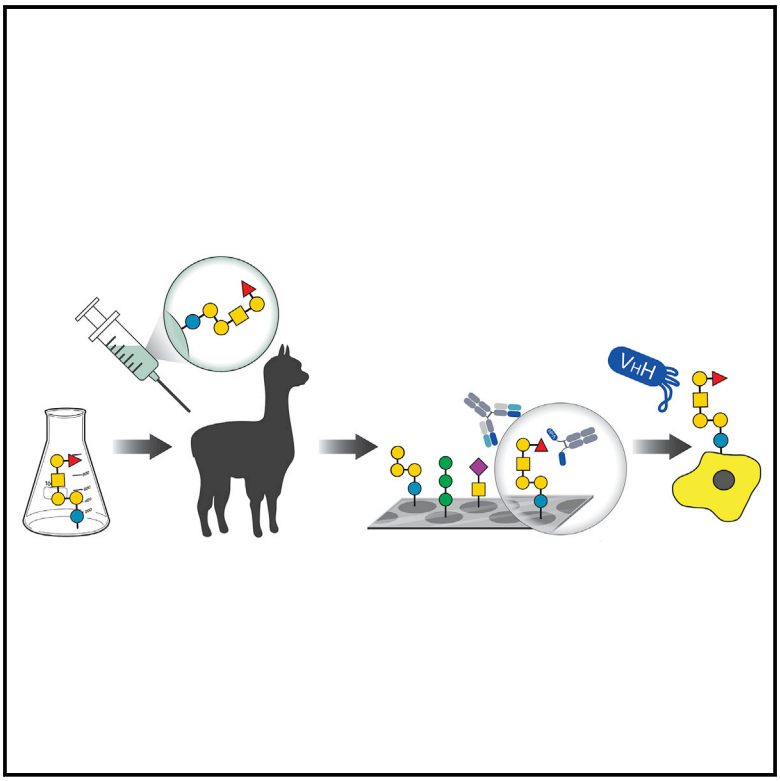
Resource

Cell Chemical Biology

Generation of glycan-specific nanobodies

Graphical abstract



Authors

Sana Khan Khilji, Felix Goerdeler, Kristin Frensemeier, ..., Juri Rappsilber, Peter H. Seeberger, Oren Moscovitz

Correspondence

peter.seeberger@mpikg.mpg.de (P.H.S.), oren.moscovitz@mpikg.mpg.de (O.M.)

In brief

By immunizing an alpaca with synthetic and well-defined carbohydrate structures, Khilji et al. developed ultra-small antibodies (termed nanobodies) that target the tumor-associated carbohydrate antigen Globo-H, expressed in multiple cancers. The results demonstrate an approach to generate glycan-specific nanobodies for biotechnological applications.

Highlights

- Alpaca immunization with different synthetic tumor-associated carbohydrate antigens
- Identification of glycan-specific heavy-chain-only antibody formation
- Isolation and characterization of Globo-H-specific nanobodies

Khilji et al., 2022, Cell Chemical Biology *29*, 1353–1361
August 18, 2022
https://doi.org/10.1016/j.chembiol.2022.05.007

CellPress

Cell Chemical Biology

CellPress

Resource

Generation of glycan-specific nanobodies

Sana Khan Khilji,[1,2,8] Felix Goerdeler,[1,2,8] Kristin Frensemeier,[1,2] David Warschkau,[1,2] Jost Lühle,[1,2] Zeinab Fandi,[1,3] Falko Schirmeister,[1] Zhuo Angel Chen,[4] Onur Turak,[5] Alvaro Mallagaray,[5] Stefan Boerno,[6] Bernd Timmermann,[6] Juri Rappsilber,[4,7] Peter H. Seeberger,[1,2,*] and Oren Moscovitz[1,9,*]

[1]Department of Biomolecular Systems, Max Planck Institute of Colloids and Interfaces, 14476 Potsdam, Germany
[2]Institute of Chemistry and Biochemistry, Freie Universität Berlin, 14195 Berlin, Germany
[3]Department of Life Sciences and Technology, Beuth University of Applied Sciences Berlin, 13347 Berlin, Germany
[4]Technische Universität Berlin, Chair of Bioanalytics, 10623 Berlin, Germany
[5]Institute of Chemistry and Metabolomics, Center of Structural and Cell Biology in Medicine (CSCM), University of Lübeck, 23562 Lübeck, Germany
[6]Sequencing Core Facility, Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany
[7]Wellcome Centre for Cell Biology, University of Edinburgh, EH9 3BF Edinburgh, UK
[8]These authors contributed equally
[9]Lead contact
*Correspondence: peter.seeberger@mpikg.mpg.de (P.H.S.), oren.moscovitz@mpikg.mpg.de (O.M.)
https://doi.org/10.1016/j.chembiol.2022.05.007

SUMMARY

The development of antibodies that target specific glycan structures on cancer cells or human pathogens poses a significant challenge due to the immense complexity of naturally occurring glycans. Automated glycan assembly enables the production of structurally homogeneous glycans in amounts that are difficult to derive from natural sources. Nanobodies (Nbs) are the smallest antigen-binding domains of heavy-chain-only antibodies (hcAbs) found in camelids. To date, the development of glycan-specific Nbs using synthetic glycans has not been reported. Here, we use defined synthetic glycans for alpaca immunization to elicit glycan-specific hcAbs, and describe the identification, isolation, and production of a Nb specific for the tumor-associated carbohydrate antigen Globo-H. The Nb binds the terminal fucose of Globo-H and recognizes synthetic Globo-H in solution and native Globo-H on breast cancer cells with high specificity. These results demonstrate the potential of our approach for generating glycan-targeting Nbs to be used in biomedical and biotechnological applications.

INTRODUCTION

Carbohydrates (or glycans) form a thick and complex layer around cell surfaces that shapes cell-cell interactions upon first contact. Heterogeneity among different cell types and species makes glycans excellent biomarkers (Adamczyk et al., 2012). At the same time, the enormous structural complexity and diversity of carbohydrates pose a serious challenge for detailed investigations. Automated glycan assembly (Pardo-Vargas et al., 2018) provides access to homogeneous glycans that have, in turn, enabled the development of anti-glycan antibodies (Broecker et al., 2016), which mostly suffer from low substrate specificity (Sterner et al., 2016).

Since the first report of aberrant glycosylation on cancer cells in 1949 (Oh-Uti, 1949), tumor-associated carbohydrate antigens (TACAs) have been identified as cancer hallmarks. Of these, Globo-series glycosphingolipids represent a sub-family with highly dynamic expression patterns throughout cell development and differentiation (Liang et al., 2010). Globo-H is a prevalent and well-studied glycosphingolipid-TACA that was first identified in human teratocarcinoma and MCF7 breast cancer cells (Kannagi et al., 1983). It consists of a lactosylceramide core as part of a hexasaccharide (Fucα(1–2)Galβ(1–3)GalNAcβ(1–3)Galα(1–4)Galβ(1–4)Glc) and is overexpressed in a variety of epithelial cancers, including breast, lung, prostate, pancreatic, ovarian, endometrial, brain, gastric, and colon cancer. Globo-H shed from tumor cells promotes immunosuppression, angiogenesis, and metastasis (Cheng et al., 2014). Finally, a recent phase II/III trial in breast cancer patients using a synthetic Globo-H vaccine demonstrated a beneficial effect in progression-free survival for patients who developed high enough levels of Globo-H-specific antibodies (Rugo et al., 2020).

More than 30 years ago, it was discovered that animals from the Camelidae family (camels, llamas, alpacas) produce unique heavy-chain-only antibodies (hcAbs) in addition to canonical Immunoglobulin G (IgG) with both heavy and light chain (Hamers-Casterman et al., 1993). Antigen-binding sites of these hcAbs are formed by single domains, designated VHHs, that can be engineered into highly stable and soluble nanobodies (Nbs). Due to their ultra-small size (~14 kDa), Nbs can penetrate small and dense microenvironments such as in tumors (Debie et al., 2020) and bind to cryptic epitopes inaccessible for canonical antibodies (De Genst et al., 2006).

Here, we report an approach to obtain glycan-binding Nbs. Our procedure entails immunization of an alpaca with well-defined, synthetic glycoconjugates to raise hcAbs. Glycan-binding Nbs

are then identified in a workflow that combines synthetic glycan array with next generation sequencing (NGS), mass spectrometry (MS), and bioinformatics. Selected Nbs are screened for binding to glycan-expressing target cells and their specificity is validated by in-solution binding assays against different synthetic glycans. We show the feasibility of our approach by successfully producing Nbs with high specificity for the TACA Globo-H. Our work provides a fast and general method for the development of glycan-binding Nbs using only small amounts of well-defined, synthetic glycans, and thereby facilitates the generation of Nbs against further glycan targets.

RESULTS

Immunization with synthetic glycoconjugates induces glycan-specific heavy-chain-only antibodies in alpaca

So far, only few studies describe glycan-binding Nbs, and substrates typically involve complex native bacterial lipopolysaccharides of undefined epitopes (Ebrahimizadeh et al., 2013). Therefore, we set out to develop routines for the production of Nbs that interact specifically with well-defined glycan epitopes. Due to the lack of established protocols for alpaca immunization with synthetic glycoconjugates, we immunized an alpaca with five different glycoconjugates. Each glycoconjugate consisted of a synthetic glycan of biological relevance conjugated to the carrier protein CRM197 (Figure S1), and all five glycoconjugates were joined and injected together to a single animal in six consecutive immunization rounds.

As alpaca serum IgGs consist of about 50% hcAbs, the majority of which being IgG3 (Maass et al., 2007), we sought to prove the development of specific anti-glycan hcAbs, by fractionating the serum after the fifth immunization round to separate conventional IgG1 and hcAbs IgG2 and IgG3 (Figure 1B). Glycan array inspection of each subclass verified the presence of anti-glycan IgG3 and, to a lesser extent, of IgG2 (Figure 1C). IgG1 was the dominant subclass for Globo-H and Sialyl-Tn, but not for KH-1, where all three subclasses exhibited similar titers (Figure 1C).

