HHS Public Access

Author manuscript

ACS Chem Biol. Author manuscript; available in PMC 2020 June 16.

Published in final edited form as:

ACS Chem Biol. 2020 April 17; 15(4): 1059–1066. doi:10.1021/acschembio.0c00074.

Engineering a proximity-directed O-GlcNAc transferase for selective protein O-GlcNAcylation in cells

Daniel H. Ramirez[1], **Chanat Aonbangkhen**[1], **Hung-Yi Wu**[1], **Jeffrey A. Naftaly**[1,^], **Stephanie Tang**[1], **Timothy R. O'Meara**[1], **Christina M. Woo**[1,*]

[1]Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford St, Cambridge, MA 02138, USA

Abstract

O-Linked *N*-acetylglucosamine (O-GlcNAc) is a monosaccharide that plays an essential role in cellular signaling throughout the nucleocytoplasmic proteome of eukaryotic cells. Strategies to selectively increase O-GlcNAc levels on a target protein in cells would accelerate studies of this essential modification. Here, we report a generalizable strategy to induce O-GlcNAc to selected target proteins in cells using a nanobody as a proximity-directing agent fused to O-GlcNAc transferase (OGT). Fusion of a nanobody that recognizes GFP (nGFP) or a nanobody that recognizes the four-amino acid sequence EPEA (nEPEA) to OGT yielded nanobody-OGT constructs that selectively delivered O-GlcNAc to a series of tagged target proteins (e.g., JunB, cJun, Nup62). Truncation of the tetratricopeptide repeat domain as in OGT(4) increased selectivity for the target protein through the nanobody by reducing global elevation of O-GlcNAc levels in the cell. Quantitative chemical proteomics confirmed the increase in O-GlcNAc to the target protein by nanobody-OGT(4). Glycoproteomics revealed that nanobody-OGT(4) or full-length OGT produced a similar glycosite profile on the target protein JunB and Nup62. Finally, we demonstrate the ability to selectively target endogenous α-synuclein for O-GlcNAcylation in HEK293T cells. These first proximity-directed OGT constructs provide a flexible strategy to target additional proteins and a template for further engineering of OGT and the O-GlcNAc proteome in the future. The use of a nanobody to redirect OGT substrate selection for glycosylation of desired proteins in cells may further constitute a generalizable strategy to control a broader array of post-translational modifications in cells.

Graphical Abstract

*Corresponding Author: Christina M. Woo, cwoo@chemistry.harvard.edu.

^Present Addresses: Laboratory of Stem Cell Biology and Molecular Embryology, The Rockefeller University, New York, NY 10065, USA

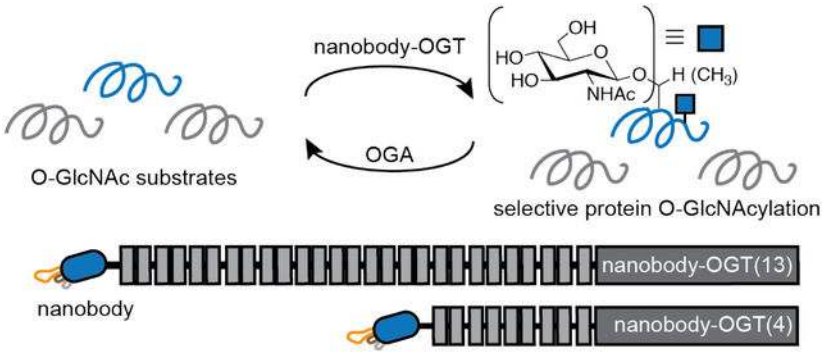
Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Supporting Information, Figures, and Methods (PDF)

Supplementary Tables (XLSX)

Mass spectrometry data is available via ProteomeXchange with identifier PXD016041.



Authors are required to submit a graphic entry for the Table of Contents (TOC) that, in conjunction with the manuscript title, should give the reader a representative idea of one of the following: A key structure, reaction, equation, concept, or theorem, etc., that is discussed in the manuscript. Consult the journal's Instructions for Authors for TOC graphic specifications.

Introduction

O-linked ß-*N*-acetyl glucosamine (O-GlcNAc) is a form of protein glycosylation installed to serine or threonine residues on thousands of nuclear, cytosolic, and mitochondrial proteins. Due to the ubiquitous nature of the modification, O-GlcNAc has been implicated in numerous biological processes and diseases, including the immune response,[1, 2] cancer progression,[3] neurodegeneration,[4, 5] and diabetes.[6, 7] However, the O-GlcNAc modification is challenging to tune on particular glycoproteins within the cell. Global alteration of O-GlcNAc levels can be achieved through manipulating gene expression or chemical inhibitors,[8–10] but relating biological effects to a specific glycoprotein requires extensive follow-up studies. Elimination of the O-GlcNAc modification and a recent strategy to introduce GlcNAc[11] at selected S-glycosites is possible following mapping and mutagenesis of the glycosite in cells. These methods have yielded defined functions for O-GlcNAc,[3, 11] but are challenging to implement for proteins carrying multiple glycosites. Mutagenesis prevents analysis of competing post-translational modification (PTM) pathways (e.g., phosphorylation,[12] ubiquitinylation),[13] and must be developed for every target protein only if the exact glycosite is known. By contrast, a method to selectively induce O-GlcNAcylation on a desired target protein in cells would enable the direct evaluation of the effects of increasing O-GlcNAcylation on that protein without disrupting PTM crosstalk or protein structure due to mutagenesis.

O-GlcNAc is installed and removed by O-GlcNAc transferase (OGT) and O-GlcNAcase (OGA), respectively (Figure 1A).[14–18] OGT is a modular protein found in three major isoforms nucleocytoplasmic OGT (ncOGT), mitochondrial (mOGT), and short (sOGT) consisting of a catalytic domain connected to a variable length tetratricopeptide repeat (TPR) domain (Figure 1B). The TPR domain is thought to primarily direct substrate and glycosite selection,[19, 20] although the parameters that dictate how OGT selects O-GlcNAc modification sites are still under investigation.[21, 22] Given the dynamic nature of O-GlcNAcylation and the large number of substrates modified by OGT, we hypothesized that controlling the O-GlcNAc modification in a protein-selective manner could be achieved

through induced proximity (Figure 1C). Of the several mechanisms to induce protein–protein interactions,[23] the defined properties of nanobodies were particularly attractive. In contrast to antibodies (~150 kDa), nanobodies are small (~12 kDa), highly-selective binding agents that are frequently used in affinity-based assays, imaging, X-ray crystallography, and recently as directing groups to recruit target proteins for degradation,[24, 25] selective phosphorylation,[26] or to a desired genomic loci.[27–29] We thus sought to evaluate the potential for nanobodies to re-direct the activity of OGT and induce O-GlcNAcylation on a series of target proteins.

Here, we report the development of nanobody-OGT fusion proteins to yield the first generation of proximity-directed OGT constructs for elevation of O-GlcNAc levels on target proteins in cells (Figure 1D). We show that fusion of a nanobody to OGT induces O-GlcNAc to expressed or endogenous target proteins in the cell and that substitution of part of the TPR domain with the nanobody increased selectivity for the target protein. Protein-selective O-GlcNAcylation was achieved using a nanobody that recognizes GFP (nGFP)[30] and a nanobody that recognizes EPEA, a four-amino acid sequence derived from α-synuclein (nEPEA).[31] Fusion of these nanobodies to full-length OGT(13) and TPR-truncated OGT(4) induced O-GlcNAcylation to several co-expressed target proteins in HEK293T cells. Quantitative proteomics and glycoproteomics revealed selective induction of O-GlcNAc to the target protein by the nanobody-OGT fusion proteins with similar glycosite selectivity on evaluated target proteins JunB and Nup62, without broad perturbation of the global O-GlcNAc proteome. We finally demonstrate targeted glycosylation of endogenous α-synuclein in HEK293T cells by nEPEA-OGT(4). Thus, the use of nanobodies to proximity-direct OGT yields a versatile platform for protein-selective O-GlcNAcylation in live cells and provides a new mechanism to engineer and control the O-GlcNAc proteome through alteration of the TPR domain.

