nature communications

Article

https://doi.org/10.1038/s41467-024-53738-9

Mammalian cell-based production of glycans, glycopeptides and glycomodules

Received: 23 August 2024

Accepted: 22 October 2024

Published online: 08 November 2024

Check for updates

Thapakorn Jaroentomeechai[1], Richard Karlsson[1], Felix Goerdeler[1], Fallen Kai Yik Teoh[1], Magnus Nørregaard Grønset[1], Dylan de Wit[1], Yen-Hsi Chen[2], Sanae Furukawa[1], Venetia Psomiadou[3], Ramon Hurtado-Guerrero[1,4,5], Elena Ethel Vidal-Calvo[6,7], Ali Salanti[6], Thomas J. Boltje[3], Leendert J. van den Bos[8], Christian Wunder[9], Ludger Johannes[9], Katrine T. Schjoldager[1], Hiren J. Joshi[1], Rebecca L. Miller[1], Henrik Clausen[1], Sergey Y. Vakhrushev[1] & Yoshiki Narimatsu[1,2]

Access to defined glycans and glycoconjugates is pivotal for discovery, dissection, and harnessing of a range of biological functions orchestrated by cellular glycosylation processes and the glycome. We previously employed genetic glycoengineering by nuclease-based gene editing to develop sustainable production of designer glycoprotein therapeutics and cell-based glycan arrays that display glycans in their natural context at the cell surface. However, access to human glycans in formats and quantities that allow structural studies of molecular interactions and use of glycans in biomedical applications currently rely on chemical and chemoenzymatic syntheses associated with considerable labor, waste, and costs. Here, we develop a sustainable and scalable method for production of glycans in glycoengineered mammalian cells by employing secreted Glycocarriers with repeat glycosylation acceptor sequence motifs for different glycans. The Glycocarrier technology provides a flexible production platform for glycans in different formats, including oligosaccharides, glycopeptides, and multimeric glycomodules, and offers wide opportunities for use in bioassays and biomedical applications.

Glycans are ubiquitous across all domains of life and play pivotal biological roles[1]. Many functions assigned to glycans are mediated through binding and recognition by glycan-binding proteins (GBPs), including lectins, receptors, toxins, microbial adhesins, antibodies, and enzymes[2,3]. More than 200 distinct human GBPs and a larger number of glycosaminoglycan (GAG) binding proteins are identified so far, and these serve in intracellular trafficking, clearance and turnover of proteins, cell adhesion, signaling, and pathogen recognition and defense[2,4,5]. A vast number of microbial and viral GBPs are involved in host-microbial interactions including decisive steps in pathogenicity and healthy homeostasis of the mucus and microbiomes[2,5,6]. Glycans exhibit great structural diversity to support these functions[3], and the

[1]Copenhagen Center for Glycomics, Departments of Cellular and Molecular Medicine, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark. [2]GlycoDisplay ApS, Copenhagen, Denmark. [3]Synthetic Organic Chemistry, Institute for Molecules and Materials, Radboud University Nijmegen, Nijmegen, The Netherlands. [4]Institute of Biocomputation and Physics of Complex Systems, University of Zaragoza, Zaragoza, Spain. [5]Fundación ARAID, Zaragoza, Spain. [6]Centre for Translational Medicine and Parasitology, Department for Immunology and Microbiology, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen University Hospital, Copenhagen, Denmark. [7]VAR2 Pharmaceuticals ApS, Copenhagen, Denmark. [8]EnzyTag BV, Nuth, The Netherlands. [9]Institut Curie, Cellular and Chemical Biology Unit, PSL Research University, U1143 INSERM, UMR3666 CNRS, Paris, France. e-mail: yoshiki@sund.ku.dk

diversity of glycan-binding epitopes is expanded when their presentation on lipids and proteins is included in the recognition by GBPs[7]. Knowledge of the natural glycan ligands for GBPs and the molecular basis for recognition is prerequisite for understanding biological functions of glycans and harnessing their potential biomedical uses.

Discovery of GBPs and their glycan-binding specificities commonly utilize glycan microarrays. Printed glycan arrays (oligosaccharides or neoglycolipids) typically display 100s of different glycans that have been isolated or produced by chemical and chemoenzymatic synthesis usually in scarce quantities sufficient for microarrays[8–13]. These printed arrays display glycans without their natural context of glycoconjugates, although synthetic glycopeptide arrays partly mitigate this void[14,15]. Cell-based glycan arrays comprised of libraries of isogenic cells genetically engineered with loss/gain of distinct glycosylation capacities represent recent alternatives that are useful to display glycans in their natural context on glycolipids, glycoproteins, and proteoglycans on the cell surface[16,17]. Studies with these glycan arrays inform of structural features of the epitope(s) bound by GBPs, and to understand the biology and support development of biomedical use of glycans, e.g., as inhibitors, further studies are needed to gain insights into how the epitope(s) is formed and recognized by GBPs.

Structural molecular studies of the recognition of glycans by GBPs require access to glycans in quantities and appropriate formats. Moreover, glycans have substantial potential as therapeutics ranging from e.g., human-milk-oligosaccharides (HMOs)[18] in infant milk formulas for prevention of infections to mucin O-glycans to prevent biofilm formation and suppress virulence of microbiota[19–21]. Currently, access to glycans is largely dependent on chemical and chemoenzymatic synthesis with the inherent challenges of complexity, labor, cost, and waste products. There is therefore a critical need to develop sustainable production methods to improve access to glycans in forms suitable for different purposes. The genetic regulation of the many different types of glycosylation pathways in mammalian cells are well established[22], and we have extensively explored and validated stable genetic glycoengineering designs with combinatorial knockout/knockin (KO/KI) of glycosyltransferase genes to customize cellular glycosylation for desirable glycosylation capacities[7,23]. Genetic glycoengineering of mammalian cells has enabled recombinant production of custom-designed glycoprotein therapeutics, which has resulted in new drugs and drug designs, e.g., IgG antibodies with improved effector functions[24], and lysosomal replacement enzymes with improved circulation and biodistribution[25,26], but also analytical opportunities for glycoproteins, e.g., native MS[25,27]. Mammalian cells with stable expression of exogenous gene(s) provide a sustainable method to produce biologics, and the Chinese hamster ovary cell (CHO) already serves as the main cell factory for production of a wide range of biologics used in the clinic[28]. Production of glycans in mammalian cells is, however, challenged by these being built almost entirely on lipids and specific proteins.

Here, we seek to develop a cost-effective and sustainable method to produce glycans in scalable quantities and flexible formats to support wider use of glycans in basic science and biomedicine. To this end, we develop artificial (i.e., without protein identity) Glycocarriers designed to acquire distinct types of glycans (GalNAc-type O-glycan, N-glycan, or glycosaminoglycan) when recombinantly expressed as secreted products in mammalian cells. Glycocarriers are designed with a flexible glycomodule comprised of tandem repeat peptide motifs for attachment of multiple glycans to amplify yields (mol/mol) and trypsin and Tobacco Etch Virus (TEV)-cleavage sites to generate flexible formats of glycan products, as released glycans, as glycans on short glycopeptides, and as glycan multimers on intact glycomodules. The Glycocarriers are designed with a globular protein module to drive high expression and secretion, and we employ glycoengineered CHO and HEK293 cells to produce Glycocarriers with distinct types and structures of attached glycans in high yields. We demonstrate that Glycocarriers acquire stoichiometric attachment of glycans, and that the glycans built on Glycocarriers are faithful to the cell engineered glycosylation design. We illustrate that stable expression of Glycocarriers in glycoengineered cells can be used to produce individual glycans homogeneously in high quantity (up to 50 mg/L in lab scale production). We provide representative examples of production of complex O-glycans, N-glycans, and GAGs, and highlight that the repertoire of glycans that can be produced by Glycocarriers is easily expandable by employing further glycoengineering of cells. We highlight examples for how the flexible glycan format provided by Glycocarriers can advance analytics (e.g., single molecule mass photometry), serve as probes (e.g., cellular uptake/trafficking), and provide building blocks (e.g., de novo synthesis of glycoproteins). We envision that the quantities and flexible format of glycans provided by the Glycocarrier platform will accelerate wider use and applications of glycans.

Results

To develop a versatile platform for producing human glycans, we designed Glycocarriers that would acquire specific types of glycans, notably without specific protein identity, when expressed as secreted proteins in glycoengineered cells (Fig. 1). We generated fusion proteins with a globular domain (super-folder green fluorescence protein (sfGFP) or monomeric human IgG Fc) and an extended glycomodule interspaced by a TEV protease site. The glycomodules were designed to contain glycosylation acceptor sites for either N-glycan, GalNAc-type O-glycan, or GAG glycosylation, and these short acceptor sequences were modified from natural substrates (N-glycosites from human erythropoietin (EPO)[29,30], O-glycosite modified from human EPO[31,32], and GAG sites based on bikunin and serglycin)[33,34] (Supplementary Table 1) to ensure high glycosylation efficiency. Glycomodules were designed to contain ~200 amino acids, as we previously found this cost-efficient for DNA synthesis, recombinant expression yields, and analysis by intact MS[35,36]. Importantly, glycomodules were designed with repeats of the short peptide acceptor sequence motifs interspaced by trypsin cleavage sites (Lys residues) (Supplementary Fig. 1), allowing for versatile conversion of the Glycocarriers into different formats, as tagged multimeric glycomodules (TEV cleavage), short glycopeptides with a single glycan (trypsin cleavage), as well as released free glycans (Fig. 1c). Note that the repeat nature of glycosylation motifs in Glycocarriers results in a significant mol/mol amplification of glycan yields (from 3 to 18× mol/mol depending on type of glycan). We previously demonstrated that stable genetic engineering of the glycosylation capacities in mammalian (CHO and HEK293) cells can be used to produce natural glycoproteins with custom-designed glycans[17,23,36], and we, therefore, employed this strategy for expression of Glycocarriers.

Production of O-glycans by O-Glycocarriers

We first used an O-Glycocarrier designed with a single O-glycosite 10-mer substrate sequence (AEAAATPAPA, flanked by K residues) that serves as substrate for most common polypeptide GalNAc-transferases (GALNTs) to ensure complete incorporation of GalNAc residues[31,32]. Figure 2a illustrates the common biosynthetic pathways of GalNAc-type O-glycosylation, the genetic regulation, and resulting O-glycan structures. The design of an O-Glycocarrier for synthesis of O-glycans is particularly important because much of the interest in O-glycans stems from their occurrence as aberrant truncated structures in cancer (also known as the pancarcinoma Tn, STn, and T antigens[37]) that are often presented in clustered motifs, and these short O-glycans without linkage to a peptide backbone present poorly in immunoassays[38–42]. Stable expression in CHO cells confirmed that an O-Glycocarrier containing 18 repeats of this substrate motif obtained full stoichiometric attachment of GalNAc residues when stably expressed in CHO$^{KO\ Cosmc}$

https://doi.org/10.1038/s41467-024-53738-9

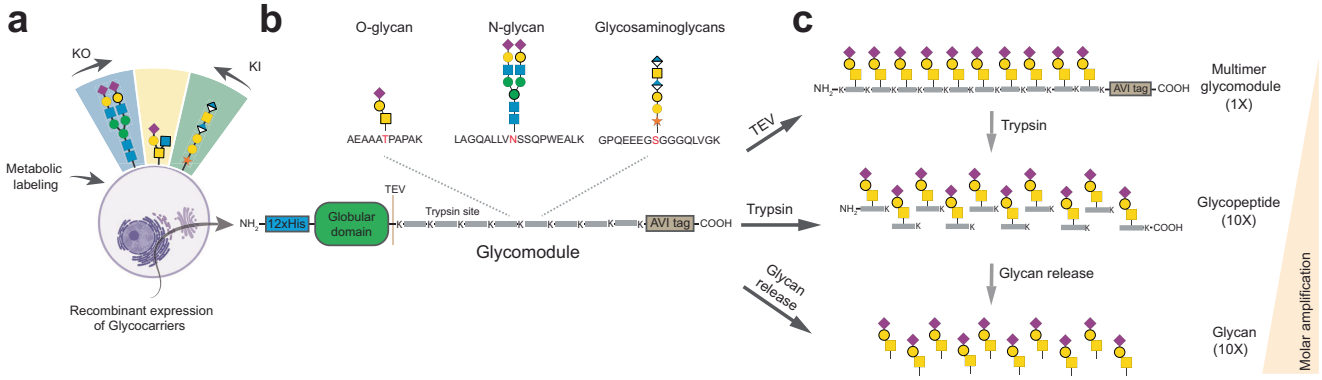

Fig. 1 | Graphic overview of the Glycocarrier technology. a A library of glycoengineered cells generated by combinatorial knock-out (KO) and knock-in (KI) of glycosyltransferase genes is used for recombinant expression of secreted Glycocarriers containing glycomodules designed to be glycosylated with different types of glycans. **b** Glycocarriers are designed to contain a folded globular module (GFP or IgG Fc) for efficient secretion, polyhistidine and site-specific biotinylation (AVI) tags for purification, and a flexible glycomodule (~200 amino acids) designed with one or more short repeat acceptor substrate sequence motifs for glycosylation interspaced by Lys residues for tryptic digestion. Intact glycodomodules can be released from the globular domain using TEV protease digestion. **c** The Glycocarrier design provides opportunities for production of glycomodules with multimeric presentation of glycans as well as production of glycopeptides and glycans in amplified yields. Structures of glycans are shown with symbols drawn according to the Symbol Nomenclature for Glycans (SNFG) format[109]. Parts of this figure were created in BioRender under agreement no. CR27EEXSTH.

