrant glycosylation as possible target for personalized-medicine. Cancer Lett. 2014; 352:102–112. DOI: 10.1016/j.canlet.2013.10.005 [PubMed: 24141190]
5. Pinho SS, Reis CA. Glycosylation in cancer: mechanisms and clinical implications. Nat Rev Cancer. 2015; 15:540–555. DOI: 10.1038/nrc3982 [PubMed: 26289314]
6. Boligan KF, Mesa C, Fernandez LE, Von Gunten S. Cancer intelligence acquired (CIA): Tumor glycosylation and sialylation codes dismantling antitumor defense. Cell Mol Life Sci. 2015; doi: 10.1007/s00018-014-1799-5
7. Kannagi R. Carbohydrate antigen sialyl Lewis a - Its pathophysiological significance and induction mechanism in cancer progression. Chang Gung Med J. 2007
8. Amon R, Reuven EM, Leviatan Ben-Arye S, Padler-Karavani V. Glycans in immune recognition and response. Carbohydr Res. 2014; doi: 10.1016/j.carres.2014.02.004
9. Engle DD, Tiriac H, Rivera KD, Pommier A, Whalen S, Oni TE, Alagesan B, Lee EJ, Yao MA, Lucito MS, Spielman B, et al. The glycan CA19-9 promotes pancreatitis and pancreatic cancer in mice. Science (80-). 2019; 364:1156–1162. DOI: 10.1126/science.aaw3145
10. Ugorski M, Laskowska A. Sialyl Lewis a: A tumor-associated carbohydrate antigen involved in adhesion and metastatic potential of cancer cells. Acta Biochim Pol. 2002; 49:303–311. [PubMed: 12362971]
11. Ballehaninna UK, Chamberlain RS. The clinical utility of serum CA 19–9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: An evidence based appraisal. J Gastrointest Oncol. 2012; 3:105–119. DOI: 10.3978/jjssn.2078-6891.2011.021 [PubMed: 22811878]
12. Huang Z, Liu F. Diagnostic value of serum carbohydrate antigen 19-9 in pancreatic cancer: a metaanalysis. Tumor Biol. 2014; doi: 10.1007/s13277-014-1995-9
13. Passerini R, Cassatella MC, Boveri S, Salvatici M, Radice D, Zorzino L, Galli C, Sandri MT. The pitfalls of CA19-9: Routine testing and comparison of two automated immunoassays in a reference oncology center. Am J Clin Pathol. 2012; doi: 10.1309/AJCPOPNPLLCYR07H
14. Bussom S, Saif MW. Methods and rationale for the early detection of pancreatic cancer. J Pancreas. 2010
15. Manimala JC, Roach TA, Li Z, Gildersleeve JC. High-throughput carbohydrate microarray profiling of 27 antibodies demonstrates widespread specificity problems. Glycobiology. 2007; doi: 10.1093/glycob/cwm047
16. Sterner E, Flanagan N, Gildersleeve JC. Perspectives on anti-glycan antibodies gleaned from development of a community resource database. ACS Chem Biol. 2016; doi: 10.1021/acschembio.6b00244
17. Amon R, Grant OC, Leviatan Ben-Arye S, Makeneni S, Nivedha AK, Marshanski T, Norn C, Yu H, Glushka JN, Fleishman SJ, Chen X, et al. A combined computational-experimental approach to define the structural origin of antibody recognition of sialyl-Tn, a tumor-associated carbohydrate antigen. Sci Rep. 2018; doi: 10.1038/s41598-018-29209-9
18. Amon R, Rosenfeld R, Perlmutter S, Grant OC, Yehuda S, Borenstein-Katz A, Alcalay R, Marshanski T, Yu H, Diskin R, Woods RJ, et al. Directed evolution of therapeutic antibodies targeting glycosylation in cancer. Cancers (Basel). 2020; doi: 10.3390/cancers12102824
19. Warszawski S, Katz AB, Lipsh R, Khmelnitsky L, Ben Nissan G, Javitt G, Dym O, Unger T, Knop O, Albeck S, Diskin R, et al. Optimizing antibody affinity and stability by the automated design of the variable light-heavy chain interfaces. PLoS Comput Biol. 2019; doi: 10.1371/journal.pcbi.1007207

Europe PMC Funders Author Manuscripts

20. VanDrisse CM, Lipsh-Sokolik R, Khersonsky O, Fleishman SJ, Newman DK. Computationally designed pyocyanin demethylase acts synergistically with tobramycin to kill recalcitrant Pseudomonas aeruginosa biofilms. Proc Natl Acad Sci U S A. 2021; 118 doi: 10.1073/pnas.2022012118
21. Koprowski H, Steplewski Z, Mitchell K, Herlyn M, Herlyn D, Fuhrer P. Colorectal carcinoma antigens detected by hybridoma antibodies. Somatic Cell Genet. 1979; 5:957–971. DOI: 10.1007/BF01542654 [PubMed: 94699]
22. Passerini R, Riggio D, Salvatici M, Zorzino L, Radice D, Sandri MT. Interchangeability of measurements of CA 19–9 in serum with four frequently used assays: An update. Clin Chem Lab Med. 2007; 45:100–104. DOI: 10.1515/CCLM.2007.003 [PubMed: 17243924]
23. Sawada R, Sun SM, Wu X, Hong F, Ragupathi G, Livingston PO, Scholz WW. Human monoclonal antibodies to sialyl-Lewis a (CA19.9) with potent CDC, ADCC, and antitumor activity. Clin Cancer Res. 2011; 17:1024–1032. DOI: 10.1158/1078-0432.CCR-10-2640 [PubMed: 21343375]
24. Houghton JL, Abdel-Atti D, Scholz WW, Lewis JS. Preloading with Unlabeled CA19.9 Targeted Human Monoclonal Antibody Leads to Improved PET Imaging with 89Zr-5B1. Mol Pharm. 2017; 14:908–915. DOI: 10.1021/acs.molpharmaceut.6b01130 [PubMed: 28191976]
25. Bashir S, Leviatan Ben Arye S, Reuven EM, Yu H, Costa C, Galinanes M, Bottio T, Chen X, Padler-Karavani V. Presentation Mode of Glycans Affect Recognition of Human Serum anti-Neu5Gc IgG Antibodies. Bioconjug Chem. 2019; 30:161–168. DOI: 10.1021/acs.bioconjchem.8b00817 [PubMed: 30500162]
26. Padler-Karavani V, Song X, Yu H, Hurtado-Ziola N, Huang S, Muthana S, Chokhawala HA, Cheng J, Verhagen A, Langereis MA, Kleene R, et al. Cross-comparison of protein recognition of sialic acid diversity on two novel sialoglycan microarrays. J Biol Chem. 2012; 287:22593–22608. DOI: 10.1074/jbc.M112.359323 [PubMed: 22549775]
27. Chao G, Lau WL, Hackel BJ, Sazinsky SL, Lippow SM, Wittrup KD. Isolating and engineering human antibodies using yeast surface display. Nat Protoc. 2006; doi: 10.1038/nprot.2006.94
28. Tasnima N, Yu H, Yan X, Li W, Xiao A, Chen X. Facile chemoenzymatic synthesis of Lewis a (Lea) antigen in gram-scale and sialyl Lewis a (sLea) antigens containing diverse sialic acid forms. Carbohydr Res. 2019; doi: 10.1016/j.carres.2018.12.004
29. Re S, Nishima W, Miyashita N, Sugita Y. Conformational flexibility of N-glycans in solution studied by REMD simulations. Biophys Rev. 2012; 4:179–187. DOI: 10.1007/s12551-012-0090-y [PubMed: 28510079]
30. Sommer R, Wagner S, Varrot A, Nycholat CM, Khaledi A, Häussler S, Paulson JC, Imberty A, Titz A. The virulence factor LecB varies in clinical isolates: Consequences for ligand binding and drug discovery. Chem Sci. 2016; 7:4990–5001. DOI: 10.1039/c6sc00696e [PubMed: 30155149]
31. Shanker S, Czako R, Sankaran B, Atmar RL, Estes MK, Prasad BVV. Structural analysis of determinants of histo-blood group antigen binding specificity in genogroup I noroviruses. J Virol. 2014; 88:6168–6180. DOI: 10.1128/jvi.00201-14 [PubMed: 24648450]
32. Kubota T, Kumagai A, Ito H, Furukawa S, Someya Y, Takeda N, Ishii K, Wakita T, Narimatsu H, Shirato H. Structural basis for the recognition of lewis antigens by genogroup I norovirus. J Virol. 2012; 86:11138–11150. DOI: 10.1128/jvi.00278-12 [PubMed: 22855491]
33. Perret S, Sabin C, Dumon C, Pokorná M, Gautier C, Galanina O, Ilia S, Bovin N, Nicaise M, Desmadril M, Gilboa-Garber N, et al. Structural basis for the interaction between human milk oligosaccharides and the bacterial lectin PA-IIL of Pseudomonas aeruginosa. Biochem J. 2005; 389:325–332. DOI: 10.1042/BJ20050079 [PubMed: 15790314]
34. Alibay I, Burusco KK, Bruce NJ, Bryce RA. Identification of rare lewis oligosaccharide conformers in aqueous solution using enhanced sampling molecular dynamics. J Phys Chem B. 2018; 122:2462–2474. DOI: 10.1021/acs.jpcb.7b09841 [PubMed: 29419301]
35. Naim JO, van Oss CJ. The effect of hydrophilicity-hydrophobicity and solubility on the immunogenicity of some natural and synthetic polymers. Immunol Invest. 1992; 21:649–662. DOI: 10.3109/08820139209069401 [PubMed: 1487323]
36. Mäkelä O, Péterfy F, Outschoorn IG, Richter AW, Seppälä I. Immunogenic properties of α(1 ⟶ 6) dextran, its protein conjugates, and conjugates of its breakdown products in mice. Scand J Immunol. 1984; 19:541–550. DOI: 10.1111/j.1365-3083.1984.tb00965.x [PubMed: 6204375]