Results

Design and expression of proximity-directed nanobody-OGT(13) constructs

We initiated our studies with a GFP-recognizing nanobody, nGFP, fused to the full-length OGT that possesses 13 TPRs [residues 1–1046, nGFP(13)] and a RFP or GFP fusion to full-length OGT [RFP(13), GFP(13)] as an untargeted control for comparison to the nanobody-OGT construct due to similarity in expression levels (Figure 2A). All fusions to OGT were connected by an alpha-helical rigid linker $(EAAAK)_4$. Expression of OGT(13) and RFP(13) was distributed throughout the cytoplasm and nucleus as shown by confocal fluorescence microscopy (Figure 2B). The nGFP(13) construct was distributed throughout the nucleocytoplasmic space in an analogous manner.

We next evaluated the proximity-directing ability of nGFP(13) in HEK293T cells co-expressing GFP-Flag-JunB-EPEA, a transcription factor carrying multiple O-GlcNAc sites,[18] as the target protein. Immunoprecipitation of GFP-Flag-JunB-EPEA and probing for O-GlcNAc revealed an increase in O-GlcNAc levels on the target protein that was dependent on the expression levels of the co-transfected nGFP(13) (Figure 2C). The O-GlcNAcylated target protein was significantly increased when co-expressed with nGFP(13) as compared to RFP(13) (Figure 2D, E). However, global O-GlcNAc levels were similarly elevated in the

presence of RFP(13) and nGFP(13), implying that although nGFP(13) elevated levels of the O-GlcNAcylated target protein JunB, the selectivity for the target protein could be further improved (Figure 2F).

Design and expression of proximity-directed nanobody-OGT(4) constructs

We next sought to evaluate whether truncation of the TPR domain would improve proximity-direction through the nanobody and increase O-GlcNAcylation selectivity for the target protein. Although truncation of the TPR domain has been reported to reduce or abolish glycosyltransferase activity for proteins, presumably through the loss of binding to the protein substrates,[21, 32] we expected to selectively recover the activity through the affinity of the nanobody for the target protein. The first 4 TPRs closest to the catalytic domain of OGT contain four asparagine and aspartate residues that in prior structural characterization were shown to be important for the binding mechanism and modification of substrates.[20] In addition, a unique advantage of using nanobodies to redirect enzymatic function is the ability to modularly fuse nanobodies that recognize target proteins within the cell. We thus evaluated a truncated OGT with 4 TPRs after fusion to nGFP and an additional nanobody, nEPEA (Figure 3A). The nEPEA nanobody was originally developed against α-synuclein (α-syn) and recognizes the four-amino acid EPEA tag as a fusion to the C-terminus of proteins.[31]

The EPEA tag was particularly appealing as the sequence itself cannot be glycosylated and is minimally perturbative to protein structure.[31] However, to remove potential competition for the target protein by endogenous α-synuclein in HEK293T cells, we generated an α-syn knocked-out (KO) HEK293T cell line for studies employing the nEPEA nanobody (Figure S3).

Expression of the OGT(4) fusion proteins in α-syn KO HEK293T cells showed a subcellular localization throughout the nucleocytoplasmic space by confocal fluorescence microscopy (Figure 3B). The activity of OGT(4), RFP(4), nGFP(4) and nEPEA(4) was evaluated against the same target protein GFP-Flag-JunB-EPEA (Figure 3C). Both nGFP(4) and nEPEA(4) significantly increased the O-GlcNAcylated target protein relative to the untargeted controls OGT(4) and RFP(4). To determine if the activity observed came from our nanobody-OGT(4) fusion, we generated two additional controls consisting of either the nanobody (nGFP, nEPEA) or a catalytically-inactive version of the nanobody-OGT(4) that is unable to bind to UDP-GlcNAc [nGFP(4,K852A), nEPEA(4,K852A)].[33] Increased levels of O-GlcNAcylated GFP-Flag-JunB-EPEA are installed directly from the active nGFP(4), but not the nanobody nGFP alone or a catalytically inactive mutant nGFP(4,K852A) (Figure 3D). Similarly, O-GlcNAcylated JunB-Flag-EPEA is selectively increased on co-expression of nEPEA(4), but not in the presence of nEPEA or the catalytically inactive nEPEA(4,K852A) (Figure 3E). The O-GlcNAcylation activity of the OGT(4) fusions was further compared to the full-length RFP(13) (Figure 3F). Truncation of the TPR domain of OGT consistently increased OGT expression levels relative to the full-length protein. However, despite higher expression levels of RFP(4), truncation of the TPR domain attenuated the increase in global O-GlcNAc levels observed with RFP(13). Installation of the nanobody as in nGFP(4) and nEPEA(4)

increased O-GlcNAc levels at the target protein, implying successful recruitment and selective O-GlcNAcylation of the target protein through the nanobody (Figure 3F).

The versatility of the nanobody-OGT(4) system to selectively increase the O-GlcNAcylated target protein was further evaluated against three targets: JunB-Flag-EPEA, cJun-Flag-EPEA, and Nup62-Flag-EPEA in HEK293T cells. In each case, the O-GlcNAcylated target protein significantly increased under proximity-direction of the matched nEPEA(4), but not the mismatched nGFP(4) (Figure 3G). Collectively, these data point to the successful increase in selective O-GlcNAcylation by replacing elements of the TPR domain with the nanobody and the modular ability of the nanobody-OGT(4) to increase O-GlcNAc levels using GFP- or EPEA-tagged target proteins.

Characterization of target protein glycosylation by chemical glycoproteomics

To better quantify the selectivity of the nanobody-OGT constructs for the target protein, we performed quantitative proteomics experiments by mass spectrometry (MS). Lysates from α-syn KO HEK293T cells expressing JunB-Flag-EPEA with or without nEPEA(4) were collected and chemoenzymatically labeled with UDP-GalNAz to introduce an azido-sugar for copper-catalyzed azide-alkyne cycloaddition (CuAAC). The O-GlcNAcylated proteins were enriched,[34] digested on-bead, and labeled with Tandem Mass Tags (TMT) for MS analysis. Glycoprotein enrichment was determined relative to the control (JunB-Flag-EPEA) for high-confidence proteins [number of unique peptides ≥2, 1% false discovery rate (FDR)] (Figure 4A, Table S1). These data show that JunB-Flag-EPEA was the only protein enriched >4-fold on co-expression of nEPEA(4), implying good target selectivity by the nEPEA nanobody (highlighted in red and blue, respectively, Figure 4A). A 2.6-fold abundance variation or less was observed with all other enriched glycoproteins.

We next sought to address whether the TPR-truncated OGT(4) exhibited glycosite selectivity that was broadly similar against protein targets in cells. As O-GlcNAc sites on JunB and Nup62 have yet to be thoroughly mapped,[18] we elected to compare glycosites produced by nEPEA(4) to full-length OGT(13). JunB-Flag-EPEA was expressed alone as the control or co-expressed with an undirected full-length OGT [RFP(13) or GFP(13)] or nEPEA(4) in α-syn KO HEK293T cells and was affinity purified, digested with chymotrypsin, and analyzed by MS in biological triplicate. Unambiguous glycosites were determined based on previously established criteria.[18] Four unambiguous glycosites were identified from JunB-Flag-EPEA. Three of the four unambiguous glycosites (S95, T153, S158) were identified in at least one nEPEA(4) sample and one reference sample [control, RFP(13)/GFP(13)] (Figure 4B, Table S2). The remaining glycosite S85 was identified only in samples co-expressing RFP(13). These results indicate that replacement of parts of the TPR domain with the nanobody results in a fusion construct that targets broadly similar glycosites as a full-length OGT(13).