SimpleCells (SC) engineered with glycosylation capacity limited to only Tn[43,44] (Fig. 2b, c). Intact mass spectrometry analysis of the TEV-cleaved, isolated O-glycomodule revealed that the most abundant product contained 18 HexNAc residues (Fig. 2b) and nano LC-MS/MS analysis of the short glycopeptide isolated after trypsin digestion by C18 HPLC purification (Supplementary Fig. 2) confirmed >90% incorporated HexNAc residues (Fig. 2c). We then generated stable expression of the O-Glycocarrier in glycoengineered CHO cells to produce the most common O-glycan structures, including disialylated core2, which required introduction of the core2 synthase GCNT1 not endogenously expressed in CHO cells. We analyzed TEV-released glycomodules by intact MS (Fig. 2b) and following trypsin digestion, the single glycopeptides with a single O-glycan by LC-MS/MS (Fig. 2c), which revealed rather homogeneous products in full agreement with the engineered glycosylation capacities. These results were confirmed by profiling of released O-glycans (Supplementary Fig. 3). Note, that intact MS of the released glycomodules with extended O-glycan structures was only possible after removal of sialic acid (Sia) residues by enzymatic desialylation, and hence sialylation was evaluated by LC-MS/MS and O-glycan profiling. We did find varying degrees of seemingly incomplete synthesis for some complex glycan structures, including variable capping efficiency with sialic acids. Incomplete synthesis can be improved by further gene engineering, where knockin to complement insufficiently expressed endogenous genes is effective and available for desirable glycans when needed. While this is beyond the scope of this study, we for example demonstrate that introducing the core2 GCNT1 gene in CHO cells ($CHO^{KI\ GCNT1}$) that do not endogenously express this enzyme, resulted in considerably more effective synthesis of core2 O-glycans (Fig. 2c) than the level of core2 O-glycans found in HEK293 cells that endogenously express GCNT1.

CHO cells have limited endogenous O-glycosylation capacities, while human HEK293 cells express a wider range of glycosyltransferases and produce both core1 and 2 O-glycans[7]. While combinatorial KO-based glycoengineering is very efficient, it is still more challenging to perform multiple KI-based glycoengineering events and introduce multiple novel glycosylation capacities[45]. We, therefore, also investigated the use of HEK293 cells and first compared transient and stable expression of the O-Glycocarrier with a single O-glycosite per repeat in $HEK293^{KO\ \textit{C1GALT1}}$ grown in suspension culture. This demonstrated that the released glycomodule and glycopeptide products were glycosylated similarly and with the same efficiency as observed for $CHO^{KO\ \textit{Cosmc}}$ cells (Supplementary Fig. 4). We then produced STn O-glycans by KI of ST6GALNAC1 in $HEK293^{KO\ \textit{C1GALT1}}$ cells (Fig. 2b, c). For unknown reasons, we consistently found relatively higher levels of residual Tn O-glycans on O-Glycocarriers expressed in HEK293 cells compared to CHO cells (Fig. 2b, c). Established human cell lines in general do not produce core3 O-glycans, including HEK293 cells, as they do not express the core3 synthase B3GNT6. To produce core3 O-glycans, we, therefore, combined KO of *C1GALT1* to abrogate core1 synthesis, that compete with core3 synthesis, with introduction of core3 synthesis by KI of B3GNT6 ($HEK293^{KO\ \textit{C1GALT1},\ KI\ B3GNT6}$), and this yielded >50% core3 trisaccharide O-glycans with some incomplete disaccharide (GlcNAcβ1-3GalNAcα) and monosaccharide components (Fig. 2b, c). The core3 pathway is restricted to the gastrointestinal tract and usually elongated with β1-3Gal residues by the B3GALT5 enzyme[46,47], however HEK293 cells only express the B4GALTs producing type 2 Galβ1-4GlcNAc linkages. Interestingly, the core3 O-glycans were also largely devoid of sialic acid capping in agreement with our previous studies[36]. We predict that this is because core3 O-glycans in gastrointestinal cells are extended by type 1 chain (Galβ1-3GlcNAc) regulated by B3GALT5 rather than the type 2 chain extension enzymes expressed in HEK293 cells. Thus, introduction of B3GALT5 is predicted to result in higher yields of core3 O-glycans with sialic acid capping. We also previously found CHO cells to be more efficient in capping both N- and O-glycans with sialic acids compared to HEK293 cells[23,36].

HEK293 cells do not express the histo-blood group glycosyltransferases, and introduction of the secretor fucosyltransferase FUT2 in cells engineered for core1 synthesis without terminal α2-3 sialylation capacity ($HEK293^{KO\ \textit{GCNT1/ST3GAL1/2},\ KI\ FUT2}$) resulted in ~50% H type 3 O-glycans (Fucα1-2Galβ1-3GalNAcα) α1-2 with some incomplete synthesis of T and Tn O-glycans (Fig. 2b, c). The incomplete synthesis of H type 3 may be because we chose the FUT2 enzyme, since it is not clarified which of the FUT1/2 isoenzymes are the most efficient to fucosylate core1 O-glycans. FUT1 may produce higher yields, however, immunohistochemical studies with mAb MBr-1 indicate that expression of type 3 chain H in epithelia is dependent on the secretor fucosyltransferase FUT2[48].

GalNAc-type O-glycans are often found in clusters comprised of multiple adjacent O-glycans and such clusters are often required for recognition by antibodies and other GBPs[3]. These clusters are difficult to produce by chemical synthesis, especially with elongated O-glycan structures. We therefore expanded the O-Glycocarrier design to include systematic clusters of 1-8 adjacent O-glycans attached to Thr residues (Supplementary Fig. 5). We first expressed these cluster

https://doi.org/10.1038/s41467-024-53738-9

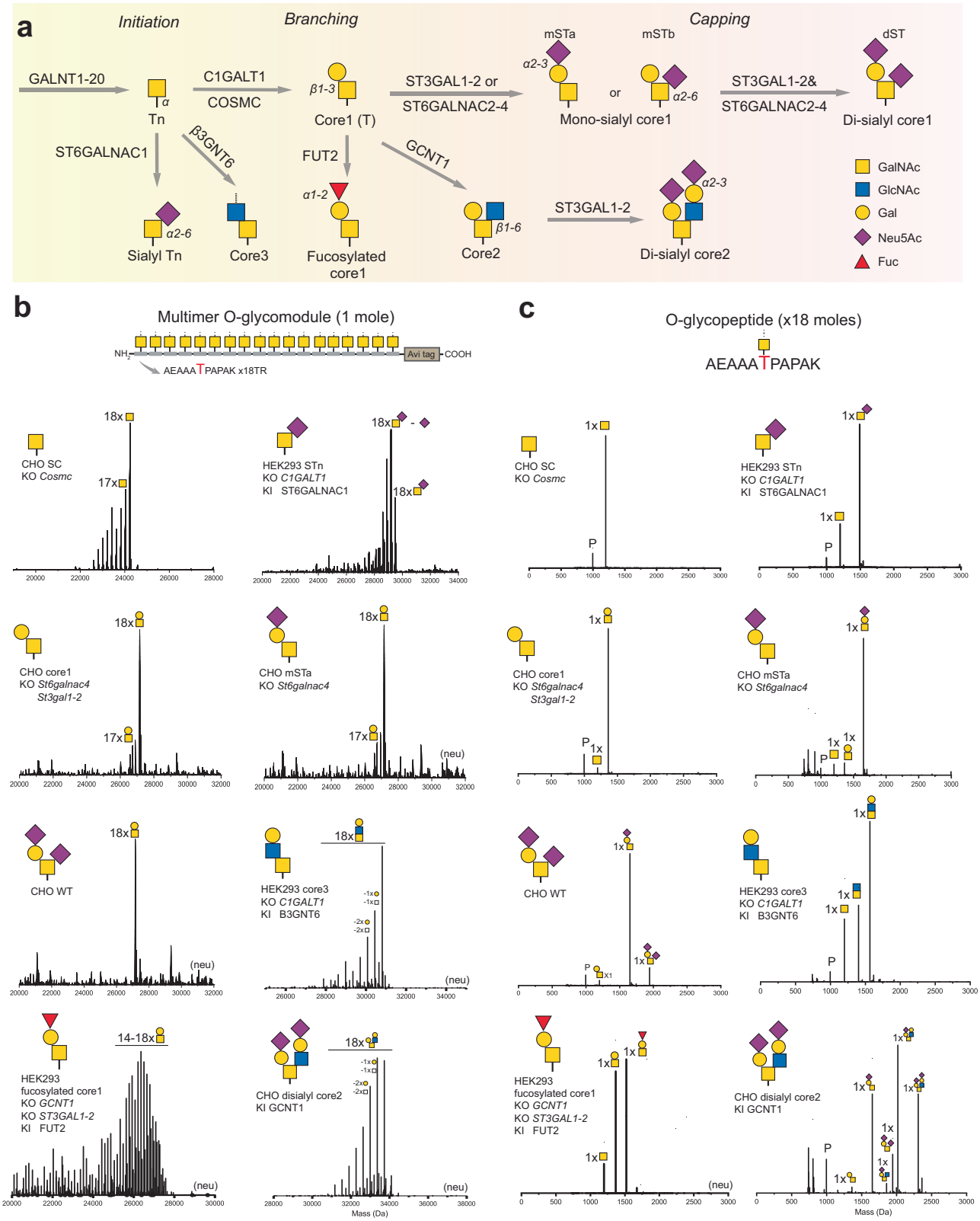

Fig. 2 | O-Glycocarriers for production of GalNAc-type O-glycans. a Graphic overview of the human O-glycosylation pathway with glycosyltransferases assigned to key biosynthetic steps. **b** Intact MS analysis of TEV released O-glycomodules expressed in wild-type and cells glycoengineered as indicated. **c** LC-MS/MS analysis of the corresponding trypsin released glycopeptides. Spectra assigned "neu" indicate that analysis was performed after pretreatment with neuraminidase. P denotes nonglycosylated peptide.

O-Glycocarriers in HEK293$^{KO\ C1GALT1}$ cells to produce Tn O-glycans. Intact MS analysis of TEV released O-glycomodules and LC-MS/MS of trypsin released glycopeptides revealed almost stoichiometric GalNAc incorporation with up to 5 consecutive Thr O-glycosites and slightly lower incorporation with larger clusters. We then focused on O-Glycocarriers with 1-3 O-glycan clusters and expressed these in HEK293 cells engineered to produce the common cancer-associated T, Tn, and STn O-glycan clusters. Intact MS analysis of the glycomodules revealed similarly high O-glycan occupancy and fidelity of O-glycan structures (Supplementary Fig. 6).

Production of N-glycans by N-Glycocarriers

The genetic regulation of N-glycan biosynthesis is well established (Fig. 3a), however, the synthesis of N-glycans may be influenced by the carrier protein with the conserved N-glycan in the Fc region (Asn297) of IgG being the most illustrative example of protein-specific restricted N-glycan branching and sialylation[49]. N-glycosylation is largely guided by the simple N-X-S/T (X ≠ P) acceptor sequence motifs, although non-conventional motifs exist[50]. We designed N-Glycocarriers with three different consensus acceptor sequences (N25, LAEAAEAENITTGCAEK; N38, LAEHASLNENITVPDTK; N83, LAGQALLVNSSQPWEALK) modified from the three natural N-glycan sequence motifs (+/- 8 residues) of the well-studied N-glycans in EPO. We first expressed these N-Glycocarriers in HEK293$^{KO\ MGAT1}$ cells to produce homogenous high mannose N-glycans and evaluate occupancy of N-glycan attachment. We found near complete N-glycan attachment for the N38 and N83 designs, while the N25 design was only partially decorated with N-glycans (40%), which may be due to the presence of a cys residue in the sequence (Supplementary Fig. 7a). We also expressed the N-Glycocarriers in HEK293^{WT} and found that N83 carried the highest degree of tetra-antennary N-glycans with sialic acid capping (Supplementary Fig. 7b). We therefore proceeded with the N83 N-Glycocarrier, and demonstrated that the TEV released N-glycomodule with 10 engineered high-Man N-glycans enabled intact MS analysis showing high occupancy and purity of N-glycans (Fig. 3b). This was confirmed by endo-H release and intact MS of the trimmed glycomodule with 10 GlcNAc monosaccharides, as well as by LC-MS/MS analysis of the trypsin-released glycopeptide with a single N-glycan and by MS analysis of the PNGaseF-released free N-glycan oligosaccharide. We also demonstrate that pronase digestion could be employed to produce N-glycans attached to short peptides of 1-4 amino acids.

Next, we expressed the N-Glycocarrier in a panel of CHO cells engineered to produce homogenous biantennary N-glycans[23,26], and we demonstrated that biantennary core fucosylated N-glycans without sialic acids or with only α2-3 or α2-6-linked sialic acids are readily produced in near homogenous forms with only minor levels of incomplete sialic acid capping (Fig. 3c). We also expressed the N-Glycocarrier in CHO^{WT} cells to demonstrate the full spectrum of N-glycans that can be produced, and importantly found that high (>60%) levels of tetraantennary N-glycans are built on the N-Glycocarrier. The N-Glycocarrier format also enables simple lectin profiling by standard ELISA (Supplementary Fig. 8). We previously performed a systematic study of genetic deconstruction and reconstruction of N-glycosylation in CHO cells[23,26], which provides genetic design matrices for production of most common N-glycan structures including different antennary structures. Here, we limited our analysis to demonstrate that the different N-glycans without core Fucose and with α2-6 instead of α2-3-linked sialic acids are easily produced with the N-Glycocarrier (Fig. 3c and Supplementary Fig. 3), however, further engineering can be used to produce more homogeneous forms of desirable N-glycans.

