Sequential Glycosylation of Proteins with Substrate-Specific *N*-Glycosyltransferases

Liang Lin[#,1,2], Weston Kightlinger[#,2,3], Sunaina Kiran Prabhu[4], Adam J. Hockenberry[5], Chao Li[4], Lai-Xi Wang[5], Michael C. Jewett[*,2,3], and Milan Mrksich[*,1,2,3,6]

#These authors contributed equally to the work.

[1]Department of Biomedical Engineering, [2]Center for Synthetic Biology, [3]Department of Chemical and Biological Engineering, [5]Interdisciplinary Biological Sciences Program, [6]Department of Chemistry

Northwestern University, 2145 Sheridan Road, Evanston, IL 60208

[4]Department of Chemistry and Biochemistry

University of Maryland, College Park, MD 20742

*Corresponding authors:

Milan Mrksich, 2145 Sheridan Road, Tech-B492, Evanston IL 60208-3109, milan.mrksich@northwestern.edu; Tel (+1) 847 467 0472; Fax: (+1) 847 491 4928
Michael C. Jewett, 2145 Sheridan Road, Tech E-136, Evanston, IL 60208-3120; m-jewett@northwestern.edu; Tel: (+1) 847 467 5007; Fax: (+1) 847 491 3728

Supplementary Information Table of Contents

Supplementary Information:

Methods

Phylogenetic analysis of putative NGTs. We first performed a preliminary phylogenetic analysis of the entire family 41 of the CAZY database[1] (access date: November 2016) by extracting 600 C-terminal residues from each protein, aligning the resulting sequences using MAFFT (v7.380)[2] and using FastTree (v2.1.10)[3] to rapidly construct a phylogenetic tree describing the relatedness of all these sequences (1,409 in total). After manual visual inspection of this large tree, we selected 41 sequences for further characterization according to criteria mentioned in the main text. We further refined the relationship between these 41 selected sequences by using MAFFT (L-INS-i method) to align the entire annotated coding sequence from each of these proteins. Next, we used RAxML (v8.1.3)[4] to perform 100 rapid bootstraps and a thorough maximum-likelihood search for the best phylogenetic tree using the LG substitution matrix, gamma distributed rate variation, and the human OGT as the outgroup. Visualization of all trees was performed using the iTol webserver[5].

Peptide synthesis. All peptide libraries were *N*-acetyl and *C*-Amide protected, and were synthesized as previously described[6]. For the unprotected peptides used in the ionization factor of trypsinized GlycTags in Im7, Arg-Wang resin (AAPPTEC, USA) was used and the synthesis steps was the same as above[6] except that there was no need for the first Arginine addition or acetic anhydride treatment for *N*-acetyl protection. These peptides were purified with HPLC and used for ionization factor analysis by LC-qTOF.

SAMDI. The preparation of SAMDI plates and peptide arrays for SAMDI profiling were completed as previously described[6]. Briefly, 50 μM peptide was reacted with the indicated concentration of NGT and 2.5 mM UDP-Glc in 100 mM HEPES (pH 8.0) and 500 mM NaCl at 30 °C for indicated time. Next, 2 μL TCEP-resin (Pierce) was added to each 10 μL reaction solution and the solution was incubated at 37 °C for 1 h. 2 μL of this solution was then added onto a 384-well SAMDI plate with 0.5 h incubation for peptide immobilization. After washing the plate, MS spectra were acquired using a MALDI-TOF (AB SCIEX 5800 TOF/TOF). All peptide specificity heatmaps presented in this study except **Fig. 2a** indicate the % glucose modification, which was calculated by adjusting the ratio of substrate and glucose modified peptide products in MS spectra by the average ionization factor (shown in **Supplementary Table 4**) and also described previously[6]. **Fig. 2a**, shows the unadjusted percent intensity of substrate and glycopeptide products, Product/(Substrate+Product). All percentage glucose modifications of less than 0.02 were regarded as background noise.

SAMDI screens for NGT sugar donor specificities were performed similarly to peptide specificity screens. Briefly, 1 mM of sugar donor UDP-Glc, UDP-Gal, UDP-Xyl, UDP-GlcN, GDP-Man, UDP-GlcNAc, or UDP-GalNAc was reacted with indicated NGTs and one of six representative peptides. The completed reactions were analyzed by SAMDI-MS. Unadjusted percent intensities of substrate and glycopeptide products are presented as heatmaps in **Supplementary Fig. 11**. All percentage intensities of ≥ 0.01 are above background noise and indicative of real modification.

Orthogonality analysis. Heatmaps showing potential orthogonality of NGTs from the analysis of HiNGT, EcNGT, ApNGT, and ApNGT^{Q469A} with X_{-1}-N-X_{+1}-S/T peptide library are shown in **Supplementary Figs. 5 and 6**. Calculations are described below. In order to more easily determine differences in specificity across NGTs with different specific activities, reaction conditions were controlled to obtain a maximum modification of ~85% across the whole substrate library for all NGTs (**Supplementary Fig. 4**). Differences in %modification from **Supplementary Fig. 4** were then used to determine potentially orthogonal peptides. For example, the %modification (HiNGT - EcNGT) was calculated and a value of greater than 50% was considered a good substrate for HiNGT over EcNGT, and if the ratio of the (HiNGT – EcNGT) value to %modification (HiNGT) was over 90% then the peptide was considered to be potentially orthogonal for HiNGT over EcNGT. HiNGT was then considered to be potentially orthogonal over EcNGT. Similarly, to determine sequences modified by HiNGT with orthogonality over EcNGT and ApNGT, the %modification of (HiNGT - EcNGT - ApNGT) was calculated, and so forth.

To identify and determine conditional orthogonality for specific NGT-substrate pairs, peptide libraries were reacted with in an order of HiNGT, EcNGT, ApNGT, and then ApNGT^{Q469A}, under optimized conditions such that each NGT would have at least one peptide having 90% more modification over preceding NGTs. First, the whole peptide library was arranged in descending order by the value of %modification (ApNGT^{Q469A} - ApNGT - EcNGT - HiNGT); then peptides with a value less than 5% for %modification (ApNGT^{Q469A} - ApNGT - EcNGT - HiNGT) were arranged in descending order by the %modification (ApNGT- EcNGT - HiNGT); then the peptides with a value less than 5% for %modification (ApNGT- EcNGT - HiNGT) were arranged in descending order by the %modification (EcNGT - HiNGT); and finally, the peptides with a value less than 5% for %modification (EcNGT - HiNGT) were arranged by the %modification (HiNGT). In this way, the peptides were divided into ApNGT^{Q469A} preferred, ApNGT preferred, EcNGT preferred, and HiNGT preferred. NGT-GlycTag pairs were then chosen from the left part of each NGT preferred region.

Plasmid construction and molecular cloning. Sources and details of plasmids used in this paper are shown in **Supplementary Tables 1 and 2**. Selected coding sequences with plasmid context are shown in **Supplementary Note 1**. DNA templates to produce

NGT homologs in CFPS were generated from codon-optimized gene fragments synthesized by Twist Bioscience, IDT, or Life Technologies which were either synthesized in or assembled into a pJL1 *in vitro* expression vector backbone[6] between NdeI and SalI restriction sites by Gibson Assembly[7]. Sequences of active NGTs YpNGT, HdNGT, MhNGT, HiNGT, EcNGT, and ApNGT were based on GenBank IDs CAL21840.1, AAP96624.1, AJE07135.1, ADO96126.1, CBL93373.1, and ABN74719.1, respectively. The ApNGT^{Q469A} mutant was generated by inverse PCR of pJL1.ApNGT followed by a one-piece Gibson assembly. In order to lessen variability in ribosome binding site strengths and therefore total CFPS expression, some putative NGTs were modified with *N*-terminal CAT-Strep-Linker fusion sequences previously shown to increase *in vitro* expression[6]. Expression vectors for purification of conditionally orthogonal NGTs were generated by PCR of pJL1 plasmids carrying each homolog followed by Gibson assembly of that PCR fragment into the pET.BCS.NS *in vivo* expression backbone[6] with *C*-terminal Strep II or Twin-Strep II tags. The expression vector for purification of 4gIm7 was synthesized by Twist Bioscience in the pET.BCS.NS vector with a *C*-terminal 6xHis-tag and four customized glycosylation sites according to SAMDI-MS characterization of the four conditionally orthogonal NGTs (full sequence shown in **Supplementary Note 1**). The expression vector for purification of SUMO-4gIm7 was constructed by PCR and assembly of the 4gIm7 gene fragment into a pET.28a vector with an *N*-terminal 6xHis-tag, SUMO sequence and Gly-Ser Linker to facilitate sequential glycosylation onto Ni-NTA magnetic beads and cleavage of the final product by Ulp1 (full sequence shown in **Supplementary Note 1**).

Cell-free protein synthesis of putative NGTs. Cell-free protein synthesis of putative NGTs was performed using the PANOx-SP crude lysate system[8] as previously described[6]. Briefly, crude extracts for CFPS were prepared from a recently described high-yielding *E. coli* strains C321.ΔA.759[9] or BL21 Star (DE3)[10] using well-established cell growth, harvest, and lysis conditions[6, 9-10]. C321.ΔA.759 *E. coli* cells were grown in 1 L of 2xYTPG media (yeast extract 10 g/L, tryptone 16 g/L, NaCl 5 g/L, K2HPO4 7 g/L, KH2PO4 3 g/L, and glucose 18 g/L, pH 7.2) in a 2.5 L Tunair flask at 34 °C and 250 r.p.m. to OD_{600} = 3.0 then pelleted by centrifugation at 5,000xg at 4 °C for 15 min. The pellets were then washed three times with cold S30 buffer (10 mM Tris–acetate pH 8.2, 14 mM magnesium acetate, 60 mM potassium acetate, and 2 mM dithiothreitol (DTT)) before being frozen on liquid nitrogen and stored at -80 °C. Cells were then thawed, resuspended in 0.8 mL of S30 buffer per gram wet weight, and lysed in 1.4 mL aliquots using a Q125 Sonicator (Qsonica) for three 45 s pulses at 50% amplitude with 59 s off between pulses. After sonication, lysed cells were supplemented with 2.8 mM DTT and centrifuged at 12,000xg. This supernatant was incubated at 37 °C at 250 r.p.m. for 1 h and centrifuged at 10,000xg at 4 °C for 10 min. The clarified lysate was flash frozen and stored at -80 °C until use. The clarified S12 lysate was used to conduct CFPS reactions as previously

described[6]. Lysates from BL21 Star (DE3) cells were prepared similarly except that T7 production in BL21 Star (DE3) cells was induced with 1 mM IPTG at OD_{600} = 0.6-0.8, frozen cells were resuspended in 1.0 mL per gram of cell weight, sonication was completed on 1.0 mL aliquots using 10 s on/off pulses at 50% amplitude until 640 J were delivered, and the lysate was not incubated at 37 °C and subsequently spun after sonication. A standard CFPS reaction contained 1.2 mM ATP; 0.85 mM each of GTP, UTP, and CTP; 34 μg/mL folinic acid; 170 μg/mL of *E. coli* tRNA mixture; 16 μg/mL purified T7 RNA polymerase; 2 mM for each of the 20 standard amino acids; 0.33 mM nicotinamide adenine dinucleotide (NAD); 0.27 mM coenzyme-A (CoA); 1.5 mM spermidine; 1 mM putrescine; 4 mM sodium oxalate; 130 mM potassium glutamate; 10 mM ammonium glutamate; 12 mM magnesium glutamate; 57 mM HEPES, pH 7.2; 33 mM phosphoenolpyruvate (PEP); 13.3 μg/mL NGT plasmid template; and 27% v/v of cell extract. *E. coli* total tRNA mixture (from strain MRE600) and phosphoenolpyruvate were purchased from Roche Applied Science. ATP, GTP, CTP, UTP, the 20 amino acids, and other materials were purchased from Sigma-Aldrich. No additional T7 was added to BL21 (DE3) Star CFPS reactions. Plasmid DNA for cell-free protein synthesis was purified from DH5-α *E. coli* strain (NEB) using ZymoPURE Midi Kit (Zymo Research). All NGTs were synthesized in 50 μL batch reactions in 2.0 mL centrifuge tubes. Putative NGT activity screening was completed using CFPS from BL21 Star (DE3) lysate and radioactive quantification of CFPS yields, autoradiograms, and preliminary $X_{-1}NX_{+1}$TRC peptide specificity screening was completed using CFPS from C321.ΔA.759 lysate. All CFPS reactions were carried out at 20 °C for 20 h.

Quantification of CFPS yields. CFPS yields of conditionally orthogonal NGTs were quantified by supplementing standard CFPS reactions with 10 μM ^{14}C leucine (PerkinElmer). Protein yields were quantified in triplicate using established protocols[6] for precipitation in 5% trichloroacetic acid (TCA) precipitation followed by radioactivity quantification using a Microbeta2 liquid scintillation counter (PerkinElmer). Soluble fractions were collected after centrifugation at 12,000xg for 15 min at 4 °C.

Autoradiogram of CFPS products. Autoradiograms of CFPS synthesized NGTs were prepared as previously described[6]. CFPS reactions supplemented with ^{14}C leucine were loaded onto a 4-12% Bolt Bis–Tris Plus SDS–PAGE gels (Invitrogen) and run in MOPS buffer at 150 V for 70 min. The gels were stained using InstantBlue (Expedeon), destained in water, dried overnight, and exposed for 48 h on a Storage Phosphor Screen (GE Healthcare). Phosphor Screens were acquired using Typhoon FLA7000 imager (GE Healthcare). Coomassie stain images of gels were acquired using a GelDoc XR+ Imager for molecular weight standard references.

Measurement of relative ionization factors for glycopeptides on LC-MS. Peptides identical to those observed following trypsinization of 4gIm7 and SUMO-4gIm7 were synthesized by SPPS, purified by HPLC, and quantified by dry weight as described above. Peptides were resuspended in water and supplemented with up to 1% formic acid and 20% ACN as necessary for soluble concentrations of 1-5 mM. Peptides were then diluted in 50 mM phosphate buffer (pH 8.0) to a working concentration of 25 μM. Glycosylated peptide samples at 12.5 μM concentration each were treated with 2 μM ApNGT^{Q469A}, 1 μM EcNGT, and 10 mM UDP-Glc for 12 h at 30 °C and then quenched at 80 °C for 20 min. After incubation and quenching, these peptides were found to be completely glycosylated and no aglycosylated peptides could be detected. This glycosylated sample was then mixed at equimolar concentration with aglycosylated samples which had been prepped similarly without NGTs and injected into LC-qTOF using peptide analysis method described below. The relative ionization factor was calculated as previously described[6].

Protein expression and purification of glycosyltransferases and targets. Purification of conditionally orthogonal NGTs and NmLgtB from living *E. coli* was performed using Strep-tacin XT Superflow resin (iba life sciences) based on manufacturer instructions. BL21 Star (DE3) cells (Thermo Fisher Scientific) were transformed with pET.BCS.NS vectors (Carb^R) bearing coding sequences for each NGT with twin or single Strep II tags or NmLgtB with a twin-strep tag. Cultures in LB media and appropriate antibiotic were inoculated from an overnight culture at an initial OD_{600} = 0.08 and grown at 37 °C at 250 r.p.m. to 0.6–0.8 OD and induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) overnight at 25 °C. The cells were pelleted by centrifugation at 7,000xg for 10 min at 4 °C, resuspended in cold Buffer 1 (50 mM NaH_2PO_4 and 300 mM NaCl at pH 8.0), and pelleted again by centrifugation at 7,000xg for 10 min, and flash frozen at -80 °C. The pellets were thawed and resuspended in 5 mL Buffer 1 per gram wet pellet weight and supplemented with 1 mg/mL lysozyme (Sigma), 25 U/mL of Benzonase (EMD Millipore), and Halt protease inhibitor before lysis by two homogenization passes (Avestin) and centrifugation at 15,000xg for 20 mins. The supernatant was added to a packed gravity flow column of 5 mL of Strep-tacin XT resin per 1 L of culture which had been equilibrated by 2 column volumes (CV) of Buffer 1. The column as then washed with 5 column volumes of Buffer 1 and then eluted in 3 CV of BXT Buffer (Buffer 1 with 50 mM biotin and 1 mM EDTA). Fractions containing at least 1 mg/mL of protein were collected and dialyzed against 50 mM HEPES, pH 7.0, 200 mM NaCl and 5% glycerol and Halt Protease overnight in a 20kDa Slide-a-lyzer cassette. After dialysis, samples were centrifuged at 12,000xg for 15 mins. NGT concentration was quantified using a Bradford assay with a BSA standard curve (Bio-rad).

Purification of engineered Im7Generalist1-4, 4gIm7, SUMO-4gIm7, SUMO-2gIm7, SUMO-2gFc, and PdST6mut proteins were performed using Qiagen Ni-NTA agarose

based on manufacturer's instructions. BL21 Star (DE3) cells (Thermo Fisher Scientific) were transformed with pETBCS.NS.Im7Generalist1-4 ($Carb^R$), pETBCS.NS.4gIm7 ($Carb^R$), pET28a.SUMO-4gIm7 (Kan^R), pETBCS.NS.SUMO-2gIm7 ($Carb^R$), pETBCS.NS.SUMO-2gFc ($Carb^R$), pJL1.PdST6mut (Kan^R) vectors. Cultures were grown, induced, harvested, and lysed identically to NGTs. Following centrifugation, lysate was applied to an Ni-NTA agarose column (Qiagen) which was equilibrated with Buffer 1 with 10 mM imidazole, washed with 10 column volumes of Buffer 3 with 20 mM imidazole, and eluted with 4 column volumes of Buffer 1 with 300 mM Imidazole. Fractions containing more than 1 mg/mL of protein by Bradford assay were pooled and exhaustively dialyzed against Buffer 1 with 5% glycerol with three or more buffer exchanges. After dialysis, samples were centrifuged at 12,000xg for 15 mins and flash frozen on liquid nitrogen and stored at -80 °C until needed. Protein concentrations in the dialyzed elution fractions were quantified by SDS-PAGE densitometry using Image Lab software version 6.0.0 using a BSA standard curve following staining with InstantBlue Coomassie stain and destaining with water.

Elaboration of glucosylated proteins by chemoenzymatic transglycosylation or enzymatic addition of monosaccarides. As shown in **Supplementary Figs. 18-20, 22, 23** and **25,** glucoses installed by NGTs onto target proteins were chemoenzymatically elaborated to human-like biantennary glycans using endoglycosidases and oxazoline glycan donors using previously reported workflows[11]. Briefly, the expression and purification of $EndoCC^{N180H}$ [12] and EndoA WT[13] was performed following reported protocols. Biantennary sialylated glycan, $Sia_2Gal_2Man_3GlcNAc$ oxazoline[14] (SCT-ox), and azido-functionalized trimannose core glycan, $Man_3GlcNAc$ oxazoline[15] (AzMan3-ox), were synthesized using published procedures. Mono or diglucosylated forms of Im7 were combined with EndoA-WT enzyme and AzMan3-ox before quenching with 0.1 M glycine buffer at pH=2.7. Monoglucosylated or previously elaborated forms of SUMO-2gFc were combined with $EndoCC^{N180H}$ and SCT-ox. Concentrations of target proteins, enzymes, and sugar donors as well as reaction durations and conditions are described in supplementary figure legends.