We further evaluated the glycosite selectivity on the highly O-GlcNAcylated protein Nup62. Despite the elevated levels of O-GlcNAc on Nup62, glycosites on Nup62 are only beginning to be measured by MS,[18] and none have been experimentally validated to our knowledge. A total of 18 unambiguous glycosites were mapped to Nup62-Flag-EPEA (Figure 4C, Table S3). Of these 18 glycosites, 17 glycosites were found in at least one of the reference samples

and in one of the samples expressing nEPEA(4). The remaining glycosite (T270) was unambiguously observed only on co-expression of nEPEA(4). Seven glycosites (T75, S100, S159, S175, T187, T306, T311) on Nup62-Flag-EPEA were found only on co-expression of an OGT construct. Taken together, nEPEA(4) displayed similar glycosite selectivity towards Nup62-Flag-EPEA and JunB-Flag-EPEA while increasing overall levels of the O-GlcNAcylated target protein.

Targeting endogenous α-synuclein for O-GlcNAcylation with nEPEA(4)

A major advantage of using nanobodies for proximity-direction of OGT is the potential to recruit OGT to endogenous target proteins. Since the nEPEA nanobody was originally developed against the C-terminal EPEA sequence of α-synuclein, we evaluated whether nEPEA(4) could increase glycosylation of endogenous α-synuclein in a selective manner in comparison to over-expression of RFP(13), which is a primary strategy for elevating O-GlcNAc levels. Endogenous α-synuclein glycosylation was visualized using a mass-shift assay, which uses a chemical reporter for O-GlcNAc to install a PEG5K tag for determination of O-GlcNAc protein stoichiometry.[35] As expected, we observed an increase in O-GlcNAc levels on α-synuclein by mass shift assay on expression of RFP(13) and nEPEA(4), but not the catalytically inactive mutant nEPEA(4,K852A) (Figure 5B). Notably, expression of the untargeted RFP(13) resulted in a dramatic increase in O-GlcNAcylation across the global HEK293T cell proteome that was significantly reduced in the nEPEA(4) sample. Taken together, these data demonstrate the high selectivity and versatility of proximity-directed nanobody-OGT(4) constructs to transfer O-GlcNAc to the desired target protein, either as a tagged or endogenous target protein, with reduced impact on global O-GlcNAc levels.

Discussion

We report here the ability to increase O-GlcNAcylation on a target protein in live cells through fusion of a nanobody to OGT. The study of O-GlcNAc in cells typically relies on global alteration of the modification, or genetic mutagenesis strategies (e.g., elimination of the modification site, S-glycosylation mutagenesis). Alternatively, an approach to selectively induce the modification on proteins in cells would introduce a mechanism to engineer O-GlcNAc on target proteins, and provide a complementary approach to recent S-glycosylation strategies for introduction of stable glycosites.[11] Proximity-directed OGT constructs were designed by fusion of nGFP or nEPEA to full-length OGT(13) or TPR-truncated OGT(4) connected by a rigid linker (EAAAK)4. These constructs were successfully recruited to target proteins carrying either GFP or the four-amino acid sequence EPEA for transfer of O-GlcNAc. Furthermore, replacement of part of the TPR domain of OGT with a nanobody can improve selectivity for the target protein as compared to an untargeted control (Figure 2, 3). Evaluation of additional nanobodies and TPR truncation constructs may further increase selectivity for the target protein.

The nEPEA-OGT(4) fusion protein was further validated by quantitative proteomics and glycoproteomics to establish the selective O-GlcNAcylation of JunB and Nup62 as target proteins without broad perturbation of the O-GlcNAc proteome (Figure 4). Glycosite maps

from nEPEA(4) were largely analogous to those derived from untransfected control or RFP(13) samples. The replacement of protein-recognition domains by a nanobody, established here by truncating the TPR domain of OGT, has broader implications for the design of additional proximity-directed enzymes for targeted installation of a range of protein modifications in live cells and illuminates a potential strategy to evaluate the glycosite selectivity of the TPR domain of OGT in cells. In vitro structural work on the mechanism of OGT substrate binding implicated a series of asparagine and aspartate residues lining the first four TPRs as mediators of substrate and glycosite selectivity, which highlights the need for further study in cells and the implications on glycosite selectivity from the OGT isoforms mOGT and sOGT that have natively truncated TPR repeats.[36, 37]

Finally, we demonstrate the selective glycosylation of endogenous α-synuclein by nEPEA(4) (Figure 5). As increased glycosylation of α-synuclein has a protective effect on aggregation,[38] the ability to induce O-GlcNAc on α-synuclein may have further implications in the treatment of Parkinson's disease. The selective delivery of O-GlcNAc directly to endogenous target proteins will expand as nanobodies that recognize additional native nucleocytoplasmic proteins are developed. The strategic incorporation of nanobodies that recognize specific protein conformations, or post-translational modification states, may be further employed to distinguish between co-translational or post-translational glycosylation and O-GlcNAc dynamics.[13] The evaluation of additional nanobodies may further lead to nanobody-OGT constructs that are protein-selective with enhanced glycosite or sub-sequence selectivity.

The use of nanobodies as proximity-directing groups has major advantages, including the relatively rapid evaluation of O-GlcNAcylation of new target proteins and the ability to target endogenous proteins, as compared to other induced proximity systems (e.g., chemically induced dimerization).[23, 39, 40] However, current nanobody systems must be controlled at the genetic level and cannot be chemically controlled after expression, though chemical stabilization of protein levels can be introduced.[41] The use of nanobodies for induced proximity may reinforce artifactual interactions with the target protein that could disrupt normal protein interactions or cellular functions. To control for these potential effects, transfection with the nanobody or a catalytically inactive construct [e.g., nEPEA(4,K852A)] should be included in the experimental design. Optimization of binding parameters to tune reversibility may be possible using the several GFP-recognizing nanobodies that have been described,[30] which may additionally provide insight to the properties of the TPR domain of OGT in cells. Co-expression of the nanobody-OGT and the target protein is also liable to spurious results simply due to artifacts from protein over-expression.[42] Furthermore, as nanobody-OGT installs O-GlcNAc to the target protein in a catalytic manner, higher expression levels of the target protein is desirable. Future development of stable cell lines, inducible expression systems, or CRISPR knock-in strategies for introduction of peptide or protein tags may address these concerns. Finally, identification of a system to demonstrate the biological effects of elevated target protein O-GlcNAcylation is required to fully establish the effectiveness of proximity-directed OGT in measuring biological phenomena, relative to alternative strategies, including OGT over-expression, OGA inhibition,[10] or mutagenic S-glycosylation.[11] Since OGT mediates several enzymatic and non-enzymatic effects, substitution of protein domains that mediate substrate

selection (e.g., the TPR domain of OGT) with a nanobody may uniquely facilitate the systematic measurement of the isolated functions of these domains in cells, such as dissecting the scaffolding and catalytic functions of OGT.