Next, we used a glycan array to delineate the serum response to different glycans over time and found that antibody titers did not correlate with glycan concentration, length, or complexity. The amount of injected glycan per immunization was pre-determined by the efficiency of conjugation to the carrier protein. Hence, the administered glycan dose varied significantly between glycoconjugates (Figure 1A). Interestingly, not all constructs elicited similar immune responses, and we observed no correlation between specific antibody titers and quantities of administrated glycan per immunization (Figure 1D). For instance, immunization with 4 μg per dose of the Globo-H hexamer glycoconjugate resulted in a peak of specific antibody levels after the third injection (28 days), which decreased over the following weeks. Maximum antibody titers after the third round of immunization are in agreement with previous trials in mice that we had performed (Broecker et al., 2016). In contrast, antibody titers were significantly lower for the charged Sialyl-Tn, and nonameric KH-1, 7.8 and 16 μg/dose, respectively, and only reached a peak after 45 days (Sialyl-Tn) and 53 days (KH-1, Figure 1D). We observed a minimal immune response against GMH7P and the Tn antigen monosaccharides. The titer of antibodies against a non-injected control glycan heparin ($GlcNSO_3$[6-OSO_3](α1-4)IdoA[2-OSO_3](α1-4)$GlcNSO_3$[6-OSO_3](α1-4)IdoA[2-OSO_3](α1-4)$GlcNSO_3$[6-OSO_3](α1-4)IdoA[2-OSO_3]) did not change during the immunization regime, confirming that serum responses were triggered by immunization (Figure 1D).

Next, we examined the specificity of antibodies elicited against KH-1 and Globo-H by testing sub-fragments of the parent glycan in glycan arrays (Figures 1E and 1F). We investigated whether antibody titers raised against KH-1 (Le^X-Le^Y) were cross-reactive against $Le^{X/Y}$ histo-blood groups and found that KH-1-specific antibodies did not bind to individual blood group components, except Le^X dimers (Figure 1E). We further detected correlations between antibody titers and glycan length, with response activities gradually increasing from the GB3 trimers to the GB5 pentamers, which are also attractive TACAs for Nb development. In contrast, the additional fucose in Globo-H resulted in a reduced antigenicity compared with GB5, although Globo-H still remained more antigenic than GB3/4 (Figure 1F).

Given the importance of Globo-H as an anti-cancer vaccine target, we focused our efforts on the development of anti-Globo-H Nbs.

Selection of Globo-H-specific Nb sequences by combining MS and next-generation sequencing data *in silico*

The Globo-H hexasaccharide constitutes a part of the well-characterized glycosphingolipid Globo-H tumor antigen that is expressed on different human cancer cells (Menard et al., 1983; Zhang et al., 1997). Several clinical trials aimed at targeting Globo-H with synthetic antibody-drug conjugates or employing Globo-H as an antigen in vaccine trials, in combination with different protein carriers and immune-stimulating adjuvants (Huang et al., 2020; Rugo et al., 2020). Our goal was to harness the structural and functional advantages of Nbs for cancer diagnosis and therapy and to develop an anti-Globo-H Nb.

Synthetic Globo-H (Werz et al., 2007) was part of the five glycoconjugates used to immunize an alpaca to generate specific hcAbs for Nb isolation as described above. To identify Globo-H-specific Nbs, we modified an MS-based approach (Fridy et al., 2014) that only requires minute amounts of synthetic glycans to isolate and identify binders. First, we immobilized synthetic Globo-H on beads to extract anti-Globo-H hcAbs from the alpaca serum extracted post immunizations. Then, we generated peptide libraries of Globo-H-specific VHH fragments. In parallel, we generated cDNA libraries for NGS, from peripheral blood mononuclear cells (PBMCs) isolated on day 53, yielding a total VH/VHH sequences database. We selected 36 different sequences of potential Globo-H-binding Nbs for recombinant expression based on different sequence coverage criteria after bioinformatic comparison of MS-derived peptide sequences and NGS library data (Figures 2A and S2B–S2E). We also performed a sequence alignment with selected sequences and found the complementarity-determining regions to be very diverse (Figure S2D).

We recombinantly expressed the 36 Nb sequences and tested them for antigen binding, for which we employed Globo-H-expressing MCF7 breast cancer cells (Bremer et al., 1984) and the commercial anti-Globo-H antibody VK9 (Kudryashov et al., 1998) as a positive control.

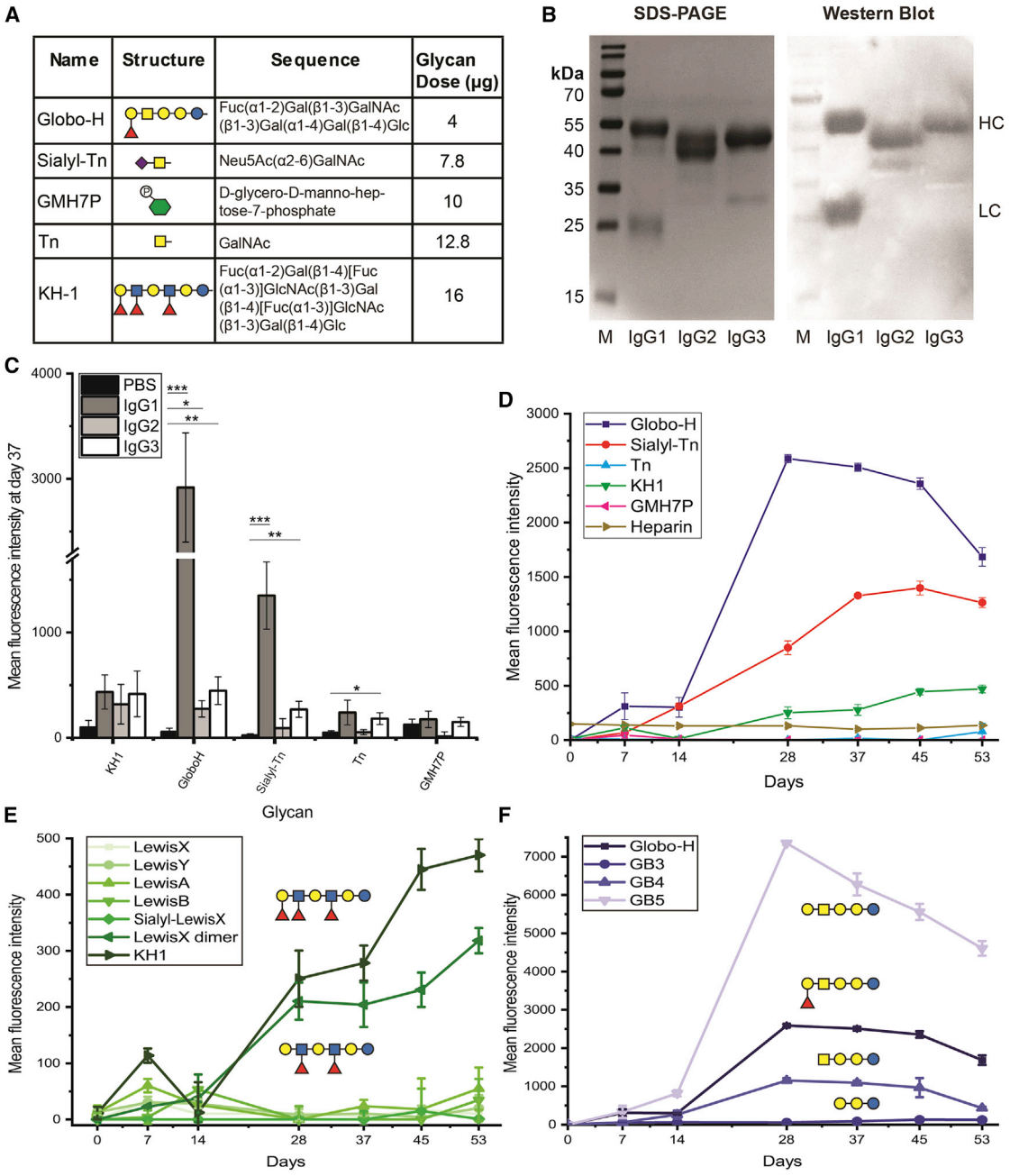


Figure 1. Glycan immunogenicity in alpaca

(A) Overview of glycan antigens coupled to CRM197 carrier protein for alpaca immunization.

(B) Fractionation of alpaca serum from day 37 to IgG1/2/3 subclasses. Purity was confirmed by SDS-PAGE (left) and western blot with anti-llama-HRP Ab (right). HC, heavy chain; LC, light chain.

(C) Antibody titer of purified IgG subclasses against the different injected glycoconjugates measured by glycan array. Means from three independent experiments were compared by two-sample t test, *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. Error bars represent SEM.

(D–F) Glycan arrays displaying the time-dependent antibody titer against synthetic glycan structures. Values represent mean ± SEM (n = 4). (D) Antibody titer against the five injected glycoconjugates and the non-injected carbohydrate antigen heparin, (E) antibody titer against Lewis blood group structures, and (F) Globo-H family members during immunization with KH1 and Globo-H, respectively. See also Figure S1.

Several Nb candidates bind to Globo-H-expressing breast cancer cells

To test binding to Globo-H-expressing cells, we purified the different Nb candidates (Figure S2F) and compared labeling of MCF7 cells in flow cytometry and confocal microscopy (Figures 2B and 2C, and 3A and 3B). Specific binding visualized by flow cytometry was defined as mean fluorescence intensity (MFI) >5, which was observed for six out of 36 Nb candidates

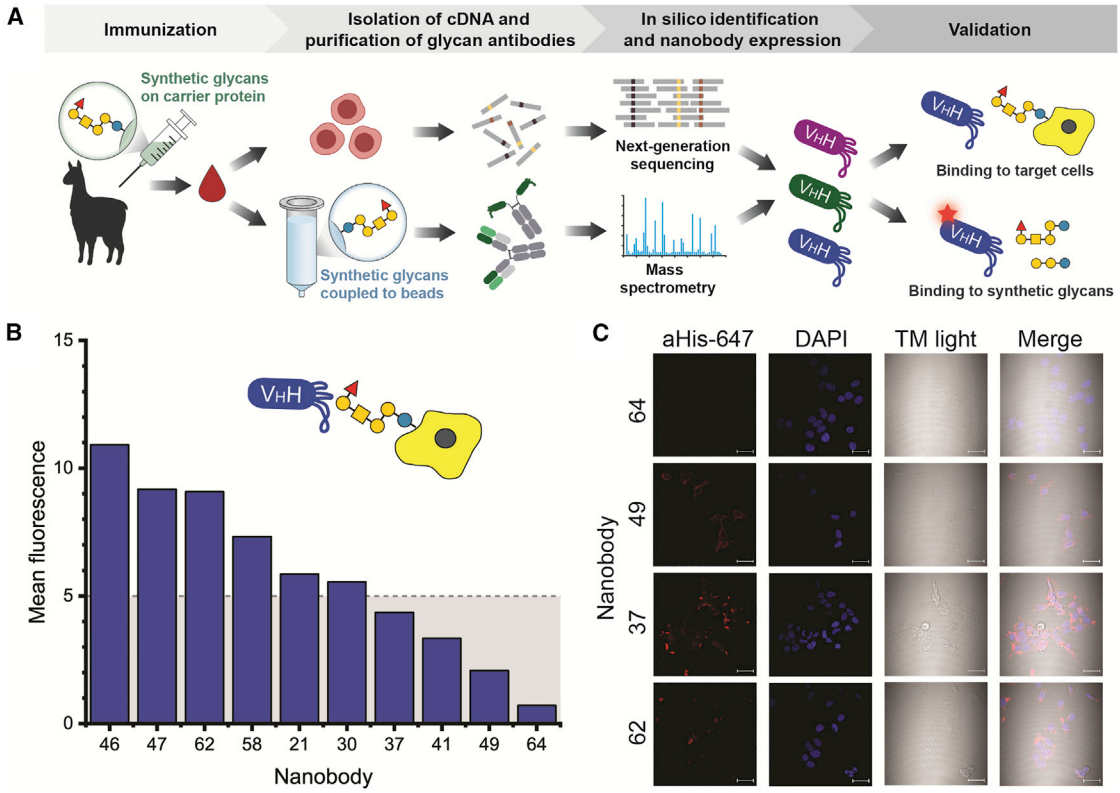


Figure 2. Globo-H-specific nanobodies

(A) Schematic overview of the workflow for generating carbohydrate-binding nanobodies (Nbs).