As shown in **Supplementary Fig. 24**, SUMO-2gFc was previously treated with EcNGT + UDP-Glc, EndoA-WT + $AzMan_3$-ox, and $ApNGT^{Q469A}$ + UDP-Glc. After the glucose installation by $ApNGT^{Q469A}$, SUMO-2gFc was then elaborated with 15 μM purified NmLgtBTrunc and 5 μM purified PdST6mut as well as 2.5 mM UDP-Gal and 2.5 mM CMP-Sia in Buffer 1 supplemented with 10 mM $MnCl_2$ and 23 mM HEPES buffer (pH=7.5). This reaction was incubated for 16 h at 30 °C.

***In vitro* glycosylation and purification of target proteins from single glycosylation reactions**. IVG reactions were performed using purified 4gIm7 (**Fig. 4**), Im7GeneralistP1-4, and SUMO-4gIm7 (**Supplementary Fig. 15**) targets and purified, conditionally

orthogonal NGTs. 10 μM purified 4gIm7 or 2 μM of purified SUMO-4gIm7 were incubated with 2.5 mM UDP-Glc and indicated concentrations of each NGT in Buffer 1 for 4 h at 30 °C then purified using Ni-NTA functionalized magnetic beads (Thermo Scientific) based on manufacturer's instructions. After incubation, the 35 μL reactions were diluted with 90 μL cold Buffer 1 and mixed with 20 μL magnetic beads which had been washed with Buffer 1. The beads were incubated at room temperature on a roller for 10 min to capture the 4gIm7 protein. The beads were then washed 2 times with 120 μL of cold Buffer 1. The beads were then eluted in 70 μL of Buffer 1 with 500 mM Imidazole by incubation on a room temperature roller for 10 mins. The eluent was then dialyzed against 1:4 diluted Buffer 1 overnight. 35 μL of the dialyzed fraction was incubated with 0.0044 μg/μL of MS Grade Trypsin (Thermo Scientific) at 37 °C overnight and analyzed using the LC-qTOF peptide method described below.

***In vitro* glycosylation and purification of target proteins for sequential glycosylation reactions**. Sequential IVG reactions and purifications of SUMO-4gIm7 (**Fig. 5**) were completed similarly to single IVG reactions but used Ni-NTA functionalized magnetic beads as a solid support for movement among sequential IVG reactions. 2 μM purified SUMO-4gIm7 was incubated with 2.5 mM UDP-Glc in Buffer 1 with 0.3 μM HiNGT for 4 h at 30 °C on a roller. After the HiNGT reaction, 10 mM imidazole was added to the sample and the sample was incubated with 42 μL of beads per 100 μL of reaction volume. After the reaction, the sample was washed two times with at least one reaction volume of Buffer 1 with 10 mM imidazole. After washing, the beads with immobilized SUMO-4gIm7 were resuspended in Buffer 1 with 10 mM imidazole, 2.5 mM UDP-Glc, and 0.3 μM EcNGT and incubated for 4 h at 30 °C. After the reaction, the beads were washed two times with at least one reaction volume of Buffer 1 with 10 mM imidazole. The reaction and washing procedures were then repeated with 0.4 μM ApNGT and 2 μM ApNGT^{Q469A}. After each enzymatic reaction, the progression of the reaction was monitored by eluting the SUMO-4gIm7 from at least 40 μL of beads following the wash steps using two times the volume of beads of Buffer 1 with 500 mM imidazole. This material was then dialyzed against 1:4 diluted Buffer 1. Diglucosylated products from 2gIm7 and 2gFc studies (**Supplementary Figs. 17 and 21)** were similarly purified using Ni-NTA beads before analysis.

Monoglucosylated or diglucosylated substrates (2gIm7 or 2gFc) required for analyses and chemoenzymatic elaboration studies (**Supplementary Figs. 17-25**) were generated by incubation with NGTs using reaction conditions specified in figure legends and then purified using Qiagen Ni-NTA agarose resin in order to terminate glucosylation reactions. Briefly, a gravity flow column of Ni-NTA was conditioned with Buffer 1 with 10 mM imidazole. Reaction products were bound to the column by incubation for 30 min, washed with 10 CV Buffer 1 with 20 mM imidazole, and then eluted using Buffer 1 with 500 mM

imidazole and dialyzed against 1:4 diluted Buffer 1. Protein concentrations were quantified using a Bradford assay with a BSA standard curve (Bio-rad).

Purified materials described above and material from subsequent transglycosylation studies (**Supplementary Figs. 18-20** and **22-25**) were prepared for intact protein analysis by neutralizing to pH=7.5 if necessary using NaOH and then supplementing with 0.0047 μg/μL Ulp1 endopeptidase (purified as previously described[16]) and incubating in 1 mM DTT for 1 h at 30 °C or overnight at 4 °C to cleave SUMO from target proteins. Glycosylation site-occupancy was quantified by incubation with trypsin and peptide analysis by LC-qTOF (described below). Intact protein mass analysis was completed by injecting eluent which was not digested by trypsin into LC-qTOF using the intact protein method described below.

LC-qTOF analysis of target protein glycosylation. Analysis of glycosylation site-occupancy was completed by injection of trypsinized 4gIm7, Im7GeneralistP1-4, or SUMO-4gIm7 onto a Bruker Elute UPLC system equipped with an ACQUITY UPLC Peptide BEH C18 Column, 300Å, 1.7 μm, 2.1 mm X 100 mm (186003686 Waters Corporation) with a 10 mm guard column of identical packing (186004629 Waters Corporation) coupled to an Impact-II UHR TOF Mass Spectrometer (Bruker Daltonics, Inc.). The chromatographic separation was completed using Solvent A of 100% H_2O with 0.1% Formic Acid and Solvent B of 100% acetonitrile (ACN) with 0.1% Formic Acid at a flow rate of 0.5 mL/min and a column temperature of 40 °C. Solvent conditions were held at 0% B for 1 min; then, the peptides of interest were eluted during a 4 min gradient from 0% to 50% B. The column was then washed and re-equilibrated using a 0.1 min gradient from 50–100% B, a 2 min hold at 100% B, a 0.1 min gradient from 100-0% B, and a 1.8 min hold at 0% B for a total run time of 9 min. LC-MS/MS of glycopeptides was performed to confirm the products of transglycosylation and enzymatic monosaccharide addition reactions (**Supplementary Figs. 18-20** and **22-25**). Theoretical glycopeptide masses were targeted for pseudo multiple reaction monitoring (MRM) MS/MS fragmentation according to detected intact protein MS peaks. A collisional energy of 30 eV and a window of ± 2 m/z from targeted m/z values was used for glycopeptide fragmentation. The scan range for MS and MS/MS of peptides was 100-3000 m/z. External calibration was performed for all datapoints.

Unless otherwise noted, analyses of intact protein masses were completed by injection of Ulp1 digested protein onto a Bruker Elute UPLC system equipped with an ACQUITY UPLC Protein BEH C4 Column, 300Å, 1.7 μm, 2.1 mm X 50 mm (186004495 Waters Corporation) with a 10 mm guard column of identical packing (186004623, Waters Corporation) coupled to an Impact-II UHR TOF-MS Mass Spectrometer (Bruker Daltonics, Inc.). The chromatographic separation was completed using Solvent A of 100% H_2O with 0.1% Formic Acid and Solvent B of 100% acetonitrile (ACN) with 0.1% Formic Acid at a flow rate of 0.5 mL/min and a column temperature of 50 °C. Solvent conditions

were held at 0% B for 1 min; then, the protein of interest was eluted during a 4 min gradient from 20% to 50% B. The column was then washed and re-equilibrated using a 0.5 min wash at 71.4% B, a 0.1 min gradient from 71.4-100% B, a 2 min wash at 100% B, a 0.1 min gradient from 100-20% B, and a 2.2 min hold at 20% B for a total run time of 10 min. The MS scan range was 100-3000 m/z with a spectral rate of 2 Hz. External calibration was performed for all datapoints.

Intact protein LC-MS analysis of the final products of **Supplementary Figs. 18-20** was performed by diluting sample in 0.1% formic acid before injection into a QExactivePlus OrbiTrap mass spectrometer in conjunction with Ultimate 3000 HPLC system (ThermoFisher) equipped with an XbridgeTM BEH300 C4 column (300 Å, 3.5 μm, 2.1 x 50 mm, Waters). Chromatography was performed at a column temperature of 23 °C and a flow rate of 0.4 ml/min using 100% water containing 0.1% Formic Acid as Solvent A and 100% acetonitrile containing 0.1% Formic Acid as Solvent B. The column was equilibrated at 5% B for 2 min, followed by protein injection and separation over a linear gradient of 5 - 95% B for 6 min. The column was washed with 95% B for 2 min and equilibrated with 5% B for 2 min. The MS scan range was 400 to 3000 m/z at a scan speed of 0.9 Hz and 140,000 resolution. The m/z profiles were deconvoluted using MagTran (Amgen).

Analysis of LC-qTOF generated target protein glycosylation data. Data from Impact-II UHR TOF Mass Spectrometer was processed using Bruker Compass Data Analysis software version 4.1. Glycosylation site occupancies, defined as $Glc_1/(Glc_0+Glc_1)$ were quantified by integrating the areas of extracted ion chromatograms (EICs) of theoretical values for monoisotopic protonated +2 charge states of glycosylated and aglycosylated peptides with adjustment for ionization efficiency by RIFs as previously described[6]. Quantification of glycoprotein modification for **Supplementary Fig. 12** was completed by integrating extracted ion chromatograms of the three highest intensity charge states of the glycosylated and aglycosylated intact protein. Deconvolution of intact protein data was performed using maximum entropy deconvolution using MS peaks within m/z range of 700–2,000 into a mass range of 14,000–18,000 Da. Raw data was then plotted and annotated using R Studio. Deconvolutions used full mass spectra averaged across the entire peak width of the proteins of interest (encompassing the full elution of the glycosylated and aglycosylated glycoforms). Theoretical intact protein masses are shown in **Supplementary Table 3**.

Quantification for correct sequential glycosylation of target protein. The purity of correct sequential glycosylation was calculated based on site-occupancy data of each GlycTag after each NGT treatment (shown in **Fig. 5c**) with the assumption that

glycosylation events at the four sites within 4gIm7 are not interdependent (supported by **Supplementary Fig. 14**). The following equation was used to calculate the overall yield:

$$\%modification\ of\ GlycTag1\ (NGT1) *$$

$$\left(\%modification\ of\ GlycTag2\ (NGT2) - \%modification\ of\ GlycTag2\ (NGT1)\right) *$$

$$\left(\%modification\ of\ GlycTag3\ (NGT3) - \%modification\ of\ GlycTag3\ (NGT2)\right) *$$

$$(\%modification\ of\ GlycTag4\ (NGT4) - \%modification\ of\ GlycTag4\ (NGT3))$$

Where %modification of GlycTagx (NGTy) is the observed site occupancy of GlycTagx after treatment with NGTy; x, y refers to 1, 2, 3, 4, which represent HiNGT, EcNGT, ApNGT and ApNGT^{Q469A} in this study.

Data availability. All data generated or analyzed during this study are included in this published article (and its supplementary files) or are available from the corresponding authors on reasonable request.

Safety Statement. No unexpected or unusually high safety hazards were encountered.

Supplementary Tables

Supplementary Table 1: Strains and plasmids used in this study. For additional descriptions of plasmids used for NGT screening see **Supplementary Table 2** and **Supplementary Note 1**. Sources [1], [2], [3], and [4] are Martin et al. *Nature Communications* 2018[9]**,** Kightlinger* and Lin* et al. *Nature Chemical Biology* 2018[6], Eshima et al. *PloS One*, 2015[17], and Takegawa et al. *Arch Biochem Biophys* 1997[18], respectively**.**

Plasmid and Strain	Relevant Characteristics	Source
Strains		
BL21 Star (DE3)	*F-ompT hsdSB (rB-, mB-) galdcmrne131 (DE3)*	Thermo Fisher Scientific
C321.ΔA.759	MG1655 C321 Derivative	[1]
Plasmids		
pJL1.HiNGT	*H. influenzae,* HiNGT, (ADO96126.1), C-term strep-tag	This study
pJL1.EcNGT	*E. coli,* EcNGT, (CBL93373.1), C-term strep-tag	This study
pJL1.ApNGT	*A. pleuroneumoniae,* ApNGT, (ABN74719.1)	[2]
pJL1.ApNGT(Q469A)	pJL1.ApNGT(Q469A), Q469A	This study
pJL1.Im7-0s	*E. coli* Im7, (IMM7_ECOLX), C-term 6xHis-tag	[2]
pETBCS.NS	pETBCS modified with Nde1 and Sal1 restriction sites, P_{T7}, KAN^R, IPTG Induction	[2]
pETBCS.NS.GeneralistP1	Im7-0s, M1_E2insRATTIYANVTLAGGR	This study
pETBCS.NS.GeneralistP2	Im7-0s, N26_D32delinsRATTWYANVTRAGG	This study
pETBCS.NS.GeneralistP3	Im7-0s, S58_S64delins RATTYMGNISRAGG	This study
pETBCS.NS.GeneralistP4	Im7-0s, G87_G88insRATTYMGNISRAGG	This study
pETBCS.NS.4gIm7	Im7-0s, M1_E2insRATTLNENVTRAGG, N26_D32delinsRATTWYANVTRAGG, S58_S64delins RATTYMGNISRAGG, G87_G88insRATTYMGNISRAGG	This study
pETBCS.NS.SUMO-2gIm7	N-term 6xHis-tag, SUMO, GS linker, Im7-0s, M1_E2insRATTWDYNLTRAGG, G87_G88insRATTYMGNISRAGG	This study
pETBCS.NS.SUMO-2gFc	N-term 6xHis-tag, SUMO, GS linker, H. sapiens Fc, (IGHG1_HUMAN), Q178_Y183delinsRATTWDYNLTRAGG, C-term RATTYMGNISRAGG	This study
pET28a	P_{T7}, KAN^R, IPTG Induction	Novagen
pET28a.SUMO-4gIm7	pET28a, N-term 6xHis-tag, SUMO, GS linker, 4gIm7	This study
pETBCS.NS.HiNGT	*H. influenzae,* HiNGT, (ADO96126.1), C-term Twin strep-tag	This study
pETBCS.NS.EcNGT	*E. coli,* EcNGT, (CBL93373.1), C-term Twin strep-tag	This study
pETBCS.NS.ApNGT	*A. pleuroneumoniae,* ApNGT, (ABN74719.1), C-term Twin Strep-tag	This study
pETBCS.NS.ApNGT(Q469A)	pETBCS.NS.ApNGT, Q469A, C-term Single Strep-tag	This study
pETBCS.NS.LgtBTrunc	N-term Twin Strep-tag, *N. meningititus* NmLgtB, LGTB_NEIMB E258_Q275delinsGDD	This study
pETBCS.NS.PdST6Mut	N-term His tag + thrombin cleavage site, *Photobacterium damselae*, O66375_9GAMM, A200Y, S232Y, ΔM1_A15 and Δ498_D675	This study
pET41b.EndoCC1(N180H)	P_{T7}, KAN^R, IPTG Induction, N-term 6xHis tag, *Coprinopsis cinerea*, Endo-β-N-acetylglucosaminidase 1, Uniprot A8P7P2_COPC7	[3]
pGEX-2T/Endo-A(WT)	P_{tac}, AMP^R, IPTG Induction N-term GST tag, *Arthrobacter protophormiae*, Endo-β-N-acetylglucosaminidase, Uniprot Q9ZB22_9MICC, Δ1-24	[4]

Supplementary Table 2: **Putative NGT sequences screened by GlycoSCORES.** Full species names and NCBI identifiers for all enzymes screened for *N*-glucosyltransferase activity in **Fig. 2**. Each sequence was identified from family 41 of the CAZY database, synthesized in the pJL1 plasmid backbone as shown in **Supplementary Note 1**, expressed in CFPS, and screened for activity using SAMDI-MS.

Abbreviation	Organism	Accession Number
Ab	*Azospirillum brasilense*	AIB12151.1
Ap	*Actinobacillus pleuropneumoniae serovar 5b str. L20*	ABN74719.1
Av	*Anabaena variabilis PCC 7937*	ABA19974.1
Bc1	*Burkholderia cepacia*	AIO47218.1
Bc2	*Burkholderia cepacia*	AKE02193.1
Bd	*Bradyrhizobium diazoefficiens*	BAR53565.1
Bm	*Burkholderia mallei*	AIO50869.1
Bs	*Burkholderia sp. CCGE1001*	ADX55793.1
Cc1	*Caulobacter crescentus NA1000*	ACL93585.1
Cc2	*Caulobacter crescentus NA1000*	ACL94629.1
Cc3	*Caulobacter crescentus NA1000*	ACL97299.1
Ec	*Escherichia coli*	CBL93373.1
Gs1	*Geobacter sulfurreducens*	AJY68899.1
Gs2	*Geminocystis sp. NIES-3709*	BAQ65624.1
Hd	*Haemophilus ducreyi*	AAP96624.1
Hh	*Herbaspirillum hiltneri N3*	AKZ61275.1
Hi	*Haemophilus influenzae*	ADO96126.1
Ls	*Leptolyngbya sp. PCC 7376*	AFY39462.1
Ma1	*Microcystis aeruginosa NIES-843*	BAG01392.1
Ma2	*Microcystis aeruginosa NIES-843*	BAG05492.1
Ma3	*Microcystis aeruginosa NIES-843*	BAG05864.1
Me	*Methylorubrum extorquens CM4*	ACK83875.1
Mh	*Mannheimia haemolytica*	AJE07135.1
Np	*Nostoc piscinale CENA21*	ALF56607.1
Pc	*Paraburkholderia caribensis MBA4*	ALL64447.1
Pf2	*Paraburkholderia fungorum*	AJZ57761.1
Pf3	*Paraburkholderia fungorum*	AJZ59790.1
Pm	*Prochlorococcus marinus str. NATL2A*	AAZ58722.1
Pp	*Pseudomonas putida*	ANC05098.1
Pphe	*Paraburkholderia phenoliruptrix BR3459a*	AFT87114.1
Pphy	*Paraburkholderia phytofirmans PsJN*	ACD18780.1
Ps	*Piscirickettsia salmonis LF-89*	AKP73554.1
Pz	*Phenylobacterium zucineum HLK1*	ACG76885.1
Rl	*Rhizobium leguminosarum bv. trifolii WSM1689*	AHF82454.1
Rp	*Ralstonia pickettii DTP0602*	AGW94945.1
Rso	*Ralstonia solanacearum*	ALF89676.1
Rsp2	*Rhodobacter sphaeroides 2.4.1*	ABA81711.1
Se	*Salmonella enterica subsp. enterica serovar Weltevreden*	CUR95075.1
Ss	*Synechocystis sp. PCC 6803*	BAM51499.1
Tt	*Thermobaculum terrenum ATCC BAA-798*	ACZ43705.1
Yp	*Yersinia pestis*	CAL21840.1

Supplementary Table 3: Theoretical masses of intact proteins. Theoretical intact protein masses of Ulp1-cleaved proteins were calculated using EXPASY Protein Parameters online tool. Observed mass values are shown in deconvoluted LC-qTOF spectra figures.