37. Kasper DL, Paoletti LC, Wessels MR, Guttormsen HK, Carey VJ, Jennings HJ, Baker CJ. Immune response to type III group B streptococcal polysaccharide-tetanus toxoid conjugate vaccine. J Clin Invest. 1996; 98:2308–2314. DOI: 10.1172/JCI119042 [PubMed: 8941648]
38. Lefranc, M-P; Lefranc, G. [accessed April 25, 2021] The immunoglobulin factsbook. 2001. https://www.elsevier.com/books/the-immunoglobulin-factsbook/lefranc/978-0-12-441351-1
39. Guo Z, Wang Q. Recent development in carbohydrate-based cancer vaccines. Curr Opin Chem Biol. 2009; 13:608–617. DOI: 10.1016/j.cbpa.2009.08.010 [PubMed: 19766052]
40. Netzer R, Listov D, Lipsh R, Dym O, Albeck S, Knop O, Kleanthous C, Fleishman SJ. Ultrahigh specificity in a network of computationally designed protein-interaction pairs. Nat Commun. 2018; 9 doi: 10.1038/s41467-018-07722-9
41. Khersonsky O, Lipsh R, Avizemer Z, Ashani Y, Goldsmith M, Leader H, Dym O, Rogotner S, Trudeau DL, Prilusky J, Amengual-Rigo P, et al. Automated design of efficient and functionally diverse enzyme repertoires. Mol Cell. 2018; 72:178–186. e5 doi: 10.1016/j.molcel.2018.08.033 [PubMed: 30270109]
42. Duffy MJ, Sturgeon C, Lamerz R, Haglund C, Holubec VL, Klapdor R, Nicolini A, Topolcan O, Heinemann V. Tumor markers in pancreatic cancer: A European Group on Tumor Markers (EGTM) status report. Ann Oncol. 2009; 21:441–447. DOI: 10.1093/annonc/mdp332 [PubMed: 19690057]
43. Staal B, Liu Y, Barnett D, Hsueh P, He Z, Gao CF, Partyka K, Hurd MW, Singhi AD, Drake RR, Huang Y, et al. The Stra plasma biomarker: Blinded validation of improved accuracy over CA19-9 in pancreatic cancer diagnosis. Clin Cancer Res. 2019; doi: 10.1158/1078-0432.CCR-18-3310
44. Dong D, Jia L, Zhang L, Ma N, Zhang A, Zhou Y, Ren L. Periostin and CA242 as potential diagnostic serum biomarkers complementing CA19.9 in detecting pancreatic cancer. Cancer Sci. 2018; 109:2841–2851. DOI: 10.1111/cas.13712 [PubMed: 29945294]
45. Kabsch W. Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr Sect D. 2010; 66:133–144. DOI: 10.1107/S0907444909047374 [PubMed: 20124693]
46. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr. 2007; 40:658–674. DOI: 10.1107/S0021889807021206 [PubMed: 19461840]
47. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010; 66:486–501. DOI: 10.1107/S0907444910007493 [PubMed: 20383002]
48. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010; 66:213–221. DOI: 10.1107/S0907444909052925 [PubMed: 20124702]
49. C.C. Project. The CCP4 suite: Programs for protein crystallography. Acta Crystallogr D Biol Crystallogr. 1994; 50:760–763. DOI: 10.1107/S0907444994003112 [PubMed: 15299374]
50. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. 1997; doi: 10.1093/nar/25.17.3389
51. Leaver-Fay A, Tyka M, Lewis SM, Lange OF, Thompson J, Jacak R, Kaufman K, Renfrew PD, Smith CA, Sheffler W, Davis IW, et al. Rosetta3: An object-oriented software suite for the simulation and design of macromolecules. Methods Enzymol. 2011; doi: 10.1016/B978-0-12-381270-4.00019-6
52. O'Meara MJ, Leaver-Fay A, Tyka MD, Stein A, Houlihan K, Dimaio F, Bradley P, Kortemme T, Baker D, Snoeyink J, Kuhlman B. Combined covalent-electrostatic model of hydrogen bonding improves structure prediction with Rosetta. J Chem Theory Comput. 2015; doi: 10.1021/ct500864r
53. Leviatan Ben-Arye S, Yu H, Chen X, Padler-Karavani V. Profiling anti-Neu5Gc IgG in human sera with a sialoglycan microarray assay. J Vis Exp. 2017; doi: 10.3791/56094
54. Wallace AC, Laskowski RA, Thornton JM. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions The LIGPLOT program automatically generates schematic 2-D representations of protein-ligand complexes from standard Protein Data Bank file input. Protein Eng. 1995