Production of glycosaminoglycans by GAG-Glycocarriers

GAGs are notoriously challenging to produce[51] and analytical options for these complex molecules largely limited to compositional disaccharide analysis[52]. Short defined GAG chains can be synthesized with great efforts, but access to single GAG chain of greater lengths is desirable for many bioassays and biomedical applications. The biosynthesis and genetic regulation of GAGs is well established (Fig. 4a), although the biosynthesis and structures of motifs on GAG chains that drive distinct bioactivities are less well understood. The GAG biosynthesis is initiated by XYLT1/2 xylosyltransferases with only XYLT2 expressed in CHO^{WT} cells[17]. The acceptor substrate specificities of XYLT1/2 are poorly understood, but a relatively common sequence motif for glycosylation has emerged, and many Xyl O-glycosites are identified in natural proteoglycans[53,54]. We therefore designed GAG-Glycocarriers encoding a single GAG glycosylation site in three different substrate sequences: (i) GAG-Glycocarrier 1 containing a sequence (GISDDYG**S**GAELADG) with a single GAG glycosylation site modified from serglycin, known to mainly carry heparan sulfate (HS) when isolated from intestinal macrophages[55]; (ii) GAG-Glycocarrier 2 containing a sequence (GPQEEEG**S**GGGQLVG) derived from alpha-1-microglobulin (bikunin) known to carry chondroitin sulfate (CS)[56]; and (iii) GAG-Glycocarrier 3 containing an artificial sequence (GLSADEG**S**GDGPDVG) derived from computational (LogoPlot) analysis of selected identified GAG glycosites[54] (Fig. 4b and Supplementary Table 1). Initial studies with the N-terminal GFP fusion design for Glycocarriers did not secrete well, and since we previously found that the full coding sequence of human serglycin containing multiple GAG glycosylation sites fused N-terminally to human IgG Fc expressed well in CHO cells[17], we used the monomeric IgG Fc design for the GAG-Glycocarrier (Fig. 4b).

We first tested the GAG-Glycocarriers in CHO cells engineered to only attach Xyl residues without further extension to simplify analysis (CHO$^{KO\ B4galt7}$). We found that all three GAG-Glycocarrier designs with a single acceptor substrate motif obtained near stoichiometric attachment of Xyl as revealed by LC-MS/MS analysis (Fig. 4c). We then introduced repeats of the bikunin GAG motif in GAG-Glycocarrier 2, and we demonstrated that the GAG-Glycocarrier with three repeated motifs obtained near full attachment of Xyl residues by both intact MS of the GAG-glycomodule and LC-MS/MS of the trypsin released glycopeptide (Supplementary Fig. 9b, c). When we expanded the GAG-Glycocarrier to contain 10 repeats we only obtained ~50% incorporation of Xyl (Supplementary Fig. 9). Since analysis of the GAG-Glycocarriers with multiple GAG sequence motifs was challenging, we continued our study with the GAG-Glycocarrier containing one GAG chain. GAG chains produced naturally on proteoglycans in mammalian cells like CHO usually comprise variable mixtures of CS/DS and HS chains with varying lengths, epimerization, and sulfation patterns, which complicate analysis of structures and defining bioactive motifs. Genetic de/reconstruction of GAG biosynthesis in cells, however, offer fruitful ways to identify and dissect structural features underlying bioactivities as clearly demonstrated with the cell-based GAGOme display platform[17,52]. To demonstrate that the GAG-Glycocarrier acquired designer CS/HS chains, we first analyzed by SDS-PAGE as elaborated GAG chains markedly alter migration in gels. When the GAG-Glycocarriers were expressed in CHO$^{KO\ B4galt7}$ without capacity for GAG biosynthesis, they migrated as rather homogeneous bands around 45 kDa corresponding to the predicted protein mass (Fig. 4d and Supplementary Fig. 10a). In contrast, when the GAG-Glycocarriers were expressed in CHO^{WT} cells, they migrated as high MW smears, clearly demonstrating that efficient processing of the attached GAG chain was obtained. To evaluate CS and/or HS synthesis, we used selective enzymatic digestion of GAG-Glycocarriers before SDS-PAGE analysis, and demonstrated that the GAG-Glycocarriers mainly acquired CS chains in WT CHO cells (Fig. 4d and Supplementary Fig. 10a). To evaluate the sulfation patterns on the GAG chains produced, we used the state-of-the-art disaccharide analysis, which showed that the sulfation patterns mirrored the general GAG biosynthesis capacities of CHO cells determined in previous studies[17]

https://doi.org/10.1038/s41467-024-53738-9

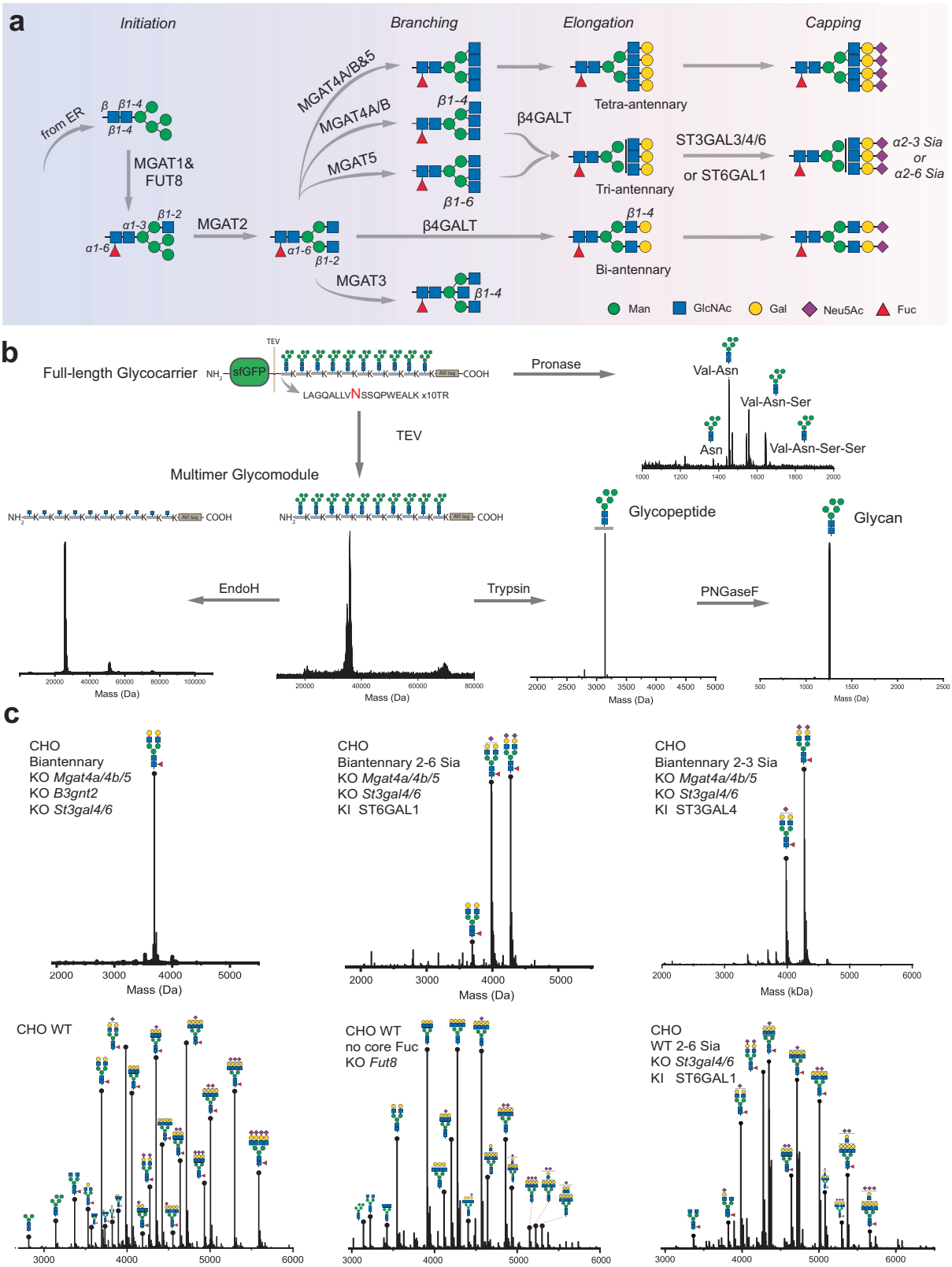

Fig. 3 | N-glycocarriers for production of N-glycans. a Graphic overview of the human N-glycosylation pathway with glycosyltransferases assigned to key biosynthetic steps. **b** Comprehensive workflow for production and analysis of multimeric (10 N-glycosites of an 18-mer acceptor sequence LAGQALLVNSSQPWEALK) N-glycomodule, N-glycopeptides, and released N-glycans illustrated with the N-Glycocarrier expressed in $CHO^{KO\ Mgat1}$ cells engineered to produce homogenous $Man_5GlcNAc_2$ N-glycan. Intact MS analysis of the TEV released N-glycomodule with and without endo-H treatment and LC-MS/MS analysis of the tryptic digested N-glycopeptides revealed complete N-glycan occupancy. PNGase F treatment of the N-glycomodule released homogeneous $Man_5GlcNAc_2$ N-glycans. Alternatively, extensive pronase digestion of the N-Glycocarrier produced N-glycans attached to a few amino acids (1-4) (upper right panel). **c** LC-MS/MS analysis of the tryptic released N-glycopeptides expressed in CHO^{WT} and CHO cells glycoengineered as indicated.

https://doi.org/10.1038/s41467-024-53738-9

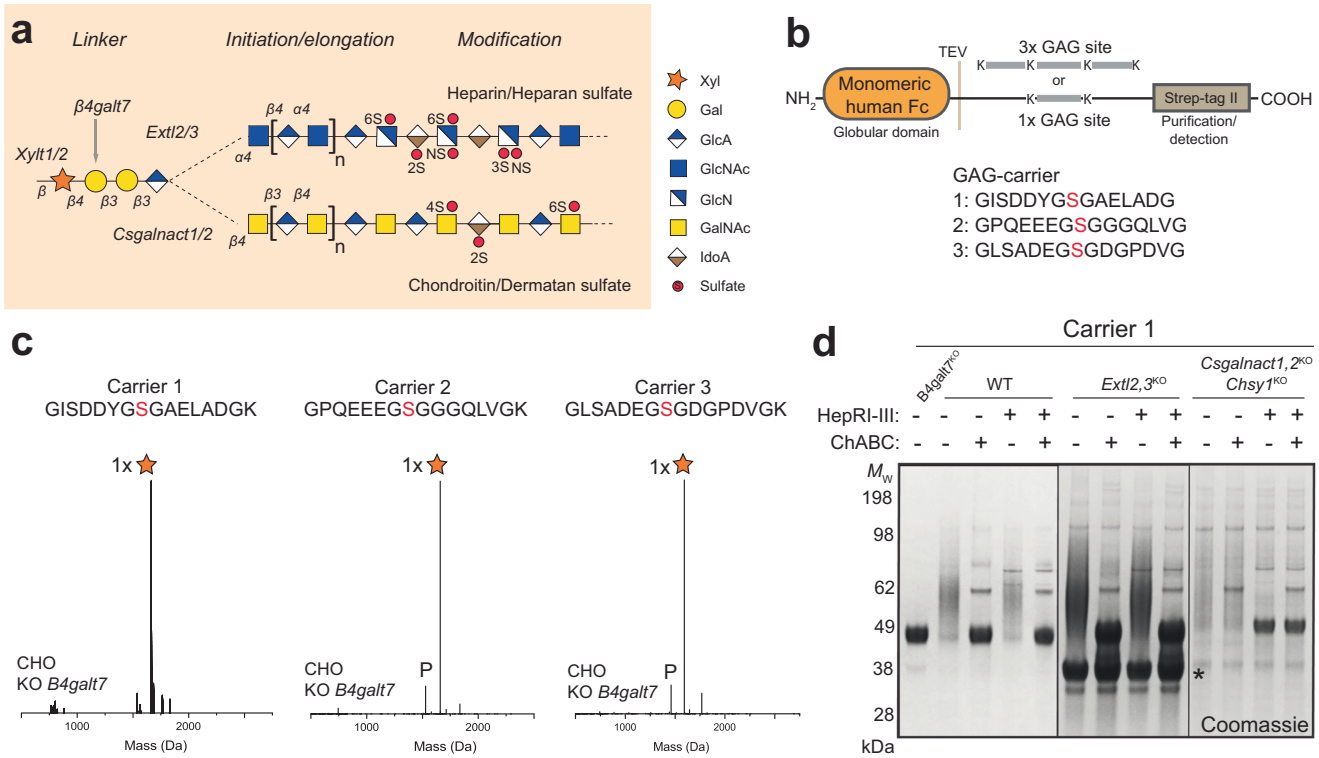

Fig. 4 | GAG-Glycocarriers for production of glycosaminoglycans. a Graphic overview of the human GAG biosynthesis pathway with glycosyltransferases assigned to key biosynthetic steps. **b** Design variants of GAG-Glycocarriers with three different acceptor peptide sequences. Monomeric human IgG-Fc was used as N-terminal globular domain, followed by 1 or 3 tandem repeats of GAG substrate sequences, separated by a TEV protease site, and with a C-terminal StrepII tag. **c** LC-MS/MS analysis of TEV and trypsin released glycopeptides from the three different GAG-Glycocarrier designs with single acceptor sequence motifs expressed in CHO$^{KO\ B4galt7}$ to evaluate initiation efficiency (xylose incorporation). **d** SDS-PAGE Coomassie analysis of the purified GAG-glycocarrier 1 expressed in CHO^{WT}, CHO$^{KO\ B4galt7}$, CHO$^{KO\ Extl2,3}$, and CHO$^{KO\ Csgalnac1/2/Chsy1}$ digested with heparinase (HepRI-III) and chondroitinase (ChABC) as indicated. Asterisk indicates non-specific contaminants. Results are the representation of three biological replications.