Protein	Glycan Modification	Theoretical (Da)
4glm7	None	15631
4glm7	Glc_1	15793
4glm7	Glc_2	15955
4glm7	Glc_3	16117
4glm7	Glc_4	16279
2glm7	None	14160
2glm7	Glc_1	14322
2glm7	Glc_2	14484
2glm7	AzMan(AzMan)ManGlcNAcGlc	15062
2glm7	AzMan(AzMan)ManGlcNAcGlc+Glc	15224
2glm7	$(AzMan(AzMan)ManGlcNAcGlc)_2$	15963
2glm7	SiaGalGlcNAcMan(SiaGalGlcNAcMan)ManGlcNAcGlc	16324
2glm7	$(SiaGalGlcNAcMan(SiaGalGlcNAcMan)ManGlcNAcGlc)_2$	18488
2gFc	None	28611
2gFc	Glc_1	28773
2gFc	Glc_2	28935
2gFc	AzMan(AzMan)ManGlcNAcGlc	29513
2gFc	AzMan(AzMan)ManGlcNAcGlc+Glc	29675
2gFc	SiaGalGlcNAcMan(SiaGalGlcNAcMan)ManGlcNAcGlc	30775
2gFc	SiaGalGlcNAcMan(SiaGalGlcNAcMan)ManGlcNAcGlc+Glc	30937
2gFc	$(SiaGalGlcNAcMan(SiaGalGlcNAcMan)ManGlcNAcGlc)_2$	32939

Supplementary Table 4: Relative ionization factors (RIFs) used in this study for SAMDI-MS. For each peptide library, approximately 16-20 representative peptides were used to calculate the Average RIF. For individual peptides, (n=3) triplicate experiments were performed for the RIF. Mean and S.D. of RIF measurements are shown.

Peptide / peptide library	Relative Ionization Factor
$X_{-1}NX_{+1}$(T/S)RC	0.540 ± 0.076
$X_{-2}(X_{-1}NX_{+1}$T/S)RC	0.570 ± 0.140
$X_{-3}(X_{-2}X_{-1}NX_{+1}$T/S)RC	0.555 ± 0.080
WYANVTRC	0.946 ± 0.036
YMGNISRC	0.789 ± 0.017
LNENVTRC	0.739 ± 0.022
WDYNLTRC	0.699 ± 0.027

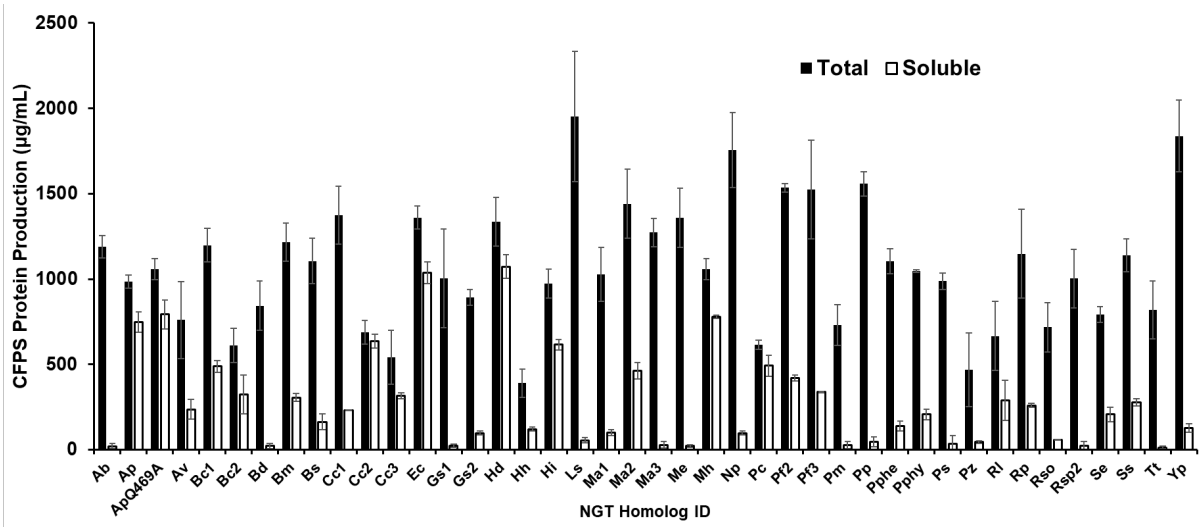
Supplementary Table 5: Relative Ionization Factors for peptides measured by LC-qTOF. RIFs were determined for glycopeptides identical to those observed following trypsinization of 4gIm7 and SUMO-4gIm7. Peptides were synthesized by SPPS, purified by HPLC and quantified by dry weight. Glycosylated peptide samples at 12.5 μM concentration were treated with 2 μM ApNGT(Q469A), 1 μM EcNGT, and 10 mM UDP-Glc overnight and then quenched at 80 °C for 20 min. After incubation and quenching, these peptides were found to be completely glycosylated and no aglycosylated peptides could be detected. This glycosylated sample was then mixed at equimolar concentration with aglycosylated samples prepared similarly without NGTs. This mixture was injected into LC-qTOF using the peptide analysis method described in the **Methods**. Mean and S.D. of RIFs were calculated from observed EIC areas from n=3 independent peptide IVGs, mixings, and LC-qTOF injections.

Peptide	Targeting NGT	RIF
ATTWYANVTR	HiNGT	0.937 ± 0.017
ATTYMGNISR	EcNGT	0.923 ± 0.010
ATTLNENVTR	ApNGT	0.871 ± 0.002
ATTWDYNLTR	ApNGT(Q469A)	0.870 ± 0.014

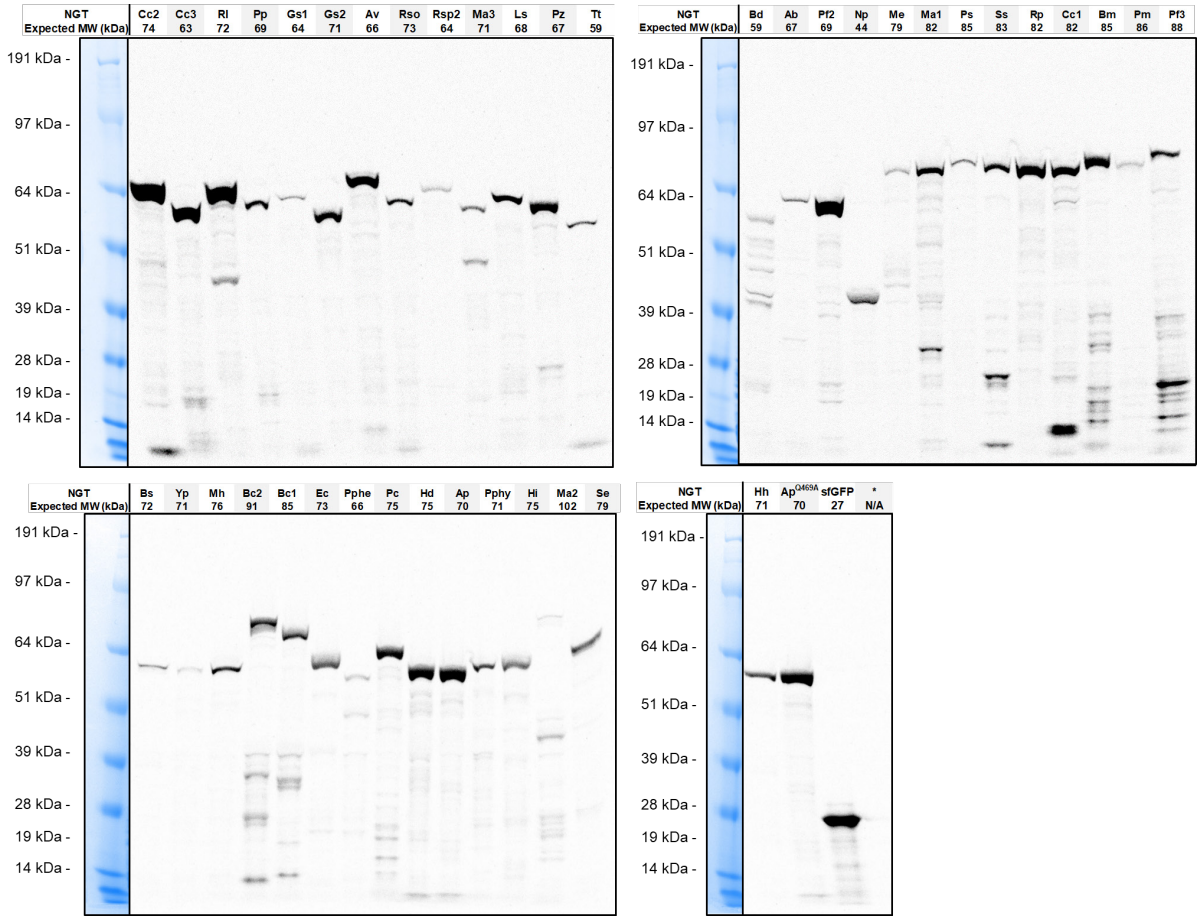
Supplementary Table 6: **Calculation of SUMO-4glm7 yields with site-specific control at various numbers of glycosylation sites.** Calculations based on site-occupancy data of SUMO-4glm7 after treatment with each NGT treatment shown in **Fig. 5** and assume that glycosylation events are not interdependent (supported by **Supplementary Fig. 13**). Yield calculations with fewer than 4 glycosylation sites assume the presence of only the indicated number of GlycTags and the use of their corresponding NGTs for sequential glycosylation. Yield of condition with no selective GlycTags assumes sequential elaboration steps with equal amounts of glycosylation at all sites at each step using perfect control of reaction time. Fold enrichment compares given scenarios with this no selectivity condition.

Sequential NGT modification with their specific GlycTags	Number of glycosylation sites	Correct glycoprotein yield	Fold enrichment
No selective NGTs and GlycTags	4	0.39%	---
HiNGT, EcNGT, ApNGT, ApNGT^{Q469A}	4	62% ± 1%	$1.6x10^2$
No selective NGTs and GlycTags	3	3.7%	---
HiNGT, EcNGT, ApNGT	3	68% ± 1%	18
HiNGT, EcNGT, ApNGT^{Q469A}	3	88% ± 2%	24
HiNGT, ApNGT, ApNGT^{Q469A}	3	75% ± 1%	20
EcNGT, ApNGT, ApNGT^{Q469A}	3	80% ± 1%	22
No selective NGTs and GlycTags	2	25%	---
HiNGT, EcNGT	2	88% ± 1%	3.5
HiNGT, ApNGT	2	82% ± 1%	3.3
HiNGT, ApNGT^{Q469A}	2	95% ± 1%	3.8
EcNGT, ApNGT	2	87% ± 1%	3.5
EcNGT, ApNGT^{Q469A}	2	100% ± 2%	4.0
ApNGT, ApNGT^{Q469A}	2	90% ± 1%	3.6

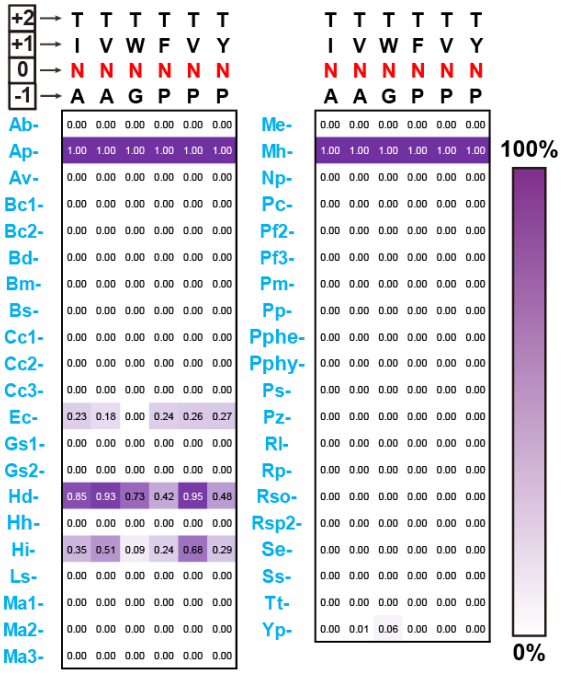
Supplementary Figures



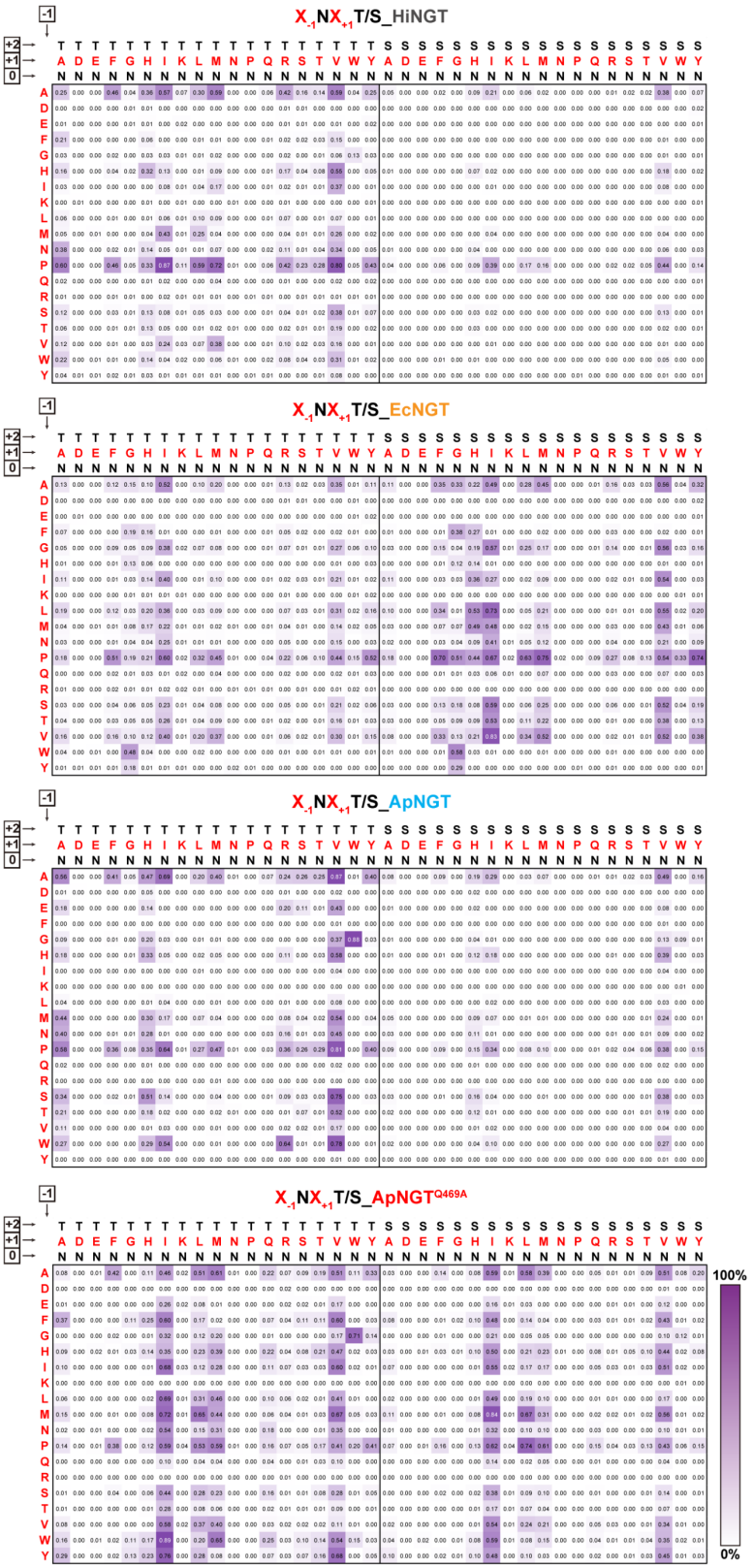
Supplementary Figure 1: CFPS expression yields of putative NGTs. Total and soluble CFPS yields of 41 NGT homologs and the ApNGT^{Q469A} variant were quantified by ^{14}C-leucine incorporation. All CFPS reactions were incubated for 20 h at 20 °C (see **Methods** for details). Average and S.D. of at least n=2 CFPS reactions are shown.



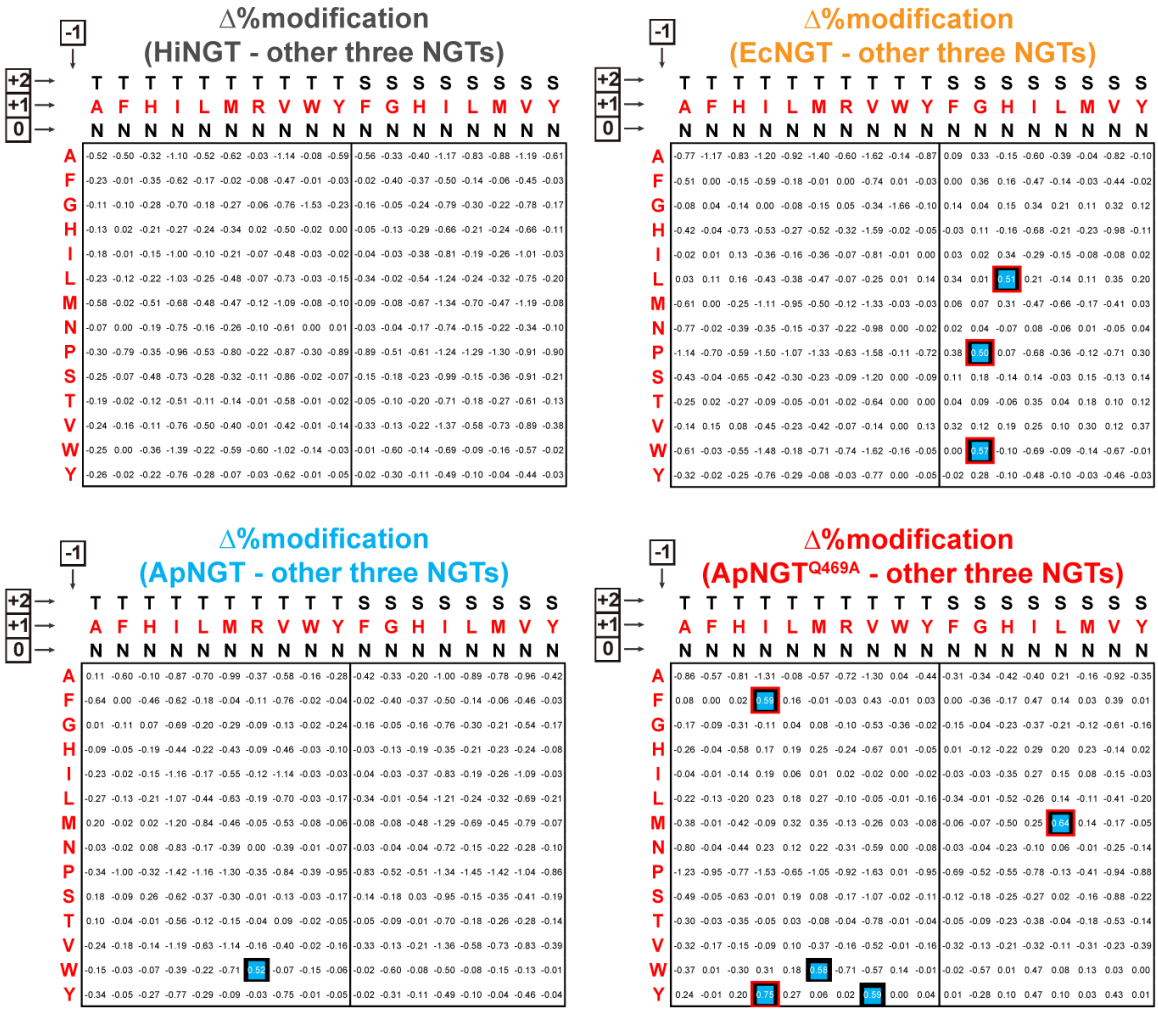
Supplementary Figure 2: Autoradiograms of putative NGTs expressed in CFPS. Soluble fractions of CFPS reactions, supernatant after centrifuged with 12,000xg for 15 min at 4 °C, containing ^{14}C-leucine were separated using a 4-12% SDS-PAGE Bis-tris gel in MOPS buffer and imaged using a 48 h phosphoscreen exposure (n=1). The same gel was Coomassie stained and aligned with autoradiogram image for molecular weight standard reference. Autoradiograms indicate the production of primarily full-length products near to expected molecular weights.