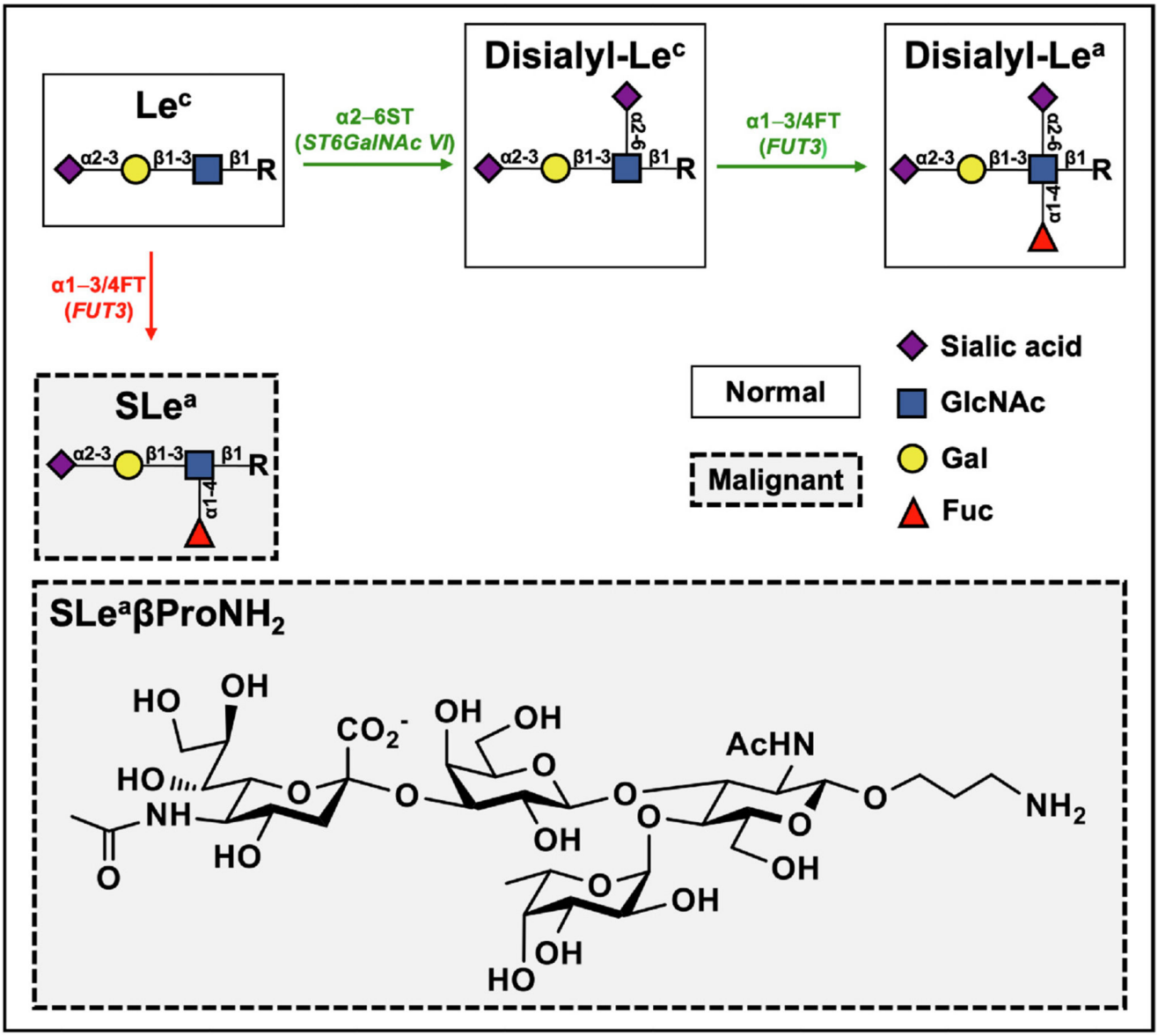

Figure 1. Biosynthetic pathway of SLe^a and Disialyl-Le^a.

SLe^a (CA19-9) is a Type-1 tetrasaccharide tumor-associated carbohydrate antigen composed of fucose (Fuc), N-acetylglucosamine (GlcNAc), galactose (Gal), and sialic acid (Sia). In the normal biosynthetic pathway, the precursor Le^c is commonly further elongated by α2 – 6-sialyltransferase and α1 – 3/4-fucosyltransferase to generate disialyl-Le^a, which has an additional sialic acid moiety compared to SLe^a. The SLe^{a}βProNH$_2$ probe, Neu5Acα2 – 3Galβ1 – 3(Fucα1 – 4)GlcNAcβO(CH$_2$)3NH$_2$, is a SLe^a antigen with a linker containing a terminal primary amine which can be used for conjugation for functional studies.