(Supplementary Fig. 10b). Thus, CS chains are mainly 4-O-sulfated, while HS chains are non-sulfated or N-sulfated with minor 2-O- and 6-O-sulfation. To demonstrate that the GAG-Glycocarriers can be used to produce CS chains without contaminating HS chains and vice versa, we expressed the GAG-Glycocarriers in CHO cells without CS synthesis (CHO$^{KO\ Csgalnact1/2/Chsy1}$) or without HS synthesis (CHO$^{KO\ Extl2/3}$)[17] (Fig. 4d and Supplementary Fig. 11). This demonstrated that GAG-Glycocarriers can be used for efficient production of CS chains as well as HS chains, albeit with lower yields. Note, that the secreted GAG-Glycocarriers expressed in CHO$^{KO\ Csgalnact1/2/Chsy1}$ without capacity for CS synthesis required considerable enrichment (6 to 15-fold compared to CS) by ion exchange chromatography before analysis, indicating that only a small fraction of the GAG-Glycocarriers attained HS chain elongation (Supplementary Fig. 11). Studies suggest that the acceptor sequence motif and potential upstream motifs hereof affect the synthesis of CS and HS, which will be the subject of future improvements in the GAG-Glycocarriers design[57]. We did not in this study employ further engineering of the sulfation capacities in CHO cells to produce CS/HS chains with different sulfation patterns, but this is clearly straightforward based on our previous engineering of GAG biosynthesis in CHO cells[17,52]. It is important to appreciate that natural GAG chains contain inherent heterogeneity in motifs (epimerization and sulfation positions and patterns) that are required for biosynthesis as well as bioactivities[58]. While the GAG-Glycocarrier technology may not be used to produce fully homogenous GAG chains as such, it can provide access to custom-designed GAG chains with/without distinct features required for distinct bioactivities in chain lengths, quantities, and different formats to widely support bioassays and applications. These GAG chains inherently may vary in lengths and sulfation patterns, but once the optimal minimum features for a particular bioactivity have been identified (e.g., by screening using the cell-based GAGOme array), it is possible to engineer CHO cells to produce such GAG chains with a minimum of heterogeneity.

Glycocarriers provide improved analytical opportunities

Glycocarriers provide access to multiple formats of glycans in one biomolecule. To illustrate the diverse opportunities provided by glycans produced by Glycocarriers, we designed proof-of-concept studies to (a) dissect interactions with glycans in different presentations (isolated, clustered, or multimeric); (b) demonstrate interactions with glycans by single molecule mass photometry; (c) generate complex glycan probes to monitor cellular trafficking and discovery of receptors; and (d) facilitate synthesis of neoglycoproteins by production of glycan building units (Fig. 5 and Supplementary Figs. 12–18). These applications are challenging or not possible with currently available methods to produce glycans.

Dissection of glycan interactions

Glycan antibodies and lectins often employ multivalent interactions to enhance binding through avidity effects (Fig. 5a), and this is particularly relevant for binding to mucins and O-glycoproteins with their display of extended regions of dense O-glycans in clusters and patterns[35]. In addition, antibodies to the cancer-associated aberrant Tn, STn, and T O-glycans often recognize clusters of O-glycans[42,59,60], but they may also recognize O-glycans in the context of a specific protein[61]. It is challenging to discern such binding properties with simple synthetic glycans and affinity and avidity effects. The O-Glycocarriers with their flexible design for presentation of isolated

https://doi.org/10.1038/s41467-024-53738-9

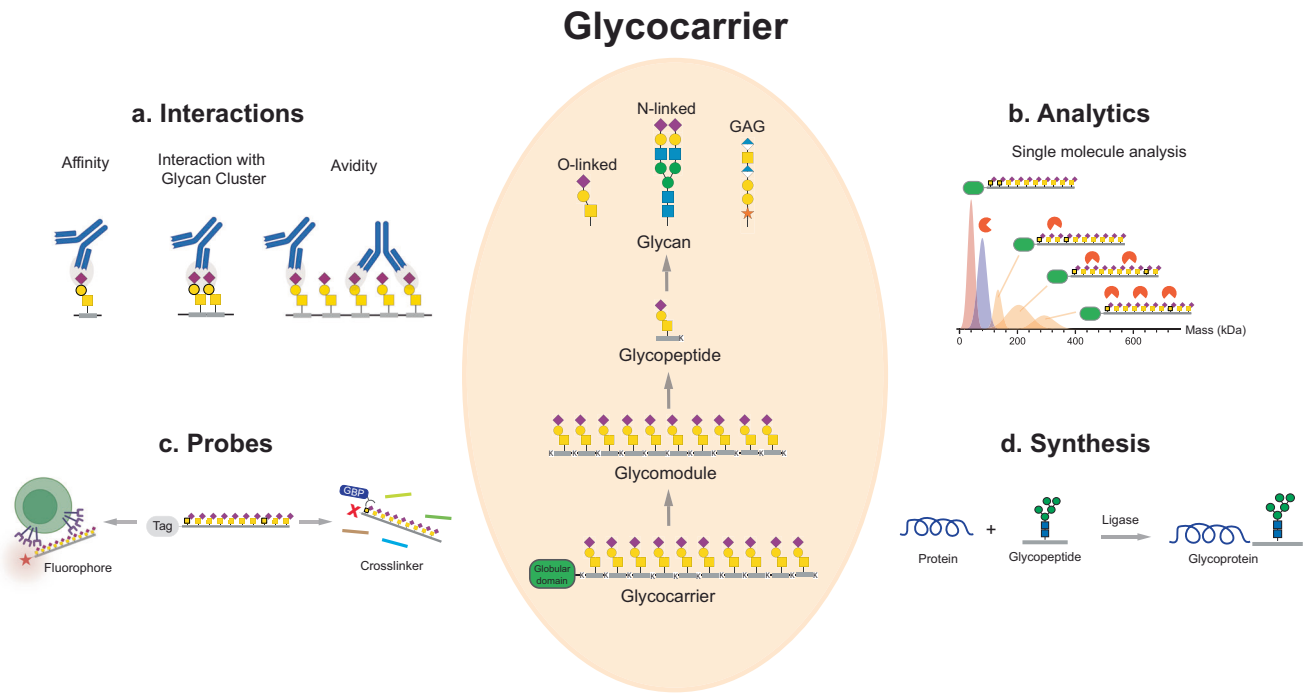

Fig. 5 | Glycocarriers expand the glycoscience toolbox. Glycocarriers produced in glycoengineered cells enable production of mammalian glycans in different formats for great diversity in applications. **a** Different presentation modes of glycans enable dissection of binding interactions with GBPs. **b** Glycocarriers provide scaffold mass that enables single molecule analytics, e.g., by mass photometry and LLPS. **c** Glycocarriers are tagged (GFP) and can easily be functionalized by chemical and metabolic methods for probing interactions and following cellular trafficking and biodistribution. **d** Glycocarriers facilitate synthesis and design of glycoproteins by supplying glycan and glycopeptide building unit. Parts of this figure were created in BioRender under agreement no. QB27EEYK24.

and clustered O-glycans on single and multimeric scaffolds without specific protein-identity, offer a simple way to explore these binding properties. To illustrate this, we used a panel of mAbs and lectins to aberrant O-glycans to test their dependence on clusters of 1-3 O-glycans on glycomodules (Supplementary Fig. 12). Most of the tested mAbs (5F4 and 1E3 to Tn, B72.3 and 3F1 to STn) only bound to clusters of two or more O-glycans, while other mAbs (TKH2 to STn, 3C9 to T) showed significant binding to isolated O-glycans, albeit at lower level than to the clustered O-glycans. We also used lectins and two of these bound equally well to isolated and clustered O-glycans (VVA to Tn, PNA to T), while one lectin (Jacalin to T) preferentially bound clustered O-glycans. These examples clearly illustrate the advantage of the flexible Glycocarriers in discerning binding specificities for clustered O-glycans versus multimeric presentation. Glycocarriers are well-defined reagents available to further dissect avidity effects and on/off rates of binding using traditional assays, e.g., SPR and real-time interaction cytometry (RT-IC).

Single molecule analysis of glycan interactions

Glycocarriers have masses (>50 kDa) suitable for single molecule analysis by mass photometry (Fig. 5b), which can provide direct view of binding partners, their relative abundance, and the complexes formed with binding affinities. Here, we illustrated the use of mass photometry to analyze the binding properties of the IgG1 mAb B72.3 using isolated O-Glycocarriers carrying clustered STn O-glycans (Supplementary Fig. 13). This confirmed that B72.3 does not bind single STn glycans even with multimeric presentation, in agreement with the results obtained by ELISA (Supplementary Fig. 12), and further revealed that B72.3 has higher affinity for a triad STn O-glycan cluster than a diad cluster. Moreover, only the triad STn O-glycan on the Glycocarrier supported binding of two IgG molecules (avidity effect). Studies with GAGs are particularly challenging and the GAG-Glycocarrier conveniently provides a single customizable GAG chain on a protein scaffold amenable for mass photometry analysis. GAGs also present cancer-associated epitopes as previously demonstrated with the malaria binding protein VAR2CSA, although the molecular nature of such epitopes is not fully characterized[62]. We recently characterized two scFv^2 (F8 and C9) with similar cancer-associated binding properties[63], and here we illustrate how the GAG-Glycocarrier enables analysis of the binding properties of these nanobodies (Supplementary Fig. 14). The F8 and C9 scFv^2 bind distinct cancer-associated CS epitopes composed with 6S or 4S sulfated CS, respectively[63]. The GAG-Glycocarrier expressed in CHO cells produces CS with 4-O-sulfate but not 6-O-sulfate (Supplementary Fig. 10), and in agreement with this, C9 but not F8 produced complexes with the GAG-Glycocarrier. Moreover, the mass photometry analysis reveals that up to three C9 scFv^2 molecules bound to the GAG-Glycocarrier, which shows that the C9 epitope is repetitively presented and confirms that the GAG chains attached are of substantial lengths. The repertoire of GAG structures on the GAG-Glycocarrier needs to be expanded, but this example clearly illustrates the opportunities provided by this technology.

Homomeric nucleation of the glycan-binding protein galectin-3

Glycocarriers can also be used to produce multimeric presentation of defined N-glycans, and we highlight the utility of these by studying nucleation of galectins. Galectin-3 (Gal3) has a single C-terminal canonical galectin carbohydrate recognition domain (CRD) and can, upon ligand-binding, induce self-association, or nucleation, through the intrinsically disordered N-terminus[64]. Oligomerization of Gal3 is important for its biological activities[65] e.g., the construction of endocytic pits[66] Gal3's self-association capacity includes the formation of biomolecular condensates through liquid-liquid phase separation (LLPS)[67–69]. Here, we demonstrated nucleation of Gal3 condensates in a LLPS assay using our GFP-tagged N-Glycocarriers with full-length human Gal3 (Gal3^{FL}), but not N-terminal truncated Gal3 (Gal3^{CRD}) (Supplementary Fig. 15a). A panel of the GFP-tagged N-Glycocarriers

with different glycoforms revealed that homogeneous biantennary N-glycans with α2-3Sia, but not α2-6Sia capping, could support LLPS and removal of sialic acids amplified this effect (Supplementary Fig. 15b). The tested N-Glycocarriers have fixed spacing (17 amino acids) between the 10 N-glycans, and it is conceivable that altering the spacing between these N-glycans may be used to further dissect the requirement for nucleation of Gal3 condensates.

Probes to target glycan receptors

Glycocarriers provide flexible scaffolds for introducing specific functional groups such as fluorophore, radiolabels, and crosslinking agents (Fig. 5c). Glycocarriers can be incorporated with e.g., unnatural monosaccharides using metabolic labeling by simply supplementing growth media during production in appropriately glycoengineered cells[70,71]. To illustrate this, we introduced functional groups for click chemistry into O-Glycocarriers by feeding cells with either Ac_5SiaNAz (azido Sia) or Ac_5SiaNPOC (propargyl alkyne Sia) and found rather efficient incorporation (30–50%) (Supplementary Fig. 16). Glycocarriers are also well suited as molecular probes for identification and tracking glycan-binding receptors. To illustrate this, we leveraged GFP fluorophore-tagged O-Glycocarriers displaying a diversity of O-glycan (Tn, αGalNAc) clusters on a multimeric scaffold and used this to study uptake by human liver HepG2 cells[72] (Supplementary Fig. 17a). Internalization of O-Glycocarriers was confirmed by confocal microscopy and confocal z-stack (Supplementary Fig. 17b–d). The asialoglycoprotein receptor (ASGPR or Ashwell-Morell receptor) on hepatocytes is a well-established target receptor for delivery of drugs including siRNA, antisense oligonucleotides, chemotherapeutic agents, and proteins to the liver by tagging these with GalNAc residues[73,74]. Using tagged O-Glycocarriers, we observed a direct positive correlation between uptake rate and size of Tn O-glycan clusters. Moreover, we demonstrate that T O-glycan clusters may provide better uptake that Tn O-glycans (Supplementary Fig. 17b). Note, the features of Glycocarriers as probes are that (1) they provide for a uniform protein scaffold to display different glycans in by-design patterns, (2) they can be tagged to probe and conjugate receptors without affecting the glycan presentation, and (3) they can be produced in unlimited quantities from glycoengineered cells.

Synthesis of neoglycoproteins

While complete chemical synthesis of glycoproteins has been achieved through great efforts[75,76], an appealing approach is to employ enzymatic ligation and short glycopeptide building units to incorporate the glycans[77] (Fig. 5d). Short glycopeptides can be synthesized, but with increasing challenges and costs as the number, size, and complexity of glycans increase. However, Glycocarriers clearly offers a much simpler and scalable way to design and sustainably produce such glycopeptide building units in cells. We illustrate this for both O- and N-glycopeptide building units using the peptiligase-based, seamless ligation platform[78](Supplementary Fig. 18).

Discussion

Here, we demonstrated that production of a diversity of human glycans in high yields and sufficient purity is feasible by use of Glycocarriers and glycoengineered mammalian cells. This cell-based method for production of glycans represents a self-renewable and sustainable source of glycans, which both complements chemical and chemoenzymatic synthesis and provides opportunities for analytical and biomedical applications. The Glycocarriers serve as uniform scaffolds for glycans and encode a flexible design to enable production of glycans in different formats (Fig. 1). The method is designed for scaled production of desirable individual glycans to support a wide range of use, however, it is possible to produce larger libraries of glycans by employing a greater panel of cells appropriately glycoengineered[16,45]. Chemical and chemoenzymatic synthesis of most glycans is possible but with increasing challenges and costs as the complexity of structures increase, and production at larger scale is often unfeasible. In contrast, the use of Glycocarriers to produce glycans in cells is facilitated by well-established bioprocessing and downstreaming protocols developed by the biopharmaceutical industry[28], and once stable expression of a Glycocarrier in a glycoengineered cell is established this provides a sustainable and cost-effective source of glycans.