In conclusion, we report a method for protein-selective O-GlcNAcylation using the first proximity-directed nanobody-OGT constructs. The approach uses engineered nanobody-OGT fusion proteins to induce O-GlcNAc to target proteins in cells. As a number of glycosyltransferases are composed of a lectin domain fused to a catalytic domain,[43] nanobody fusion proteins possess a great potential for proximity-induced glycosylation of a broader array of glycan structures. The strategy reported here could drive new insights to the substrate selection mechanisms and functions of the glycoproteome and design principles for extension to the control of additional PTMs in a live cell environment.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENT

We thank P. Schwein, Y. Ge, and B. Yang for helpful discussions; and B. Budnik, the director of the Harvard University Proteomics Facility, and D. Richardson, the director of Harvard Center of Biological Imaging (HCBI). Support from the National Institutes of Health (U01CA242098-01, C.M.W.), Burroughs Welcome Fund, Career Award at the Scientific Interface (C.M.W.), Milton Fund (C.M.W.), Mizutani Foundation for Glycoscience (C.M.W.), Sloan Foundation (C.M.W.), Merck Fellowship Fund, Corning Fund for Faculty Development, and Harvard University is gratefully acknowledged.

REFERENCES

[1]. Golks A, Tran TT, Goetschy JF, and Guerini D (2007) Requirement for O-linked N-acetylglucosaminyltransferase in lymphocytes activation, Embo J 26, 4368–4379. [PubMed: 17882263]

[2]. Lund PJ, Elias JE, and Davis MM (2016) Global Analysis of O-GlcNAc Glycoproteins in Activated Human T Cells, J Immunol 197, 3086–3098. [PubMed: 27655845]

[3]. Yi W, Clark PM, Mason DE, Keenan MC, Hill C, Goddard WA III, Peters EC, Driggers EM, and Hsieh-Wilson LC (2012) Phosphofructokinase 1 Glycosylation Regulates Cell Growth and Metabolism Science 337, 975–980. [PubMed: 22923583]

[4]. Yuzwa SA, Shan X, Macauley MS, Clark T, Skorobogatko Y, Vosseller K, and Vocadlo DJ (2012) Increasing O-GlcNAc slows neurodegeneration and stabilizes tau against aggregation, Nat Chem Biol 8, 393–399. [PubMed: 22366723]

[5]. Pravata VM, Muha V, Gundogdu M, Ferenbach AT, Kakade PS, Vandadi V, Wilmes AC, Borodkin VS, Joss S, Stavridis MP, and van Aalten DMF (2019) Catalytic deficiency of O-GlcNAc transferase leads to X-linked intellectual disability, Proc Natl Acad Sci U S A 116, 14961–14970. [PubMed: 31296563]

[6]. Lagerlöf O, Slocomb JE, Hong I, Aponte Y, Blackshaw S, Hart GW, and Huganir RL (2016) The nutrient sensor OGT in PVN neurons regulates feeding, Science 351, 1293–1296. [PubMed: 26989246]

[7]. Gambetta MC, and Muller J (2015) A critical perspective of the diverse roles of O-GlcNAc transferase in chromatin, Chromosoma 124, 429–442. [PubMed: 25894967]

[8]. Gloster TM, Zandberg WF, Heinonen JE, Shen DL, Deng L, and Vocadlo DJ (2011) Hijacking a biosynthetic pathway yields a glycosyltransferase inhibitor within cells, Nat Chem Biol 7, 174–181. [PubMed: 21258330]

[9]. Martin SES, Tan Z-W, Itkonen HM, Duveau DY, Paulo JA, Janetzko J, Boutz PL, Törk L, Moss FA, Thomas CJ, Gygi SP, Lazarus MB, and Walker S (2018) Structure-Based Evolution of Low Nanomolar O-GlcNAc Transferase Inhibitors, J Am Chem Soc 140, 13542–13545. [PubMed: 30285435]

[10]. Yuzwa SA, Macauley MS, Heinonen JE, Shan X, Dennis RJ, He Y, Whitworth GE, Stubbs KA, McEachern EJ, Davies GJ, and Vocadlo DJ (2008) A potent mechanism-inspired O-GlcNAcase inhibitor that blocks phosphorylation of tau in vivo, Nat Chem Biol 4, 483–490. [PubMed: 18587388]

[11]. Gorelik A, Bartual SG, Borodkin VS, Varghese J, Ferenbach AT, and van Aalten DMF (2019) Genetic recoding to dissect the roles of site-specific protein O-GlcNAcylation, Nat Struct Mol Biol 26, 1071–1077. [PubMed: 31695185]

[12]. Leney AC, El Atmioui D, Wu W, Ovaa H, and Heck AJR (2017) Elucidating crosstalk mechanisms between phosphorylation and O-GlcNAcylation, Proc Natl Acad Sci 114, E7255–E7261. [PubMed: 28808029]

[13]. Zhu Y, Liu TW, Cecioni S, Eskandari R, Zandberg WF, and Vocadlo DJ (2015) O-GlcNAc occurs cotranslationally to stabilize nascent polypeptide chains, Nat Chem Biol 11, 319–325. [PubMed: 25774941]

[14]. Trinidad JC, Barkan DT, Gulledge BF, Thalhammer A, Sali A, Schoepfer R, and Burlingame AL (2012) Global identification and characterization of both O-GlcNAcylation and phosphorylation at the murine synapse, Mol Cell Proteomics 11, 215–229. [PubMed: 22645316]

[15]. Hahne H, Sobotzki N, Nyberg T, Helm D, Borodkin VS, van Aalten DM, Agnew B, and Kuster B (2013) Proteome wide purification and identification of O-GlcNAc-modified proteins using click chemistry and mass spectrometry, J Proteome Res 12, 927–936. [PubMed: 23301498]

[16]. Wang X, Yuan Z-F, Fan J, Karch KR, Ball LE, Denu JM, and Garcia BA (2016) A novel quantitative mass spectrometry platform for determining site-specific protein O-GlcNAcylation dynamics, Mol Cell Proteomics 15, 2462–2475. [PubMed: 27114449]

[17]. Wang S, Yang F, Petyuk VA, Shukla AK, Monroe ME, Gritsenko MA, Rodland KD, Smith RD, Qian WJ, Gong CX, and Liu T (2017) Quantitative proteomics identifies altered O-GlcNAcylation of structural, synaptic and memory-associated proteins in Alzheimer's disease, J Pathol 243, 78–88. [PubMed: 28657654]

[18]. Woo CM, Lund PJ, Huang AC, Davis MM, Bertozzi CR, and Pitteri SJ (2018) Mapping and Quantification of Over 2000 O-linked Glycopeptides in Activated Human T Cells with Isotope-Targeted Glycoproteomics (Isotag), Mol Cell Proteomics 17, 764–775. [PubMed: 29351928]

[19]. Haltiwanger RS, Blomberg MA, and Hart GW (1992) Glycosylation of nuclear and cytoplasmic proteins. Purification and characterization of a uridine diphospho-N-acetylglucosamine:polypeptide beta-N-acetylglucosaminyltransferase, J Biol Chem 267, 9005–9013. [PubMed: 1533623]

[20]. Lazarus MB, Nam Y, Jiang J, Sliz P, and Walker S (2011) Structure of human O-GlcNAc transferase and its complex with a peptide substrate, Nature 469, 564–567. [PubMed: 21240259]

[21]. Iyer SPN, and Hart GW (2003) Roles of the Tetratricopeptide Repeat Domain in O-GlcNAc Transferase Targeting and Protein Substrate Specificity, J Biol Chem 278, 24608–24616. [PubMed: 12724313]

[22]. Pathak S, Alonso J, Schimpl M, Rafie K, Blair DE, Borodkin VS, Schuttelkopf AW, Albarbarawi O, and van Aalten DM (2015) The active site of O-GlcNAc transferase imposes constraints on substrate sequence, Nat Struct Mol Biol 22, 744–750. [PubMed: 26237509]

[23]. Stanton BZ, Chory EJ, and Crabtree GR (2018) Chemically induced proximity in biology and medicine, Science 359.