Europe PMC Funders Author Manuscripts

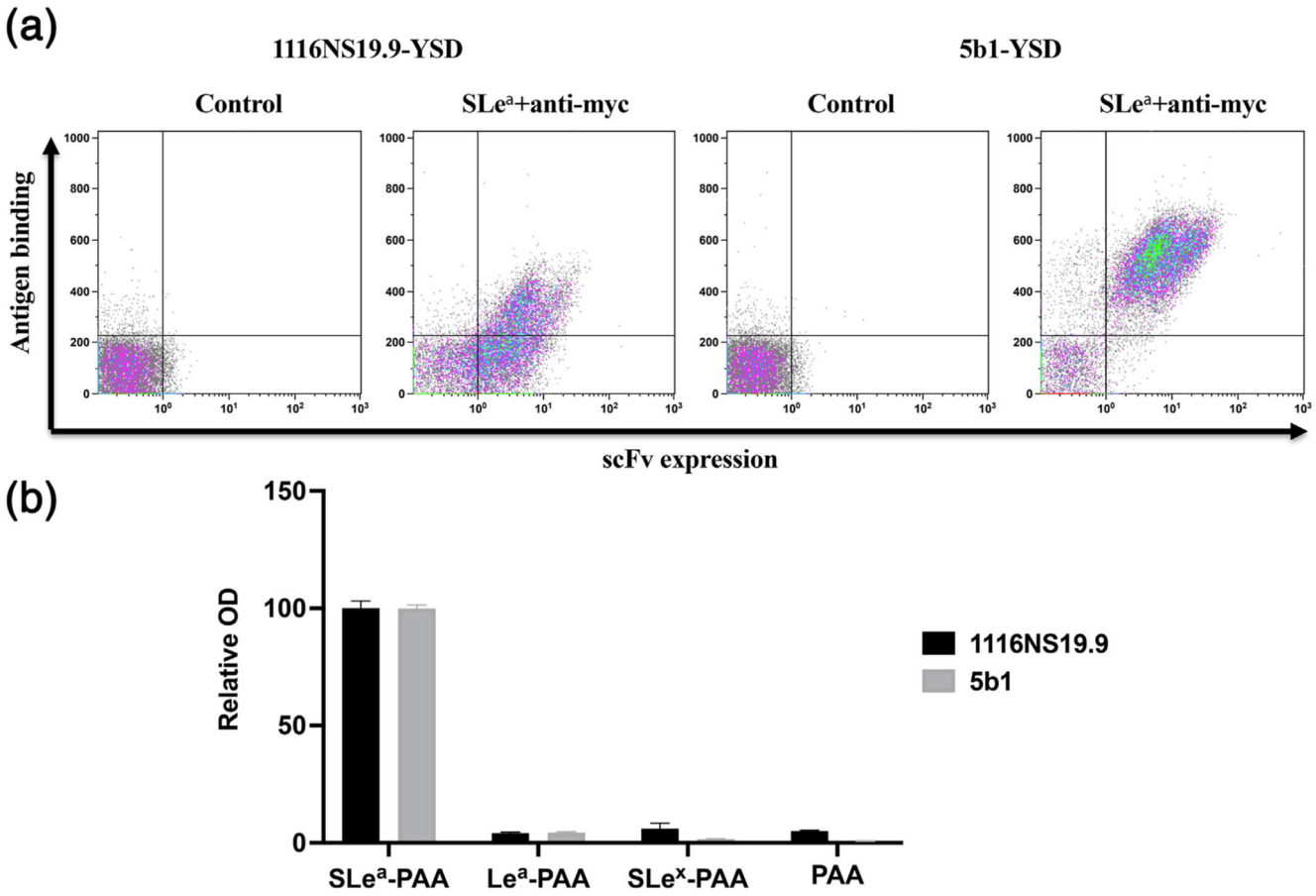

Figure 2. Expression of 1116NS19.9 and 5b1 antibodies and their antigen recognition.

(a) FACS analyses of yeast cells expressing either 1116NS19.9, 5b1, or an empty vector as negative controls. The x-axis indicates the surface display levels (by detecting C'-c-Myc tag) and the y-axis indicates antigen binding (using 1 μm SLe^a-PAA-Bio – multivalent SLe^a-polyacrylamide polymers tagged with biotin). These are representative plots from at least two independent repeats. (b) Specificity of antigen binding. Binding of 1116NS19.9 and 5b1 formatted as full-length IgG at 10 mg/ml to a solid surface coated with SLe^a-PAA-Bio, Le^a-PAA-Bio, SLe^x-PAA-Bio, and PAA was evaluated by ELISA. Relative optical density (OD) was calculated as percentage of maximal binding of each antibody (mean ± SEM; Representative of at least two independent repeats).

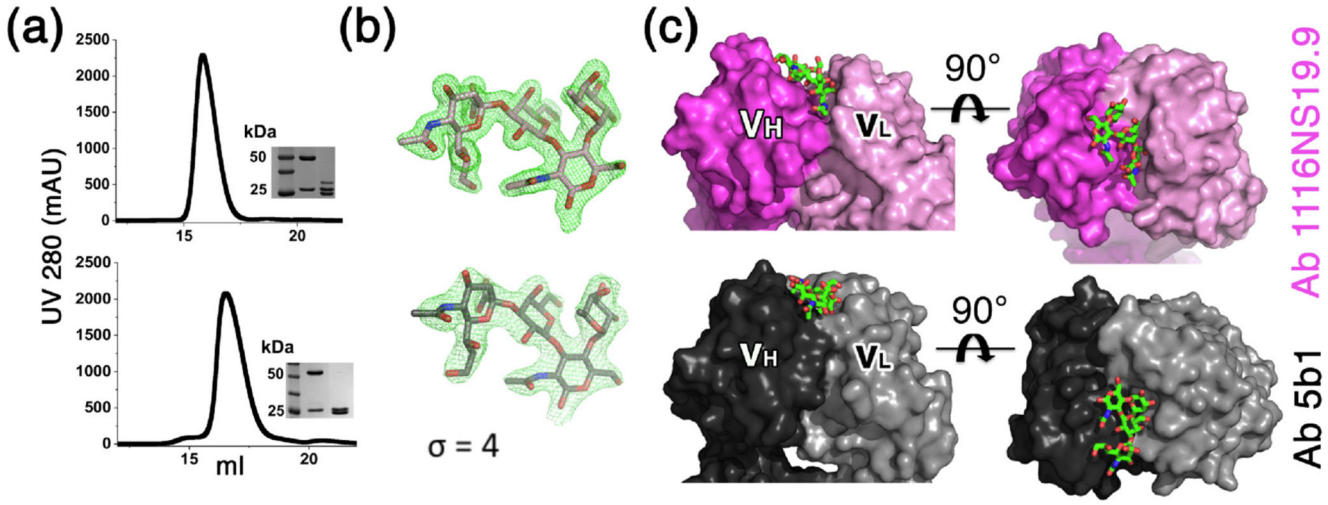

Figure 3. Holo-structures of 1116NS19.9 and 5b1 Fabs.

(a) Production and purification of Fab fragments. Chromatograms for the Fab fragments of 1116NS19.9 (top) and 5b1 (bottom) indicate their migration profile in the size-exclusion chromatography column. Images of Coomassie-stained SDS-PAGE are shown in the insets, loaded with a size marker, the IgG fraction, and the digested Fab fraction (left to right). (b) CA19-9 in Fo-Fc (difference) omit maps. CA19-9 structures from 1116NS19.9 (top) and from 5b1 (bottom) are shown in Fo-Fc maps (green mesh) at σ = 4. These difference maps were calculated from the final models after omitting the CA19-9 molecules. (c) The binding pockets for CA19-9. The 1116NS19.9 and 5b1 Fabs are shown using surface representations. The chains are colored in dark pink and light pink for the heavy and light chains of 1116NS19.9, respectively, and in dark gray and light gray for the heavy and light chain of 5b1. The CA19-9 molecules are shown using stick representation.