The Glycocarrier method of producing glycans was not designed with glycan arrays in mind. The printed[8–11] and cell-based[16] glycan arrays are designed for screening, discovery and dissection of glycan-binding specificities, and their power of use lies in the huge diversity in glycan epitopes they display for interrogation. While it is possible in principle to establish similarly large libraries of cells to produce Glycocarriers with such repertoires of glycans, there is little need as the cell-based glycan arrays provides a simpler, higher throughput, and more cost-effective platform for discovery and dissection of glycan interactions. The real value of the Glycocarrier method lies in the steps following discovery and characterization of glycan interactions, where access to defined glycans (and critical variations hereof) are needed for further studies of GBPs and their binding mode, as exemplified in Fig. 5. These needs are not merely access to oligosaccharides, but also glycans in formats that support the highest affinity interaction (e.g., clustered on a short peptide or multimeric on an inert glycomodule) and that enable use of advanced analytics (e.g., mass photometry, X-ray crystallography and CryoEM). We envision that the Glycocarrier technology will stimulate ideas and applications beyond our current imagination.

The Glycocarrier technology was designed to produce glycans on peptide scaffolds without protein identity or sequence context. Most GBPs including plant lectins recognize and interact with simple often terminal saccharide (mainly mono- and disaccharides) epitopes without particular protein context. Studies with printed glycan arrays displaying isolated oligosaccharides have highlighted this[79], while studies with the cell-based glycan array also highlight that clustered and context interactions exists[35,70]. The key challenge in design of the Glycocarriers was to identify short acceptor sequences that would obtain stoichiometric attachment of glycans, even when placed in tandem repeats, and also allow biosynthesis of the full diversity of glycan structures produced in cells without constraints by individual proteins. We did achieve complete occupancy for all types of glycans on Glycocarriers, although for GAG-Glycocarriers only with three repeat GAG chains. We confirmed that the produced glycans on the different Glycocarriers acquired glycan structures consistent with the general glycosylation capacities in cells using WT cells as well as a number of glycoengineered cells. This included core2 O-glycans, tetraantennary N-glycans, and extended sulfated CS chains. A remarkable advantage of the Glycocarriers is that these in one produced biomolecule provide flexible access not only to the oligosaccharide of interest, but also to the same oligosaccharide in the form of a monomeric glycopeptide (trypsin released or e.g., pronase digested) and the corresponding glycomodule (TEV released) with multimeric presentation. This is not provided by other methods of synthesis of glycans and can only be achieved by elaborate and impractical procedures.

The Glycocarrier technology can be used to produce homogenous glycans. We demonstrated this for core1/2/3 based O-glycans and biantennary N-glycans, while we did not pursue this for GAG chains as these encode inherent heterogeneity that cannot be fully discerned, even with the most advance analytics in the field. We illustrated examples of how distinct glycan features, e.g., core fucosylation of N-glycans, can be controlled, and production of most glycans in sufficient purity will be possible by employing further combinatorial glycosyltransferase gene KO/KI, as previously used for cell-based glycan arrays[7]. This will require step-by-step engineering of individual biosynthetic steps as previously discussed[16], and with more complex

structures where incomplete synthesis of some steps may be challenging, post-production purification can be employed as this is already widely used in chemical and chemoenzymatic syntheses protocols. The ability to produce large glycomodules with repeat homogenous glycans is also illustrated by our recent study of a mucin O-glycomodule with multiple core3 trisaccharide O-glycans using single molecule imaging by scanning tunneling microscopy[80].

Glycans have antimicrobial potential as demonstrated for example with simple O-glycans that can disperse *Pseudomonas aeruginosa* biofilm[21,81] and inhibit virulence traits and quorum-sensing[19,20,82,83]. In these studies, the O-glycans used were released from isolated natural mucins with inherent heterogeneities, but a more recent study showed that a synthetic core2 trisaccharide (Galβ1-3[GlcNAcβ1-6]GalNAc) had similar effects[83]. O-Glycocarriers are ideal to produce large amounts of such O-glycans needed for clinical studies. In our study we reached production of ~20-100 mg/L Glycocarrier (equivalent to 5–50 mg released oligosaccharides/L) with stably engineered CHO cells using common shaker flasks (Supplementary Table 6), and these yields can be dramatically improved by advanced bioprocessing. Human milk oligosaccharides (HMOs) are also in demand as prebiotics in milk formulas[84]. While HMOs are lactose based (Galβ1-4Glc), the structural repertoire of extended sialylated and fucosylated structures are essentially identical to those of N- and O-glycans and can be produced and released using the Glycocarrier technology. Moreover, it is noteworthy that current prebiotic approaches only focus on lactose-based HMOs, while mucins and mucin O-glycans in fact, also are substantial components in human milk[85]. The GalNAc monosaccharide is an important metabolite for the microbiome[86,87], which may be supplied in HMOs only from blood group A individuals, while GalNAc it is abundantly available from O-glycans.

A challenging type of glycans to produce are GAGs that inherently exhibit great structural heterogeneity in chain length and sulfation/modification patterns[88,89]. Shorter, well-defined GAG oligosaccharides can be synthesized and synthetic GAG arrays have been developed[90,91], however, access to a single (or multiple) GAG chain on a short peptide produced in genetically glycoengineered cells with distinct sulfation capacities opens for other analytical opportunities. Here, we demonstrated that single CS or HS GAG chains can be produced by the GAG-Glycocarriers, and it is possible to produce distinct GAG chains by use of cells with more extensive glycoengineering[17]. We were unable to analyze GAG-glycocarriers with repeat GAG motifs in detail due to challenges with trypsin cleavage. However, GAG-Glycocarriers with multiple GAG chains may still be useful for high yield production of GAGs, e.g., following extensive pronase digestion. Thus, Glycocarriers may eventually be used for cell-based production of designer GAGs, and for example, a next generation of heparins with improved safety profile[52].

The Glycocarrier technology fulfils important voids in analytics. Short O-glycans such as the prevalent cancer-associated truncated O-glycans (Tn, STn, T) are poorly presented on printed glycan arrays, and the O-Glycocarriers make these O-glycans readily available on short glycopeptides to support studies of natural and vaccine-induced immunity to these glycan epitopes[38–42,79,92]. O-Glycocarriers are useful for characterization of the large number of mAbs generated to aberrant O-glycans and dissection of their dependence on clusters of O-glycans and multivalent binding[42,59,60] (Supplementary Fig. 12), and they may complement studies of mAbs that recognize truncated O-glycans in context of peptide/protein motifs[93–96]. Glycocarriers provide scaffold modules that enable studies of interactions with glycans at the single molecule level by e.g., mass photometry (Supplementary Figs. 13, 14) and the nucleation of galectin condensates on N-glycan lattices (Supplementary Fig. 15). The Glycocarrier design provides wide options to generate tagged probes for discovery and dissection of biological interactions with glycans in monomeric and multimeric formats. We illustrated this by metabolic labeling of Glycocarriers during the cell-based production and by chemical conjugation of isolated Glycocarriers (Supplementary Fig. 16), and we employed tagged Glycocarriers to probe binding and uptake by the ASGPR in liver cells (Supplementary Fig. 17). Finally, Glycocarriers can supply short glycopeptide building blocks for use e.g., in chemical synthesis of glycoproteins using enzymatic ligation strategies as demonstrated with both N- and O-glycopeptides (Supplementary Figs. 18)[97].

There are limitations with the Glycocarrier technology. The Glycocarrier strategy may work for most common types of protein glycosylation, however, for the expanding group of protein domain-specific types of glycosylation[22], including the O-Man, O-Fuc, O-Glc, O-GlcNAc (extracellular types), and C-Man types, the design of Glycocarriers will have to take the domain structure of the substrates into account in order to achieve glycosylation. Introduction of unnatural glycan structures may be challenging, but Glycocarriers may complement chemical and chemoenzymatic strategies[98,99], e.g., by supplying substrates for further chemical modification[100,101]. Genetic glycoengineering may be challenging when KI of multiple genes is required to assemble more complex novel glycosylation features[45], however, this may be alleviated use of alternative strategies such as gene activation. In general, glycoengineered cells faithfully produce glycans according to the designed biosynthetic capacities, however, not all glycans can be produced as individual single homogenous structures. The glycoengineering design strategies to obtain a particular glycan aim to limit and streamline the cellular glycosylation capacities in the particular glycosylation pathway only towards this glycan structure. This involves (1) eliminating unwanted and potential competitive glycosylation capacities by KO of relevant glycosyltransferase genes and (2) introduction of additional glycosylation capacities by KI of glycosyltransferase genes that are not endogenously expressed (or expressed at sufficient levels). While elimination of glycosylation capacities usually provides uniform loss of unwanted glycan features, the endogenous and introduced glycosylation capacities are prone to variable degrees of heterogeneous outcomes due to inefficiency in individual glycosylation steps (e.g., sialylation of O- and N-glycans (Figs. 2, 3)) and in some cases competition (e.g., branching of O-glycans (Fig. 2)). The stable site-specific KI strategy using safe-harbor sites[23] results in expression of the introduced glycosyltransferases at protein levels comparable to those of the most abundant endogenously expressed enzymes[102,103]. KI of glycosyltransferase genes can in many cases improve efficiency of a biosynthetic step if the endogenous enzyme is insufficient, but clearly KI of glycosyltransferase genes does not ensure completion of reaction steps as illustrated here (e.g., sialylation of N-glycans by KI of ST3GAL4/ST6GAL1 (Fig. 3) and fucosylation of O-glycans by KI of FUT2 (Fig. 2)). However, it is important to note that such incomplete synthesis of biosynthetic steps mainly results in heterogeneity from related precursor glycan structures, which in most cases are amenable to simple postproduction purification to obtain homogeneous glycans (e.g., fucosylated O-glycans from non-fucosylated precursor (Supplementary Fig. 19)). The Glycocarries are well secreted (20–200 mg/L) and can yield substantial amounts of released glycans (5–50 mg/L) using CHO cells grown in simple shaker flasks as shown here, however, once stably engineered cell lines are generated these represent sustainable sources and the glycan yields from these cells can likely be increased at least 10-fold by controlled bioprocessing at scale. It is also worth noting that glycosylation and especially degree of sialylation is affected by growth conditions as well as degradation in the culture media, e.g., by secreted neuraminidases, and most of these effects can be controlled by consistent bioprocessing protocols not employed in the present study[28].

In summary, the Glycocarrier technology combined with glycoengineered cells provides a sustainable source of mammalian glycans in flexible formats. We provide proof-of-concept for this with the most abundant types and structures of glycans, and the design of

Glycocarriers and the repertoire of glycoengineered cells can be expanded to produce most structures of glycans. The flexible Glycocarrier format supports various analytical approaches and provides a stable and scalable source for biomedical use.

Methods

Cell culturing

Adherent HEK293 cells (WT from ATCC: CRL-1573) were grown in DMEM/F-12 (1:1 v/v) (Gibco) supplemented with 10% heat-inactivated fetal bovine serum (Sigma-Aldrich) and 2 mM GlutaMAX (Gibco) at 37 °C and 5% CO_2. Suspension culture of HEK293 cells was performed in serum-free F17 media (Invitrogen) supplemented with 4 mM GlutaMAX (Gibco) and 0.1% Kolliphor P188 (Sigma) under agitation (120 rpm). Suspension CHO cells (WT from Sigma-Aldrich no. 85050302) were cultured in EX-CELL CHO CD Fusion serum-free media (Sigma) and BalanCD CHO Growth A media (FUJIFILM) (1:1 v/v) supplemented with 4 mM GlutaMAX (Gibco). HepG2 cells were cultured in DMEM media (Gibco) supplemented with 10% FBS, 1% glutamine, and 1% non-essential amino acids solution (Gibco). Authentication of each cell line used in this study included PCR assays with species-specific primers as previously described[7,23], and routinely mycoplasma analysis.

Genetic engineering of HEK293 and CHO cells

Genetic engineering of HEK293 cells was performed using validated gRNAs libraries from the GlycoCRISPR resource[104]. Briefly, HEK293 cells grown in 6-well plates at ~80% confluency were transfected with 1 μg each of gRNA and GFP-tagged Cas9-PBKS using lipofectamine 3000 reagents (ThermoFisher Scientific), following the manufacture's protocol. 24 h post-transfection, GFP-expressing cells were bulk-sorted by FACS (SONY SH800) and cultured for a week and cells were then single cell-sorted into 96-well plates. KO clones were screened by Indel Detection by Amplicon Analysis using primers specific to gRNA targeting sites[105]. Final clones were further verified by Sanger sequencing. For site-directed knock-in (KI), a modified ObLiGaRe targeted KI strategy utilizing two inverted ZFN binding sites flanking the full coding region of human glycosyltransferase genes in donor plasmids was used[106]. KI was performed as described for targeted KO with 1 μg of each ZFN tagged with GFP/Crimson and 3 μg donor plasmid. KI clones were screened by PCR with primers specific for the junction area between the donor plasmid and the AAVS1 locus, and a primer set flanking the targeted KI locus was used to characterize the allelic insertion status. All gRNAs and primers used in this study were provided in Supplementary Table 4. A detailed list of engineered HEK293 and CHO used in this study are provided in Supplementary Table 1 and 5.

Glycocarrier construct designs

Secreted N- and O-Glycocarriers were designed by fusion of human CD33 signal peptide (amino acids 1-16, Uniprot P20138) with 12xHis, Alpha-tag, sfGFP, TEV cleavage site, multiple cloning site, and C-terminal Avi-tag. For GAG-Glycocarrier, monomeric human Fc (amino acids 106-328, Uniprot P01857) and C-terminal Strep tag II were used instead. The Glycocarrier scaffold was subcloned into either pIRES-puro or pGS vectors[23,35]. Exchangeable Glycocarrier gene inserts were synthesized with BamHI/NotI cloning sites for N- and O-Glycocarriers and BglII/BamHI for GAG-Glycocarriers (Genscript, USA) and inserted into the vector using standard restriction enzyme-based molecular cloning. Full sequences of glycomodules in Glycocarriers are shown in Supplementary Fig. 1.