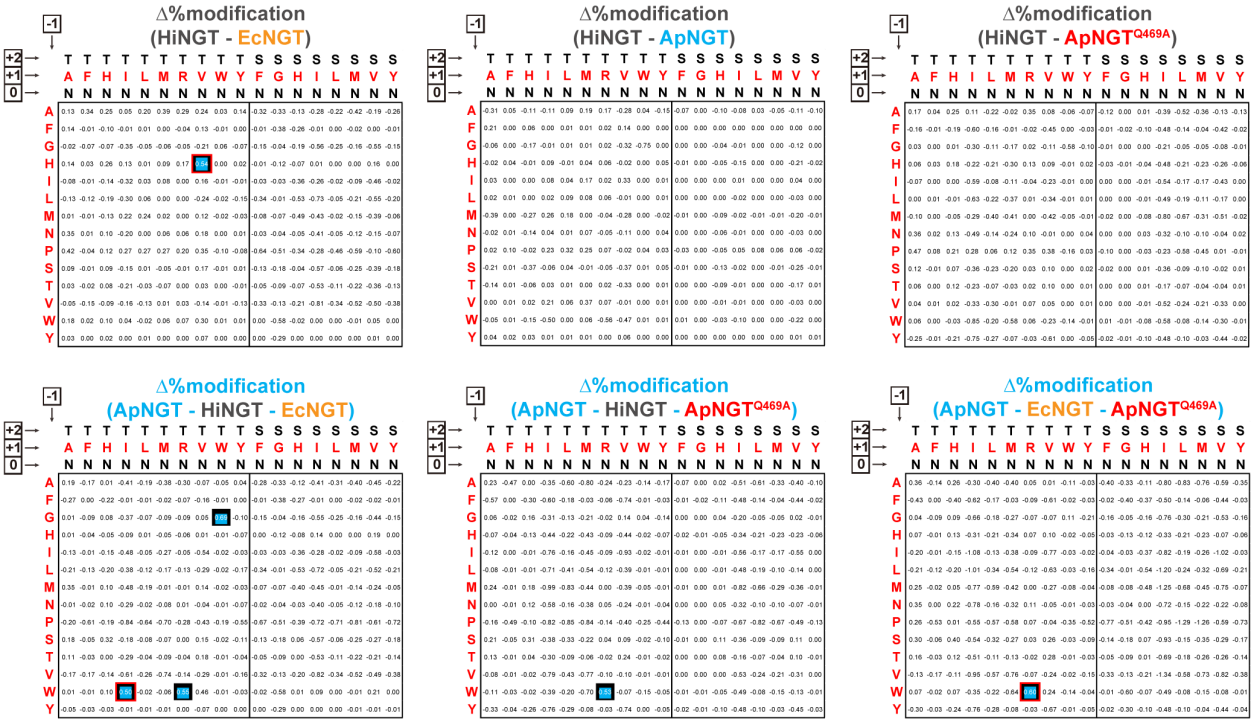
Supplementary Figure 3: Activity of 41 putative NGTs analyzed with six representative NGT substrate peptides. Five NGTs showed good activity on most peptides, and YpNGT only showed low levels of modification on one peptide “GNWT”. Relative intensities of peptide substrates and glycosylated products observed in mass spectra $Glc_1/(Glc_1+Glc_0)$ are shown when treated with 5% (v/v) CFPS NGTs and 2.5 mM UDP-Glc at 30 °C for 21 h (n=1).


X-1NX+1T/S_HiNGT
X-1NX+1T/S_EcNGT
X-1NX+1T/S_ApNGT
X-1NX+1T/S_ApNGTQ469A
100%
0%

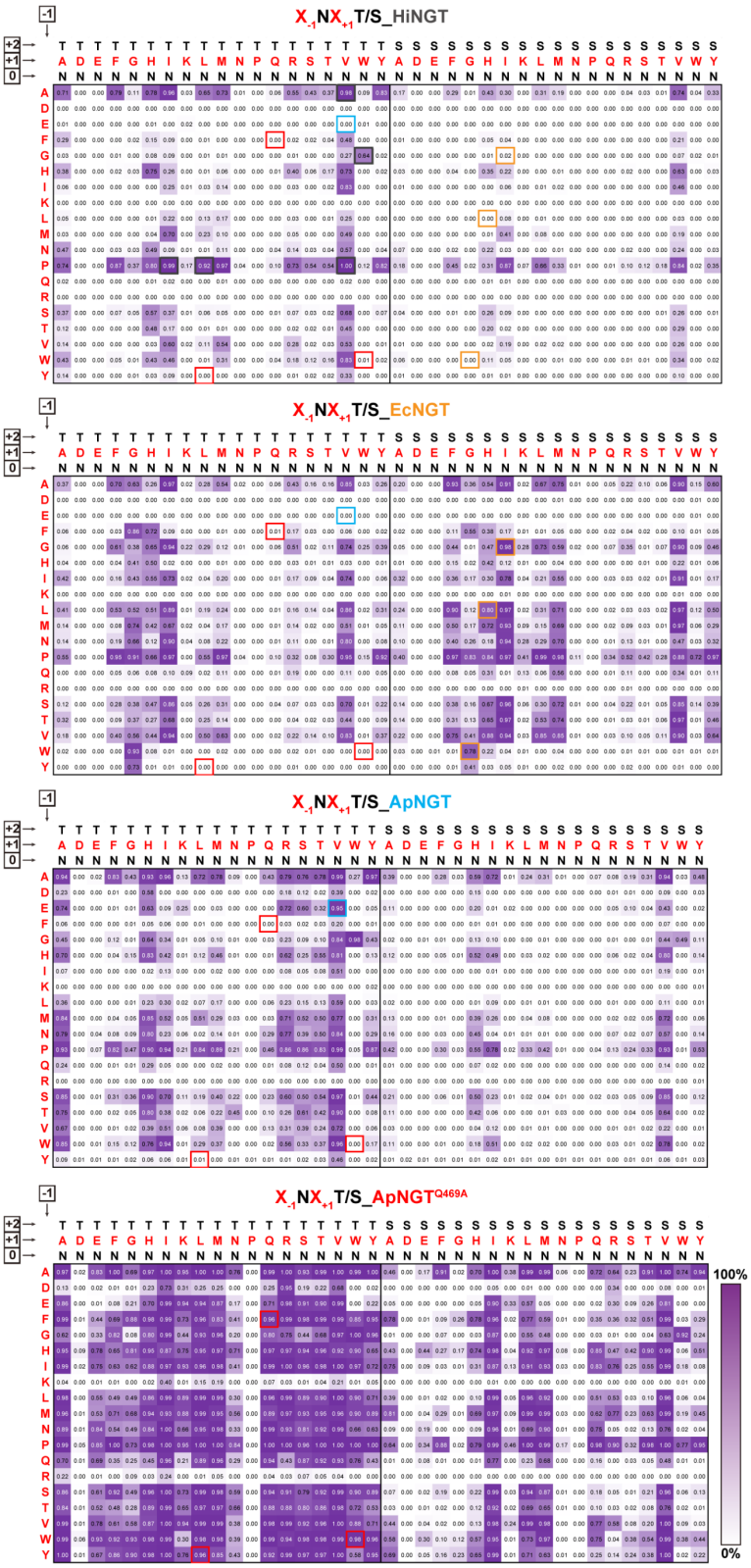
Supplementary Figure 4: NGTs show differences in peptide selectivity across $X_{-1}NX_{+1}$(T/S)RC peptide library. Percent glucose modification of peptides when treated with the four indicated NGTs. Conditions from n=1 experiment shown in **Fig. 2**.



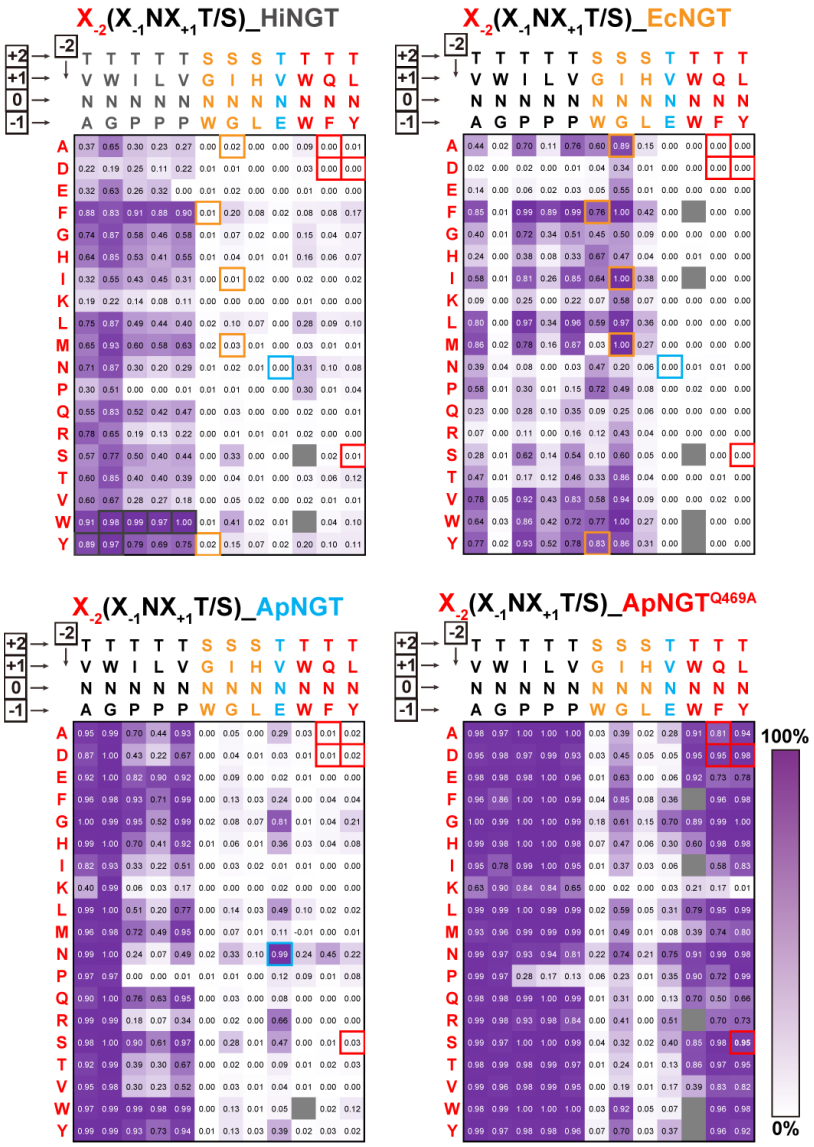
Supplementary Figure 5: Peptide sequences which were uniquely preferred by each NGT and could provide orthogonality to other NGTs. Differences in percent glucose modification (each NGT - other three NGTs) using the data in **Supplementary Fig. 4** showed the potential orthogonality of each NGT. For clarity, rows and columns in which all peptides and NGT combinations less than 50% modification are not shown. When the value is higher than 50%, the value is highlighted in “white” and a mini bar graph (blue bar in black background) is presented to show the ratio of %modification (NGT - other three NGTs) to %modification (NGT) 0-1 range. If the ratio is higher than 90%, the peptide is highlighted again with red squares. These peptides are potentially orthogonal to the other three NGTs. Only EcNGT and ApNGT^{Q469A} had several potentially orthogonal peptides when compared to the other three NGTs.



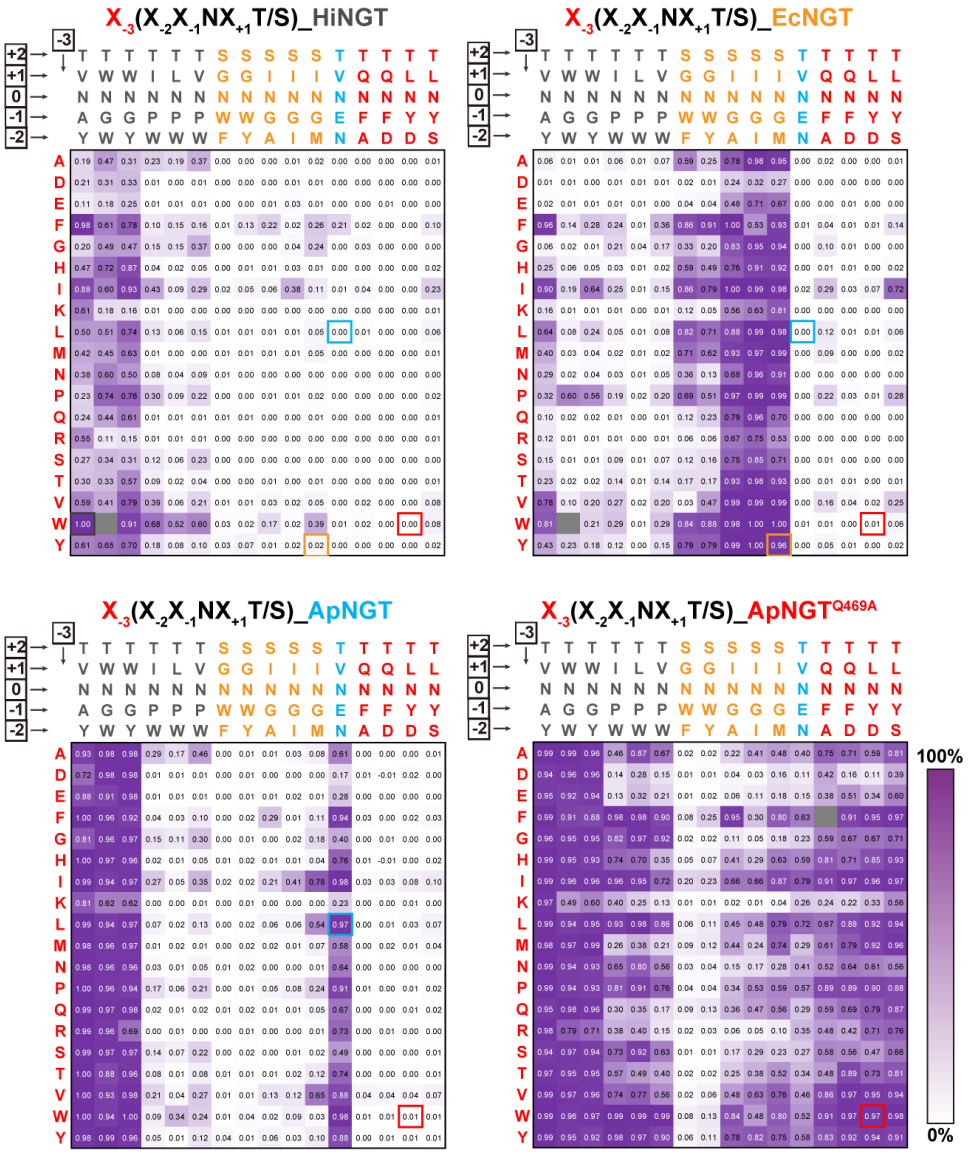
Supplementary Figure 6: Sequences showing potential orthogonality of HiNGT and ApNGT to other one or two NGTs. The analysis is similar to **Supplementary Fig. 5**. HiNGT only showed potential orthogonality to EcNGT while ApNGT showed potential orthogonality to HiNGT and EcNGT, or EcNGT and ApNGT^{Q469A}.


X-1NX+1T/S_HiNGT
X-1NX+1T/S_EcNGT
X-1NX+1T/S_ApNGT
X-1NX+1T/S_ApNGTQ469A
100%
0%

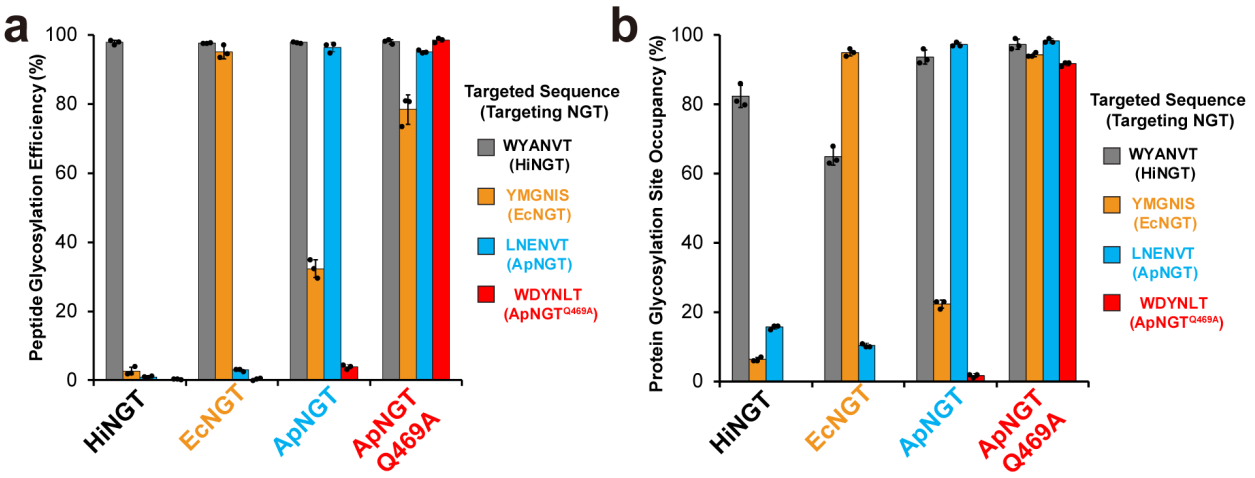
Supplementary Figure 7: Screening the conditional orthogonality of four NGTs using a $X_{-1}NX_{+1}$(T/S)RC peptide library. The $X_{-1}NX_{+1}$(T/S)RC peptide library was screened with HiNGT, EcNGT, ApNGT, and then ApNGT^{Q469A}, respectively, to find preferred peptides over preceding NGTs, in which the modification of each NGT is >90% more than the modification by preceding NGTs. Several GlycTags for each NGT were selected for the subsequent screens (colored boxes). Experimental conditions (n=2): 0.4 μM purified HiNGT or 0.7 μM purified EcNGT, 30 °C for 21h; 0.45 μM purified ApNGT or 0.25 μM purified ApNGT^{Q469A}, 30 °C for 3h.