(B) Screening potential glycan-binding Nbs by flow cytometric analysis of Nb binding to MCF7 cells. The best 10 binders out of 36 candidates are shown. The y axis is mean fluorescence intensity (MFI), and x axis is different Nbs. Nbs with MFI <5 were considered negative (gray box). Based on these results, anti-Globo-H Nb46 (GH46) was selected for further characterization.

(C) Confocal microscopy of MCF7 cells incubated with different Nbs (red) from (B) and DAPI (blue). Scale bar, 5 μm. See also Figures S2 and S3.

(Figure 2B). To exclude binding to the lower-molecular-weight Globo-family glycans, Nb binding to human embryonic kidney cells (HEK293) and the non-tumorigenic breast cell line MCF10A was examined as these cells express GB2, GB3, and GB4 but not GB5 and Globo-H (Fujitani et al., 2013; Huang et al., 2017). Two Nb candidates (Nb30 and Nb62) showed high binding to both MCF7 (Figure 2B) and HEK293 cells (Figure S3A), suggesting a lower-molecular-weight Globo-family member as their minimal epitope. However, our most potent binder to MCF7 cells (Figures 3B and S3F–S3G), herein GH46, did not recognize MCF10A and HEK293 cells (Figures S3B–S3E). In addition, we excluded non-specific binding of Atto 647 anti-His secondary antibody by measuring concentration-dependent increase in Nb binding to MCF7 cells (Figure S3H). We further validated binding of GH46 to MCF7 cells by immunostaining and confocal microscopy and found the cell surface uniformly labeled with GH46, compared with no staining when using one of the non-binding Nb candidates as a control (Figure 3A). Interestingly, the commercial Ab VK9 seems to recognize a different part of the Globo-H structure as MCF7 cells co-incubated with VK9 and GH46 appear double stained and no competition between the two proteins was observed (Figure S4).

GH46 specifically recognizes Globo-H

Based on the cell-binding assays, we hypothesized that the minimal binding epitope of GH46 is either GB5 or Globo-H and attempted to test this hypothesis by glycan array. However, even though a range of concentrations and conditions was examined, we could not detect Nb binding on glycan array, presumably because GH46 is washed away during the numerous washing steps in the protocol due to relatively low-affinity binding. Therefore, we sought to determine the minimal binding epitope of GH46 in solution via microscale thermophoresis (MST) (Wienken et al., 2010) using synthetic Globo-family glycans and GH46. We observed specific binding of GH46 to synthetic Globo-H with a K_D value of (212 ± 89) μM (Figure 3C). GH46 cross-reactivity against GB5, GB4, and GB3 was significantly lower ($K_D \geq 1$ mM; Figure 3C), indicating that the terminal fucose of Globo-H forms part of the GH46 recognition motif.

Considering the single-domain nature of Nbs and the generally lower affinity of protein-glycan interactions (Collins and Paulson, 2004), we postulated that a multivalent GH46 construct would bind Globo-H-expressing cells with improved apparent affinity due to higher avidity. We therefore designed a head-to-tail fusion of three GH46 domains (herein $GH46_3$) interspaced by $(GGGGS)_2$ linkers. Although we did not observe binding of

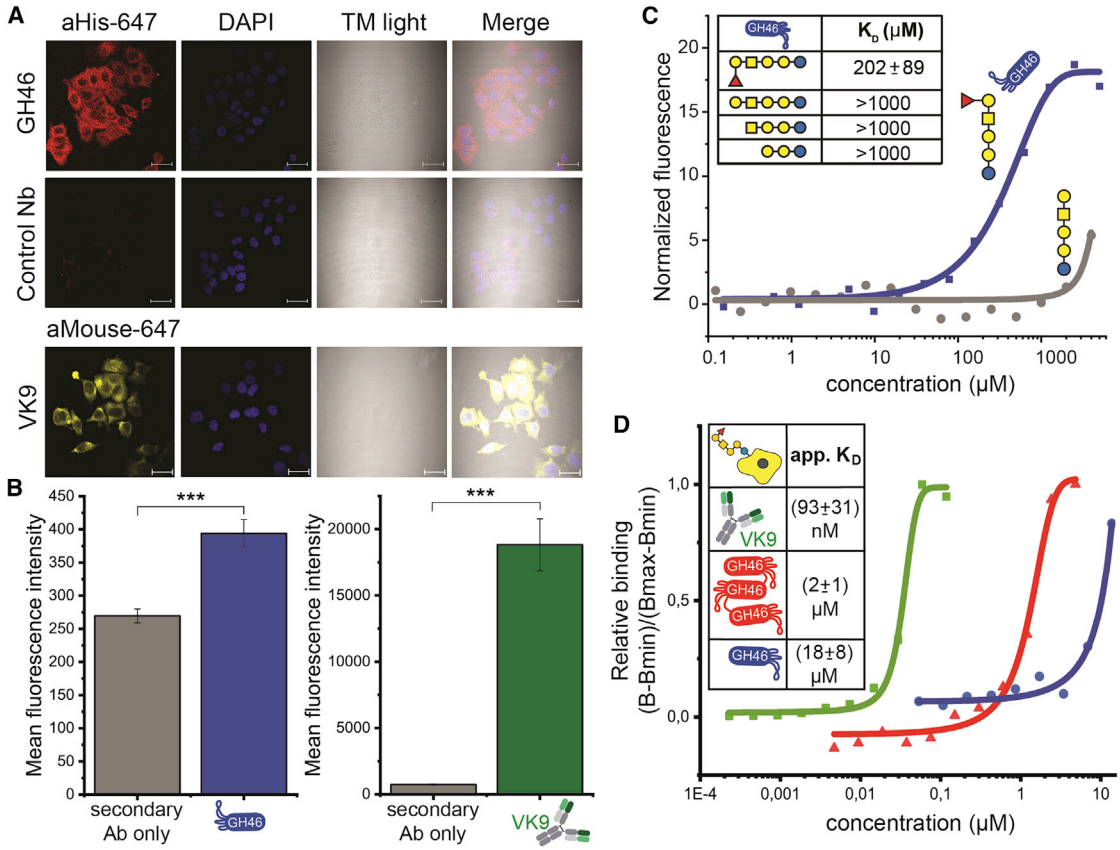


Figure 3. GH46 binding to synthetic and native Globo-H

(A) Immunostaining of MCF7 cells with GH46 (red), an unrelated control Nb (red), or VK9 (yellow), and DAPI (blue) (scale bar, 5 μm). Same pseudocolor reflects same acquisition and display settings.

(B) Flow cytometry binding assay with MCF7 cells. Left: staining of MCF7 cells with anti-6xHis-Atto647N Ab in the absence (gray) or presence (blue) of GH46. Right: staining of MCF7 cells with anti-mouse-IgG-635 Ab in the absence (gray) or presence (green) of VK9. The y axis is MFI. Values represent mean ± SEM (n = 7). Differences were tested for significance via t test. ***$p < 0.001$.

(C) Measured in-solution K_D of GH46 and synthetic glycans (table) and comparative affinity curves for GB5 (gray) and Globo-H (blue). Values represent mean ± SEM ($n \geq 3$).

(D) Measured apparent K_D on MCF7 cells (table) and representative affinity curves for GH46 (blue), $GH46_3$ (red), and VK9 (green). Values represent mean ± SEM (n = 3). See also Figures S3 and S4.

$GH46_3$ in glycan array, again presumably due to the low-affinity interaction, we were able to determine the apparent K_D of GH46, $GH46_3$, and VK9 in a flow cytometric binding assay with MCF7 cells. Interestingly, trimerization improved the on-cell affinity of GH46 9-fold, from K_D^{app}(GH46) = 18 ± 8 μM to K_D^{app}($GH46_3$) = 2 ± 1 μM (Figure 3D). These data confirm that multivalent $GH46_3$ shows improved avidity compared with monovalent GH46, at a molecular weight that is still one-third of a conventional IgG.