Transient expression of Glycocarriers

HEK293 cells were seeded at 0.8×10^6 cells/mL 1 day before transfection, and 3×10^7 cells harvested and resuspended in 30 mL fresh media and transfected with 30 μg plasmid DNA and 160 μg polyethylenimine hydrochloride MAX (MW 40,000, Polysciences) in 300 μL Opti-MEM media (Gibco) pre-incubated at room temperature for 10 min. Culture supernatants were harvested 5 days after transfection and processed as described below.

Production and purification of recombinant N- and O-Glycocarriers

Secreted Glycocarriers were stably expressed in glycoengineered HEK293 and CHO cells. Stably expressing HEK293 cells were selected by culturing in the presence of 1 μg/mL puromycin (InvivoGen) for 2 weeks, while stably expressing CHO cells were selected by culturing in a CHO-media omitted GlutaMAX. Stable cells seeded at 0.2×10^6 cells/mL were cultured for 5–7 days on an orbital shaker and media harvested ($1000 \times g$, 3 min, $10{,}000 \times g$, 20 min), diluted (100 mM sodium phosphate, pH 7.4, 2.0 M NaCl, 40 mM imidazole), and incubated with nickel-nitrilotriacetic acid (Ni-NTA) affinity resin (ThermoFisher Scientific) overnight at 4 °C. Resin was collected into 5 mL gravity column (ThermoFishcer Scientific), washed (25 mM sodium phosphate, pH 7.4, 500 mM NaCl, 20 mM imidazole), and Glycocarriers eluted with 300 mM imidazole. Yields were quantified by NuPAGE Coomassie and Pierce™ BCA Protein Assay Kit (ThermoFisher Scientific).

Metabolic labeling of O-Glycocarriers

HEK293 cells were grown in serum-free media with 100 μM Ac_5SiaNAz or Ac_5SiaNPOC for at least 24 h following replenishment with fresh supplemented media. Culture was grown in an orbital shaker for 3 days and the purification of Glycocarriers from media performed as described above. To label SiaNAz with fluorochrome, 1 μg purified O-Glycocarrier (azido STn) was incubated for 1 h at RT in 20 μL of 100 mM 2-iodoacetamide. DBCO-Cy5 (Vector Laboratories Inc.) dissolved in DMSO was added to a final concentration of 100 μM and the reaction incubated for 2 h at 37 °C, blocked by 4x LDS sample buffer (Invitrogen), heat-denatured for 5 min, and analyzed by SDS-PAGE and Cy5 fluorescence.

Production and purification of recombinant GAG-Glycocarriers

For stably expressing of GAG-Glycocarrier CHO cells (1.5×10^6) were transfected with 4 μg plasmids by electroporation using Amaxa kit V and Amaxa Nucleofector 2B (Lonza), plated in 6-well plates (3 mL, 2% GlutaMAX), grown 2 days, and plated in 96-well plates (mini-pools of 250 cells/well) in 100 μL Cellvento 4CHO-C cloning medium (Sigma-Aldrich) without GlutaMAX. Minipools with high expression of GAG-Glycocarriers were selected by ELISA for StrepII tag and SDS-PAGE. GAG-Glycocarriers were purified from media by dialysis (20 mM NaOAc, pH 6.0, 100 mM NaCl) and DEAE-sepharose (Sigma-Aldrich) chromatography with stepwise elution (100 mM NaCl increments) and desalted (10 kDa MWCO Ultra Centrifugal Filters (Millipore)). GAG-Glycocarriers produced in $CHO^{KO\ B4galt7}$ cells without GAG chains were affinity purified with Strep-Tactin (IBA Lifesciences) following the manufacturer's instructions. Briefly, cleared culture media was loaded onto Strep-Tactin beads (2 mL) equilibrated in 100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA, and the GAG-Glycocarriers eluted by 2.5 mM desthiobiotin and desalted.

Enzymatic digestion of CS and HS

GAG-Glycocarriers (4.8 μg) was mixed with 10 mU chondroitinase ABC (Sigma-Aldrich) and/or 5 mU heparinases I, II, and III (IBEX pharmaceuticals) in 40 mM NaOAc for 4 h at 37 °C. Reaction was stopped by LDS sample buffer (ThermoFisher Scientific) with 10% 2-mercaptoethanol, heated at 90 °C for 10 min, and analyzed by SDS-PAGE Coomassie.

HPLC Disaccharide analysis of GAGs

Disaccharide analysis was performed based on published protocol[52]. Briefly, GAG-Glycocarriers purified from 250 μL media were treated

with 20 mU chondroitinase ABC in 40 mM NaOAc and 5 mM $CaCl_2$ or 15 mU heparinases I, II, and III in 50 mM NaOAc and 5 mM $Ca(OAc)_2$ overnight at 37 °C, heat inactivated at 98 °C for 10 min, and lyophilized. Disaccharides were labeled with 2-aminoacridone (AMAC) by dissolving samples in 10 μL 0.1 M AMAC in 3:17 (vol/vol) acetic acid:-DMSO, incubating for 15 min in RT, addition of 10 μL 1 M $NaCNBH_3$ with incubation at 45 °C for 3 h, and lyophilization. Labeled disaccharides were precipitated twice using 500 μL acetone and dissolved in 2% acetonitrile and analyzed on a Waters Acquity UPLC system with a BEH C18 column (2.1 × 150 mm, 1.7 μm, Waters) at 525 nm. For CS analysis, 80 mM ammonium acetate (pH 5.5) was used as mobile phase A and acetonitrile as mobile phase B. For HS analysis, 150 mM ammonium acetate (pH 5.6) was used as mobile phase A and acetonitrile as mobile phase B. Disaccharides were separated by 3–13% B over 30 min at 0.2 mL/min. AMAC-labeled disaccharide standards (20 pmol) were always analyzed immediately before samples.

Isolation of multimer glycomodules and glycopeptides

100 μg purified Glycocarriers were digested with 2 μg in-house purified TEV protease at 30 °C for 18 h in 50 mM ammonium bicarbonate buffer (pH 8.0). For samples requiring desialylation, 40 mU *C. perfringens* neuraminidase (Sigma-Aldrich) was sequentially added to the reaction mixture and incubated overnight at 37 °C. Enzymes were heat inactivated at 98 °C for 15 min and the glycomodules were separated by C4 HPLC (Aeris™ WIDEPORE C4, 3.6 μm, 200 Å, 250 × 2.1 mm, Phenomenex) using a 0–100% gradient of 90% acetonitrile in 0.1% TFA. Fractions containing the released glycomodules were verified by ELISA using glycoform-specific lectins or mAbs, dried and resuspended in 20 μl of 0.1% FA for intact mass analysis. For bottom-up analysis, the multimer glycomodule samples (50 μg) were further digested with MS-grade trypsin (Roche) at a 1:35 ratio for 18 h at 37 °C in 50 mM ammonium bicarbonate buffer (pH 8.0). After inactivation by addition of TFA to 0.1% v/v concentration, samples were separated by C18 HPLC (Aeris™ WIDEPORE C18, 3.6 μm, 200 Å, 250 × 2.1 mm, Phenomenex) or PGC HPLC (Hypercarb PGC, 3.0 μm, 250 Å,150 × 2.1 mm, ThermoFisher Scientific) using a 0–100% gradient of 90% acetonitrile in 0.1% TFA before being analyzed by LC-MS/MS.

Intact mass analysis

Purified glycomodules were analyzed by either direct injection in the Synapt G2 system equipped with a T-wave ion mobility cell (Waters Ltd) or with EASY-nLC 1200 HPLC (ThermoFisher Scientific) interfaced via nanoSpray Flex ion source to an OrbitTrap Fusion/Lumos instrument (ThermoFisher Scientific). In the Synapt G2, samples were mixed in 50:50 (v/v) methanol:water and directly infused into Synapt G2 mass spectrometer using in-house fabricated Pd/Pt-coated borosilicate tip. Samples were ionized using a capillary voltage of 1.0 kV, sample cone voltage of 45 V and an extraction cone voltage of 5 V. All samples were acquired in positive ion mode in mass range setting in m/z rang 500–5000. For LC-MS analysis, OrbitTrap Fusion/Lumos instrument was set using "high" mass range setting in m/z range 700–4000. The instrument was operated in a low-pressure mode to provide optimal detection of intact protein masses. MS parameters settings spray voltage 2.2 kV, source fragmentation energy 35 V. All ions were detected in OrbiTrap at the resolution of 7500 (at m/z 200). The number of microscans was set to 20. The nLC was operated in a single analytical column set up using PicoFrit Emitters (New Objectives, 75 mm inner diameter) packed in-house with C4-phase. Each sample was injected onto the column and eluted in gradients from 5 to 30% B in 25 min, from 30 to 100% B in 20 mins, and 100% B for 15 mins at 300 nL/min (Solvent A 100% water; Solvent B 80% acetonitrile; both containing 0.1% (v/v) formic acid).

LC-MS/MS analysis

LC-MS/MS analysis was performed on EASY-nLC 1200 UHPLC (ThermoFisher Scientific) interfaced via nanoSpray Flex ion source to an Orbitrap Fusion Lumos MS (ThermoFisher Scientific). Briefly, the nLC was operated in a single analytical column set up using PicoFrit Emitters (New Objectives, 75 mm inner diameter) packed in-house with Reprosil-Pure-AQ C18 phase (Dr. Maisch, 1.9-mm particle size, 19–21 cm column length). Each sample was injected onto the column and eluted in gradients from 3 to 32% B for glycopeptides, and 10–40% for released and labeled glycans in 45 min at 200 nL/min (Solvent A, 100% water; Solvent B, 80% acetonitrile; both containing 0.1% (v/v) formic acid). A precursor MS1 scan (m/z 350–2000) of intact peptides was acquired in the Orbitrap at the nominal resolution setting of 120,000, followed by Orbitrap HCD-MS2 and ETD-MS2 at the nominal resolution setting of 60,000 of the five most abundant multiply charged precursors in the MS1 spectrum; a minimum MS1 signal threshold of 50,000 was used for triggering data-dependent fragmentation events. Targeted MS/MS analysis was performed by setting up a targeted MSn (tMSn) Scan Properties pane.

Glycoprofiling

O-glycoprofiling was performed as described with minor modifications[23,35]. Briefly, 20 μg purified O-Glycocarriers were incubated in 0.1 M NaOH/1.0 M $NaBH_4$ solution at 50 °C for 16 h. Released O-glycans were purified by passing through in-house C18 stage tip (Empore 3 M) and desalted on Dowex (AG 50 W 8X). Glycans were permethylated at RT for 1 h using 150 μL DMSO, 20 mg NaOH powder and 30 μL methyl iodide. 200 μL ice-cold MQ water was added to quench the reaction, followed by the addition of 200 μL chloroform. The organic phase was washed 5 times with 1 mL MQ water before evaporated under nitrogen stream. Dried permethylated O-glycans were resuspended in 50% (v/v) methanol, of which 1 μL was co-crystallized with 1 μL DHB matrix (10 mg/mL in 70% acetonitrile, 0.1% TFA, 0.5 mM sodium acetate) before being analyzed by MALDI-TOF MS (Autoflex Speed, Bruker Daltonics) in positive mode with data acquisition at 1000 shots/spot in the 400-4000 Da mass range. N-glycoprofiling was performed as previously described[107]. Briefly, 10 μg purified N-Glycocarrier was incubated with 2 U of PNGase F from *Elizabethkingia meningoseptica* (Sigma) in PBS supplemented with 0.8% NP-40 at 37 °C for 16 h, and released N-glycans labeled with aminobenzoic acid (2-AB) in the presence of 2-Methylpyridine borane complex (PB) in 90:10 (v/v) methanol acetic acid solvent and purified using in-house HILIC cotton column eluted in MiliQ water, followed by LC-MS/MS.

ELISA

MaxiSorp 96-well plates (Nunc) were coated with purified Glycocarriers overnight in 50 μl carbonate-bicarbonate buffer (pH 9.6) at 4 °C, blocked (PO_4, Na/K, 1% Triton-X100, 1% BSA, pH 7.4) for 1 h at RT, and incubated with mAbs or biotinylated-lectins for 1 h at RT. Plates were washed in PBS containing 0.05% Tween-20, incubated with 50 μl of 1 μg/mL HRP conjugated anti-mouse Ig (Dako) or 1 μg/mL streptavidin-conjugated HRP (Dako) for 1 h at RT, followed by development with TMB (Dako), termination with 0.5 M H_2SO_4, and measurement (450 nm) with Synergy LX (BioTek). All mAbs and lectins used in the study are provided in the Supplementary Table 3.

Peptiligase-based glycoprotein synthesis

Peptides were dissolved in Tricine buffer (200 mM, pH 8.3, 4 mM TCEP) to a concentration of 20 mM ester fragment, 1 mM amine fragment and 20 μM omniligase-1. The reaction was performed at room temperature for 4 h and followed via RP-HPLC/MS (Agilent 1100, LC-MSD SL) using an MeCN/H_2O gradient on a C-18 column (Phenomenex 5 μm EVO C18 100 Å 150 × 4.6 mm). After completion of the ligation, the product was purified by preparative HPLC (Waters 2545, 2998, Aquity QDA) using a MeCN/H_2O gradient on a C18 column (XBridge C18 5 μm OBD 30 × 100 mm). Finally, the product was lyophilized (Christ alpha 2-4 LSCbasic) and checked for purity via HPLC/MS (Agilent 1100, LC-MSD SL). Peptiligases, such as omniligase-1, are a

patented technology by EnzyPep B.V. and was obtained through EnzyTag B.V. (www.enzytag.com). The enzyme is produced by recombinant expression in Bacillus subtilis and purified via a C-terminal hexahistidine tag.