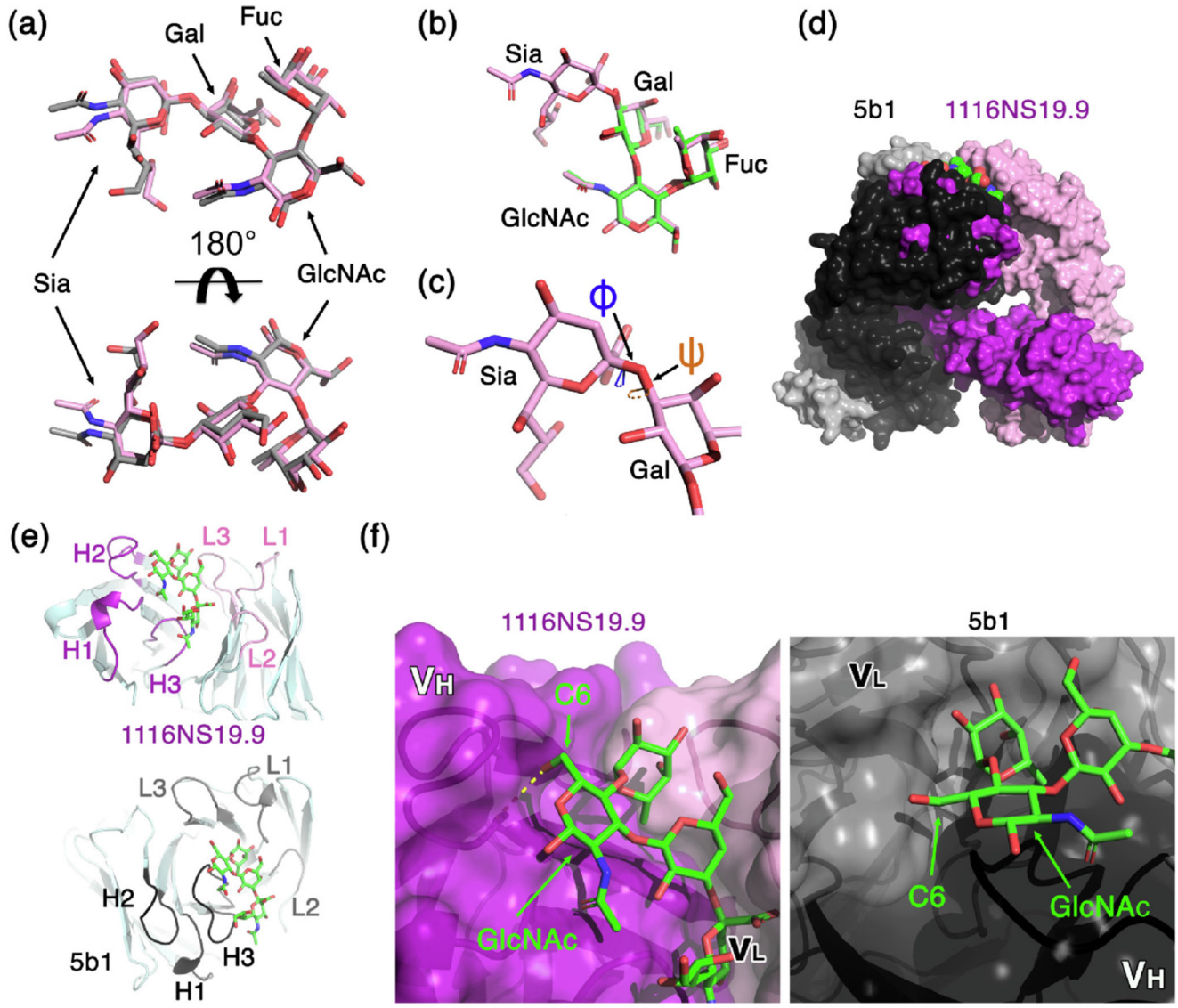

Figure 4. Molecular recognition of a similar extended CA19-9 conformer.

(a) CA19-9 assumes a simh a dashed yellow lineilar extended conformation in both structures. Superimposition of CA19-9 from the 1116NS19.9 structure (pink) and of CA19-9 from the 5b1 structure (gray) is shown in two different orientations, as indicated. The saccharides labeled are L-fucose (Fuc), N-acetylglucosamine (GlcNAc), galactose (Gal), and sialic acid (Sia) (b) The Le^a core of CA19-9 (pink) assumes a similar conformation to isolated Lewis antigen. A representative Lewis a (PDB-ID: 1W8H, green) superimposed with the core of CA19-9 (all-atom RMSD = 0.64 Å) (c) The / (defined by O6 – C2 – O3 – C3) and W (defined by C2 – O3 – C3 – C2) angles of the Sia a(2-3) Gal bond are indicated. (d) The relative binding orientations of 1116NS19.9 and 5b1 to CA19-9. The two structures are superimposed based on the CA19-9 antigen. 5b1 and 1116NS19.9 are shown in gray and pink surfaces, using dark and light colors for the heavy and light chains, respectively. (e) The CDRs' role in binding CA19-9. Ribbon diagrams of 1116NS19.9 (top) and 5b1 (bottom)

illustrate the organization of the CDRs around CA19-9, which is shown in the same orientation in both images. CDRs are highlighted and labeled. (f) Molecular basis for CA19-9 selectivity. Surface representations of 1116NS19.9 (left) and 5b1 (right) show the C6 position of the GlcNAc. The hydrogen bond that the C6 hydroxyl is forming with 1116NS19.9 is indicated with a dashed yellow line.

Europe PMC Funders Author Manuscripts

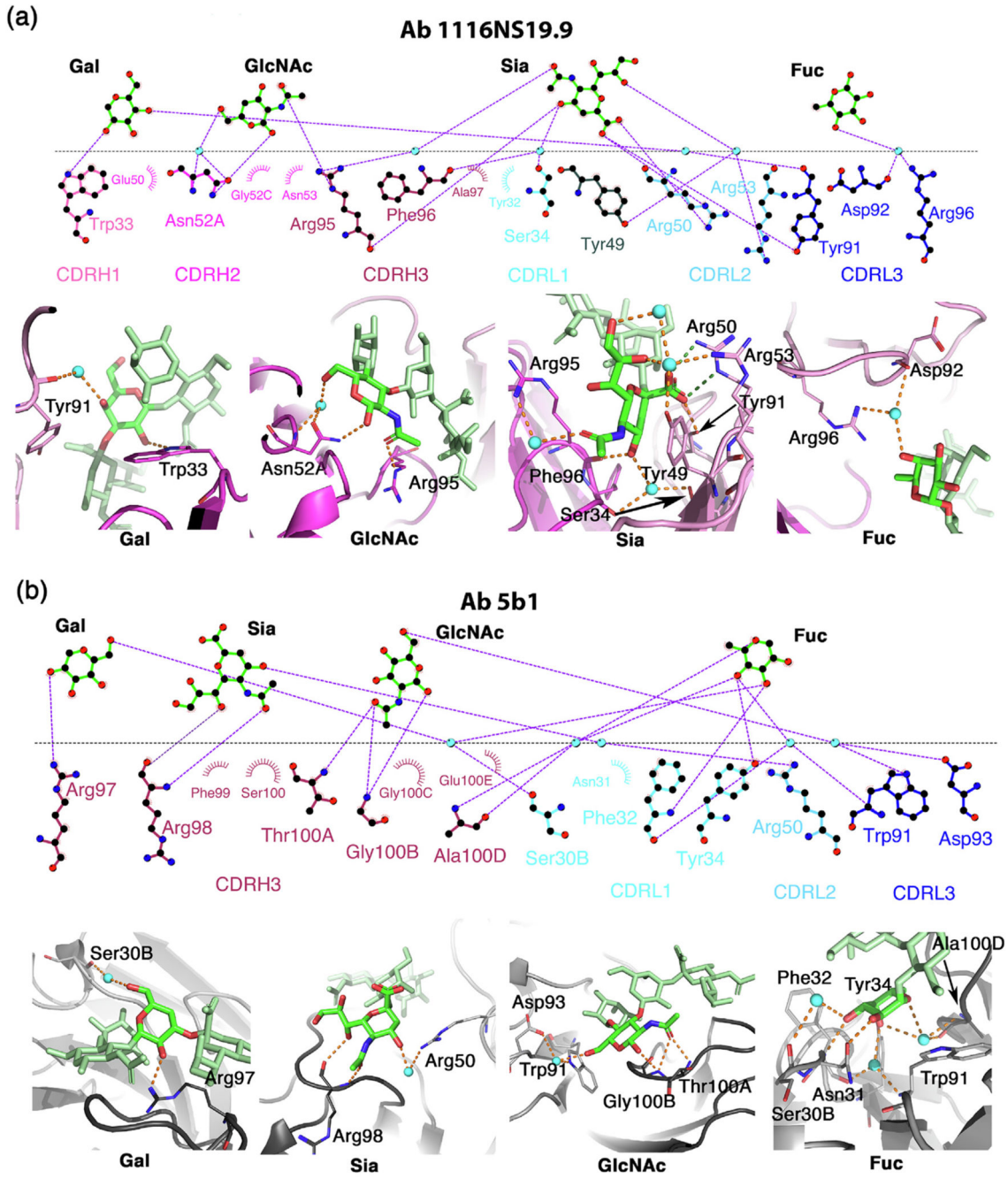

Figure 5. The hydrophilic binding sites of CA19-9.