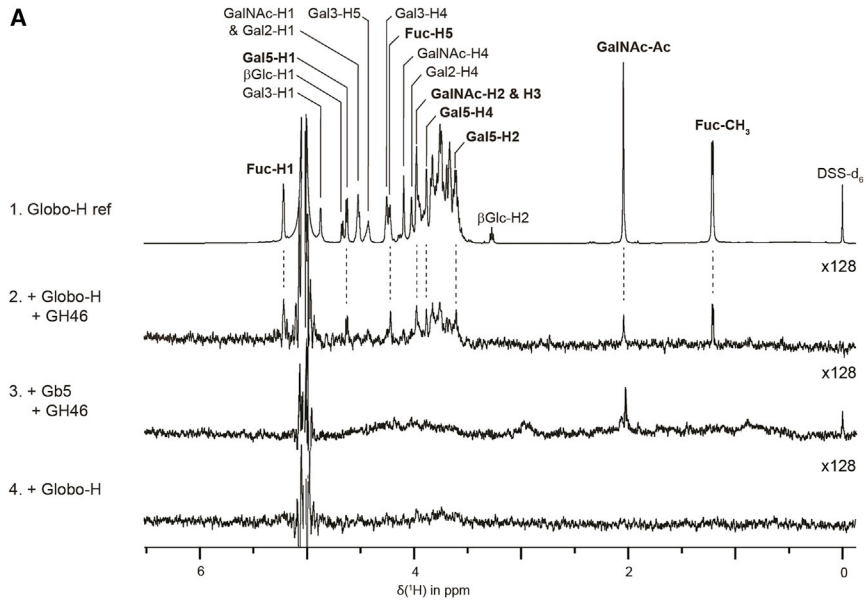
The terminal fucose of Globo-H is required for GH46 binding

To validate MST results and to map GH46's binding epitope, we employed saturation transfer difference (STD) nuclear magnetic resonance (NMR), which is particularly well suited to probing interactions with dissociation constants in the micromolar range (Meyer and Peters, 2003). Due to signal overlap, we only obtained STD amplification factors for a subset of Globo-H proton signals (Figure 4A), which nonetheless established that primary contacts between GH46 and Globo-H were mediated by the Fucα(1–2)Galβ moiety (Figure 4B). By contrast, we did not observe STD signals with GB5 and GH46, suggesting that fucose is indeed required for ligand recognition. This observation is in good agreement with the results delivered by MST, where the lack of the terminal fucose in the Globo series impaired carbohydrate recognition by GH46 (Figure 3C).

To further support GH46 binding to Globo-H, we recorded Globo-H transferred nuclear Overhauser effect (trNOE) signals in the presence and absence of GH46. We detected significantly faster NOE build-up rates in the presence of GH46, further confirming the protein-glycan interaction (Kumar et al., 1981) (Figure 4C and S4E and F).

DISCUSSION

Growth and progress in glycobiology are limited mainly by the lack of appropriate glycan-specific monoclonal antibodies. The main reason for the shortage of highly specific glyco-tools is



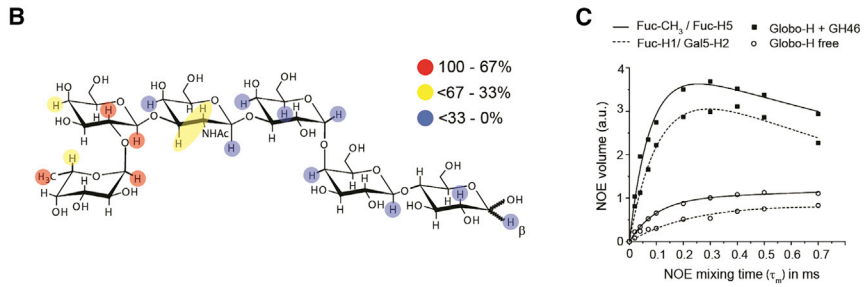


Figure 4. GH46 – Globo-H interaction by STD NMR

(A) Reference spectrum of Globo-H (1). Isolated signals showing magnetization transfer are highlighted in bold. (2) STD spectrum of 4 mM Globo-H in the presence of 60 μM GH46. (3) STD spectrum of 4 mM GB5 in the presence of 110 μM GH46. (4) STD spectrum of 4 mM Globo-H. Residual water signal was not suppressed to allow for the observation of Fuc and Gal3 anomeric protons.

(B) Binding epitope of Globo-H bound to GH46 from a single saturation time of 4 s. STD amplification factors for protons without color coding were not available due to signal overlap. NMR experiments were acquired at 600 MHz and 277 K.

(C) NOE build-up curves of Globo-H with and without GH46 versus mixing time (milliseconds) at 600 MHz and 298 K. Curves represent intra- (Fuc-CH3/Fuc-H5) and inter-pyranose (Fuc-H1/ Gal5-H2) connectivities. Nb to carbohydrate ratio is 1:16.4. See also Figures S4E and S4F.

the inherent heterogeneity and structural complexity of natural glycans. Recent developments in synthetic glycan assembly enable rapid access to sufficient quantities of increasingly complex glycan structures (Hahm et al., 2017) that provide the basis for developing specific anti-glycan antibodies (Broecker et al., 2015).

Nevertheless, the quantity, quality, and availability of glycan-specific monoclonal antibodies are substantially lower than of protein-targeting antibodies. Currently, there are only 428 commercially available mAbs listed in the Database for Anti-Glycan Reagents (DAGR). These antibodies target fewer than 70 different glycans structures in total (Sterner et al., 2016) and indicate the need for highly characterized anti-glycan mAbs that target defined glycan epitopes.

Due to their ultra-small size, highly versatile functionalization, and straightforward genetic manipulation, Nbs are ideal tools for biomedical applications. Protein-targeting Nbs are used for applications ranging from basic research to imaging and therapeutics (Duggan, 2018). So far, very few glycan-binding Nbs have been reported. These Nbs were produced by immunization with whole organisms or native glycoproteins and bind to heterogeneous and uncharacterized bacterial lipopolysaccharides (Ebrahimizadeh et al., 2013) or glycoproteins rather than pure glycan epitopes (Behar et al., 2009; Stijlemans et al., 2004). We set out to demonstrate that immunizing with well-defined synthetic glycans triggers the formation of glycan-specific hcAbs, which can be exploited to produce Nbs that recognize well-defined glycan epitopes.

We found that simultaneous immunization with several glycoconjugates is applicable. In addition, our findings suggest that three consecutive immunizations are sufficient for the highly immunogenic glycans and that PBMCs could be collected at days 28 or 35. However, less immunogenic antigens such as KH-1 and Tn might benefit from additional boosting steps to reach higher antibody titers. The low predictability of glycan immunogenicity indicates the need for an individual immunization regime rather than a general immunization protocol. However, pre-immunization studies are hardly practical in large animals. Therefore, we recommend multiple collections of PBMCs, both after the third immunization round and at least 2 weeks after the sixth final boosting.

Alpaca immunization with synthetic glycoconjugates has not been described before. Avci et al. previously showed in mice that a T cell-dependent response, and thus affinity-matured IgGs, can be induced by immunization with glycoconjugates but not with glycans alone (Avci et al., 2011). Here, we prove that immunization with synthetic glycoconjugates leads to the formation of glycan-specific IgG2/3 in the alpaca serum, alongside canonical IgG1.

While monovalent GH46 binds to synthetic Globo-H with relatively low affinity, we showed that covalent trimerization improved the apparent K_D of GH46 9-fold. Compared with a conventional IgG Ab, $GH46_3$ maintains a small size of approximately 50 kDa and is still easily expressible in bacterial systems. This can be an advantage as previous work revealed that dimers of HER2-specific Nb show not only improved apparent affinity compared with Nb monomers but also more homogeneous tumor penetration compared with conventional IgGs due to their smaller size (Debie et al., 2020). Taken together, these findings

suggest multimerization as a helpful and viable approach, especially when aiming to boost the avidity of glycan-binding Nbs.

Anti-glycan Nbs can form the basis for multispecific tools with increased avidity and specificity. Multispecific Nb constructs that target tumor-associated glycans and proteins simultaneously will allow for more efficient discrimination of healthy and cancer cells, thus decreasing off-target binding. Furthermore, conjugation to antibody-recruiting molecules such as 2,4-dinitrophenol, α-Gal, or L-rhamnose can enable GH46 to activate immune effector cells in the immunosuppressive tumor microenvironment.

The recent approval of a therapeutic Nb (Duggan, 2018) against the von Willebrand factor protein will likely render Nbs an equal or even preferred alternative to canonical antibodies. Combined with synthetic glycans, we anticipate that the future of glycan-targeting Nbs is promising. Here we provide the first report making use of synthetic glycans throughout the entire process of anti-glycan Nb development, from target identification to binding analysis.

Limitations of the study

Compared with the commercially available bivalent anti-Globo-H Ab VK9, the affinity of $GH46_3$ for Globo-H is still lower. VK9 was produced in hybridoma cells derived from a mouse previously immunized with Globo-H keyhole limpet hemocyanin (KLH) conjugate and QS21 adjuvant (Kudryashov et al., 1998). Comparing different carrier proteins and adjuvants for alpaca immunization was outside the scope of this work, but it would be interesting to test whether alternative formulations facilitate the purification of glycan-binding Nbs with higher affinity.

Another potential leverage point for affinity improvement resides in the method for Nb candidate selection. The few existing glycan-binding Nbs were derived by phage display in multiple panning rounds (Behar et al., 2009; Ebrahimizadeh et al., 2013; Stijlemans et al., 2004). While these Nbs bind to their respective targets with nanomolar affinity, the identity and precise structure of their epitopes is not clear as animals were immunized with glycoproteins or heterogeneous glycans from natural sources. The combination of immunization with well-defined synthetic glycans and phage, bacteria, or yeast display thus seems promising for the selection of high-affinity Nbs with defined epitopes in future studies.

SIGNIFICANCE

Cell surface glycosylations are critical for proper cell function and the organism's well-being. Nevertheless, our understanding of how exactly glycan structure translates into function is limited. The enormous heterogeneity and structural complexity of native glycans hamper the development of glycan-epitope-specific antibodies that are crucial for deciphering the glycan code. Well-defined synthetic glycans emerge as a key tool for acquiring anti-glycan antibodies. Despite their attractiveness, the number of anti-glycan antibodies that bind pure glycan epitopes is still meager and insufficient.

Using synthetic glycans throughout the entire workflow, we developed single-domain antibodies (nanobodies) against specific glycan epitopes. We induced heavy-chain-only Abs (hcAbs) against different tumor-associated carbohydrate antigens by simultaneously immunizing an alpaca with multiple synthetic glycoconjugates. Glycan-specific hcAbs were isolated by using synthetic glycans, and nanobodies based on these hcAbs were expressed and characterized. Our nanobodies could bind with high specificity to synthetic and native glycan structures on cancer cells.

Our work contradicts the long-standing prevailing opinion that glycan-specific nanobodies cannot be obtained. In addition, our strategy opens a new avenue for further production of glycan-specific tools that are urgently needed in the glycobiology field and can be extended toward functionalized probes for therapeutic application.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Alpaca immunization
  - Bacterial strains and growth conditions for nanobody expression
  - Mammalian cells
- METHOD DETAILS
  - Glycan synthesis
  - Serum and peripheral blood mononuclear cell (PBMC) isolation
  - Alpaca IgG fractionation
  - Glycan array
  - VHH cDNA library construction and high throughput sequencing
  - Preparation of synthetic glycan-coupled beads
  - Isolation of alpaca IgGs using protein A/G beads
  - Isolation of globo-H-specific antibodies using synthetic glycan-coupled beads
  - Papain cleavage of globo-H-specific antibodies
  - Mass spectrometry analysis of affinity-purified VHH fragments
  - Identification of potentially globo-H-specific VHH sequences using mascot
  - Nanobody purification
  - Microscale thermophoresis
  - Nuclear magnetic resonance
  - Flow cytometry
  - Confocal microscopy
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.chembiol.2022.05.007.