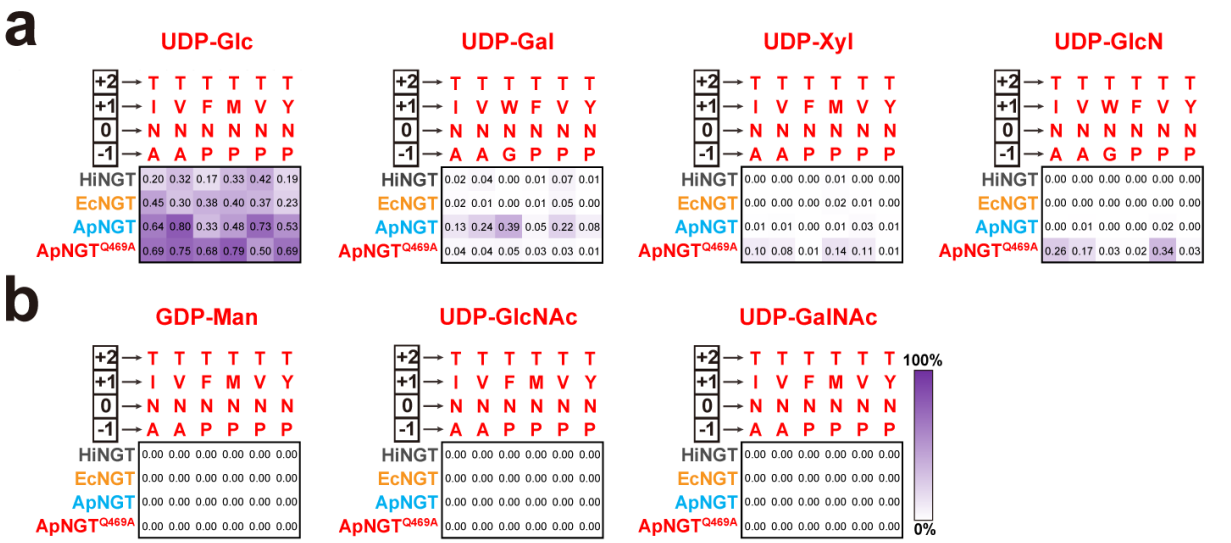
Supplementary Figure 8: Screening the conditional orthogonality of four NGTs using peptide library $X_{-2}(X_{-1}NX_{+1}T/S)$RC derived from the selected peptides in the $X_{-1}NX_{+1}T/S$ screens. The peptide library was screened in the order of HiNGT, EcNGT, ApNGT and then ApNGT^{Q469A}, individually, to find peptides preferred over preceding NGTs, in which the modification of each NGT is >90% more than the sum of the modifications by preceding NGTs. Several GlycTags for each NGT were selected for the next screen (colored boxes). Experimental conditions (n=1): 0.4 μM purified HiNGT or 1.2 μM purified EcNGT, 30 °C for 21h; 0.8 μM purified ApNGT or 0.2 μM purified ApNGT^{Q469A}, 30 °C for 3h.



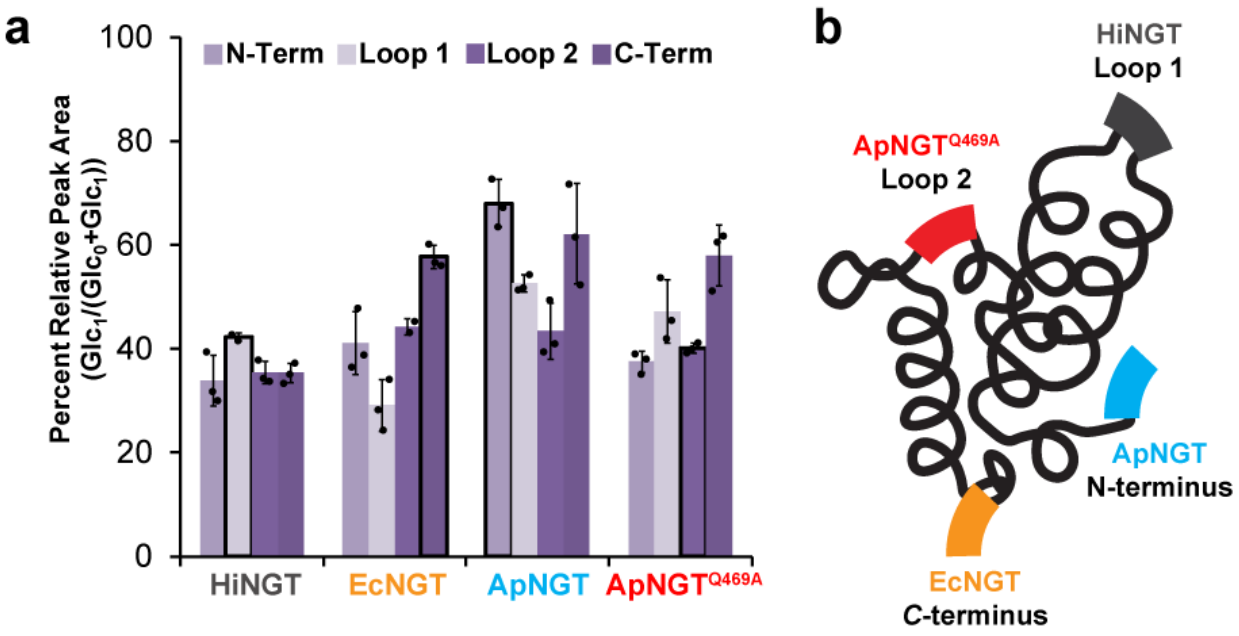
Supplementary Figure 9: Screening the conditional orthogonality of four NGTs using peptide library X$_{-3}$(X$_{-2}$X$_{-1}$NX$_{+1}$T/S)RC derived from the selected peptides in the X$_{-2}$(X$_{-1}$NX$_{+1}$T/S)RC screens. The peptide library was screened in the order of HiNGT, EcNGT, ApNGT and then ApNGT^{Q469A}, respectively, to identify preferred peptides over preceding NGTs, in which the modification of a given NGT is >95% more than the sum of the modifications by preceding NGTs. A set of four GlycTags was selected to engineer into proteins for sequential glycosylation. Experimental conditions (n=1): 0.15 μM purified HiNGT or 0.8 μM purified EcNGT, 30 °C for 21h; 0.3 μM purified ApNGT or 0.05 μM purified ApNGT^{Q469A}, 30 °C for 3h.



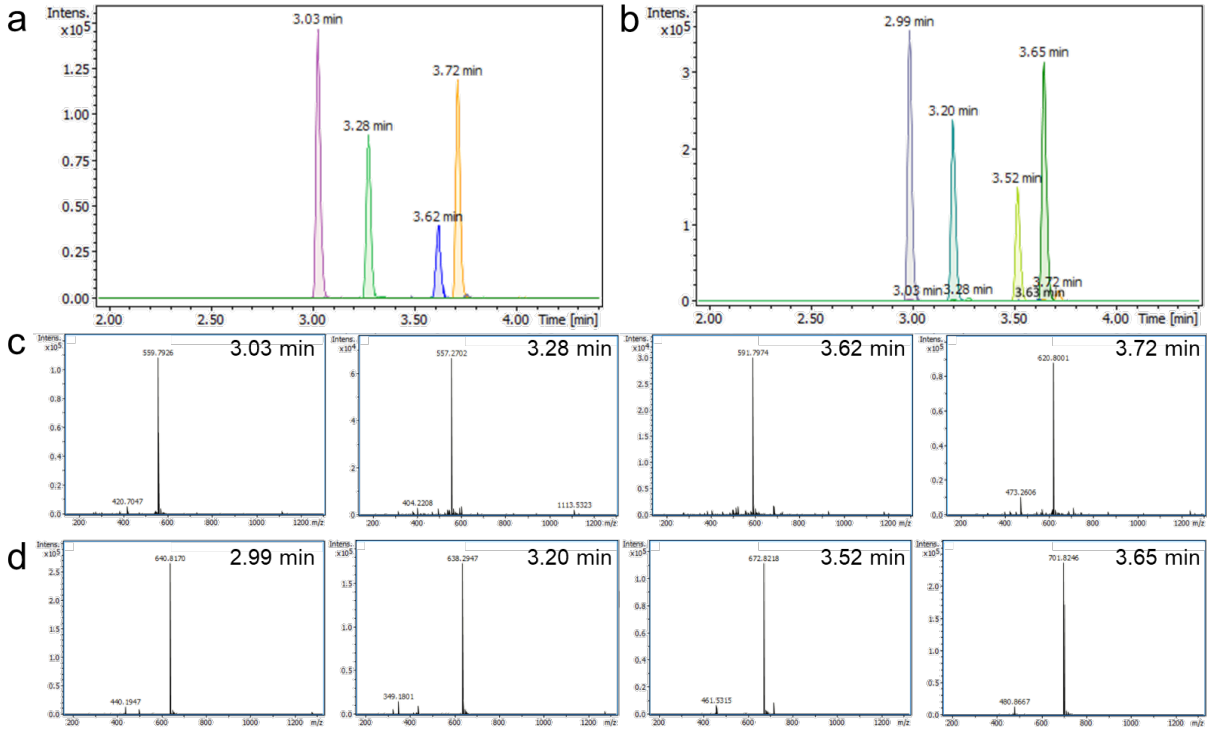
Supplementary Figure 10: Bar graphs showing conditional orthogonality of optimized GlycTags at the peptide level and differential targeting of protein glycosylation sites. (a) Bar graph showing conditional orthogonality of optimized 6-mer GlycTag peptides by HiNGT, EcNGT, ApNGT and ApNGT^{Q469A} (data shown as heatmap in **Fig. 4a**). Mean and S.D. of peptide glycosylation efficiencies determined by n=3 SAMDI-MS experiments are shown. **(b)** Bar graph showing differential targeting of optimized glycosylation sites within 4gIm7 protein under optimized conditions (data shown as heatmap in **Fig. 4c**). Mean and S.D. of n=3 independent protein IVG reactions and quantification of trypsinized peptide glycosylation efficiencies by LC-qTOF are shown. Full reaction condition information described in **Fig. 4** and **Methods**. Bar graphs show consistent glycosylation results and similar modification patterns for peptide sequences and GlycTags within 4gIm7 acceptor protein.



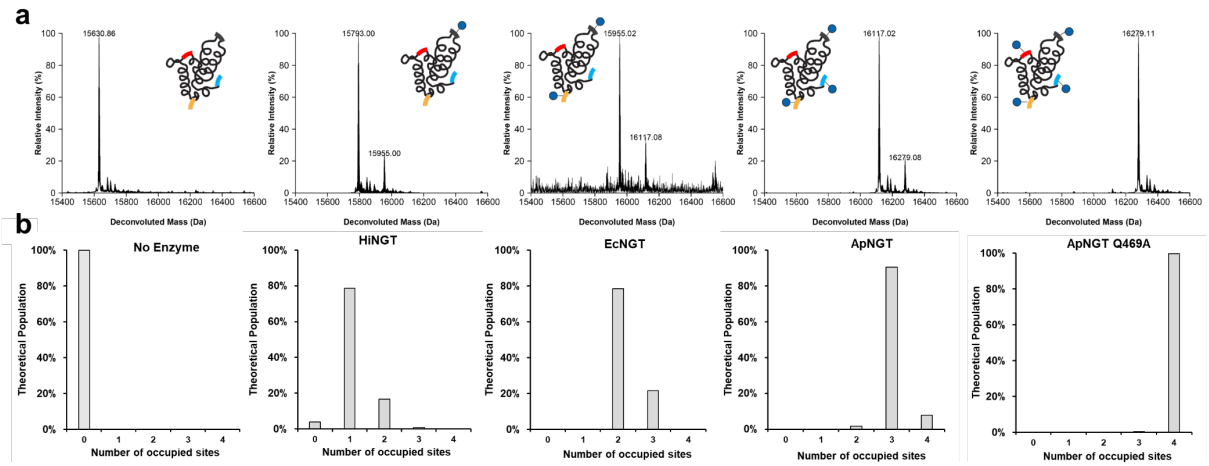
Supplementary Figure 11: Sugar donor specificities of the selected four NGTs. (a) Percentage intensities showing monosaccharide modification of 6 representative peptides after treatment with HiNGT, EcNGT, ApNGT, or ApNGT^{Q469A} as well as sugar donor UDP-Glc, UDP-Gal, UDP-Xyl, or UDP-GlcN. All experiments completed with NGTs produced in CFPS except for reactions with UDP-Gal, which were screened with purified NGTs to prevent possible conversion of UDP-Gal to UDP-Glc by endogenous *E. coli* enzymes within CFPS. All four NGTs showed modification with Glc, Gal and Xyl. ApNGT and ApNGT^{Q469A} also showed modification with GlcN. **(b)** Percentage intensities showing monosaccharide modification of 6 representative peptides after treatment with HiNGT, EcNGT, ApNGT, or ApNGT^{Q469A} as well as sugar donor GDP-Man, UDP-GlcNAc, or UDP-GalNAc. No NGTs showed modification on these three sugar donors. The following experimental conditions were used. In experiments with UDP-Glc donor, 0.42 μM HiNGT or 0.75 μM EcNGT produced in CFPS was reacted at 30 °C for 21 h, or 0.1 μM ApNGT or 0.05 μM ApNGT^{Q469A} produced in CFPS was reacted at 30 °C for 1 h. In experiments with UDP-Xyl, GDP-Man, UDP-GlcNAc, and UDP-GalNAc, 0.42 μM HiNGT, 0.75 μM EcNGT, 0.1 μM ApNGT, or 0.05 μM ApNGT^{Q469A} produced in CFPS was reacted at 30 °C for 21 h. In experiments with UDP-Gal, 0.84 μM purified HiNGT, 1.5 μM purified EcNGT, 0.2 μM purified ApNGT, or 0.1 μM purified ApNGT^{Q469A} was reacted at 30 °C for 21 h. In experiments with UDP-GlcN, 0.84 μM HiNGT, 1.5 μM EcNGT, 1.15 μM ApNGT or 1 μM ApNGT^{Q469A} produced in CFPS was reacted at 30 °C for 21 h. All experiments were performed once (n=1). All percentage intensities of ≥ 0.01 are above background noise and indicative of real modification.



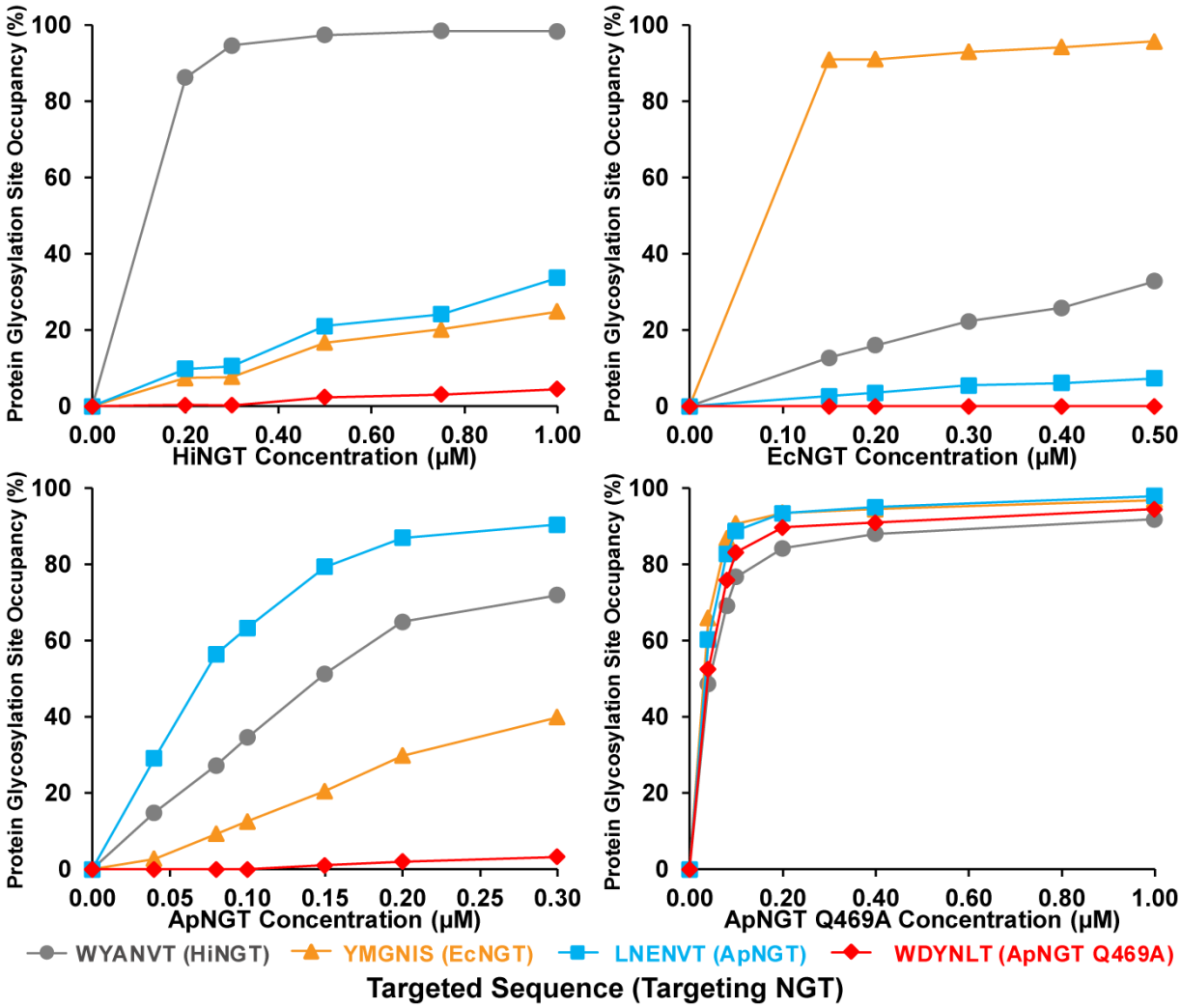
Supplementary Figure 12: Optimization of GlycTag placement within 4gIm7 protein. (a) Bar graph comparing the preferences of HiNGT, EcNGT, ApNGT, and ApNGT^{Q469A} for GlycTags placed into various locations of 4gIm7. Preferences were determined by placing the promiscuous sequence "IYANVTL" (which our previous analyses indicated could be easily modified by all four NGTs) at the *N*-terminus, *C*-terminus, replacing N26_D32 in Loop 1, or replacing S58_S64 in Loop 2 of wildtype Im7. Each of these versions of Im7 containing only one glycosylation site were purified from overexpression in cells using plasmids pET.BCS.Generalist1-4, reacted with each NGT and UDP-Glc, purified again using magnetic Ni-NTA beads, and then analyzed by LC-qTOF intact protein analysis. Glycosylation was quantified as previously described[6, 19] by integrating extracted ion chromatograms of the three highest intensity charge states of the glycosylated and aglycosylated intact protein using Bruker Compass Data Analysis software. Average and S.D. of percent relative peak areas of the glycosylated and aglycosylated products from n=3 independent IVGs, purification, and LC-qTOF analyses are shown. IVG reactions were performed by combining 5 μM of the purified Im7 variant; 2.5 mM UDP-Glc; and 0.4 μM HiNGT, 0.25 μM EcNGT, 0.05 μM ApNGT, or 0.025 μM ApNGT^{Q469A} and incubating for 4 h at 30 °C. While we did not observe drastic differences in the positional preferences of each NGT in **a**, we determined that Loop 1, *C*-terminus, and the *N*-terminus were preferred by HiNGT, EcNGT, and ApNGT, respectively. Because ApNGT^{Q469A} is last in the sequential glycosylation reaction, its GlycTag was assigned the remaining Loop2 site. Outlined bars indicate assigned positions for each NGT. **(b)** Positioning of NGT-GlycTag pairs in final design of 4gIm7 protein based on results in **a**.