[24]. Caussinus E, Kanca O, and Affolter M (2012) Fluorescent fusion protein knockout mediated by anti-GFP nanobody, Nat Struct Mol Biol 19, 117–121.

[25]. Fulcher LJ, Hutchinson LD, Macartney TJ, Turnbull C, and Sapkota GP (2017) Targeting endogenous proteins for degradation through the affinity-directed protein missile system, Open Biol 7, 170066. [PubMed: 28490657]

[26]. Roubinet C, Tsankova A, Pham TT, Monnard A, Caussinus E, Affolter M, and Cabernard C (2017) Spatio-temporally separated cortical flows and spindle geometry establish physical asymmetry in fly neural stem cells, Nat Commun 8, 1383. [PubMed: 29123099]

[27]. Kirchhofer A, Helma J, Schmidthals K, Frauer C, Cui S, Karcher A, Pellis M, Muyldermans S, Casas-Delucchi CS, Cardoso MC, Leonhardt H, Hopfner KP, and Rothbauer U (2010) Modulation of protein properties in living cells using nanobodies, Nat Struct Mol Biol 17, 133–138. [PubMed: 20010839]

[28]. Dmitriev OY, Lutsenko S, and Muyldermans S (2016) Nanobodies as Probes for Protein Dynamics in Vitro and in Cells, J Biol Chem 291, 3767–3775. [PubMed: 26677230]

[29]. Anton T, and Bultmann S (2017) Site-specific recruitment of epigenetic factors with a modular CRISPR/Cas system, Nucleus 8, 279–286. [PubMed: 28448738]

[30]. Kubala MH, Kovtun O, Alexandrov K, and Collins BM (2010) Structural and thermodynamic analysis of the GFP:GFP-nanobody complex, Protein Sci 19, 2389–2401. [PubMed: 20945358]

[31]. De Genst EJ, Guilliams T, Wellens J, O'Day EM, Waudby CA, Meehan S, Dumoulin M, Hsu ST, Cremades N, Verschueren KH, Pardon E, Wyns L, Steyaert J, Christodoulou J, and Dobson CM (2010) Structure and properties of a complex of alpha-synuclein and a single-domain camelid antibody, J Mol Biol 402, 326–343. [PubMed: 20620148]

[32]. Lubas WA, and Hanover JA (2000) Functional Expression of O-linked GlcNAc Transferase: DOMAIN STRUCTURE AND SUBSTRATE SPECIFICITY, J Biol Chem 275, 10983–10988. [PubMed: 10753899]

[33]. Jiang J, Lazarus MB, Pasquina L, Sliz P, and Walker S (2011) A neutral diphosphate mimic crosslinks the active site of human O-GlcNAc transferase, Nat Chem Biol 8, 72–77. [PubMed: 22082911]

[34]. Thompson JW, Griffin ME, and Hsieh-Wilson LC (2018) Chapter Four - Methods for the Detection, Study, and Dynamic Profiling of O-GlcNAc Glycosylation, In Meth Enzymol (Imperiali B, Ed.), pp 101–135, Academic Press.

[35]. Rexach JE, Rogers CJ, Yu S-H, Tao J, Sun YE, and Hsieh-Wilson LC (2010) Quantification of O-glycosylation stoichiometry and dynamics using resolvable mass tags, Nat Chem Biol 6, 645–651. [PubMed: 20657584]

[36]. Levine ZG, Fan C, Melicher MS, Orman M, Benjamin T, and Walker S (2018) O-GlcNAc Transferase Recognizes Protein Substrates Using an Asparagine Ladder in the Tetratricopeptide Repeat (TPR) Superhelix, J Am Chem Soc 140, 3510–3513. [PubMed: 29485866]

[37]. Joiner CM, Levine ZG, Aonbangkhen C, Woo CM, and Walker S (2019) Aspartate Residues Far from the Active Site Drive O-GlcNAc Transferase Substrate Selection, J Am Chem Soc 141, 12974–12978. [PubMed: 31373491]

[38]. Marotta NP, Lin YH, Lewis YE, Ambroso MR, Zaro BW, Roth MT, Arnold DB, Langen R, and Pratt MR (2015) O-GlcNAc modification blocks the aggregation and toxicity of the protein α-synuclein associated with Parkinson's disease, Nat Chem 7, 913–920. [PubMed: 26492012]

[39]. Spencer DM, Wandless TJ, Schreiber SL, and Crabtree GR (1993) Controlling signal transduction with synthetic ligands, Science 262, 1019–1024. [PubMed: 7694365]

[40]. Fegan A, White B, Carlson JCT, and Wagner CR (2010) Chemically Controlled Protein Assembly: Techniques and Applications, Chem Rev 110, 3315–3336. [PubMed: 20353181]

[41]. Banaszynski LA, Chen L.-c., Maynard-Smith LA, Ooi AGL, and Wandless TJ (2006) A Rapid, Reversible, and Tunable Method to Regulate Protein Function in Living Cells Using Synthetic Small Molecules, Cell 126, 995–1004. [PubMed: 16959577]

[42]. Moriya H (2015) Quantitative nature of overexpression experiments, Mol Biol Cell 26, 3932–3939. [PubMed: 26543202]

[43]. Lira-Navarrete E, de las Rivas M, Compañón I, Pallarés MC, Kong Y, Iglesias-Fernández J, Bernardes GJL, Peregrina JM, Rovira C, Bernadó P, Bruscolini P, Clausen H, Lostao A, Corzana F, and Hurtado-Guerrero R (2015) Dynamic interplay between catalytic and lectin domains of GalNAc-transferases modulates protein O-glycosylation, Nat Commun 6, 6937. [PubMed: 25939779]