ACKNOWLEDGMENTS

The authors thank Prof. Phil Selenko for his careful and critical reading of the manuscript and Katrin Sellrie for her technical support. A.M. thanks Prof. Thomas Peters and the state of Schleswig-Holstein for supplying NMR infrastructure (European Funds for Regional Development, LPW-E/1.1.2/857). This work was supported by Deutsche Forschungsgemeinschaft (RTG2046) (F.G. and J.L.), the Wellcome Trust through core funding for The Wellcome Center for Cell Biology (203149), a Senior Research Fellowship (103139), and the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy, EXC 2008,390540038, UniSysCat. (J.R.), Alexander von Humboldt Foundation (O.M.), and the Max Planck Society.

AUTHOR CONTRIBUTIONS

S.K.K., F.G., K.F., D.W., J.L., and Z.F. produced and characterized the Nbs described in this work. F.G. prepared glycoconjugates and performed glycan arrays. Z.A.C. and J.R. performed the MS analysis and database construction. Next-generation sequencing and cDNA library assembly were done by S.B. and B.T. Bioinformatic analysis was done by F.S. and O.T. A.M. characterized binding by nuclear magnetic resonance spectroscopy. F.G. and O.M. wrote the manuscript, with contributions from S.K.K., A.M. and P.H.S. to the final version. O.M. conceived the original idea and designed the concept and experiments. O.M. and P.H.S. supervised the project.

DECLARATIONS OF INTERESTS

O.M. and P.H.S. are co-founders of Tacalyx GmbH, where P.H.S. is a board member. O.M. and P.H.S. have a significant financial interest in the company and field a patent for the Nbs described in this manuscript (EP3799881A1).

Received: April 26, 2021
Revised: February 21, 2022
Accepted: May 23, 2022
Published: June 14, 2022

SUPPORTING CITATIONS

The following references appear in the supplemental information: Coira et al. (2016); Madeira et al. (2019).

REFERENCES

Adamczyk, B., Tharmalingam, T., and Rudd, P.M. (2012). Glycans as cancer biomarkers. Biochim. Biophys. Acta Gen. Subj. *1820*, 1347–1353. https://doi.org/10.1016/j.bbagen.2011.12.001.

Angulo, J., Langpap, B., Blume, A., Biet, T., Meyer, B., Krishna, N.R., Peters, H., Palcic, M.M., and Peters, T. (2006). Blood group B galactosyltransferase: insights into substrate binding from NMR experiments. J. Am. Chem. Soc. *128*, 13529–13538. https://doi.org/10.1021/ja063550r.

Avci, F.Y., Li, X., Tsuji, M., and Kasper, D.L. (2011). A mechanism for glycoconjugate vaccine activation of the adaptive immune system and its implications for vaccine design. Nat. Med. *17*, 1602–1609. https://doi.org/10.1038/nm.2535.

Behar, G., Chames, P., Teulon, I., Cornillon, A., Alshoukr, F., Roquet, F., Pugnière, M., Teillaud, J.L., Gruaz-Guyon, A., Pèlegrin, A., and Baty, D. (2009). Llama single-domain antibodies directed against nonconventional epitopes of tumor-associated carcinoembryonic antigen absent from nonspecific cross-reacting antigen. FEBS J. *276*, 3881–3893. https://doi.org/10.1111/j.1742-4658.2009.07101.x.

Bremer, E.G., Levery, S.B., Sonnino, S., Ghidoni, R., Canevari, S., Kannagi, R., and Hakomori, S. (1984). Characterization of a glycosphingolipid antigen defined by the monoclonal antibody MBr1 expressed in normal and neoplastic epithelial cells of human mammary gland. J. Biol. Chem. *259*, 14773–14777. https://doi.org/10.1016/s0021-9258(17)42669-x.

Broecker, F., Anish, C., and Seeberger, P.H. (2015). Generation of monoclonal antibodies against defined oligosaccharide antigens. Methods in Mol. Biol. *1331*, 57–80.

Broecker, F., Hanske, J., Martin, C.E., Baek, J.Y., Wahlbrink, A., Wojcik, F., Hartmann, L., Rademacher, C., Anish, C., and Seeberger, P.H. (2016). Multivalent display of minimal Clostridium difficile glycan epitopes mimics antigenic properties of larger glycans. Nat. Commun. *7*, 11224. https://doi.org/10.1038/ncomms11224.

Bushnell, B., Rood, J., and Singer, E. (2017). BBMerge – accurate paired shotgun read merging via overlap. PLoS One *12*, e0185056. https://doi.org/10.1371/journal.pone.0185056.

Cheng, J.Y., Wang, S.H., Lin, J., Tsai, Y.C., Yu, J., Wu, J.C., Hung, J.T., Lin, J.J., Wu, Y.Y., Yeh, K.T., and Yu, A.L. (2014). Globo-H ceramide shed from cancer cells triggers translin- Associated factor X-dependent angiogenesis. Cancer Res. *74*, 6856–6866. https://doi.org/10.1158/0008-5472.can-14-1651.

Coira, E., Barmettler, P., Giamarchi, T., and Kollath, C. (2016). Temperature dependence of the NMR spin-lattice relaxation rate for spin- 12 chains. Phys. Rev. B *94*, 144408.

Collins, B.E., and Paulson, J.C. (2004). Cell surface biology mediated by low affinity multivalent protein-glycan interactions. Curr. Opin. Chem. Biol. *8*, 617–625. https://doi.org/10.1016/j.cbpa.2004.10.004.

Debie, P., Lafont, C., Defrise, M., Hansen, I., van Willigen, D.M., van Leeuwen, F.W.B., Gijsbers, R., D'Huyvetter, M., Devoogdt, N., Lahoutte, T., et al. (2020). Size and affinity kinetics of nanobodies influence targeting and penetration of solid tumours. J. Control. Release. *317*, 34–42.

Disney, M.D., and Seeberger, P.H. (2004). The use of carbohydrate microarrays to study carbohydrate-cell interactions and to detect pathogens. Chem. Biol. *11*, 1701–1707. https://doi.org/10.1016/j.chembiol.2004.10.011.

Duggan, S. (2018). Caplacizumab: first global approval. Drugs *78*, 1639–1642.

Ebrahimizadeh, W., Mousavi Gargari, S., Rajabibazl, M., Safaee Ardekani, L., Zare, H., and Bakherad, H. (2013). Isolation and characterization of protective anti-LPS nanobody against V. cholerae O1 recognizing Inaba and Ogawa serotypes. Appl. Microbiol. Biotechnol. *97*, 4457–4466. https://doi.org/10.1007/s00253-012-4518-x.

Fridy, P.C., Li, Y., Keegan, S., Thompson, M.K., Nudelman, I., Scheid, J.F., Oeffinger, M., Nussenzweig, M.C., Fenyö, D., Chait, B.T., and Rout, M.P. (2014). A robust pipeline for rapid production of versatile nanobody repertoires. Nat. Methods *11*, 1253–1260. https://doi.org/10.1038/nmeth.3170.

Fujitani, N., Furukawa, J. ichi, Araki, K., Fujioka, T., Takegawa, Y., Piao, J., Nishioka, T., Tamura, T., Nikaido, T., Ito, M., et al. (2013). Total cellular glycomics allows characterizing cells and streamlining the discovery process for cellular biomarkers. Proc. Natl. Acad. Sci. U. S. A. *110*, 2105–2110. https://doi.org/10.1073/pnas.1214233110.

De Genst, E., Silence, K., Decanniere, K., Conrath, K., Loris, R., Kinne, J., Muyldermans, S., and Wyns, L. (2006). Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. Proc. Natl. Acad. Sci. U. S. A. *103*, 4586–4591. https://doi.org/10.1073/pnas.0505379103.

Guberman, M., Bräutigam, M., and Seeberger, P.H. (2019). Automated glycan assembly of Lewis type I and II oligosaccharide antigens. Chem. Sci. *10*, 5634–5640. https://doi.org/10.1039/c9sc00768g.

Hahm, H.S., Schlegel, M.K., Hurevich, M., Eller, S., Schuhmacher, F., Hofmann, J., Pagel, K., and Seeberger, P.H. (2017). Automated glycan assembly using the glyconeer 2.1 synthesizer. Proc. Natl. Acad. Sci. U. S. A. *114*, E3385–E3389. https://doi.org/10.1073/pnas.1700141114.

Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hammers, C., Songa, E.B., Bendahman, N., and Hammers, R. (1993). Naturally occurring antibodies devoid of light chains. Nature *363*, 446–448. https://doi.org/10.1038/363446a0.

Huang, C.S., Yu, A.L., Tseng, L.M., Chow, L.W.C., Hou, M.F., Hurvitz, S.A., Schwab, R.B., L Murray, J., Chang, H.K., Chang, H.T., et al. (2020). Globo H-KLH vaccine adagloxad simolenin (OBI-822)/OBI-821 in patients with metastatic breast cancer: phase II randomized, placebo-controlled study. J. Immunother. Cancer *8*, e000342. https://doi.org/10.1136/jitc-2019-000342.

Huang, X., Schurman, N., Handa, K., and Hakomori, S. (2017). Functional role of glycosphingolipids in contact inhibition of growth in a human mammary epithelial cell line. FEBS Lett. *591*, 1918–1928. https://doi.org/10.1002/1873-3468.12709.

Kannagi, R., Levery, S.B., Ishigami, F., Hakomori, S., Shevinsky, L.H., Knowles, B.B., and Solter, D. (1983). New globoseries glycosphingolipids in human teratocarcinoma reactive with the monoclonal antibody directed to a developmentally regulated antigen, stage-specific embryonic antigen 3. J. Biol. Chem. *258*, 8934–8942. https://doi.org/10.1016/s0021-9258(18)32147-1.

Kaplonek, P., Khan, N., Reppe, K., Schumann, B., Emmadi, M., Lisboa, M.P., Xu, F.F., Calow, A.D.J., Parameswarappa, S.G., Witzenrath, M., et al. (2018). Improving vaccines against Streptococcus pneumoniae using synthetic glycans. Proc. Natl. Acad. Sci. U. S. A. *115*, 13353–13358. https://doi.org/10.1073/pnas.1811862115.