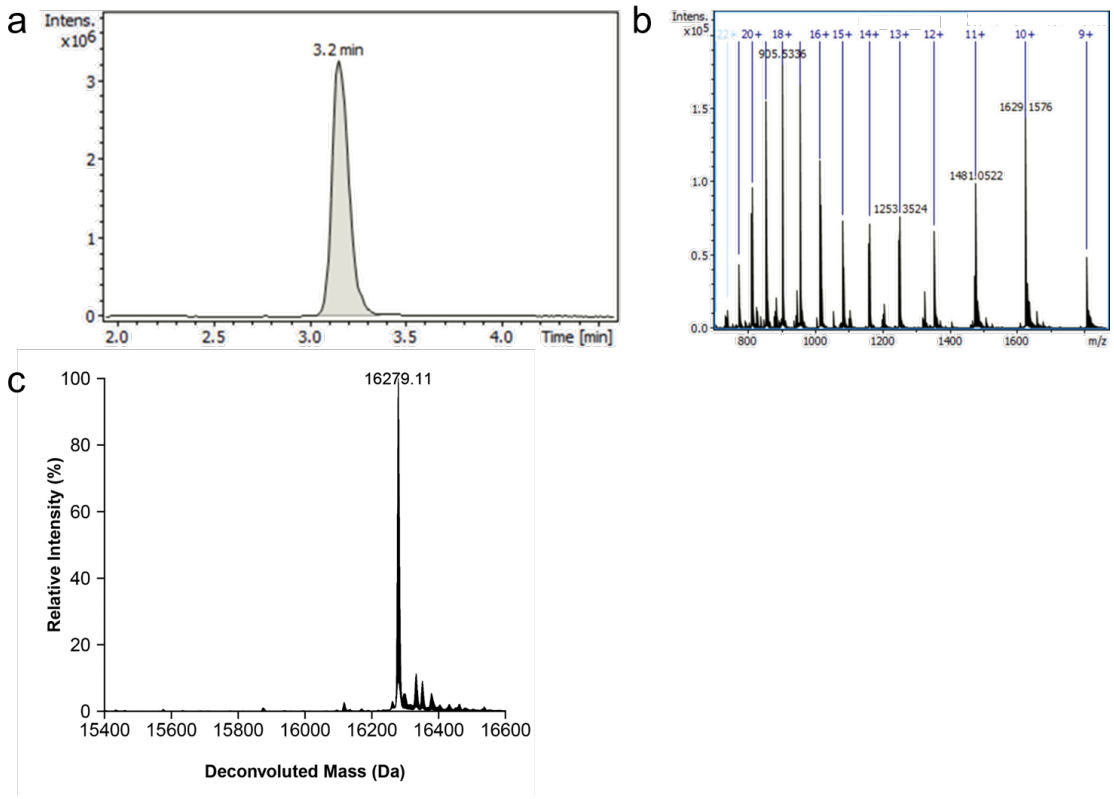
Supplementary Figure 13: Representative LC-qTOF peptide chromatograms MS spectra for trypsinized 4glm7. (a) Representative extracted ion chromatograms (EICs) from LC-qTOF glycopeptide analysis of aglycosylated, trypsinized 4glm7. EICs based on theoretical m/z values for +2 charged peptide products arising from trypsinization of unglycosylated 4glm7, each containing one engineered GlycTag. Trypsinized peptides, targeting NGT, theoretical m/z values, and observed retention time are: ATTYMGNISR, EcNGT, 557.27 m/z, 3.28 min; ATTLNENVTR, ApNGT, 559.79 m/z, 3.03 min; ATTWYANVTR, HiNGT, 591.80 m/z, 3.62 min; ATTDYNLTR, ApNGT^{Q469A}, 620.80 m/z, 3.72 min. **(b)** Representative extracted ion chromatograms (EICs) from LC-qTOF glycopeptide analysis of trypsinized 4glm7 following purification from an *in vitro* glycosylation reaction containing 10 μM 4glm7, 2.5 mM UDP-Glc, and 0.4 μM ApNGT^{Q469A} incubated for 4 h at 30 °C. EICs based on theoretical m/z values for +2 charged peptide products arising from trypsinization of glycosylated 4glm7, each containing one engineered GlycTag with glucose modification. Trypsinized glycopeptides, targeting NGT, theoretical m/z values, and observed retention time are: ATTYMGN(Glc)ISR, EcNGT, 638.30 m/z, 3.20 min; ATTLLNEN(Glc)VTR, ApNGT, 640.82 m/z, 2.99 min; ATTWYAN(Glc)VTR, HiNGT, 672.82 m/z, 3.52 min; ATTDYN(Glc)LTR, ApNGT^{Q469A}, 701.83 m/z, 3.65 min. We observed slightly decreased retention time for glycosylated peptides compared to their aglycosylated counterparts, likely due to the addition of a hydrophilic sugar moiety. **(c)** MS spectra from apex of each EIC shown in **a** at indicated retention time. **(d)** MS spectra from apex of each EIC shown in **b** at indicated retention time. All ion chromatograms extracted with ± 0.01 m/z range around theoretical m/z values. All chromatograms and spectra representative of at least n=3 independent experiments.



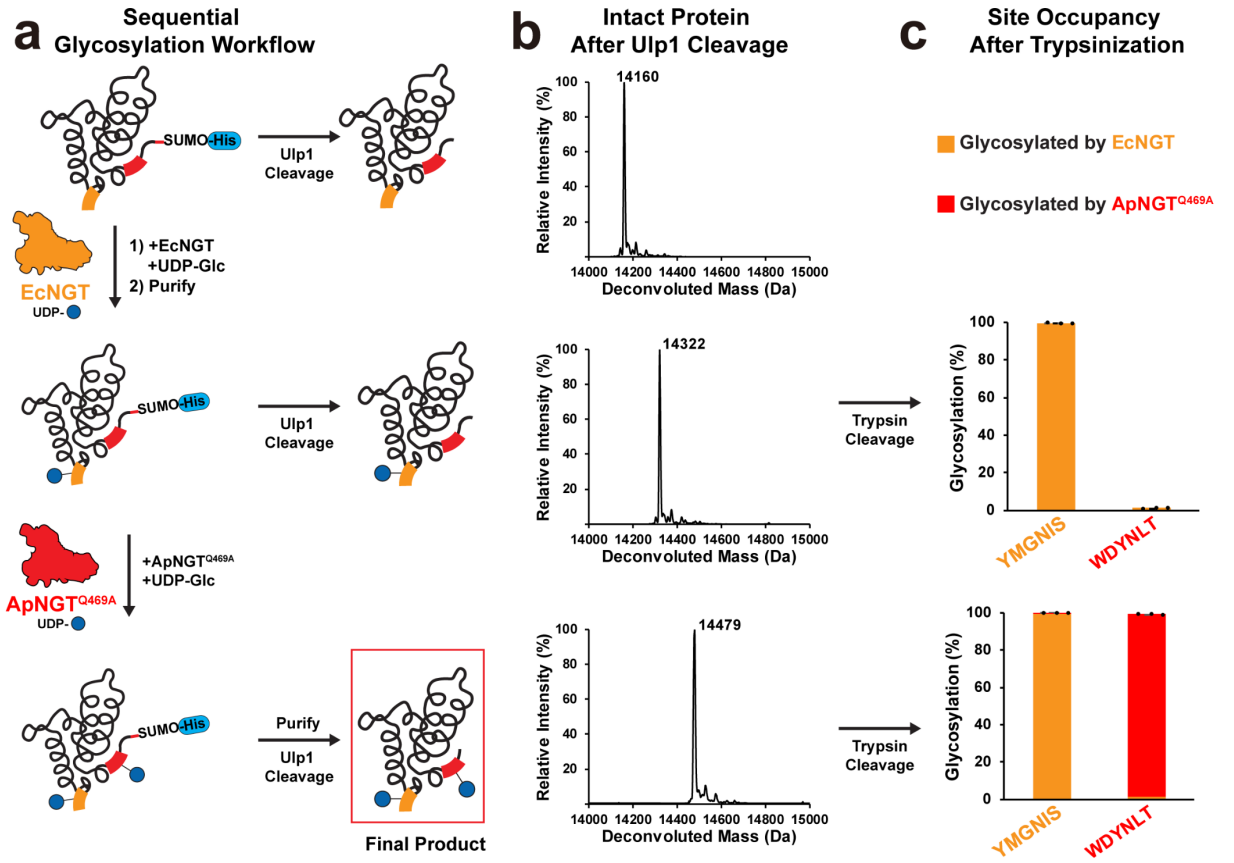
Supplementary Figure 14: **Sequential glycosylation events are not interdependent. (a)** Deconvoluted intact protein spectra of Ulp1 cleaved SUMO-4glm7 after each NGT treatment, analyzed by LC-qTOF; representative of n=3 independent experiments. Same as **Fig. 5b**. **(b)** Calculation of the relative amount of the SUMO-4glm7 population with different numbers of occupied glycosylation sites based on quantification of site-occupancy in **Fig. 5c** after each indicated NGT treatment, with the assumption of that each glycosylation event is independent. Probability calculations completed in Microsoft Office Excel 2016. The correlation between spectra in **a** and relative amounts in **b** does not suggest any interdependence of glycosylation sites.



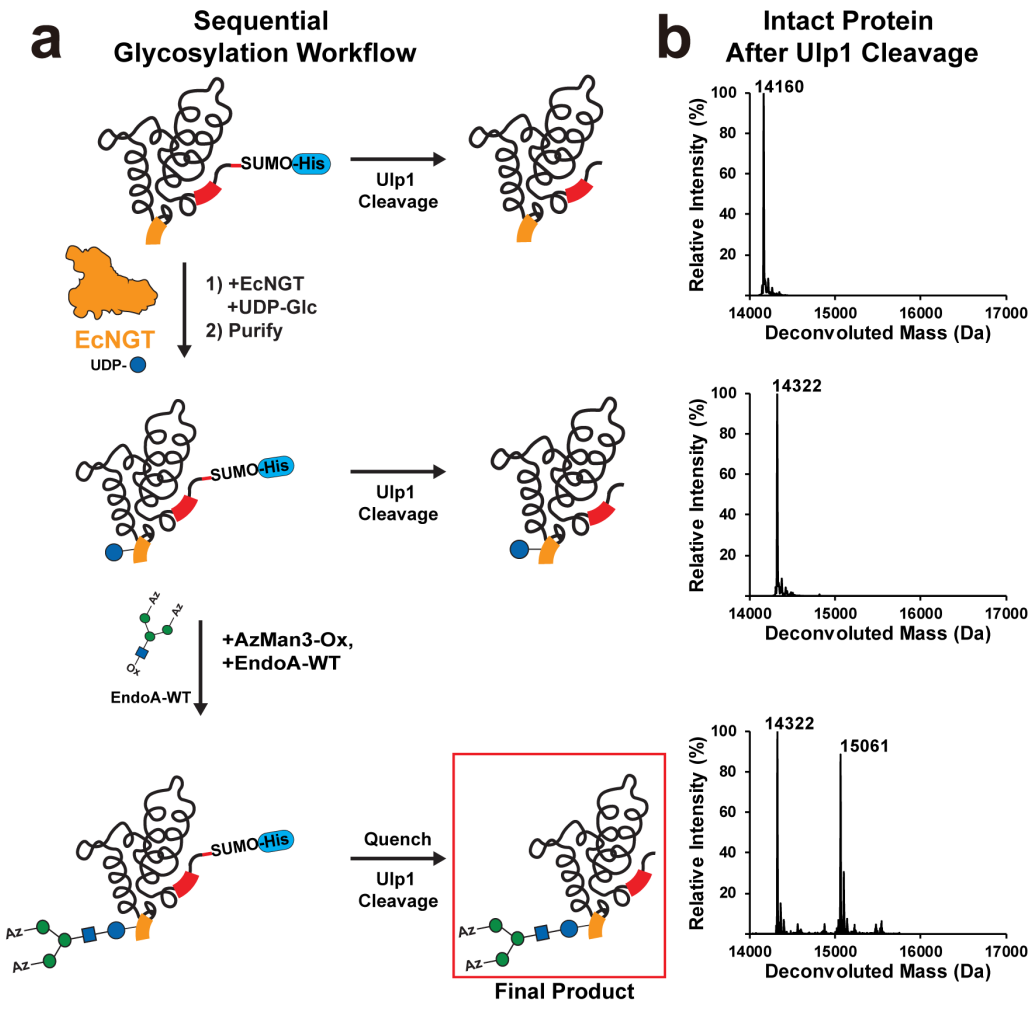
Supplementary Figure 15: Four NGTs show differential modification of optimized glycosylation sites within SUMO-4glm7. Purified SUMO-4glm7 bearing four optimized 6-mer GlycTags from **Fig. 3** was reacted with 2.5 mM UDP-Glc and various concentrations of purified HiNGT, EcNGT, ApNGT, or ApNGT^{Q469A} for 4 h. After the IVG reaction, SUMO-4glm7 was purified using Ni-NTA magnetic beads, treated with trypsin, and the occupancy of each glycosylation site was determined by LC-qTOF peptide analysis. Each datapoint was determined from n=1 IVG reaction, purification, and measurement of protein glycosylation site occupancy using LC-qTOF.



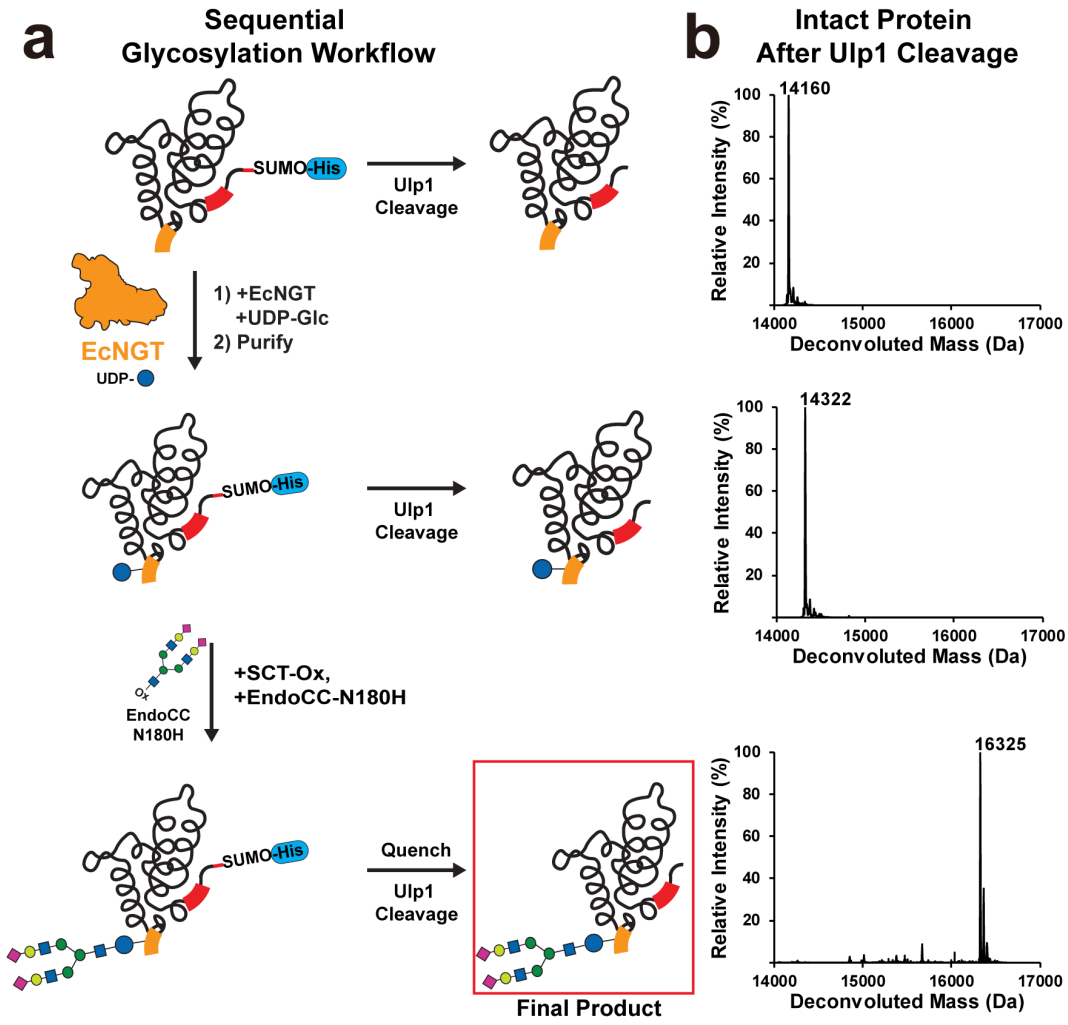
Supplementary Figure 16: Representative LC-qTOF intact protein chromatogram, MS spectra, and deconvoluted MS spectra of SUMO-4glm7 after Ulp1 cleavage. (a) Representative EIC from LC-qTOF intact protein analysis of SUMO-4glm7 following sequential *in vitro* glycosylation as described in **Fig. 5**. EIC extracted based on theoretical m/z value for most abundant charge state (+18) of fully glycosylated, intact protein product, 905.39 m/z ± 0.1 m/z range. **(b)** MS spectra from entire elution of EIC in **a**. **(c)** Deconvolution of MS spectra in **b** using Bruker Compass Data Analysis maximum entropy deconvolution of MS peaks within m/z range 700-2000 into mass range of 14,000-18,000 Da. Raw deconvolution data was then plotted and annotated using R Studio. Chromatogram, MS spectra, and deconvoluted spectra representative of n=3 independent sequential glycosylation experiments. Data analysis method representative of LC-qTOF intact protein analysis resulting in deconvoluted spectra in **Fig. 5** and described in **Methods**.