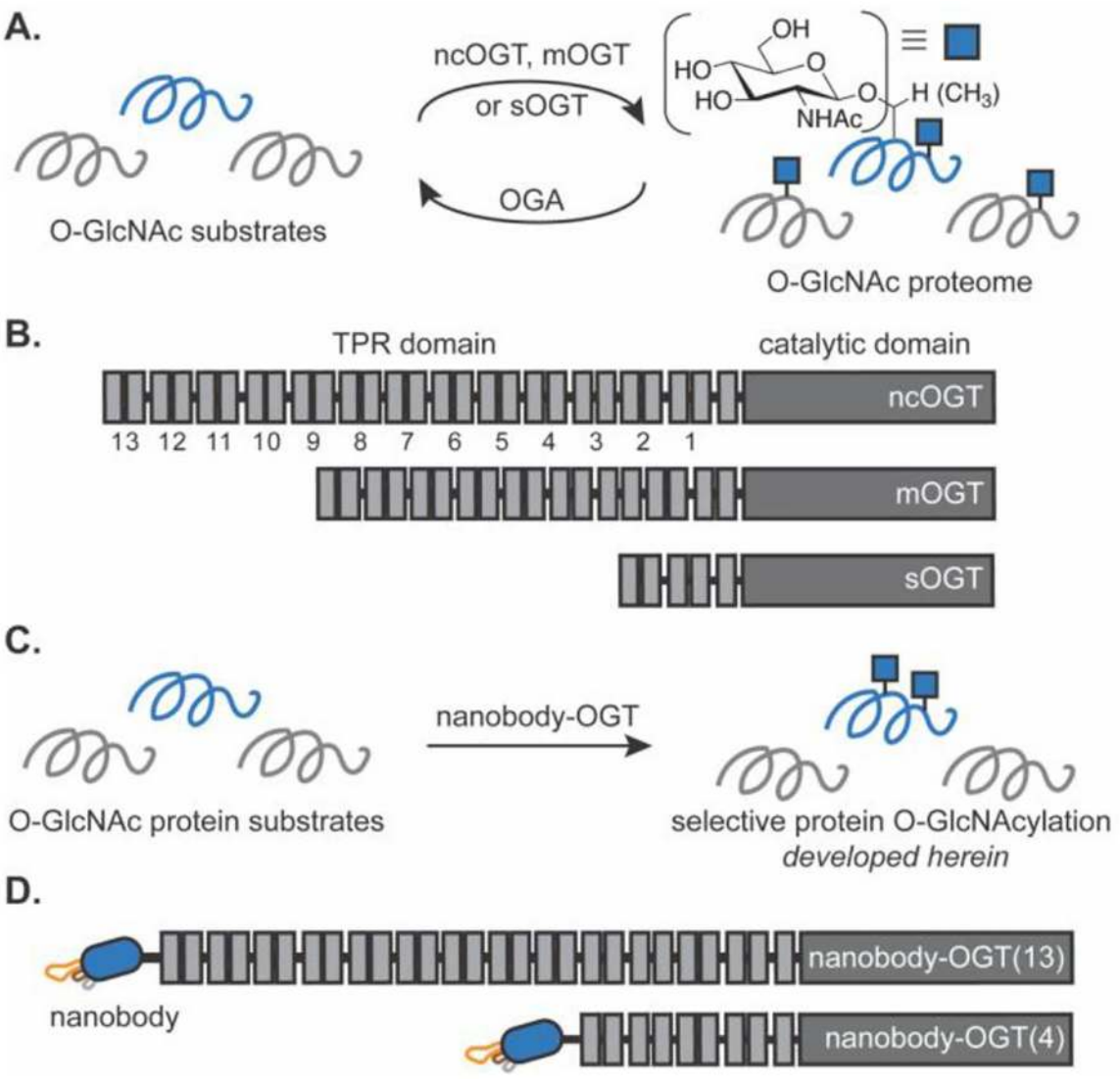


Figure 1.

Overview of proximity-directed OGT strategy. **A**. Schematic of dynamic O-GlcNAc modification of protein substrates. **B**. Linear representation of natural OGT isoforms ncOGT, mOGT, and sOGT. **C**. Strategy for selective induction of O-GlcNAc using a proximity-directed nanobody-OGT to transfer O-GlcNAc to the target protein. **D**. Linear representation of nanobody-OGT(13) and nanobody-OGT(4) fusion proteins.

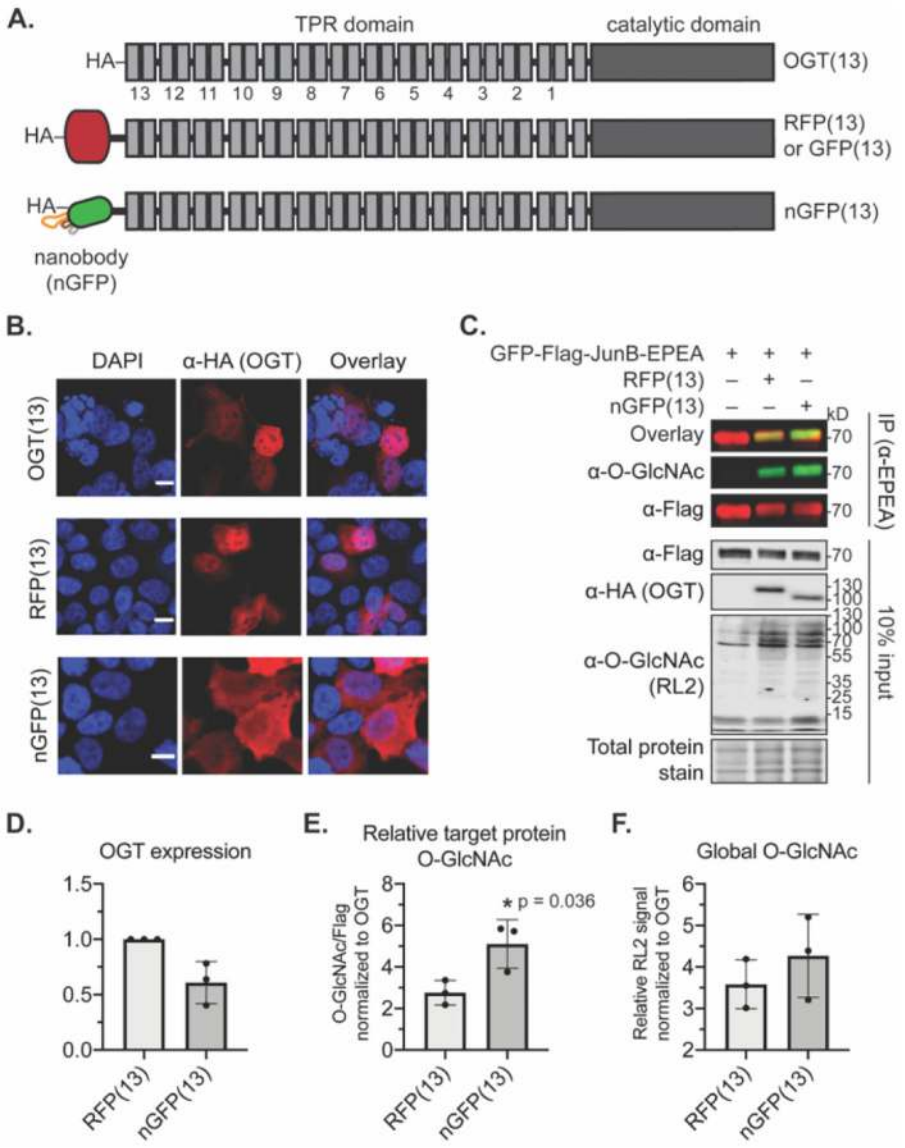


Figure 2.

Characterization of nanobody-OGT(13) as proximity-directing glycosyltransferase constructs. **A**. Linear representation of full-length OGT(13), RFP(13), and nGFP(13). **B**. Subcellular localization of OGT(13), RFP(13), and nGFP(13) constructs expressed in HEK293T cells by confocal fluorescent microscopy. Scale bars represent 10 μm. **C**. Western blot of O-GlcNAc levels on GFP-Flag-JunB-EPEA after immunoprecipitation with EPEA-beads from HEK293T cells. The expression of the various constructs was verified by Western blot analysis (10% input). **D**. Quantification of OGT expression. **E**. Quantification of O-GlcNAc levels on GFP-Flag-JunB-EPEA after normalization to OGT expression. **F**. Quantification of O-GlcNAc levels in whole cell lysate after normalization to OGT expression. Data are representative of three biological replicates per experiment. Error bars represent standard deviation, * represents a p-value <0.05 under a two-tailed t-test. Full blot and confocal images can be found in Figure S1–S2.