Kudryashov, V., Ragupathi, G., Kim, I.J., Breimer, M.E., Danishefsky, S.J., Livingston, P.O., and Lloyd, K.O. (1998). Characterization of a mouse monoclonal IgG3 antibody to the tumor-associated globo H structure produced by immunization with a synthetic glycoconjugate. Glycoconj. J. *15*, 243–249. https://doi.org/10.1023/a:1006992911709.

Wuethrich, K., Kumar, A., Wagner, G., Wüthrich, K., Kumar, A., and Ernst, R.R. (1981). Buildup rates of the nuclear overhauser effect measured by two-dimensional proton magnetic resonance spectroscopy: implications for studies of protein conformation. J. Am. Chem. Soc. *103*, 3654–3658. https://doi.org/10.1021/ja00403a008.

Lai, C.H., Hahm, H.S., Liang, C.F., and Seeberger, P.H. (2015). Automated solid-phase synthesis of oligosaccharides containing sialic acids. Beilstein J. Org. Chem. *11*, 617–621. https://doi.org/10.3762/bjoc.11.69.

Liang, Y.J., Kuo, H.H., Lin, C.H., Chen, Y.Y., Yang, B.C., Cheng, Y.Y., Yu, A.L., Khoo, K.H., and Yu, J. (2010). Switching of the core structures of glycosphingolipids from globo- and lacto- to ganglio-series upon human embryonic stem cell differentiation. Proc. Natl. Acad. Sci. U. S. A. *107*, 22564–22569. https://doi.org/10.1073/pnas.1007290108.

Maass, D.R., Sepulveda, J., Pernthaner, A., and Shoemaker, C.B. (2007). Alpaca (Lama pacos) as a convenient source of recombinant camelid heavy chain antibodies (VHHs). J. Immunol. Methods *324*, 13–25. https://doi.org/10.1016/j.jim.2007.04.008.

Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D., and Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. *47*, W636–W641. https://doi.org/10.1093/nar/gkz268.

Mayer, M., and Meyer, B. (2001). Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor. J. Am. Chem. Soc. *123*, 6108–6117. https://doi.org/10.1021/ja0100120.

Menard, S., Tagliabue, E., Canevari, S., Fossati, G., and Colnaghi, M.I. (1983). Generation of Monoclonal Antibodies Reacting with Normal and Cancer Cells of Human Breast. Cancer Res *43*, 1295–1300.

Meyer, B., and Peters, T. (2003). NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. Angew. Chemie - Int. Ed. *42*, 864–890. https://doi.org/10.1002/anie.200390233.

Oh-Uti, K. (1949). Polysaccharides and a glycidamin in the tissue of gastric cancer. Tohoku J. Exp. Med.

Pardo-Vargas, A., Delbianco, M., and Seeberger, P.H. (2018). Automated glycan assembly as an enabling technology. Curr. Opin. Chem. Biol. *46*, 48–55.

Rugo, H.S., Chow, L.W.C., Cortes, J., Fasching, P.A., Hsu, P., Huang, C.-S., Kim, S.-B., Lu, Y.-S., Melisko, M.E., Nanda, R., et al. (2020). Phase III, randomized, double-blind, placebo-controlled study to evaluate the efficacy and safety of adagloxad simolenin (OBI-822) and OBI-821 treatment in patients with early-stage triple-negative breast cancer (TNBC) at high risk for recurrence. J. Clin. Oncol. *38*, TPS599. https://doi.org/10.1200/jco.2020.38.15_suppl.tps599.

Sterner, E., Flanagan, N., and Gildersleeve, J.C. (2016). Perspectives on anti-glycan antibodies gleaned from development of a community resource database. ACS Chem. Biol. *11*, 1773–1783. https://doi.org/10.1021/acschembio.6b00244.

Stijlemans, B., Conrath, K., Cortez-Retamozo, V., Van Xong, H., Wyns, L., Senter, P., Revets, H., De Baetselier, P., Muyldermans, S., and Magez, S. (2004). Efficient targeting of conserved cryptic epitopes of infectious agents by single domain antibodies: african trypanosomes as paradigm. J. Biol. Chem. *279*, 1256–1261. https://doi.org/10.1074/jbc.m307341200.

Werz, D.B., Castagner, B., and Seeberger, P.H. (2007). Automated synthesis of the tumor-associated carbohydrate antigens Gb-3 and Globo-H: incorporation of α-galactosidic linkages. J. Am. Chem. Soc. *129*, 2770–2771. https://doi.org/10.1021/ja069218x.

Wienken, C.J., Baaske, P., Rothbauer, U., Braun, D., and Duhr, S. (2010). Protein-binding assays in biological liquids using microscale thermophoresis. Nat. Commun. *1*, 100. https://doi.org/10.1038/ncomms1093.

Xu, L., Song, X., and Jia, L. (2017). A camelid nanobody against EGFR was easily obtained through refolding of inclusion body expressed in Escherichia coli. Biotechnol. Appl. Biochem. *64*, 895–901. https://doi.org/10.1002/bab.1544.

Zhang, S., Zhang, H.S., Cordon-Cardo, C., Reuter, V.E., Singhal, A.K., Lloyd, K.O., and Livingston, P.O. (1997). Selection of tumor antigens as targets for immune attack using immunohistochemistry: II. Blood group-related antigens. Int. J. Cancer *73*, 50–56. https://doi.org/10.1002/(sici)1097-0215(19970926)73:1<50::aid-ijc9>3.0.co;2-0.

Zou, X.P., Qin, C.J., Hu, J., Fu, J.J., Tian, G.Z., Moscovitz, O., Seeberger, P.H., and Yin, J. (2020). Total synthesis of D-glycero-D-mannno-heptose 1β, 7-bisphosphate with 3-O-amyl amine linker and its monophosphate derivative. Chin. J. Nat. Med. *18*, 628–632. https://doi.org/10.1016/s1875-5364(20)30075-3.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-Llama FITC	Invitrogen Thermo Fisher Scientific	Cat# A16061; RRID: AB_2534734
VK9	Thermo Fisher Scientific	Cat# 14-9700-82; RRID: AB_2572961
Anti-6xHis-Atto647N	Rockland	Cat# 200-356-382; RRID: AB_2611811
anti-mouse-IgG-AlexaFluor635	Invitrogen Thermo Fisher Scientific	Cat# A-31574; RRID: AB_2536184
anti-mouse-IgG3 (γ3)-AlexaFluor488	Invitrogen Thermo Fisher Scientific	Cat# A21151; RRID: AB_2535784
Bacterial and virus strains		
E.coli SHuffle	New England Biolabs Inc.	Cat# C3026J
Biological samples		
Alpaca Serum	This paper	N/A
Peripheral Blood Mononuclear Cells (PBMCs)	This paper	N/A
Chemicals, peptides, and recombinant proteins		
CRM-197	Reagent Proteins	Cat# 50-198-13
Alhydrogel adjuvant	Brenntag	Cat# 5499
L-cysteine	Sigma-Aldrich	Cat# 168149
Trypsin MS grade	Pierce™ Thermo Fisher Scientific	Cat# 90058
Protein A columns	Pierce™ Thermo Fisher Scientific	Cat# 20356
Protein G columns	Pierce™ Thermo Fisher Scientific	Cat# 89927
A/G beads	Pierce™ Thermo Fisher Scientific	Cat# 88802
Papain-immobilized beads	Thermo Fisher Scientific	Cat# 20341
Kanamycin	Carl Roth	Cat# T832.2
IPTG	Carl Roth	Cat# CN08.4
DMEM	Pan Biotech	Cat# P04-03500
Ham's F12 Nutrient Mix	Gibco Thermo Fisher Scientific	Cat# 21765-029
Fetal calf serum	Pan Biotech	Cat# P30-3031
Horse serum	Invitrogen	Cat# 16050-122
Glutamine	Pan Biotech	Cat# P04-82100
Penicillin/Streptomycin	Pan Biotech	Cat# P06-07100
Epidermal growth factor	Peprotech	Cat# AF-100-15
Insulin	Sigma Aldrich	Cat# I-1882
Hydrocortisone	Sigma Aldrich	Cat# H-0888
Cholera toxin	Sigma Aldrich	Cat# C-8052
Trypsin/EDTA	Pan Biotech	Cat# P10-020100
SulfoLink Coupling Resin	Thermo Fisher Scientific	Cat# 20404
Superdex 200 Increase 10/300	GE Healthcare	Cat# 28-9909-44
Superdex 75 16/600	GE Healthcare	Cat# 28-9893-33
Monolith RED-tris-NTA 2nd Generation dye	Nanotemper Tech, Munich, Germany	Cat# MO-L018
Several nbs (incl. GH46)	This paper	N/A
Critical commercial assays		
RNeasy MiniKit	Qiagen	Cat# 74104
SuperScriptII reverse transcriptase kit	Invitrogen	Cat# 18064-014
Experimental models: Cell lines		
Human: MCF7	Purchased and Authenticated by ATCC	ATCC HTB-22
Human: HEK293T	ATCC	ATCC CRL-3216
Human: MCF10A	Purchased and Authenticated by ATCC	ATCC CRL-10317

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Organisms/strains		
Alpaca (*Vicugna pacos*)	Zadik-Lamas, Märkischer Lamahof	N/A
Oligonucleotides		
Oligo-dT primers	Thermo Fisher Scientific	Cat# SO131
CALL001: GTCCTGGCTGCTCTTCTACAAGG	Fridy et al., 2014	N/A
CALL002: GGTACGTGCTGTTGAACTGTTCC	Fridy et al., 2014	N/A
VHH_For: GGACTAGTGCGGCCGCTGGAGACGGTGA CCTGG	Fridy et al., 2014	N/A
VHH_back: GATGTGCAGCTGCAGGAGTCTGGRGGAGG	Fridy et al., 2014	N/A
Deposited data		
Next generation sequencing	This paper	Sequence read archive, NCBI Accession number: PRJNA830746
Recombinant DNA		
pET28b(+) with VHH sequences	Synbio Technologies (New Jersey, USA)	N/A
Software and algorithms		
Clustal Omega	EMBL-EBI	www.ebi.ac.uk/tools/msa/clustalo
GenePix Pro7	Molecular Devices	www.moleculardevices.com
BBMerge	Bushnell et al., 2017	https://sourceforge.net/projects/bbmap/
Mascot ^{TM} search engine	Matrix Science, MA, USA	www.matrixscience.com
Llama Magic v1.0	Fridy et al., 2014	http://github.com/fenyolab/llama-magic
Python 3.5.4	Python^{TM}	www.python.org
Biopython 1.68	Python^{TM} Package Index	http://pypi.org/
PythonORF	Daofeng Li	http://Github.com/lidaof/PythonORF
Pandas 0.22.0	pandas 0.22.0 documentation	http://pandas.pydata.org
MO.Affinity Analysis	Nanotemper Tech	http://nanotempertech.com
TopSpin 4.0.6	Bruker	www.bruker.com
FlowJo v7.6.5	FlowJo LLC, Ashland, OR, USA	www.flowjo.com
OriginPro v2021.b	OriginLab	www.originlab.com
ZEN	Zeiss, Germany	https://www.zeiss.com/microscopy/int/products/microscope-software/zen.html
Other		
Globo-H	Werz et al., 2007	N/A
KH-1	Guberman et al., 2019	N/A
Tn	Lai et al., 2015	N/A
Sialyl-Tn	Lai et al., 2015	N/A
GMH7P	Zou et al., 2020	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Oren Moscovitz (oren.moscovitz@mpikg.mpg.de).