(a) An overview of the CA19-9 binding site on 1116NS19.9. The saccharides are labeled L-fucose (Fuc), N-acetylglucosamine (GlcNAc), galactose (Gal), and sialic acid (Sia) (b) An overview of the CA19-9 binding site on 5b1. For both 1116NS19.9 and 5b1, upper panels show 2D representations of the interactions between CA19-9 and the antibodies that were generated using LIGPLOT.[54] The monosaccharide moieties of CA19-9 were separated for clarity. Water molecules are indicated with spheres in cyan, and polar bonds are indicated with dashed purple lines. CDRs are labeled and differentially colored. In the lower panels,

the interactions of the antibodies with each of the sugar moieties of CA19-9 are shown. The heavy and light chains are colored in purple/pink and black/grey for 1116NS19.9 and 5b1, respectively. Water molecules are indicated with spheres in cyan. Orange and green dashed lines indicate hydrogen bonds and salt-bridge interactions, respectively.

Europe PMC Funders Author Manuscripts

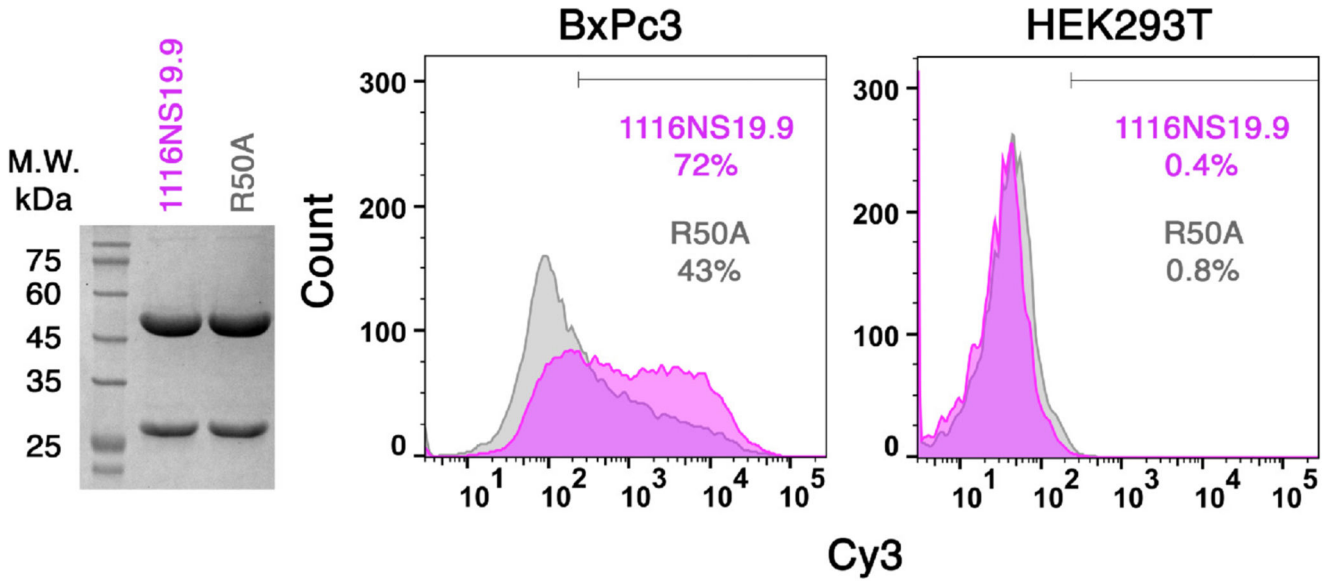

Figure 6. Arg50 of the 1116NS19.9 light chain is important for binding CA19-9 on cells.

Left panel shows analysis of 1116NS19.9 and its R50A mutant using SDS-PAGE stained with Coomassie blue. The right panel shows FACS analyses of the pancreatic cancer BxPc3 and HEK293T cells stained with 0.1 μg/ml of 1116NS19.9 (magenta) or with the R50A mutant (grey). The percent of positive cells according to the indicated gating are specified.

Europe PMC Funders Author Manuscripts

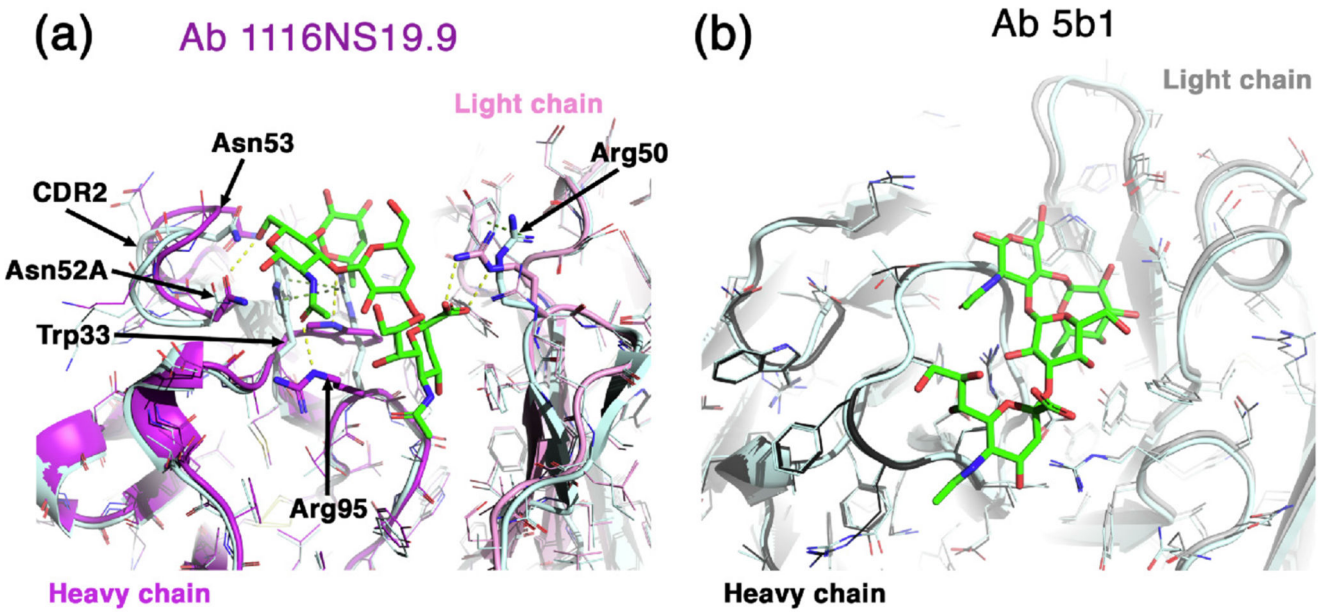

Figure 7. Conformational changes following binding to CA19-9.