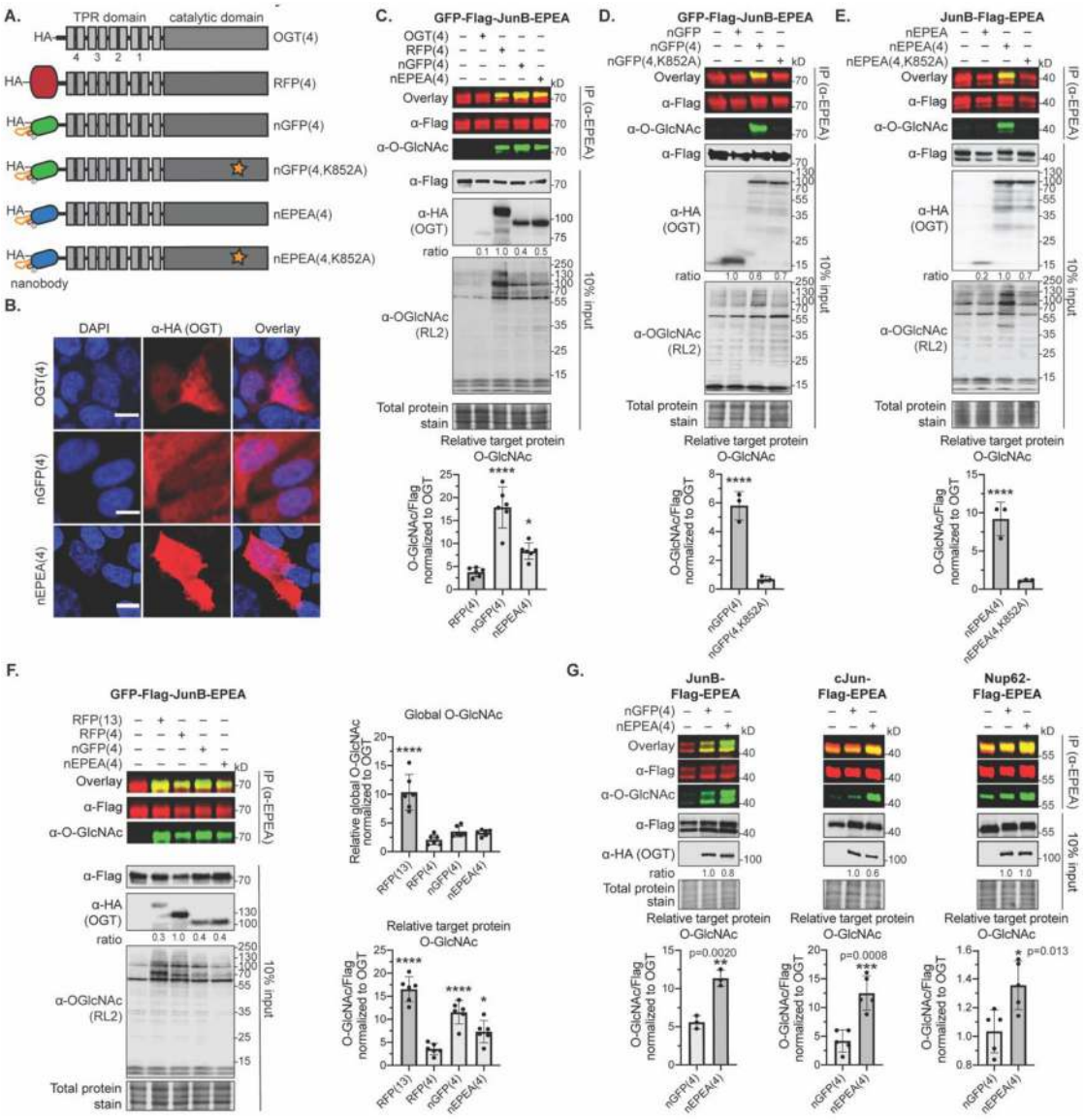
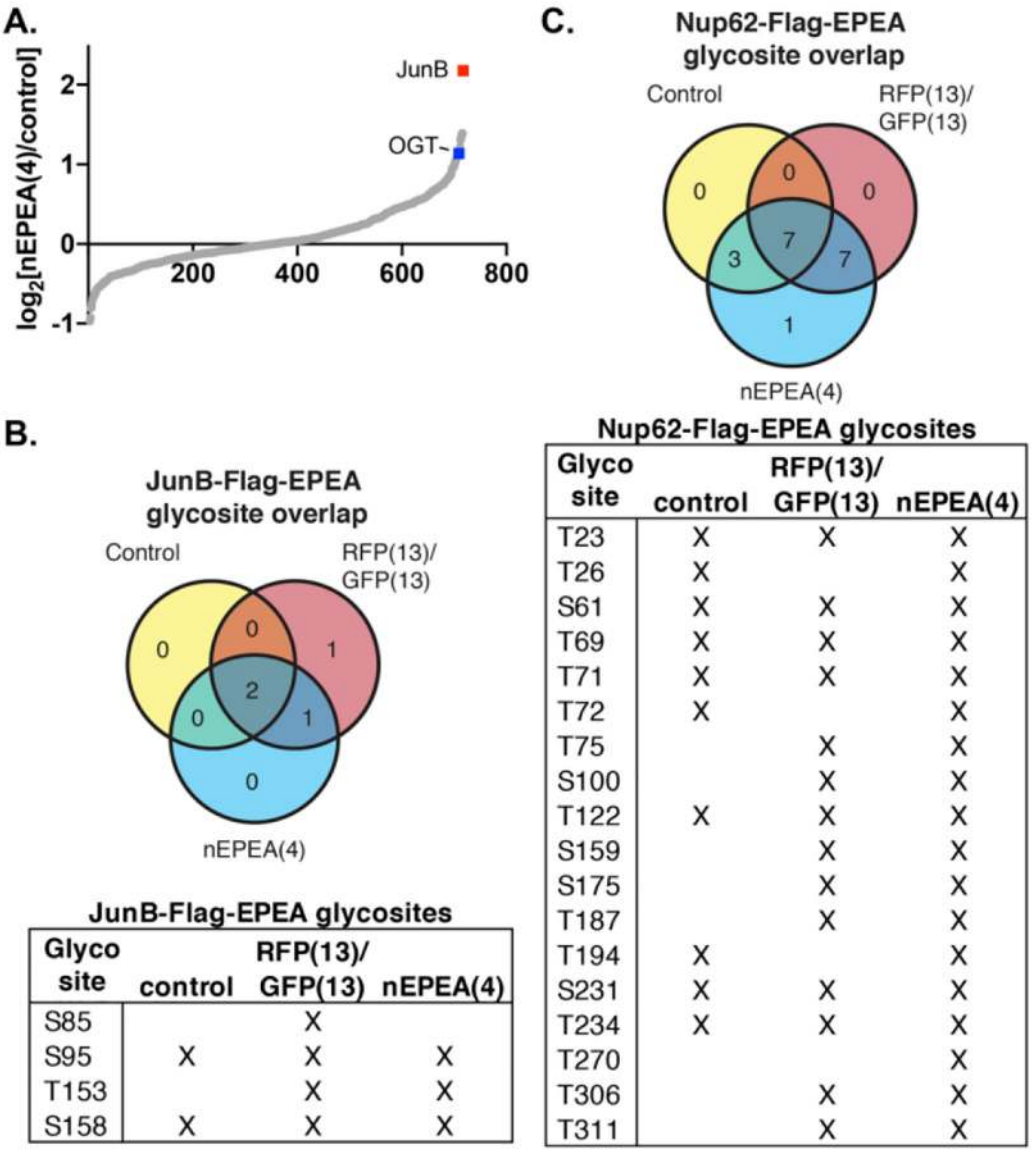


Figure 3.