Materials availability

Nanobodies or nanobody plasmids generated in this work can be made available only after completion of an MTA. Requests for nanobodies or plasmids should be directed to the lead contact.

Data and code availability

All analyzed data supporting the findings of this study are available within the paper and the supplemental information. Next generation sequencing data is deposited at the Sequence Read Archive (NCBI, accession number: PRJNA830746). Additional raw data will be shared by the lead contact upon request. No original code was generated in this study. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Alpaca immunization

CRM glycoconjugates were prepared and characterized as described previously (Kaplonek et al., 2018) (Figure S1). The alpaca was treated strictly according to German and European Law. The Office for Health and Social Affairs Potsdam, Brandenburg (LAVG) approved the experiment (Permit 2347-A-30-1-2018). An adult male alpaca at the age of 10 was immunized by simultaneous injection of five different glycoconjugates in phosphate-buffered saline, mixed 1:1 (v:v) with aluminum hydroxide adjuvant. The alpaca was injected subcutaneously along the base of the neck for a total of six consecutive immunization rounds (days 0, 7, 14, 28, 37, 45). Whole blood (300mL) was extracted for serum and PBMC isolation eight days after the last immunization (day 53). After each immunization, 10 mL of blood were drawn for assessment of specific anti-glycan antibodies development by glycan array.

Bacterial strains and growth conditions for nanobody expression

E. coli SHuffle cells (NEB, Ipswich, MA, USA) were used for expression of nanobody plasmids. Transformed bacteria were grown at 37°C, 250 rpm in terrific broth in the presence of 50 μg/mL kanamycin to OD = 2 before induction with 0.5 mM IPTG overnight at 18°C.

Mammalian cells

MCF7 cells and HEK cells were cultured in DMEM + 10% fetal calf serum + 2 mM glutamine + 1x penicillin/streptomycin at 37°C, 5% CO_2. MCF10A cells were cultured in DMEM/F12 (1:1) + 5% horse serum +20 ng/mL EGF +0.5 mg/mL hydrocortisone +100 ng/mL cholera toxin +10 μg/mL insulin +2 mM glutamine + 1x penicillin/streptomycin at 37°C, 5% CO_2. All cell lines were passaged and diluted every 3–4 days when confluent using trypsin/EDTA. All cell lines were tested for mycoplasma contamination on a monthly basis. All cell lines were originally derived from female donors.

METHOD DETAILS

Glycan synthesis

All linker-bound synthetic glycans used in this work were synthesized from protected building blocks using established solution-phase protocols or automated glycan assembly (Guberman et al., 2019; Lai et al., 2015; Werz et al., 2007; Zou et al., 2020).

Serum and peripheral blood mononuclear cell (PBMC) isolation

Whole blood for serum isolation was collected into 50 mL tubes and incubated at RT for 1 h followed by centrifugation at 2,000 g/10 min/RT. Serum was aliquoted into sterile 15 mL tubes (7.5 mL each), and flash-frozen for further storage at −80°C. For PBMC isolation, 100 mL whole blood was collected into 10 mL EDTA coated tubes (BD Vacutainer™) and inverted ten times. PBMCs from undiluted whole blood were isolated using SepMate™-50 (STEMCELL Technologies. Inc, Vancouver, Canada), dissolved in RNAlater™ to $5*10^7$ cells/mL and stored at −80 °C.

Alpaca IgG fractionation

Six mL of glycan-immunized alpaca serum were mixed in a 1:1 ratio with 20 mM sodium phosphate buffer and fractionated by differential absorption on Protein A and Protein G columns (Pierce) (Hamers-Casterman et al., 1993).The different IgG subtypes were polished and buffer-exchanged to PBS by size exclusion chromatography with a Superdex 200 10/300 column (Cytiva).

Glycan array

Synthetic glycans (0.2 mM) were printed in the lab on glass slides coated with epoxy for amine/thiol-functionalized glycans (Disney and Seeberger, 2004) or *N*-hydroxyl succinimide ester for amine-functionalized glycans (codelink®, SurModics Inc, USA), using a piezoelectric spotting device (S3; Scienion, Germany). After 24 h in a humid chamber at room temperature, slides were quenched with 50 mM ethanolamine solution (pH 9) for 1 h at 50°C. Next, the slides were assembled onto microplate holders and wells were blocked in PBS +1% BSA for 1 h. Blocked wells were incubated with 50 μL alpaca serum (1:50) or PBS for 1 h at RT. For examination of IgG subclasses, wells were incubated with 25 μL of fractions of IgG1, 2 and 3 (0.2 mg/mL) in duplicates in three independent experiments. Wells were washed three times with PBS and incubated with 25 μL anti-llama FITC Ab (1:200, Invitrogen) for 1 h at RT in the dark. Wells were washed three times with PBST, once with ddH_2O, and dried by centrifugation (100 g, 2 min). After removal of the multi-well grid, the slide was scanned with an Axon GenePix® 4300A scanner (Molecular Devices, USA) and results were analyzed using GenePix Pro7 (Molecular Devices).

VHH cDNA library construction and high throughput sequencing

RNA was isolated from alpaca PBMCs using RNeasy MiniKit (Qiagen, Hilden, Germany). cDNA was created using oligo-dT primers and the SuperScriptII reverse transcriptase kit (Invitrogen, Carlsbad, CA, USA). The VHH variable regions were specifically amplified from the cDNA in a two-step nested PCR. In the first step, primers CALL001: GTCCTGGCTGCTCTTCTACAAGG (leader sequence specific) and CALL002: GGTACGTGCTGTTGAACTGTTCC (CH2 specific) were used. After gel purification and size selection of the 600 bp region followed by MinElute cleanup (Qiagen), the second PCR was performed to extract the VHH region using the primers VHH_back: GATGTGCAGCTGCAGGAGTCTGGRGGAGG and VHH_For: GGACTAGTGCGGCCGCTGGAGACGGTGACCTGG followed by gel extraction of the 400 bp band. Illumina sequencing libraries were created following the TruSeq protocol without shearing. Each library was sequenced with 2 × 250 bp chemistry on an Illumina MiSeq instrument yielding 6.9/6.6 million read pairs per sample. BBmerge (Bushnell et al., 2017) was used to overlap forward and reverse reads to reconstruct the original PCR fragments.

Preparation of synthetic glycan-coupled beads

Thiol-functionalized Globo-H (1 mg) was reduced with 0.6 eq. of resin-bound TCEP in 120 μL water for 1 h at 1,500 rpm/RT. The resin was removed by filtration and the filter (0.22 μm) was washed five times with 50 μL water. The filtrate and all wash fractions were combined and lyophilized. Reduced Globo-H was conjugated to 1 mL SulfoLink® Coupling Resin (Thermo Fisher Scientific) according to manufacturer instructions.

Isolation of alpaca IgGs using protein A/G beads

To isolate conventional IgGs and hcAbs from alpaca serum, 7.5 mL of serum were diluted with 67.5 mL sodium phosphate (20 mM, pH 7) and filtered using 0.22 μm filter. Protein A/G Magnetic Beads (3 mL, Pierce, Waltham, MA, USA) were pre-washed with 20 mM sodium phosphate (pH 7). Thereafter, these samples were incubated for 1 h at RT with spinning. After collecting the flow-through, the beads were washed with 100 mL of 20 mM sodium phosphate (pH 7). Bound antibodies were eluted using 39 mL of 100 mM citric acid (pH 2.75) into 50 mL tubes pre-filled with 11 mL of 1.5 M Tris (pH 8.8) for neutralization of eluted antibodies. Elution samples were united, concentrated, and dialyzed overnight with 5 L sodium phosphate (20 mM, pH 7.4).

Isolation of globo-H-specific antibodies using synthetic glycan-coupled beads

Globo-H-coupled beads or empty beads (control) were washed with water and sodium phosphate (20 mM, pH 7). Antibodies eluted from protein A/G beads were then evenly distributed between columns containing pre-equilibrated empty and glycan-coupled beads and rotated for 1 h at RT. Flow-through from each column was switched between columns for 1 h at RT. Elution from both columns was performed using 20 mM sodium phosphate (pH 7) and increased NaCl concentration (100 mM/500 mM/1 M/3 M), followed by washing with sodium phosphate (20 mM, pH 5.2). 75 mL of each washing step from both columns were concentrated using an Amicon Ultra Centrifugal Filter (MWCO = 10 kDa) and dialyzed overnight at 4°C against papain digestion buffer (20 mM sodium phosphate/10 mM EDTA pH 7.1).

Papain cleavage of globo-H-specific antibodies

Affinity-purified antibodies in 20 mM sodium phosphate/10 mM EDTA buffer were incubated with 10–20 μL of papain-immobilized beads in solution (Thermo Fisher Scientific, Waltham, MA, USA) 37°C/300 rpm for up to 90 min with the addition of 5 mM L-cystein (Sigma-Aldrich). The mixture was loaded on 4–20% acrylamide SDS-PAGE and 13–15 kDa bands corresponding to the VHH domains were excised.

Mass spectrometry analysis of affinity-purified VHH fragments

VHH fragments from SDS-PAGE separation were excised from the gel and digested with trypsin. Tryptic peptides were extracted and de-salted using C18 StageTips. LC-MS/MS analysis was performed on an Orbitrap Fusion Lumos Tribrid™ mass spectrometer (Thermo Fisher Scientific) coupled on-line to Ultimate 3000 RSLCano Systems™ (Dionex, Thermo Fisher Scientific). The eluted peptides were ionized by an EASY-Spray source (Thermo Fisher Scientific). The MS data was acquired in the data-dependent mode with the top-speed option using the Lumos instrument with high resolution for both MS1 and MS2 spectra. HCD was applied for fragmentation. Peak lists of MS2 data were generated using MS convert software. Database search against the high-throughput cDNA sequencing library was conducted using a combination of the Mascot™ search engine (Matrix Science, MA, USA) modified and Llama Magic v1.0 software with default parameters.