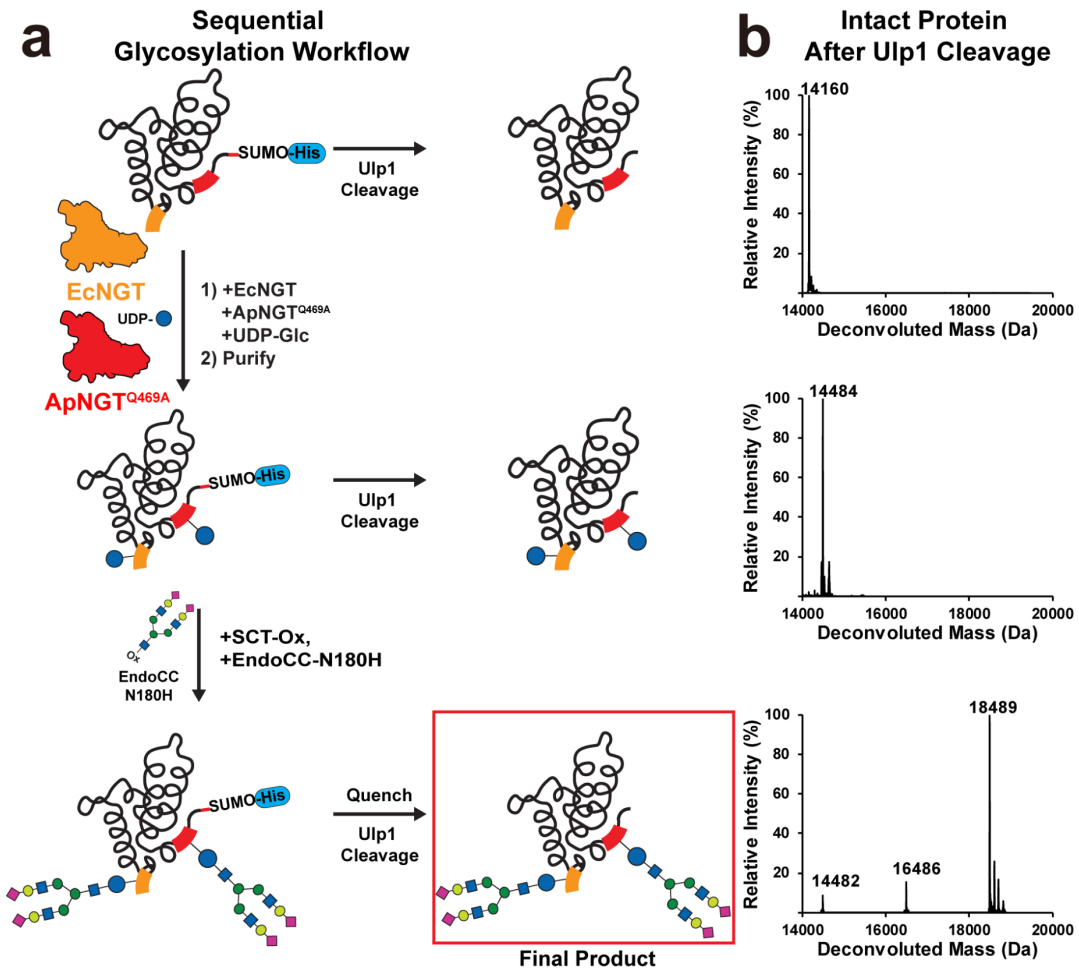
Supplementary Figure 17: **Quantitative, site-specific control of glycosylation at two GlycTags within a single Im7 protein**. **(a)** Workflow for control of glycosylation at two sites within 2gIm7 by sequential addition of NGTs. **(b)** Deconvoluted intact 2gIm7 MS spectra showing SUMO-2gIm7 cleaved by Ulp1 and analyzed by LC-qTOF before NGT treatment, after EcNGT treatment, and after treatment with ApNGT^{Q469A}. 10 μM purified SUMO-2gFc was incubated with 1 μM EcNGT for 4 h at 30°C containing 2.5 mM UDP-Glc and then purified using Qiagen Ni-NTA resin, dialyzed into buffer 1, then treated with 10 μM ApNGT^{Q469A} and 2.5 mM UDP-Glc for 16 h at 30°C, and purified using Ni-NTA magnetic beads. The intact spectra show the addition of one glucose by EcNGT and then another by ApNGT^{Q469A}. **(c)** A bar graph of site-occupancy at each of the two GlycTags after the same experiments shown in **b** after NGT treatment, cleavage by trypsin and analysis by LC-qTOF. Average and S.D. of n=3 separate IVG reactions and LC-qTOF runs are shown. Nearly complete conversion of each GlycTag by its intended enzyme was observed with very little off-target modification. Approximately 98% of the 2gIm7 was converted correctly during both steps.



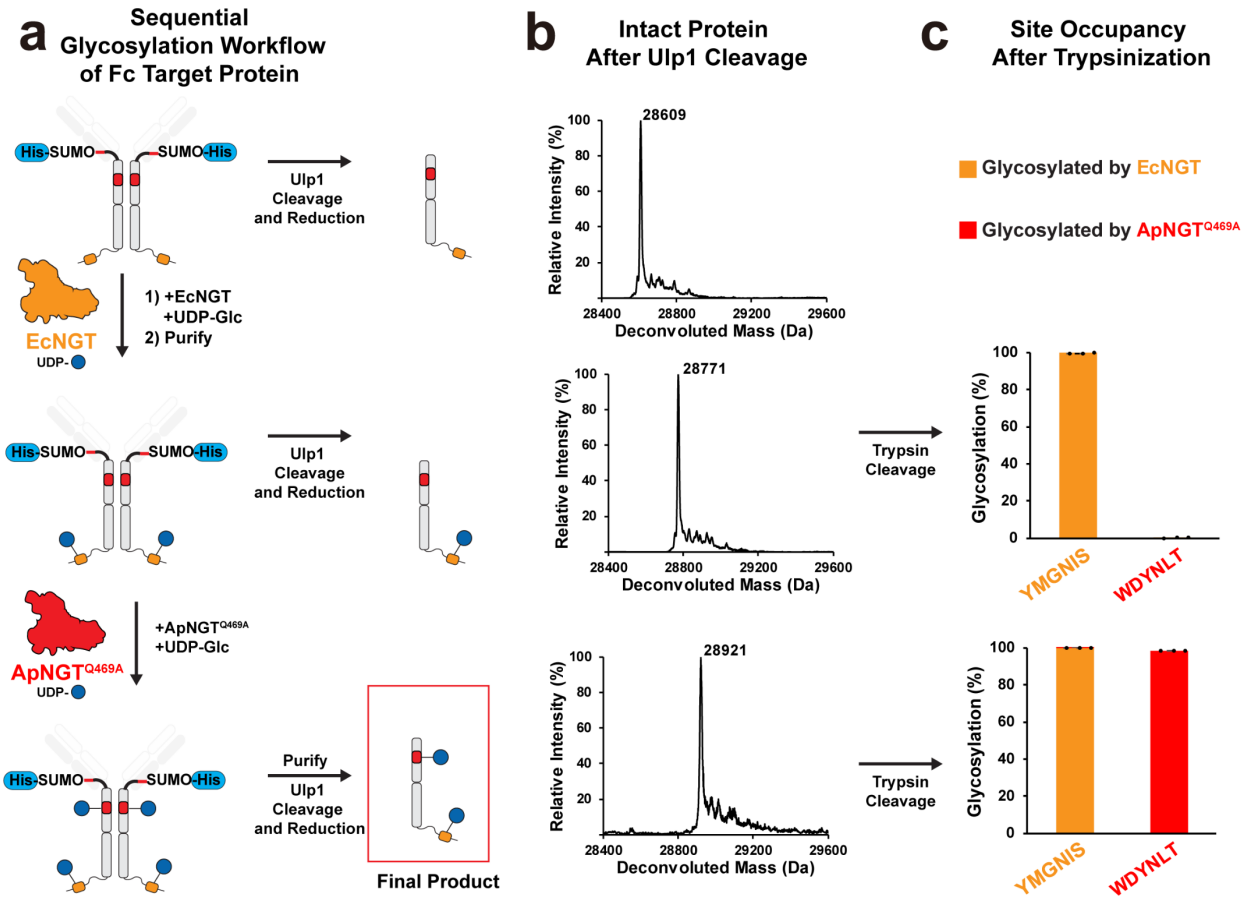
Supplementary Figure 18: **Site-specific installation of AzMan3 onto Im7 protein. (a)** Workflow for selectively modifying EcNGT-targeted site at the *C*-term of 2gIm7 by installing with human-like AzMan2ManGlcNAcGlc glycan ($AzMan_3$). **(b)** Deconvoluted intact 2gIm7 MS spectra showing 2gIm7 before NGT treatment, after EcNGT treatment, and after transglycosylation. 10 μM SUMO-2gIm7 was monoglycosylated by treating with 1 μM EcNGT for 4 h at 30 °C before purification with Qiagen Ni-NTA resin. 50 μg (0.0018 μmoles) of monoglucosylated SUMO-2gIm7 and 120 molar equivalents of AzMan3-ox (160 μg, 0.218 μmoles) were incubated with 9 μg of EndoA WT at a final enzyme concentration of 0.2 μg/μl in PBS (pH=7.4) at 30 °C. An additional 120 molar equivalents of AzMan3-ox were added at 1 h. The reaction was quenched with 0.1 M glycine buffer (pH=2.7) at 2 h. Samples were analyzed by LC-qTOF after pH neutralization and Ulp1 cleavage (see **Methods**). Spectra of aglycosylated and monoglucosylated 2gIm7 are also shown in **Supplementary Fig. 17**. Transglycosylation reactions were performed once (n=1) with a yield of approximately 47% based on relative areas of substrate and product peaks in deconvoluted protein spectra.



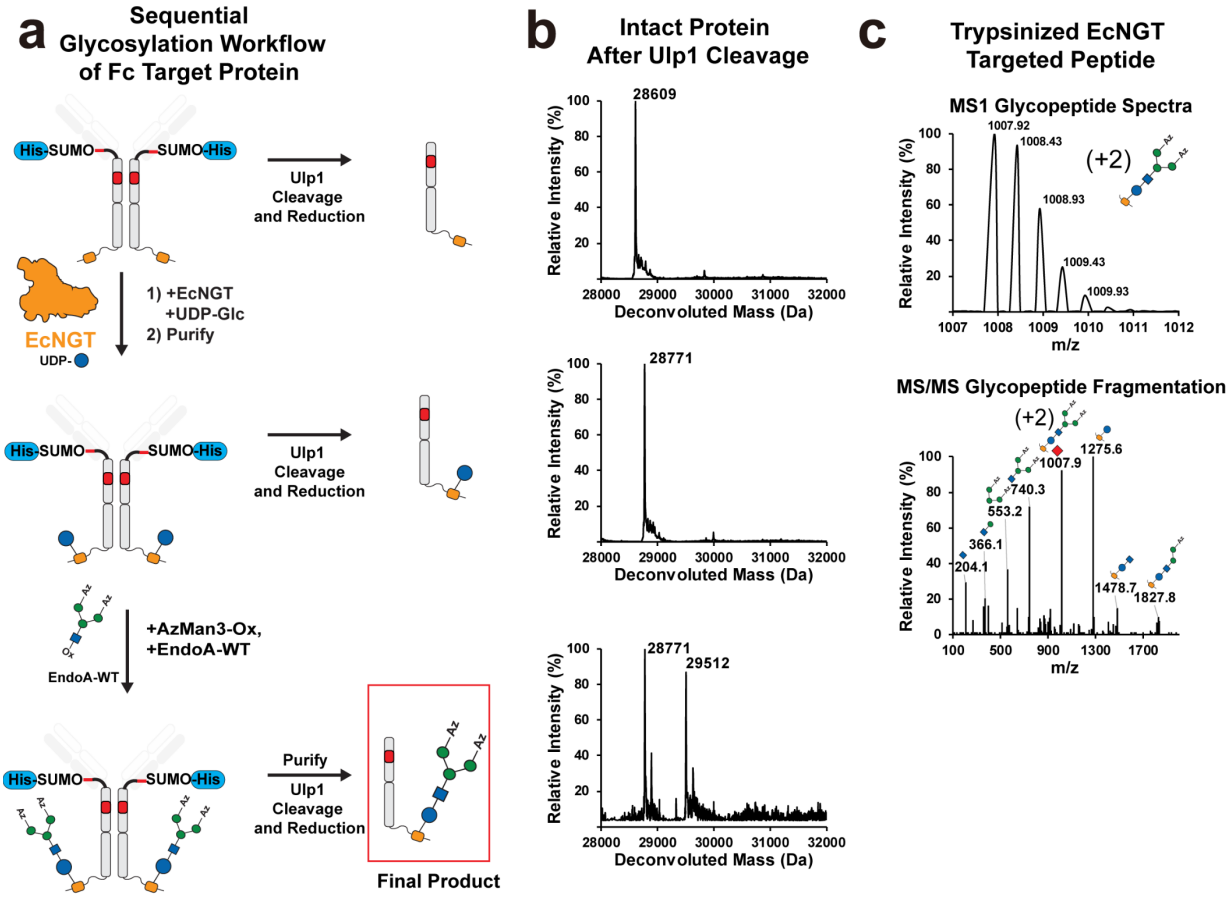
Supplementary Figure 19: Site-specific installation of biantennary SCT onto Im7. (a) Workflow for selectively modifying EcNGT-targeted site at the *C*-term of 2gIm7 with human-like $Sia_2Gal_2GlcNAc_2Man_3GlcNAcGlc$ glycan (SCT). **(b)** Deconvoluted intact 2gIm7 MS spectra showing 2gIm7 before NGT treatment, after EcNGT treatment, and after transglycosylation. 10 μM SUMO-2gIm7 was monoglycosylated by treating with 1 μM EcNGT for 4 h at 30 °C before purification with Qiagen Ni-NTA resin. 50 μg (0.0018 μmoles) of monoglucosylated SUMO-2gIm7 and 120 molar equivalents of SCT-ox (440 μg, 0.218 μmoles) were incubated with 0.8 μg of $EndoCC^{N180H}$ at a final enzyme concentration of 0.019 μg/μl in 50 mM Tris-HCl (pH=7.5) and 150 mM NaCl. The reaction was carried out at 30 °C for 30 minutes followed by quenching with 0.1 M glycine buffer (pH=2.7) for 15 minutes. pH was neutralized with Tris buffer (pH=8.5) and incubated with Ulp1 endopeptidase in 1 mM DTT at 30 °C for 1 h to cleave SUMO tag from 2gIm7. The sample was analyzed by LC-qTOF (see **Methods**). Spectra of aglycosylated and monoglucosylated 2gIm7 are also shown in **Supplementary Fig. 17**. Transglycosylation reactions were performed once (n=1) with apparent quantitative yield based on deconvoluted intact protein spectra.



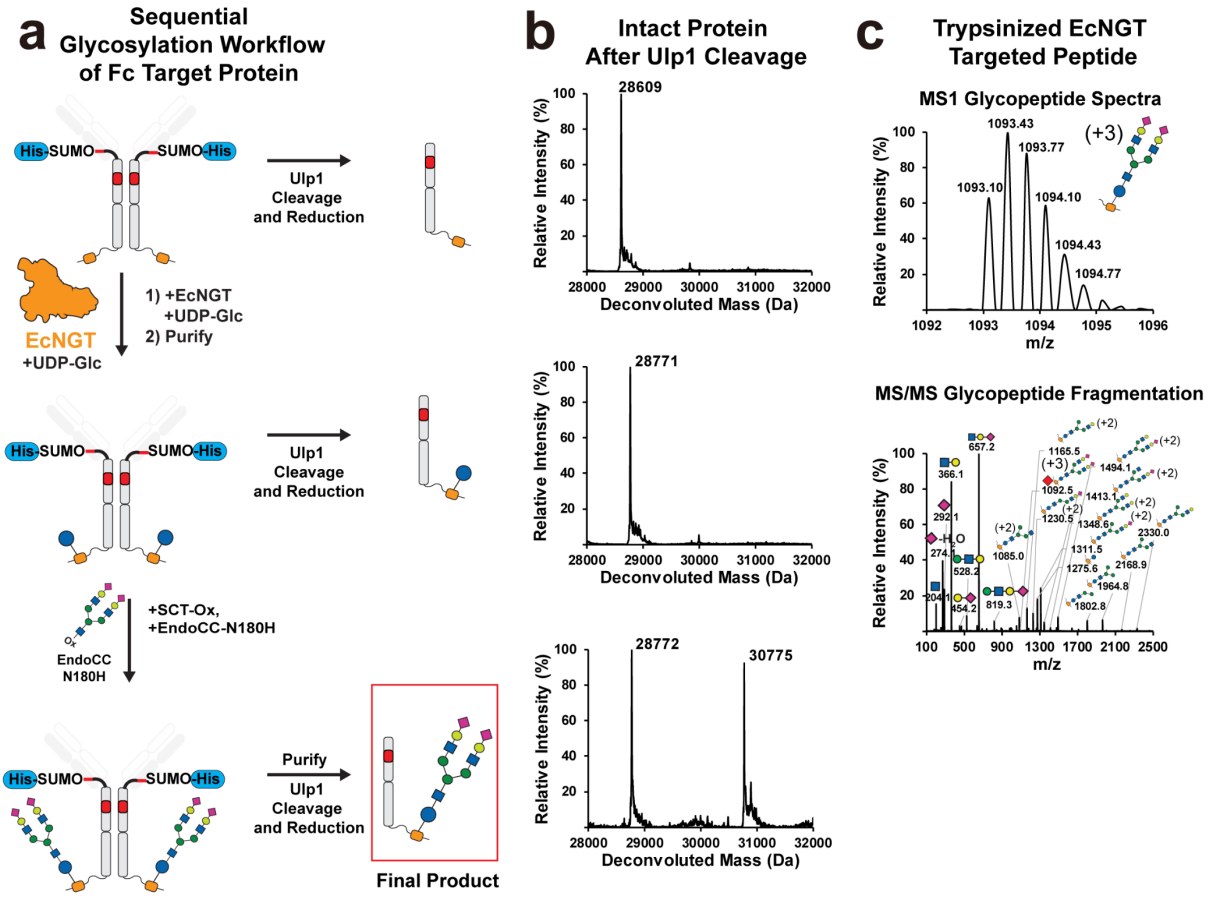
Supplementary Figure 20: Installation of biantennary SCT glycan onto two sites within Im7. **(a)** Workflow for modifying sites targeted by EcNGT and ApNGT^{Q469A} site at the *N*-term and *C*-term of 2gIm7 with human-like SCT glycan. **(b)** Deconvoluted intact 2gIm7 MS spectra showing 2gIm7 before NGT treatment, after NGTs treatment, and after transglycosylation. 10 μM SUMO-2gIm7 was glycosylated by treating with 1 μM EcNGT and 10 μM ApNGT^{Q469A} at 30°C for 16 h before purification with Qiagen Ni-NTA resin. 30 μg (0.0011 μmoles) of diglucosylated SUMO-2gIm7 and 240 molar equivalents of SCT-ox (530 μg, 0.265 μmoles) were incubated with 0.8 μg of EndoCC^{N180H} at a final enzyme concentration of 0.019 μg/μl in 50 mM Tris-HCl (pH=7.5) and 150 mM NaCl. The reaction was carried out at 30°C. Additional 150 molar equivalent of SCT-ox was added at 90 minutes. Reaction was quenched with 0.1 M glycine buffer (pH=2.7) at 140 minutes. Samples were analyzed by LC-MS (**Methods**). The spectrum of aglycosylated is also shown in **Supplementary Fig. 17**. Transglycosylation reactions were performed once (n=1) show approximately 81% conversion to the di-SCT-modified final product based on relative intensities of substrate and product peaks in deconvoluted protein spectra.