HepG2 uptake assay

For flow cytometry analysis, HepG2 cells cultured in 6-well plates at ~60% confluence were seeded (50,000 cells per well) in 96-well plates 12 hrs prior to assays, and cells replenished with fresh media containing purified Glycocarriers and incubated at 37 °C for 1 h before harvesting by TrypLE (Gibco). Cells were washed twice with PBA (PBS supplemented with 1% FBS), fixed with 4% formaldehyde for 10 min at 4°C, washed once with PBA, and analyzed by flow cytometry (Fortessa X20, BD Biosciences). Results were analyzed using FlowJo software (Flowjo LLC) and expressed as geometric mean fluorescence of the activated gated cells. For confocal cytology analysis, HepG2 cells (0.5×10^6) were seeded in a μ-Dish 35 mm (IBIDI) for 24 hrs, media replenished with fresh media containing 5 μg/mL Glycocarrier and grown at 37 °C for 2 h. Cells were washed trice with cold PBS and stained with 0.5 μg/mL AF647-WGA lectin (Vector Laboratory) at 4 °C for 15 min, followed by twice washing with cold PBS and fixing with 4% formaldehyde for 15 min at RT. After washed twice with cold PBS, cells were mounted in ProLong Gold antifade reagent (Invitrogen) and imaged using Zeiss LSM900 microscope equipped with AxioCam 702 mono camera and a laser scanning confocal.

Galectin biomolecular condensate formation assay

Human recombinant wild-type Gal3 and Gal3ΔN-ter (CRD) with C-terminal 6xHis tags were prepared as previously described[66]. Briefly, for cloning, Gal3 (full length) or Gal3ΔNter-6xHis (aa115-250) of BC053667 and 6xHis were separated by a Leu- Glu linker at the C-terminus and cloned in pHis-Parallel2. Both proteins were expressed overnight at 20 °C in Rossetta2-pLysS using 3 L LB-media with 60 μM IPTG, and purified with cobalt resin (ThermoFisher Scientific) affinity chromatography and gel filtration (Superdex75 16 × 60) in PBS at pH 7.4. Small aliquots for single use were snap-frozen and stored at −80 °C. Final protein yield was around 15 mg/L culture. Immediately before assays, N-Glycocarriers and Gal3 were individually diluted in PBS to 10 μM stocks, and serial dilutions of N-Glycocarriers were mixed with 7.5 μM Gal3 or CRD in Eppendorf tubes and incubated at RT for 15 min. Reaction mixtures (2 μL) were spotted on a 10 well, 6.7 mm diagnostic glass slide (ThermoFisher Scientific) and capped with 12 mm diameter microscope cover slip (VWR international). LLPS droplets were visualized using a Zeiss Axioskop 2 plus with an AxioCam MR3 in a white field and wide field (Ex/Em 475/509 nm) modes.

Mass photometry analysis

Microscope cover glasses (No 1.5H, 24 × 50 mm, Paul Marienfeld GmbH) were prepared by sequential washing with MQ-water and HPLC-grade isopropanol (5 times each) and dried under nitrogen stream. A silicon gasket was fixed on clean cover glasses by gently pressing with forceps. Native protein marker (InvitroGen) was used to generate mass calibration curves (66, 146, 480, and 1048 kDa). Immediately prior measurements, protein stocks were diluted to 400 nM in PBS, and analyzed in mass photometry at 5 nM. For each acquisition, a new well in a silicon gasket was used and sample containing antigen and binder, at specific concentration, was introduced into a well. Following autofocus stabilization, a movie was recorded for 90 s at RT. All data acquisition was performed using AcquireMP software (Refeyn Ltd) and data was analyzed using DiscoverMP (Refeyn Ltd). Data was presented as kernel density estimates with a 5 kDa bandwidth.

Data analysis

Glycopeptide compositional analysis was performed from m/z features extracted from LC-MS data using SysBioWare software[108]. For m/z feature recognition from full MS scans Minora Feature Detector Node of the Proteome discoverer 2.2 (ThermoFisher Scientific) was used. The list of precursor ions (m/z, charge, peak area) was imported as ASCII data into SysBioWare and compositional assignment within 3 ppm mass tolerance was performed. The main building blocks used for the compositional analysis were: NeuAc, Hex, HexNAc, dHex and the theoretical mass increment of the most prominent peptide corresponding to each potential glycosites. Upon generation of the potential glycopeptide list each glycosite was rank for the top 10 most abundant candidates and each candidate structure was confirmed by doing targeted MS/MS analysis followed by manual interpretation of the corresponding MS/MS spectrum. For intact mass analysis, raw spectra were deconvoluted to zero-charge by BioPharma Finder Software (ThermoFisher Scientific) using default settings. Glycoproteoforms were annotated by in-house written SysBioWare software using average masses of Hexose, N-acetylhexosamine, and the known backbone mass of Glycocarrier reporter sequence.

Statistics and reproducibility

No statistical method was used to predetermine sample size, and no data were excluded from the analyses. The experiments were not randomized, and the investigators were not blinded to allocation during experiments and outcome assessment.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data generated in the study are included in this article and its supplementary information files/Source Data file. The mass spectrometry proteomics data are available via ProteomeXchange with identifier PXD056690. Source data are provided with this paper.

References

1. Varki, A. Biological roles of glycans. *Glycobiology* **27**, 3–49 (2017).
2. Taylor, M. E. & Drickamer, K. Mammalian sugar-binding receptors: known functions and unexplored roles. *FEBS J.* **286**, 1800–1814 (2019).
3. Cummings, R. D. The repertoire of glycan determinants in the human glycome. *Mol. Biosyst.* **5**, 1087–1104 (2009).
4. Varki, A. et al. In *Essentials of Glycobiology* 3rd edn (eds Varki, A. et al.) 493–502 (Cold Spring Harbor, 2015).
5. Schnaar, R. L. Glycans and glycan-binding proteins in immune regulation: a concise introduction to glycobiology for the allergist. *J. Allergy Clin. Immunol.* **135**, 609–615 (2015).
6. Poole, J., Day, C. J., von Itzstein, M., Paton, J. C. & Jennings, M. P. Glycointeractions in bacterial pathogenesis. *Nat. Rev. Microbiol.* **16**, 440–452 (2018).
7. Narimatsu, Y. et al. An atlas of human glycosylation pathways enables display of the human glycome by gene engineered cells. *Mol. Cell* **75**, 394–407.e395 (2019).
8. Feizi, T. & Chai, W. Oligosaccharide microarrays to decipher the glyco code. *Nat. Rev. Mol. Cell Biol.* **5**, 582–588 (2004).
9. Blixt, O. et al. Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc. Natl. Acad. Sci. USA* **101**, 17033–17038 (2004).
10. Rillahan, C. D. & Paulson, J. C. Glycan microarrays for decoding the glycome. *Annu. Rev. Biochem.* **80**, 797–823 (2011).
11. Gao, C. et al. Glycan microarrays as chemical tools for identifying glycan recognition by immune proteins. *Front. Chem.* **7**, 833 (2019).
12. Pardo-Vargas, A., Delbianco, M. & Seeberger, P. H. Automated glycan assembly as an enabling technology. *Curr. Opin. Chem. Biol.* **46**, 48–55 (2018).

13. Jaroentomeechai, T. et al. Cell-free synthetic glycobiology: designing and engineering glycomolecules outside of living cells. *Front. Chem.* **8**, 645 (2020).
14. Kracun, S. K. et al. Random glycopeptide bead libraries for seromic biomarker discovery. *J. Proteome Res.* **9**, 6705–6714 (2010).
15. Mehta, A. Y. et al. Parallel glyco-SPOT synthesis of glycopeptide libraries. *Cell Chem. Biol.* **27**, 1207–1219.e1209 (2020).
16. Bull, C., Joshi, H. J., Clausen, H. & Narimatsu, Y. Cell-based glycan arrays-a practical guide to dissect the human glycome. *STAR Protoc.* **1**, 100017 (2020).
17. Chen, Y. H. et al. The GAGOme: a cell-based library of displayed glycosaminoglycans. *Nat. Methods* **15**, 881–888 (2018).
18. Bode, L. The functional biology of human milk oligosaccharides. *Early Hum. Dev.* **91**, 619–622 (2015).
19. Werlang, C. A. et al. Mucin O-glycans suppress quorum-sensing pathways and genetic transformation in Streptococcus mutans. *Nat. Microbiol.* **6**, 574–583 (2021).
20. Wheeler, K. M. et al. Mucin glycans attenuate the virulence of Pseudomonas aeruginosa in infection. *Nat. Microbiol.* **4**, 2146–2154 (2019).
21. Wagner, S. et al. Covalent lectin inhibition and application in bacterial biofilm imaging. *Angew. Chem. Int. Ed. Engl.* **56**, 16559–16564 (2017).
22. Schjoldager, K. T., Narimatsu, Y., Joshi, H. J. & Clausen, H. Global view of human protein glycosylation pathways and functions. *Nat. Rev. Mol. Cell Biol.* **21**, 729–749 (2020).
23. Yang, Z. et al. Engineered CHO cells for production of diverse, homogeneous glycoproteins. *Nat. Biotechnol.* **33**, 842–844 (2015).
24. Shields, R. L. et al. Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fcgamma RIII and antibody-dependent cellular toxicity. *J. Biol. Chem.* **277**, 26733–26740 (2002).
25. Chen, Y. H. et al. A universal GlycoDesign for lysosomal replacement enzymes to improve circulation time and biodistribution. *Front. Bioeng. Biotechnol.* **11**, 1128371 (2023).
26. Tian, W. et al. The glycosylation design space for recombinant lysosomal replacement enzymes produced in CHO cells. *Nat. Commun.* **10**, 1785 (2019).
27. Caval, T., Tian, W., Yang, Z., Clausen, H. & Heck, A. J. R. Direct quality control of glycoengineered erythropoietin variants. *Nat. Commun.* **9**, 3342 (2018).
28. Walsh, G. & Walsh, E. Biopharmaceutical benchmarks 2022. *Nat. Biotechnol.* **40**, 1722–1760 (2022).
29. Sasaki, H., Ochi, N., Dell, A. & Fukuda, M. Site-specific glycosylation of human recombinant erythropoietin: analysis of glycopeptides or peptides at each glycosylation site by fast atom bombardment mass spectrometry. *Biochemistry* **27**, 8618–8626 (1988).
30. Lai, P. H., Everett, R., Wang, F. F., Arakawa, T. & Goldwasser, E. Structural characterization of human erythropoietin. *J. Biol. Chem.* **261**, 3116–3121 (1986).
31. Wandall, H. H. et al. Substrate specificities of three members of the human UDP-N-acetyl-alpha-D-galactosamine:Polypeptide N-acetylgalactosaminyltransferase family, GalNAc-T1, -T2, and -T3. *J. Biol. Chem.* **272**, 23503–23514 (1997).
32. Yoshida, A., Suzuki, M., Ikenaga, H. & Takeuchi, M. Discovery of the shortest sequence motif for high level mucin-type O-glycosylation. *J. Biol. Chem.* **272**, 16884–16888 (1997).
33. Briggs, D. C. & Hohenester, E. Structural basis for the initiation of glycosaminoglycan biosynthesis by human xylosyltransferase 1. *Structure* **26**, 801–809.e803 (2018).
34. Roch, C., Kuhn, J., Kleesiek, K. & Gotting, C. Differences in gene expression of human xylosyltransferases and determination of acceptor specificities for various proteoglycans. *Biochem. Biophys. Res. Commun.* **391**, 685–691 (2010).
35. Nason, R. et al. Display of the human mucinome with defined O-glycans by gene engineered cells. *Nat. Commun.* **12**, 4070 (2021).
36. Konstantinidi, A. et al. Exploring the glycosylation of mucins by use of O-glycodomain reporters recombinantly expressed in glycoengineered HEK293 cells. *J. Biol. Chem.* **298**, 101784 (2022).
37. Springer, G. F. T and Tn, general carcinoma autoantigens. *Science* **224**, 1198–1206 (1984).
38. Ju, T., Otto, V. I. & Cummings, R. D. The Tn antigen-structural simplicity and biological complexity. *Angew. Chem. Int. Ed. Engl.* **50**, 1770–1791 (2011).
39. Luetscher, R. N. D. et al. Unique repertoire of anti-carbohydrate antibodies in individual human serum. *Sci. Rep.* **10**, 15436 (2020).
40. Blixt, O. et al. Analysis of Tn antigenicity with a panel of new IgM and IgG1 monoclonal antibodies raised against leukemic cells. *Glycobiology* **22**, 529–542 (2012).
41. Dobrochaeva, K. et al. Specificity of human natural antibodies referred to as anti-Tn. *Mol. Immunol.* **120**, 74–82 (2020).
42. Mazal, D. et al. Monoclonal antibodies toward different Tn-amino acid backbones display distinct recognition patterns on human cancer cells. Implications for effective immuno-targeting of cancer. *Cancer Immunol. Immunother.* **62**, 1107–1122 (2013).
43. Yang, Z. et al. The GalNAc-type O-glycoproteome of CHO cells characterized by the simplecell strategy. *Mol. Cell Proteom.* **13**, 3224–3235 (2014).
44. Steentoft, C., Bennett, E. P. & Clausen, H. Glycoengineering of human cell lines using zinc finger nuclease gene targeting: SimpleCells with homogeneous GalNAc O-glycosylation allow isolation of the O-glycoproteome by one-step lectin affinity chromatography. *Methods Mol. Biol.* **1022**, 387–402 (2013).
45. Narimatsu, Y. et al. Genetic glycoengineering in mammalian cells. *J. Biol. Chem.* **296**, 100448 (2021).
46. Zhou, D., Berger, E. G. & Hennet, T. Molecular cloning of a human UDP-galactose:GlcNAcbeta1,3GalNAc beta1, 3 galactosyltransferase gene encoding an O-linked core3-elongation enzyme. *Eur. J. Biochem.* **263**, 571–576 (1999).
47. Holgersson, J. & Lofling, J. Glycosyltransferases involved in type 1 chain and Lewis antigen biosynthesis exhibit glycan and core chain specificity. *Glycobiology* **16**, 584–593 (2006).
48. Mandel, U. et al. Simple mucin-type carbohydrates in oral stratified squamous and salivary gland epithelia. *J. Investig. Dermatol.* **97**, 713–721 (1991).
49. Schulz, M. A. et al. Glycoengineering design options for IgG1 in CHO cells using precise gene editing. *Glycobiology* **28**, 542–549 (2018).
50. Zielinska, D. F., Gnad, F., Wisniewski, J. R. & Mann, M. Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. *Cell* **141**, 897–907 (2010).
51. Zhang, X., Lin, L., Huang, H. & Linhardt, R. J. Chemoenzymatic synthesis of glycosaminoglycans. *Acc. Chem. Res.* **53**, 335–346 (2020).
52. Karlsson, R. et al. Dissecting structure-function of 3-O-sulfated heparin and engineered heparan sulfates. *Sci. Adv.* **7**, eabl6026 (2021).
53. Kuhn, J. et al. Xylosyltransferase II is the predominant isoenzyme which is responsible for the steady-state level of xylosyltransferase activity in human serum. *Biochem. Biophys. Res. Commun.* **459**, 469–474 (2015).
54. Noborn, F., Nilsson, J. & Larson, G. Site-specific glycosylation of proteoglycans: a revisited frontier in proteoglycan research. *Matrix Biol.* **111**, 289–306 (2022).
55. Kolset, S. O. & Tveit, H. Serglycin-structure and biology. *Cell Mol. Life Sci.* **65**, 1073–1085 (2008).
56. Pugia, M. J., Valdes, R. Jr. & Jortani, S. A. Bikunin (urinary trypsin inhibitor): structure, biological relevance, and measurement. *Adv. Clin. Chem.* **44**, 223–245 (2007).