(a) Superimposition of apo 1116NS19.9 (pale cyan) and CA19-9-bound holo 1116NS19.9 (purple and pink for the heavy and light chains, respectively). Important residues that change conformation are highlighted as sticks. Hydrogen and cation-π bonds are indicated with dashed yellow and green lines, respectively. (b) Superimposition of apo 5b1 (pale cyan) and CA19-9-bound holo 5b1 (black and grey for the heavy and light chains, respectively).

Europe PMC Funders Author Manuscripts

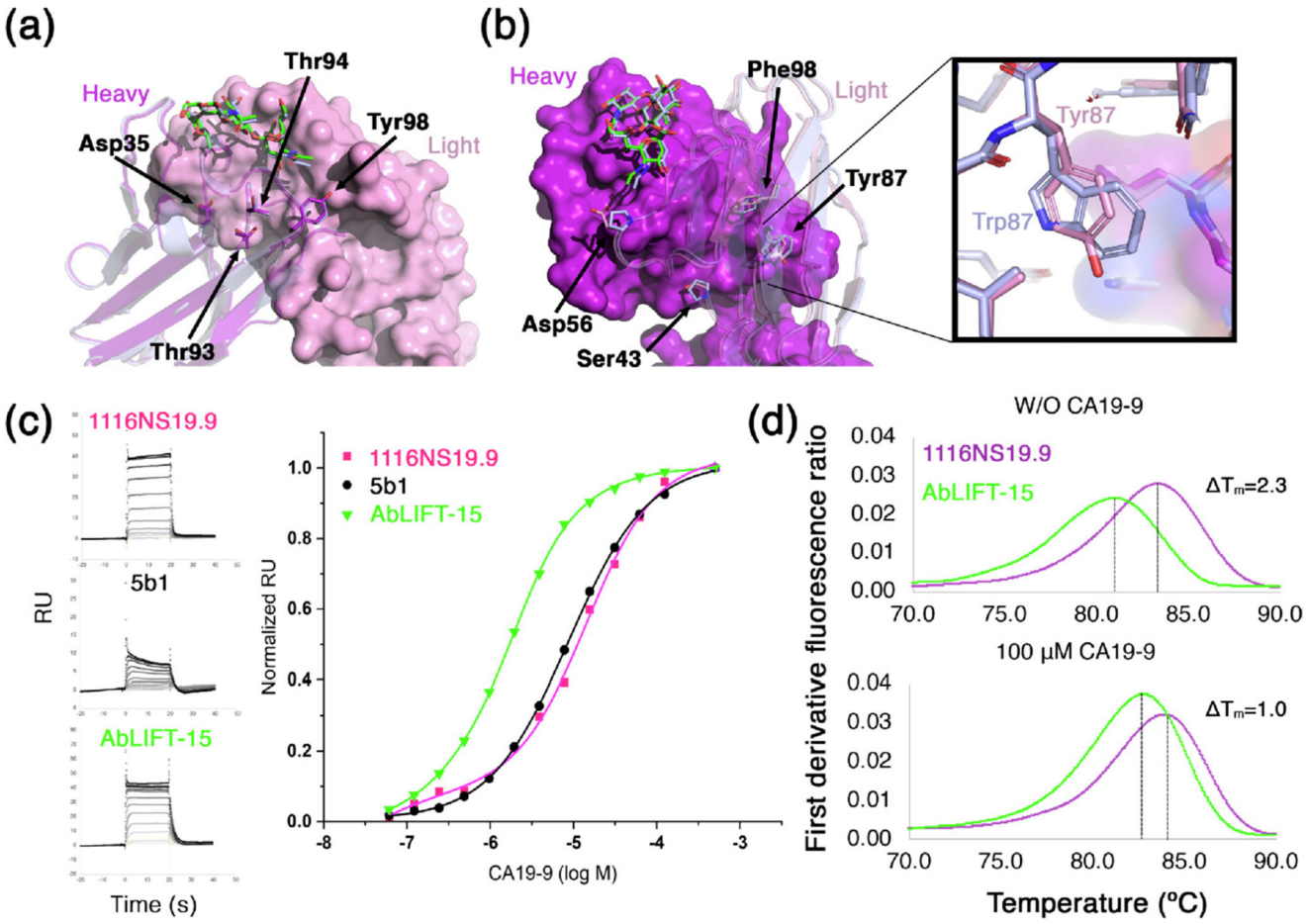

Figure 8. Recognition of CA19-9 and enhanced binding affinity of AbLIFT-15.

(a) Superimposition of the crystallographic structures of 1116NS19.9 (purple and pink for the heavy and light chains, respectively) and of AbLIFT-15 (pale cyan). The light chain is shown using surface representation, and the heavy chains are shown as ribbons. The four heavy-chain positions that were designed by the AbLIFT design protocol are indicated. (b) Superimposition of 1116NS19.9 (purple and pink for the heavy and light chains, respectively) and of AbLIFT-15 (pale cyan). The heavy chain is shown using surface representation, and the light chains are shown as ribbons. The four light-chain residues that were indicated by the AbLIFT design protocol are indicated. Inset shows an enlarged view of the Y87W mutation of AbLIFT-15. (c) Steady-state SPR analyses of 1116NS19.9, 5b1, and AbLIFT-15 using CA19-9 in a twofold dilution series starting at 500 μM. On the left, raw sensorgrams and on the right are fitted curves to normalized response data. Binding experiments were repeated three times, and representative results are shown. (d) Thermostability of 1116NS19.9, 5b1 and AbLIFT-15. Melting curves of 1116NS19.9, 5b1 and AbLIFT-15 with or without 100 μM of CA19-9. The melting midpoints are indicated by the first derivative of the ratio of tryptophan-fluoresce signal at 330 nm and 350 nm. The differences in Tm for 1116NS19.9 and AbLIFT-15 are indicated. These assays were repeated twice and representative curves are shown.