Identification of potentially globo-H-specific VHH sequences using mascot

A Nb protein sequence library was generated from the cDNA sequencing data and utilized in the MascotServer version 2.3 for MS/MS Ions Search. Nucleic acid (NA) sequences were filtered (max 10% unknown residues; min 300 bp length) and duplicates were removed. All open reading frames were identified and translated into their corresponding amino acid (AA) sequences. The presence of (Q/H)VT or (Q/H)XTV after sequence position 80 was used to detect valid Nb sequences and any C-terminal artefact TVX was replaced with TVS according to the consensus Nb sequence. The results were filtered (max one unknown residue; min 105 AA length) and duplicates were removed.

Mascot search parameters comprised monoisotopic mass, 6 ppm tolerance, charge 2+ and 3+, Trypsin cleavage, and two variable modifications (Carbamidomethylation of C; Oxidation of M). The identified peptides and protein hits were obtained in CSV file format for Nb candidate identification.

The individual Mascot result files were grouped into the respective sample of interest and all others (other samples and controls), respectively. All protein hits identified in controls or other samples were removed from the actual sample, because they were potentially not specific to the sample of interest. Any duplicate sequences were removed. The generated candidate sequences were further analyzed and annotated for manual inspection. Peptide matches, the three CDRs, and all cysteine residues were highlighted. The CDR identification and Llama Magic score calculation was implemented with the algorithm established by (Fridy et al., 2014). The annotated protein sequence, protein score from Mascot, Llama Magic score, protein coverage, protein length, CDR coverages, CDR lengths, number of NA and AA sequence duplicates, and Cysteine residue count were included in the final table. The data was manually inspected (recommended view filter to start: min 20% protein coverage; min protein score of 80; greater than 0% CDR3 coverage; min Llama Magic score of 10; sort by decreasing CDR3 length). The final set of candidates was selected for synthesis based on sequence diversity, highest CDRs coverage, and maximal CDR3 length. Python 3.5.4 with biopython 1.68, PythonORF and pandas 0.22.0 was used as environment for sequence data preparation and analysis.

Nanobody purification

VHH sequences containing C-terminal His tags were synthesized by Synbio Technologies (New Jersey, USA), and plasmids were transformed into *E. coli* SHuffle cells (NEB, Ipswich, MA, USA). After Nb expression (as mentioned above), the bacteria were lysed using a French press and Nbs were purified using standard nickel beads and size exclusion chromatography on an ÄKTApurifier with a Superdex S200 or S75 column (GE Healthcare). Nb expression and purity once purified were examined via SDS-PAGE. Trimeric $GH46_3$ expressed in inclusion bodies and was isolated, purified and refolded as described elsewhere (Xu et al., 2017).

Microscale thermophoresis

Non-linked synthetic Globo-series glycans were purchased from Elicityl Oligotech (Crolles, France). Nb (200 μL of 100 nM) was incubated with 10 pmol of Monolith RED-tris-NTA 2nd Generation dye (Nanotemper Tech, Munich, Germany) for 30 min in the dark. A dilution series of the glycan was created in PBS-T (16 different dilutions, ranging from 4 mM to 122 nM, 5 μL per concentration). Each glycan dilution was mixed with 5 μL of labeled Nb and loaded into a Monolith NT.115 capillary (Nanotemper Tech). MST signals were recorded at 25°C, 40% MST power and 70–100% LED power in a Monolith NT.115 (Nanotemper Tech) and analyzed using MO.Affinity Analysis (v2.3, Nanotemper Tech). For K_D determination, measurements were repeated in at least three independent experiments.

Nuclear magnetic resonance

NMR experiments were acquired on a Bruker Avance III HD 600 MHz NMR spectrometer equipped with a TCI cryogenic probe. NMR spectra were processed in TopSpin 4.0.6. Peak volumes were extracted using CCPNMR Analysis 2.4.2 software suite for the quantification of TRNOE build-up rates. Chemical shifts were measured relative to 3-(trimethylsilyl)-1-propanesulfonic acid-d_6 (DSS-d_6).

Preliminary STD NMR (Mayer and Meyer, 2001) experiments of GH46:Globo-H complex with 800 scans and selective saturation of protein resonances at 7 and -0.1 ppm showed that GH46 can be saturated uniformly without significant direct irradiation over the carbohydrate at 7 ppm. Investigation of the time dependence of the saturation transfer with saturation times from 1.0 to 4.0 s showed that 4 s afforded the best results. Therefore, STD NMR final spectra were acquired at 277 K using a train of 50 ms Gaussian-shaped radio frequency pulses separated by 1 ms for a total duration of 4 s. An attenuation of 45 dB was chosen, resulting in a 380.0 flip angle. Protein signals were attenuated applying a 30 ms spinlock filter before acquisition. The acquisition time was set at 1.95 s with an additional relaxation delay of 5 s. On- and off-resonances were set at 7 ppm and 200 ppm, respectively, with a total of 6000 scans. Ligand epitopes were determined by calculating STD amplification factors (AF) from Equation 1:

$$STD - AF = \frac{I_0 - I_{sat}}{I_0} \quad \text{(Equation 1)}$$

where I_0 and I_{sat} are the signal intensities in the off- and on-resonance spectra, respectively. STD-AFs were normalized using the largest STD effect (anomeric proton H1 of the fucose unit, 100%) as reference. Normalized STD-AFs were color-coded according to 3 categories: 100 - 67% strong (red), <67 - 33% medium (yellow) and <33 - 0% weak (blue). NMR samples were prepared in 40 mM sodium phosphate buffer (pH 8.0), 150 mM NaCl, 200 μM DSS-d_6 and 0.02% NaN_3 in D_2O (>99.9%) and contained 4 mM Globo-H or 4 mM GB5 in the presence or absence of 60 μM or 110 μM GH46, respectively.

Transferred NOE experiments for the determination of trNOE build-up rates were recorded with 1024 points and 128 t1 increments, a relaxation delay of 2.3 s and with a total of 8 scans. The frequency offset and spectral width were set at 4.70 ppm and 10.0 ppm for the both dimensions, and the residual water signal was suppressed via excitation sculpting. Data processing were performed by zero filling to 256 points in F1 to give a final 1024 · 256 matrix. The optimal conditions for the transferred NOESY measurements were determined by considering three factors: a) Temperatures ranging from 277 k to 315 K. At 298 K NOESY cross-peaks from the free Globo-H are minimized relative to the complex, as Globo-H approaches the $\omega\tau_c = 1$ condition (Figures S4E and S4F), and the protein remains stable during the acquisition of the NMR spectra. Therefore, 298 K was selected as experimental temperature for transferred NOESY experiments. b) The GH46:Globo-H molar ratio, which ranged from 1:4.5 to 1:16.4 (Angulo et al., 2006). c) Mixing

times (τ_m) were 20, 40, 70, 100, 200, 300, 400, 500 and 700 ms. The faster rate of build up for the pairs (Fuc-CH_3/Fuc-H5 and Fuc-H1/Gal5-H2) in the bound state *vs.* the free state, as shown in (Figure 4C), indicated that the effects are representative of the bound state (Kumar et al., 1981). NMR samples contained 3 mM Globo-H and 180 μM GH46 (ligand:protein ratio 16.4:1) or 20 mM Globo-H in the same buffer as described for STD NMR experiments.

Flow cytometry

MCF7 or MCF10A cells were trypsinized and 0.5 × 10^6 cells per sample were transferred into 1.5 mL tubes. After centrifugation for 5 min at 300 *g* the cell pellet was incubated with Nb (0.4 mg/mL in PBS) at RT for 45 min VK9 (5 μg/mL VK9) was included as control. Afterwards, cells were washed twice with ice-cold PBS, pellets were incubated with α6xHis-Atto647N Ab (1:500, Rockland) or anti-mouse-IgG-AlexaFluor635 Ab (1:400, for VK9 sample, Invitrogen). After three more washes with PBS, the cells were analyzed by flow cytometry. Data was collected on a FACSCantoII (BD Biosciences, San Jose, CA, USA) and analyzed using FlowJo (v7.6.5, FlowJo LLC, Ashland, OR, USA), first selecting intact cells (FSC-A over SSC-A), then gating for single cells (FSC-A over FSC-H). Nb binding was examined in histograms of the APC channel.

For co-staining experiments, cells were incubated with Nb and VK9 simultaneously at the same concentrations used in single-staining experiments. For detection, a mixture of α6xHis-Atto647N Ab (1:500, Rockland) and anti-mouse-IgG-AlexaFluor488 Ab (1:400, for VK9 sample, Invitrogen) was used and double staining was examined in 2D dot plots.

For apparent K_D measurements, different concentrations of GH46, $GH46_3$ or VK9 were incubated with MCF7 cells. Based on the mean fluorescence intensity, relative binding was calculated and plotted over the logarithmic protein concentration. Apparent K_D values were determined by sigmoidal fitting in Origin (v2021.b, OriginLab).

Confocal microscopy

To immunostain MCF7 cells, 20,000 cells were seeded on 12 mm coverslips in 24-well plates 48 h before the assay. Cells were rinsed with PBS and fixed with 4% PFA. After blocking for 1 h in 1% BSA, coverslips were placed onto drops of Nb solution or VK9 on parafilm in a humidified chamber. After 1 h, they were placed back and incubated with 200 μL of Atto 647 anti-His Ab (1:5,000 in PBS +1% BSA, Rockland Inc.) for 1 h at RT. The VK9 positive control was stained with mouse-IgG specific AlexaFluor635 secondary Ab (1:400) in PBS + 1 % BSA. Then, coverslips were flipped onto drops of mouting medium on a microscopy slide. All microscopy was performed using an Axio Imager.M2 confocal LSM 800 (Zeiss). All images were acquired at 22% laser power, pinhole = 80 μm. Images of GH46 and the control nb are displayed using the same brightness/contrast. Different brightness/contrast settings were chosen for the display of VK9 images due to the high intensity of the staining. A different pseudocolor was selected for VK9 to emphasize the distinct display settings and the fact that VK9 and GH46/control nb were detected with different secondary Abs.

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analysis was carried out using OriginPro (version 2021b). For each experiment where analysis was performed, statistical details can be found in the figure legends.