57. Zhang, L., David, G. & Esko, J. D. Repetitive Ser-Gly sequences enhance heparan sulfate assembly in proteoglycans. *J. Biol. Chem.* **270**, 27127–27135 (1995).
58. Esko, J. D. & Selleck, S. B. Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annu. Rev. Biochem* **71**, 435–471 (2002).
59. Matsumoto, Y. et al. Identification of Tn antigen O-GalNAc-expressing glycoproteins in human carcinomas using novel anti-Tn recombinant antibodies. *Glycobiology* **30**, 282–300 (2020).
60. Matsumoto-Takasaki, A. et al. Surface plasmon resonance and NMR analyses of anti Tn-antigen MLS128 monoclonal antibody binding to two or three consecutive Tn-antigen clusters. *J. Biochem* **151**, 273–282 (2012).
61. Tarp, M. A. & Clausen, H. Mucin-type O-glycosylation and its potential use in drug and vaccine development. *Biochim. Biophys. Acta* **1780**, 546–563 (2008).
62. Salanti, A. et al. Targeting human cancer by a glycosaminoglycan binding malaria protein. *Cancer Cell* **28**, 500–514 (2015).
63. Vidal-Calvo, E. E. et al. Tumor-agnostic cancer therapy using antibodies targeting oncofetal chondroitin sulfate. *Nat. Commun.* **15**, 7553 (2024).
64. Lepur, A., Salomonsson, E., Nilsson, U. J. & Leffler, H. Ligand induced galectin-3 protein self-association. *J. Biol. Chem.* **287**, 21751–21756 (2012).
65. Johannes, L., Jacob, R. & Leffler, H. Galectins at a glance. *J. Cell Sci.* **131**, jcs208884 (2018).
66. Lakshminarayan, R. et al. Galectin-3 drives glycosphingolipid-dependent biogenesis of clathrin-independent carriers. *Nat. Cell Biol.* **16**, 595–606 (2014).
67. Lin, Y. H. et al. The intrinsically disordered N-terminal domain of galectin-3 dynamically mediates multisite self-association of the protein through fuzzy interactions. *J. Biol. Chem.* **292**, 17845–17856 (2017).
68. Chiu, Y. P. et al. Liquid-liquid phase separation and extracellular multivalent interactions in the tale of galectin-3. *Nat. Commun.* **11**, 1229 (2020).
69. Zhao, Z. et al. Galectin-3 N-terminal tail prolines modulate cell activity and glycan-mediated oligomerization/phase separation. *Proc. Natl. Acad. Sci. USA* **118**, e2021074118 (2021).
70. Bull, C. et al. Steering siglec-sialic acid interactions on living cells using bioorthogonal chemistry. *Angew. Chem. Int. Ed. Engl.* **56**, 3309–3313 (2017).
71. Dube, D. H. & Bertozzi, C. R. Metabolic oligosaccharide engineering as a tool for glycobiology. *Curr. Opin. Chem. Biol.* **7**, 616–625 (2003).
72. Biessen, E. A. et al. Targeted delivery of oligodeoxynucleotides to parenchymal liver cells in vivo. *Biochem J.* **340**, 783–792 (1999).
73. Debacker, A. J., Voutila, J., Catley, M., Blakey, D. & Habib, N. Delivery of oligonucleotides to the liver with GalNAc: from research to registered therapeutic drug. *Mol. Ther.* **28**, 1759–1771 (2020).
74. D'Souza, A. A. & Devarajan, P. V. Asialoglycoprotein receptor mediated hepatocyte targeting—strategies and applications. *J. Control Release* **203**, 126–139 (2015).
75. Wang, P. et al. Erythropoietin derived by chemical synthesis. *Science* **342**, 1357–1360 (2013).
76. Shivatare, S. S., Shivatare, V. S. & Wong, C. H. Glycoconjugates: synthesis, functional studies, and therapeutic developments. *Chem. Rev.* **122**, 15603–15671 (2022).
77. Kulkarni, S. S., Sayers, J., Premdjee, B. & Payne, R. J. Rapid and efficient protein synthesis through expansion of the native chemical ligation concept. *Nat. Rev. Chem.* **2**, 0122 (2018).
78. Schmidt, M. et al. Design of a substrate-tailored peptiligase variant for the efficient synthesis of thymosin-alpha(1). *Org. Biomol. Chem.* **16**, 609–618 (2018).
79. Marglous, S., Brown, C. E., Padler-Karavani, V., Cummings, R. D. & Gildersleeve, J. C. Serum antibody screening using glycan arrays. *Chem. Soc. Rev.* **53**, 2603–2642 (2024).
80. Anggara, K. et al. Direct observation of glycans bonded to proteins and lipids at the single-molecule level. *Science* **382**, 219–223 (2023).
81. Co, J. Y. et al. Mucins trigger dispersal of Pseudomonas aeruginosa biofilms. *NPJ Biofilms Microbiomes* **4**, 23 (2018).
82. Wang, B. X. et al. Mucin glycans signal through the sensor kinase RetS to inhibit virulence-associated traits in Pseudomonas aeruginosa. *Curr. Biol.* **31**, 90–102.e107 (2021).
83. Wang, B. X. et al. Host-derived O-glycans inhibit toxigenic conversion by a virulence-encoding phage in Vibrio cholerae. *EMBO J.* **42**, e111562 (2023).
84. Bode, L. Human milk oligosaccharides: every baby needs a sugar mama. *Glycobiology* **22**, 1147–1162 (2012).
85. Liu, B. & Newburg, D. S. Human milk glycoproteins protect infants against human pathogens. *Breastfeed. Med.* **8**, 354–362 (2013).
86. Yang, H. et al. ABO genotype alters the gut microbiota by regulating GalNAc levels in pigs. *Nature* **606**, 358–367 (2022).
87. Zhernakova, D. V. et al. Host genetic regulation of human gut microbial structural variation. *Nature* **625**, 813–821 (2024).
88. Xu, D. & Esko, J. D. Demystifying heparan sulfate-protein interactions. *Annu. Rev. Biochem.* **83**, 129–157 (2014).
89. Mizumoto, S., Yamada, S. & Sugahara, K. Molecular interactions between chondroitin-dermatan sulfate and growth factors/receptors/matrix proteins. *Curr. Opin. Struct. Biol.* **34**, 35–42 (2015).
90. Shipp, E. L. & Hsieh-Wilson, L. C. Profiling the sulfation specificities of glycosaminoglycan interactions with growth factors and chemotactic proteins using microarrays. *Chem. Biol.* **14**, 195–208 (2007).
91. Horton, M. et al. Construction of heparan sulfate microarray for investigating the binding of specific saccharide sequences to proteins. *Glycobiology* **31**, 188–199 (2021).
92. Rosenbaum, P. et al. The fully synthetic glycopeptide MAG-Tn3 therapeutic vaccine induces tumor-specific cytotoxic antibodies in breast cancer patients. *Cancer Immunol. Immunother.* **69**, 703–716 (2020).
93. Reis, C. A. et al. Development and characterization of an antibody directed to an alpha-N-acetyl-D-galactosamine glycosylated MUC2 peptide. *Glycoconj. J.* **15**, 51–62 (1998).
94. Sorensen, A. L. et al. Chemoenzymatically synthesized multimeric Tn/STn MUC1 glycopeptides elicit cancer-specific anti-MUC1 antibody responses and override tolerance. *Glycobiology* **16**, 96–107 (2006).
95. Schietinger, A. et al. A mutant chaperone converts a wild-type protein into a tumor-specific antigen. *Science* **314**, 304–308 (2006).
96. Steentoft, C. et al. A strategy for generating cancer-specific monoclonal antibodies to aberrant O-glycoproteins: identification of a novel dysadherin-Tn antibody. *Glycobiology* **29**, 307–319 (2019).
97. Abrahmsen, L. et al. Engineering subtilisin and its substrates for efficient ligation of peptide bonds in aqueous solution. *Biochemistry* **30**, 4151–4159 (1991).
98. Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* **77**, 521–555 (2008).
99. Pratt, M. R. & Bertozzi, C. R. Synthetic glycopeptides and glycoproteins as tools for biology. *Chem. Soc. Rev.* **34**, 58–68 (2005).
100. Hudak, J. E. & Bertozzi, C. R. Glycotherapy: new advances inspire a reemergence of glycans in medicine. *Chem. Biol.* **21**, 16–37 (2014).
101. Hong, S. et al. Glycoengineering of NK cells with glycan ligands of CD22 and selectins for B-cell lymphoma therapy. *Angew. Chem. Int. Ed. Engl.* **60**, 3603–3610 (2021).

102. Hintze, J. et al. Probing the contribution of individual polypeptide GalNAc-transferase isoforms to the O-glycoproteome by inducible expression in isogenic cell lines. *J. Biol. Chem.* **293**, 19064–19077 (2018).
103. Steentoft, C. et al. A validated collection of mouse monoclonal antibodies to human glycosyltransferases functioning in mucin-type O-glycosylation. *Glycobiology* **29**, 645–656 (2019).
104. Narimatsu, Y. et al. A validated gRNA library for CRISPR/Cas9 targeting of the human glycosyltransferase genome. *Glycobiology* **28**, 295–305 (2018).
105. Lonowski, L. A. et al. Genome editing using FACS enrichment of nuclease-expressing cells and indel detection by amplicon analysis. *Nat. Protoc.* **12**, 581–603 (2017).
106. Pinto, R. et al. Precise integration of inducible transcriptional elements (PrIITE) enables absolute control of gene expression. *Nucleic Acids Res.* **45**, e123 (2017).
107. de Haan, N. et al. Sensitive and specific global cell surface N-glycoproteomics shows profound differences between glycosylation sites and subcellular components. *Anal. Chem.* **95**, 17328–17336 (2023).
108. Vakhrushev, S. Y., Dadimov, D. & Peter-Katalinic, J. Software platform for high-throughput glycomics. *Anal. Chem.* **81**, 3252–3260 (2009).
109. Varki, A. et al. Symbol nomenclature for graphical representations of glycans. *Glycobiology* **25**, 1323–1324 (2015).

Acknowledgements

We thank Katarina Madunic for help with mass spectrometry data interpretation and Stan Van der Beelen for help with glycopeptide ligation. We acknowledge the Core Facility for Flow Cytometry and Single Cell Analysis and Core Facility for Integrated Microscopy, Faculty of Health and Medical Sciences, University of Copenhagen, and the CurieCoreTech facility for Recombinant Proteins. This work was supported by the NEYE Foundation and the European Molecular Biology Organization (EMBO) postdoctoral fellowship (NEYE Foundation ID: 24030064 and ALTF 336-2021 to T.J), Novo Nordisk Foundation (NNF24OC0088218 to Y.N., NNF22OC0073736 to R.L.M., NNF22OC0073793 to K.T.S., NNF22OC0076899 to H.J.J., NNF22OC0076055 and NNF21OC0068192 to A.S., and NNF21OC0071658 to H.C.), Lundbeck Foundation (R223-2016-563 to Y.N. and H.C.), Independent Research Fund Denmark (2066-00043B to KTS), the Human Frontier Science Program grant RGEC31/2023 (to R.L.M.), Innovation Fund Denmark (9065-00216A to E.E.V-.C.), ERC Proof of Concept "Lectibodies" (project 101062030 to L.J.), Foundation pour la Recherche Médicale (EQU202103012926 to L.J.), EMBO postdoctoral fellowship (ALTF 105-2023 to F.G.). Carlsberg Foundation (CF22-0354 to R.K. and CF20-0412 to R.L.M.), the EU Marie Sklodowska-Curie (101106296 to F.K.Y.T), and the MSCA-ITN-2020 grant (Glytunes, 956758 to T.J.B).

Author contributions

T.J. conceived the project, designed and performed all research, analyzed all data, and wrote the manuscript. R.K. contributed to the GAG studies. F.G., F.K.Y.T., M.N.G., D.W. Y.H.C, and S.F. contributed to cell-based binding assays and glycan analyses. R.H-G. contributed to the antibody studies. E.E.V-C. and A.S. contributed with GAG antibodies. V.P. and T.J.B. contributed to the metabolic labeling studies. C.W. and L.J. contributed to the galectin studies. K.T.S., H.J.J., and L.J.v.d.B. contributed to data interpretation. R.M. and S.Y.V. performed mass spectrometry analysis and aided in data interpretation. Y.N. and H.C. conceived the project, designed the experiments, and contributed to writing of the manuscript. All authors read and approved the final manuscript.

Competing interests

Y.N., Y.H.C., and H.C. have a financial interest in GlycoDisplay Aps, Y.N., Y.H.C., and H.C.'s interests are reviewed and managed by the University of Copenhagen in accordance with their conflict of interest policies. L.J.v.d.B. has a financial interest in EnzyTag BV. A.S. and E.E.V-C. have a financial interest in VAR2Pharmaceuticals. A.S. has financial interest in VARCT Diagnostics. All other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-53738-9.

Correspondence and requests for materials should be addressed to Yoshiki Narimatsu.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.