Europe PMC Funders Author Manuscripts

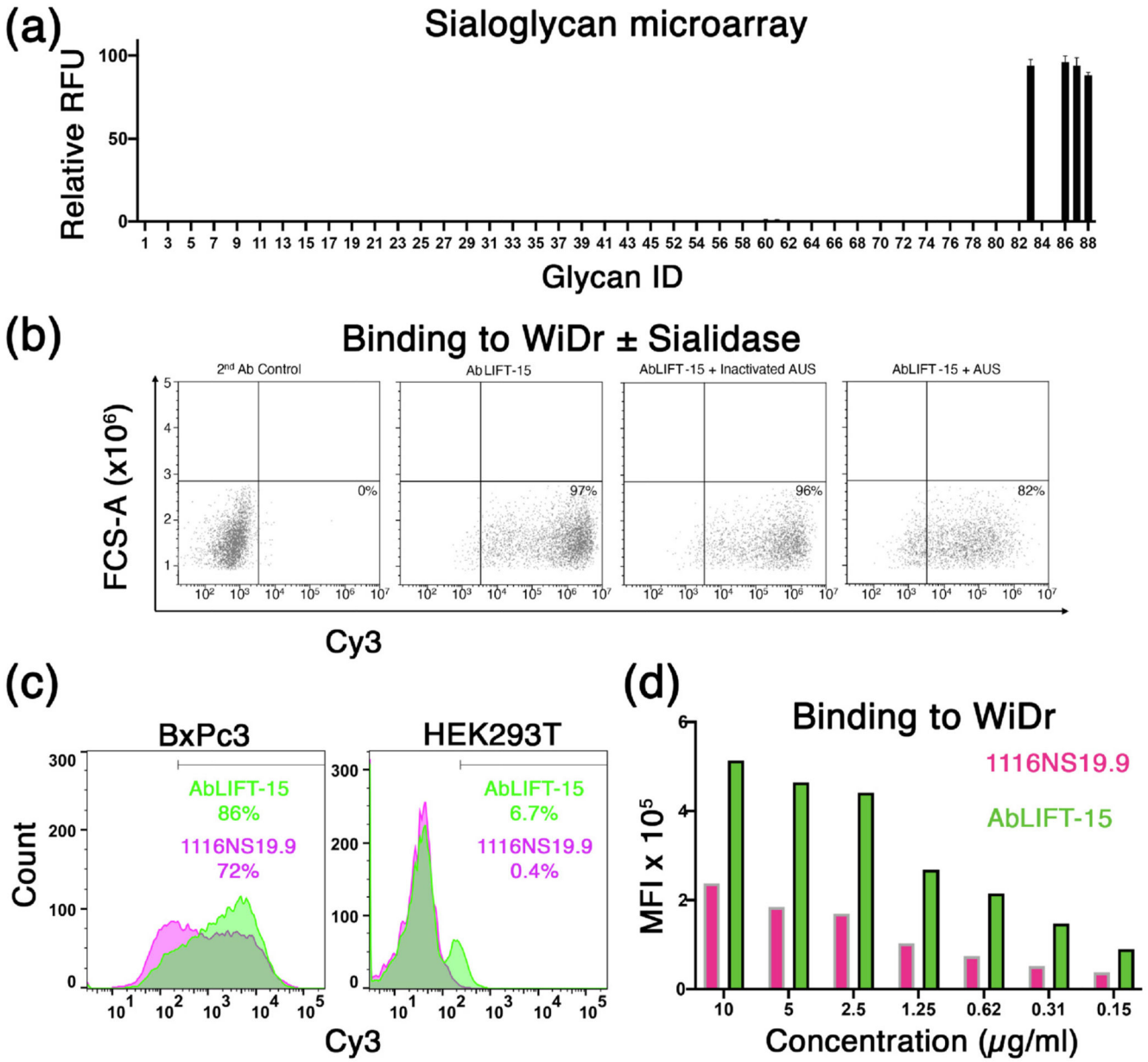

Figure 9. AbLIFT-15 exhibits specificity and high reactivity toward CA19-9.

(a) Binding of AbLIFT-15 IgG against diverse glycans was examined by nano-printed sialoglycan microarray at serial three concentrations ranging at $1.92 \times 10^{-2} – 4.81 \times 10^{-3}$ μg/ml (List of glycans in Supplementary Table 2). Relative fluorescence units (RFU) was calculated as percentage of maximal binding at each concentration, followed by averaging the relative RFU rank of the three tested antibody concentrations for each glycan (mean ± SEM). (b) Sialic-acid dependent specificity of AbLIFT-15. Binding of AbLIFT-15 to colorectal adenocarcinoma WiDr cells following a treatment with *Arthrobacter ureafaciens Sialidase* (AUS) was determined by FACS. Secondary antibody only and treatment with heat-inactivated AUS are shown as controls. Representative of two independent experiments. (c) Cell specific binding of AbLIFT-15. FACS analyses of pancreatic cancer BxPc3 cells and

HEK293T cells stained by 0.1 mg/ml of 1116NS19.9 or AbLIFT-15. Percent of Cy3-positive cells are indicated (gating is shown above). Representative of two independent repeats. (d) Binding of 1116NS19.9 and AbLIFT-15 at various concentrations (10–0.15 μg/ml) to WiDr cells was examined by FACS. Mean fluorescent intensity (MFI) for binding is indicated. Representative of two independent experiments.

Europe PMC Funders Author Manuscripts

Table 1
Data collection and refinement statistics

Parameter	Holo 1116NS19.9	Apo 1116NS19.9	Holo 5b1	Apo 5b1	AbLIFT-15
PDB ID	6XTG	6XUD	6XUN	6XUL	6XUK
Wavelength (Å)	0.97242	0.97242	0.87313	0.87313	0.97371
Space group	$P2_1\,2_1\,2$	$P4_1\,2_1\,2$	$P3_1$	$P\,3_2$	$P4_3\,2_1\,2$
Cell dimensions					
a, b, c (Å)	89.17 60.20 84.44	64.92 64.92 242.6	152.54 152.54 60.89	155.01 155.01 121.78	64.93 64.93 244.9
α, β, γ °	90 90 90	90 90 90	90 90 120	90 90 120	90 90 90
Resolution (Å)	44.58-1.55	45.9-1.51	47.58-2.41	46.84-2.41	62.75-1.42
	(1.60-1.55)[a]	(1.56-1.51)[a]	(2.48-2.41)[a]	(2.50-2.41)[a]	(1.47-1.42)[a]
R_{meas} (%)	5.9 (158)[a]	13.4 (103.1)[a]	17.42 (90.17)[a]	19.24 (103.8)[a]	10.85 (111.8)[a]
$^{CC}1/2$	99.9 (56.8)[a]	99.6 (79.2)[a]	99.1 (67.9)[a]	98.3 (53.3)[a]	99.6 (57.7)[a]
$I/\sigma I$	13 (1.1)[a]	9.4 (2.3)[a]	6.62 (1.81)[a]	5.58 (1.35)[a]	9.4 (2.3)[a]
Completeness (%)	95 (95.2)[a]	99.9 (100)[a]	99.76 (98.2)[a]	99.70 (99.9)[a]	99.89 (99.6)[a]
Multiplicity	2.9 (3)[a]	9.1 (9.4)[a]	5.1 (5.2)[a]	3.5 (3.5)[a]	6.1 (5.6)[a]
Reflections	359,539	753,716	309,763	446,316	610,565
Unique reflections	64,876	82,652	61,168	126,376	99,646
Refinement					
Resolution (Å)	44.58-1.55	45.32-1.51	47.58-2.41	46.84-2.41	62.75-1.42
No. of reflections	64,837	82,525	61,030	126,037	99,619
R_{work}/R_{free}(%)	17.39/20.13	14.69/ 17.32	16.99/21.34	16.78/22.95	14.9/18.49
No. of atoms					
Protein	3356	3361	9978	19,904	3312
Ligand/ion	57	-	192	-	68
Water	382	643	485	1022	661
B factors					
Protein	38.67	21.43	45.35	49.5	20.16
Ligand/ion	33.85	-	44.92	-	23.93
Water	45.09	35.81	44.41	46.33	34.20
Ramachandran					
Favored (%)	98.12	98.83	97.60	96.54	98.12
Allowed (%)	1.88	1.17	2.32	3.34	1.41
Outlier (%)	0.0	0.0	0.08	0.12	0.47
RMSD					
Bond length (Å)	0.006	0.008	0.006	0.008	0.013
Bond angles °	0.733	0.995	1.17	1.22	1.65

[a]Values in parentheses are for the highest resolution-shell.