Characterization of nanobody-OGT(4) as proximity-directing glycosyltransferase constructs. **A**. Linear representation of TPR truncated OGT(4), RFP(4), nGFP(4), nEPEA(4), and catalytically inactive mutants. **B**. Subcellular localization of OGT(4), nGFP(4), and nEPEA(4) in HEK293T cells by confocal fluorescent microscopy. Scale bars represent 10 μm. **C**. Western blot and quantification of O-GlcNAc levels on GFP-Flag-JunB-EPEA after immunoprecipitation with EPEA-beads. The expression of the various constructs was verified by Western blot analysis (10% input). **D**. Western blot and quantification of O-GlcNAc levels on GFP-Flag-JunB-EPEA after immunoprecipitation with EPEA-beads. The expression of the various constructs was verified by Western blot analysis (10% input). **E**. Western blot and quantification of O-GlcNAc levels on JunB-Flag-EPEA after immunoprecipitation with EPEA-beads. The expression of the various constructs was verified by Western blot analysis (10% input). **F**. Western blot and quantification of O-

GlcNAc levels on GFP-Flag-JunB-EPEA after immunoprecipitation with EPEA-beads. The expression of the various constructs was verified by Western blot analysis (10% input). **G**. Western blot and quantification of O-GlcNAc levels on JunB-Flag-EPEA, cJun-Flag-EPEA, and Nup62-Flag-EPEA after immunoprecipitation with EPEA-beads from α-syn knocked-out (KO) HEK293T cells co-transfected with the indicated nanobody-OGT fusion protein and target protein. The expression of the various constructs was verified by Western blot analysis (10% input). At least three biological replicates were performed per experiment. Error bars represent standard deviation, * represents $p \leq 0.05$, ** represents $p \leq 0.01$, *** represents $p \leq 0.001$, **** represents $p \leq 0.0001$ under a two-tailed t-test or one-way ANOVA. Full blot and confocal images can be found in Figures S4–S9.



JunB-Flag-EPEA glycosites

Glyco site	control	RFP(13)/ GFP(13)	nEPEA(4)
S85		X	
S95	X	X	X
T153		X	X
S158	X	X	X

Nup62-Flag-EPEA glycosites

Glyco site	control	RFP(13)/ GFP(13)	nEPEA(4)
T23	X	X	X
T26	X		X
S61	X	X	X
T69	X	X	X
T71	X	X	X
T72	X		X
T75		X	X
S100		X	X
T122	X	X	X
S159		X	X
S175		X	X
T187		X	X
T194	X		X
S231	X	X	X
T234	X	X	X
T270			X
T306		X	X
T311		X	X

Figure 4.

Quantitative proteomics and glycoproteomics of the global O-GlcNAc proteome from α-syn KO HEK293T cells. **A**. Quantitative proteomics of enriched O-GlcNAcylated proteins from α-syn KO HEK293T cells after co-expression of nEPEA(4) (highlighted in blue) and JunB-Flag-EPEA (highlighted in red) compared to expression of JunB-Flag-EPEA alone (control). **B**. Venn diagram and list of unambiguous glycosite assignments for JunB-Flag-EPEA. **C**. Venn diagram and list of unambiguous glycosite assignments for Nup62-Flag-EPEA. The target protein was co-expressed with the indicated OGT fusion protein in α-syn KO HEK293T cells, immunoprecipitated, and analyzed by MS. Only mono-glycosylated peptides with unambiguous assignments and a PSM count > 2 across biological replicates are shown. Three biological replicates were performed per experiment.

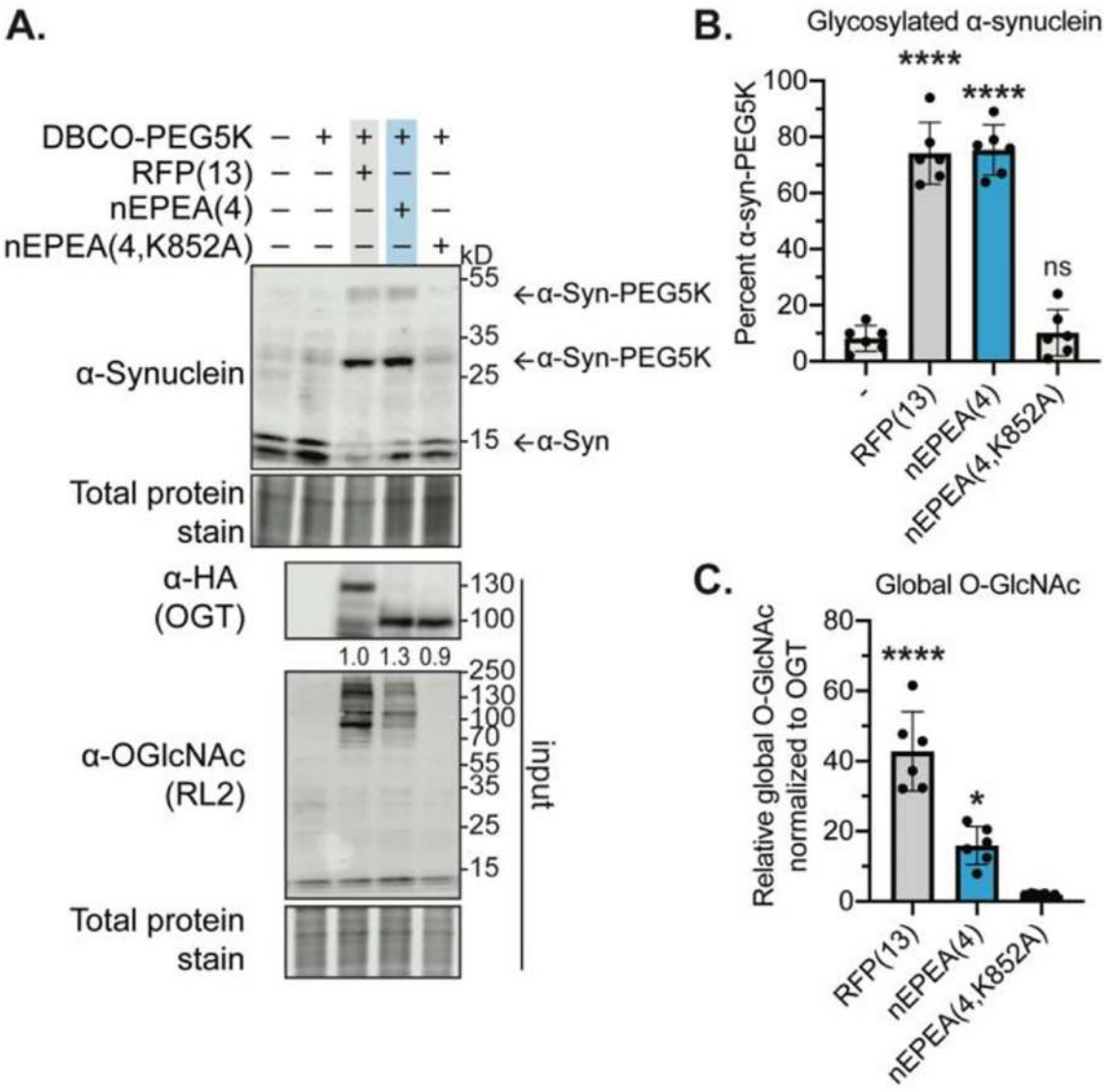


Figure 5.

Western blot and quantification of O-GlcNAc induced to α-synuclein by a mass shift assay. The indicated nanobody-OGT construct was expressed in HEK293T cells, the cells were lysed, chemoenzymatically labeled, and analyzed by mass shift assay. Global O-GlcNAc levels and the expression of the nanobody-OGT constructs was verified by Western blot analysis (10% input). Six biological replicates were performed per experiment. Error bars represent standard deviation, ns represents $p \geq 0.05$, * represents $p \leq 0.05$, ** represents $p \leq 0.01$, **** represents $p \leq 0.0001$ under a two-tailed t-test or one-way ANOVA. Full blots can be found in Figure S10.