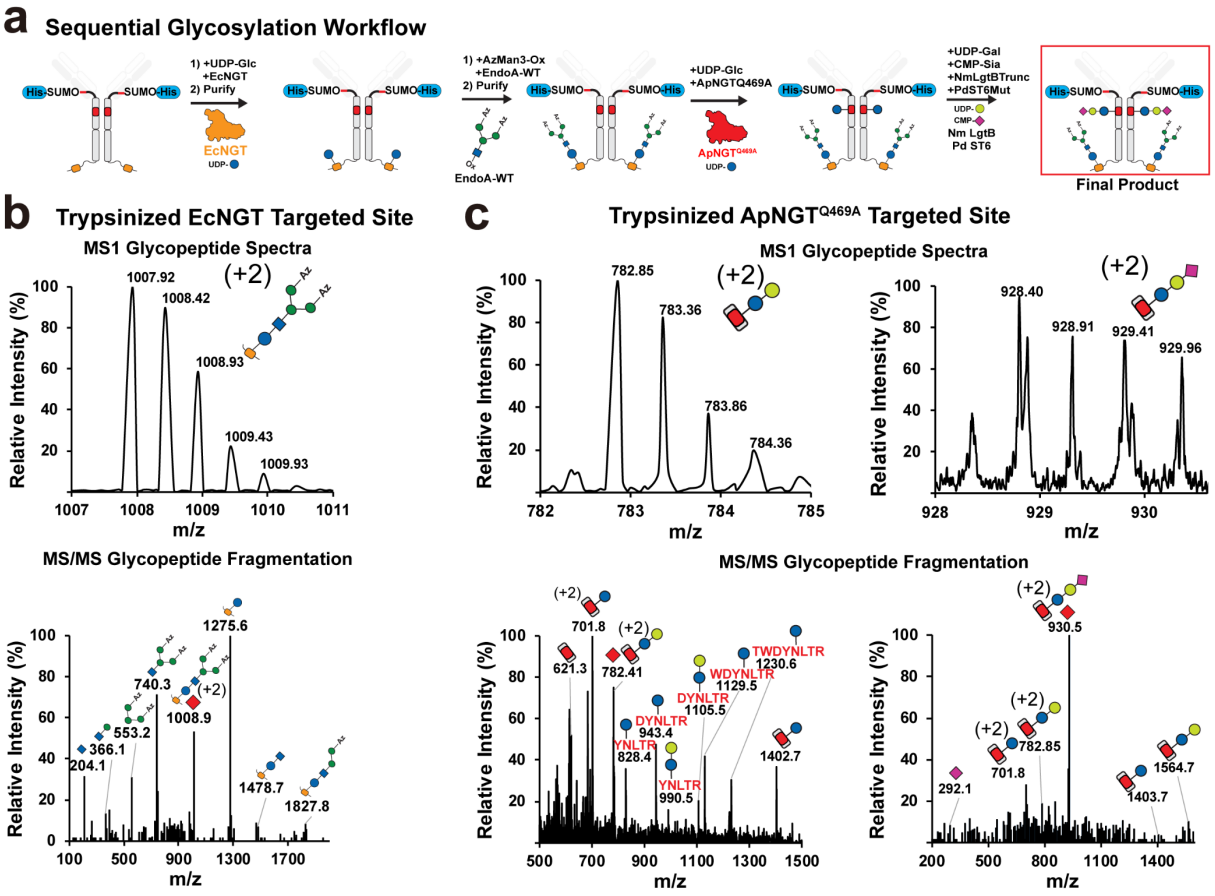
Supplementary Figure 21: Quantitative, site-specific control of glycosylation at two GlycTags within a single human Fc protein. **(a)** Workflow for control of glycosylation at two sites within 2gFc by sequential addition of NGTs. **(b)** Deconvoluted intact 2gFc MS spectra showing SUMO-2gFc cleaved by Ulp1 and analyzed by LC-qTOF before NGT treatment, after EcNGT treatment, and after treatment with ApNGT^{Q469A}. 10 μM purified SUMO-2gFc was incubated with 1 μM EcNGT for 4 h at 30 °C containing 2.5 mM UDP-Glc and then purified using Qiagen Ni-NTA resin, dialyzed into buffer 1, and then treated with 10 μM ApNGT^{Q469A} and 2.5 mM UDP-Glc for 16 h at 30 °C, and purified using Ni-NTA magnetic beads. The intact LC-qTOF spectra show the addition of one glucose by EcNGT and then another by ApNGT^{Q469A}. **(c)** A bar graph of site-occupancy at each of the two GlycTags after the same experiments shown in **b** after NGT treatment, cleavage by trypsin and analyzed by LC-qTOF. Average and S.D. of n=3 separate IVG reactions are shown. Nearly complete conversion of each GlycTag by its intended enzyme was observed with very low off-target modification. Approximately 98% of the 2gFc was correctly converted for both sites.



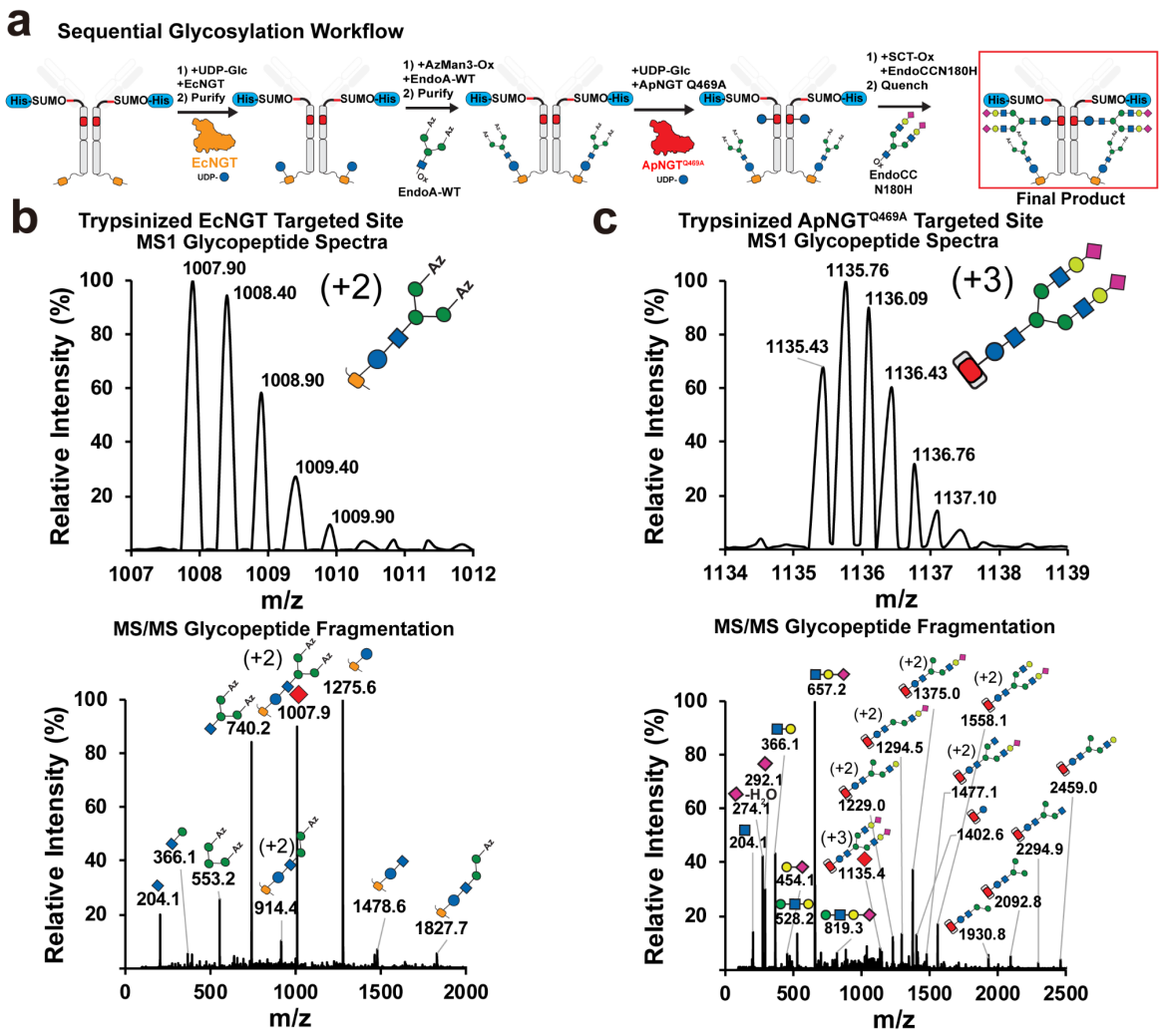
Supplementary Figure 22: Site-specific installation of AzMan3 onto human Fc protein. (a) Workflow for selectively modifying EcNGT-targeted site at the *C*-term of 2gFc with AzMan3 glycan. **(b)** Deconvoluted intact 2gFc MS spectra showing SUMO-2gFc cleaved by Ulp1 and analyzed by LC-qTOF before NGT treatment, after EcNGT treatment, and transglycosylation. 10 μM SUMO-2gFc was monoglycosylated by treating with 1 μM EcNGT for 4 h at 30°C before purification with Qiagen Ni-NTA resin. 355 μg (0.009 μmoles) of monoglucosylated SUMO-2gFc and 60 molar equivalents of AzMan3-ox (380 μg, 0.514 μmoles) were incubated with 70 μg of EndoA WT at a final enzyme concentration of 0.2 μg/μl in 50 mM Tris-HCl (pH=7.5) and 150 mM NaCl. The reaction was carried out at 30 ºC. An additional 60 molar equivalents of AzMan3-ox were added at 30 and 60 minutes. Reaction was quenched by adding 0.1 M glycine buffer (pH=2.7) at 90 minutes. Samples were analyzed by LC-qTOF (see **Methods**). Spectra of aglycosylated and monoglucosylated 2gFc are also shown in **Supplementary Fig. 21**. Transglycosylation reactions were performed once (n=1) with a yield of approximately 47% based on relative intensities of substrate and product peaks in deconvoluted protein spectra. **(c)** MS1 and MS/MS spectra of EcNGT-targeted glycopeptide after transglycosylation reaction and trypsinization. MS1 spectra shows dominant +2 charge state of ATTMGN(AzMan3)ISR glycopeptide (theoretical monoisotopic m/z = 1007.92). MS/MS spectra show glycan and glycopeptide fragments confirming glycopeptide identity. Fragments are +1 charged unless otherwise noted.



Supplementary Figure 23: Site-specific installation of SCT onto human Fc protein. (a) Workflow for selectively modifying EcNGT-targeted site at the *C*-term of 2gFc with SCT glycan. **(b)** Deconvoluted intact 2gFc MS spectra showing SUMO-2gFc cleaved by Ulp1 and analyzed by LC-qTOF before NGT treatment, after EcNGT treatment, and transglycosylation. 10 μM SUMO-2gFc was monoglycosylated by treating with 1 μM EcNGT for 4 h at 30°C before purification with Qiagen Ni-NTA resin. 30 μg of monoglucosylated SUMO-2gFc and 120 molar equivalents of SCT-ox (150 μg, 0.075 μmoles) and 0.5 μg of EndoCC^{N180H} at a final enzyme concentration of 0.02 μg/μl in 50 mM Tris-HCl (pH=7.5) and 150 mM NaCl, at 30 ºC for 30 minutes. Samples were analyzed by LC-qTOF (see **Methods**). Spectra of aglycosylated and monoglucosylated 2gFc are also shown in **Supplementary Fig. 21**. Transglycosylation reactions were performed once (n=1) with a yield of approximately 48% based on relative intensities of substrate and product peaks in deconvoluted protein spectra. **(c)** MS1 and MS/MS spectra of EcNGT-targeted glycopeptide after transglycosylation reaction and trypsinization. MS1 spectra shows dominant +3 charge state of ATTMGN(SCT)ISR glycopeptide (theoretical monoisotopic m/z = 1093.06). MS/MS spectra show glycan and glycopeptide fragments confirming MS1 glycopeptide identity. Fragments are labeled and +1 charged unless otherwise noted.



Supplementary Figure 24: Sequential installation of AzMan3 and sialyllactose onto human Fc protein. (a) Workflow for sequential modification of Fc protein with AzMan3 at EcNGT-targeted site at the *C*-term of 2gFc followed by sequential elaboration of an engineered ApNGT^{Q469A}-targeted sequence at the naturally occurring Asn297 glycosylation position to Siaα2-6Galβ1-4Glc-Asn. **(b)** MS1 and MS/MS spectra of EcNGT-targeted, trypsinized glycopeptide in the final product. 100 μg (4.9 μM) of the product from **Supplementary Fig. 22** was incubated with 10 μM ApQ and 2.5 mM UDP-Glc in Buffer 1 containing 1 mM DTT at 30 °C for 16 h. MS1 spectra show dominant +2 charge state of ATTMGN(AzMan3)ISR (theoretical monoisotopic m/z = 1007.92). **(c)** MS1 and MS/MS spectra of ApNGT^{Q469A}-targeted, trypsinized glycopeptide in the final product. SUMO-2gFc product from **b** was combined with 15 μM purified NmLgtBTrunc and 5 μM purified PdST6mut as well as 2.5 mM UDP-Gal and 2.5 mM CMP-Sia in Buffer 1 supplemented with 10 mM $MnCl_2$ and 23 mM HEPES buffer (pH = 7.5). This reaction was incubated for 16 h at 30 °C. Spectra and previously documented specificities of NmLgtB and PdST6 suggest the installation of the glycan Siaα2-6Galβ1-4Glc-Asn at the ApNGT^{Q469A}-targeted site. MS1 spectra show dominant +2 charge state of ATTWDYN(-GlcGalSia)LTR (theoretical monoisotopic m/z = 782.64 for -GlcGal and 928.39 for -GlcGalSia). MS/MS spectra show glycan, peptide and glycopeptide fragments confirming glycopeptide identity. Fragments are +1 charged unless otherwise noted.



Supplementary Figure 25: Sequential installation of AzMan3 and SCT onto human Fc protein. (a) Workflow for sequential modification of Fc protein with AzMan3 at EcNGT-targeted site at the *C*-term of 2gFc followed by installation of SCT glycan at an engineered ApNGT^{Q469A}-targeted sequence at the naturally occurring Asn297 glycosylation. **(b)** MS1 and MS/MS spectra of EcNGT-targeted, trypsinized glycopeptide in the final product. 100 μg (4.9 μM) of the product from **Supplementary Fig. 22** was exchanged into Buffer 1 and then incubated with 10 μM ApQ and 2.5 mM UDP-Glc in Buffer 1 containing 1 mM DTT at 30 °C for 16 h. 70 μg of the SUMO-2gFc product from this glucosylation reaction was incubated with 120 molar equivalents of SCT-ox (340 μg, 0.170 μmoles) and 1.2 μg of EndoCC^{N180H} at a final enzyme concentration of 0.02 μg/μl in Buffer 1 at 30 °C for 30 minutes. MS1 spectrum shows dominant +2 charge state of ATTMGN(AzMan3)ISR (theoretical monoisotopic m/z = 1007.92). **(c)** MS1 and MS/MS spectra of ApNGT^{Q469A}-targeted, trypsinized glycopeptide in the final product. MS1 spectra show dominant +3 charge state of ATTWDYN(SCT)LTR (theoretical monoisotopic m/z = 1135.44). MS/MS spectra show glycan and glycopeptide fragments confirming MS1 glycopeptide identity. Fragments are labeled and +1 charged unless otherwise noted.

Supplementary Note 1: Example DNA sequences encoding engineered Im7 protein glycosylation targets, active NGTs expressed by CFPS, as well as NGTs and Endoglycosidases expressed and purified from *E. coli*.

Key:
TRANSLATED REGION
untranslated region
Promoter
Terminator
Sequences Flanking Glycosylation Sites
Engineered Glycosylation Sequence
PROTEIN REGION REMOVED AFTER ULP1 CLEAVAGE

DNA sequence for HiNGT in pJL1 plasmid context:
gaaattaatacgactcactataggagaccacaacggtttccctctagaaataattttgtttaactttaagaaggagatatacat
ATGACCAAAGAGAACCTGCAGAGCGTTCCGCAGAATACCACCGCAAGCCTGGTTGAAAGC
AATAATGATCAGACCAGCCTGCAGATTCTGAAACAGCCTCCGAAACCGAATCTGCTGCGTCT
GGAACAGCATGTTGCAAAAAAAGATTATGAACTGGCATGCCGTGAACTGATGGCAATTCTGG
AAAAAATGGATGCCAATTTTGGTGGCGTGCACGATATTGAATTTGATGCACCGGCACAGCTG
GCATATCTGCCGGAAAAACTGCTGATTCATTTTGCAACCCGTCTGGCAAATGCAATTACCAC
CCTGTTTAGCGATCCGGAACTGGCAATTAGCGAAGAGGGTGCACTGAAAATGATTAGCCTG
CAGCGTTGGCTGACCCTGATTTTTGCAAGCAGCCCGTATGTTAATGCAGATCATATCCTGAA
CAAATACAACATCAATCCGGATAGCGAAGGTGGTTTTCATCTGGCAACCGATAATAGCAGCA
TTGCCAAATTTTGCATCTTTTATCTGCCTGAGAGCAACGTTAATATGAGCCTGGATGCACTGT
GGGCAGGTAATCAGCAGCTGTGTGCAAGCCTGTGTTTTGCACTGCAGAGCAGCCGTTTTAT
TGGCACCGCCAGCGCATTTCATAAACGTGCAGTTGTTCTGCAGTGGTTTCCGAAAAAACTG
GCAGAAATTGCAAATCTGGATGAACTGCCTGCAAACATTCTGCATGATGTTTATATGCATTGC
AGCTATGATCTGGCCAAAAACAAACATGATGTTAAACGTCCGCTGAATGAACTGGTGCGTAA
ACATATTCTGACCCAGGGTTGGCAGGATCGTTATCTGTATACCCTGGGTAAAAAAGATGGTA
AACCGGTTATGATGGTTCTGCTGGAACATTTTAACAGCGGTCATAGCATTTATCGTACCCATA
GCACCAGCATGATTGCAGCACGTGAAAAATTCTATCTGGTTGGTCTGGGTCATGAAGGCGT
TGATAATATTGGTCGTGAAGTGTTTGATGAGTTCTTTGAGATTAGCAGCAACAACATTATGGA
ACGCCTGTTTTTTATCCGCAAACAGTGTGAAACCTTTCAGCCTGCCGTGTTTTATATGCCGA
GCATTGGTATGGATATCACCACCATTTTTGTTAGCAATACACGTCTGGCACCGATTCAGGCA
GTTGCACTGGGTCATCCGGCAACCACCCATAGCGAATTTATTGATTATGTGATCGTCGAGGA
TGATTACGTTGGTAGCGAAGATTGTTTTAGCGAAACCCTGCTGCGCCTGCCGAAAGATGCA
CTGCCGTATGTTCCGAGCGCACTGGCACCGCAGAAAGTGGATTATGTTCTGCGTGAAAATC
CGGAAGTTGTGAATATTGGTATTGCAGCAACCACCATGAAACTGAATCCGGAATTTCTGCTG
ACACTGCAAGAAATTCGTGATAAAGCCAAAGTGAAAATTCACTTCCATTTTGCGCTGGGTCA
GTCAACCGGTCTGACCCATCCGTATGTGAAATGGTTTATTGAAAGCTATCTGGGTGATGATG

CAACCGCACATCCGCATGCACCGTATCATGATTATCTGGCCATTCTGCGCGATTGTGATATG
CTGCTGAATCCGTTTCCGTTTGGTAATACCAATGGCATTATTGATATGGTTACCCTGGGCCTG
GTTGGTGTTTGTAAAACCGGTGATGAAGTGCATGAACATATTGATGAAGGTCTGTTTAAACG
TCTGGGCCTGCCGGAATGGCTGATTGCAGATACCCGTGAAACCTATATTGAATGTGCACTGC
GCCTGGCGGAAAATCATCAAGAACGCCTGGAACTGCGTCGTTATATCATTGAAAATAATGGC
CTGCAGAAACTGTTTACCGGTGATCCGCGTCCGCTGGGCAAAATTCTGCTGAAAAAAACCA
ATGAGTGGAAACGCAAACATCTGAGCAAAAAAGGATCCTGGAGCCATCCGCAGTTCGAAAA
ATAAgtcgaccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataactagcataaccccttg
gggcctctaaacgggtcttgaggggttttttgctgaaag

DNA sequence for EcNGT in pJL1 plasmid context:

gaaattaatacgactcactataggagaccacaacggtttccctctagaaataattttgtttaactttaagaaggagatatacat
ATGATGAGCCACAAAACCGATACCGCACCGGTTCAAGAACAGGCAGGCCTGACCTTTCGTC
TGGAAACCTTTGAATGGCAGGTTCATCAGGGTCTGAATGAAGAAGCAGCACGTAGCCTGAT
TAGCCTGCTGCAGCTGCTGGATCGTCATTATGCACAGTGGGGTGAAAGCTTTAGCGCATGG
GCACCGGGTATGACCGCAGAAGAAATTAATCCGCATCTGTGTACCCGTATTGCCGGTGCAAT
TACCGCACTGTTTAGCCGTCCGGGTTTTCGTGTTAGTGATGGTGGTTTTGCAGAACTGATG
GATTATCATCGTTGGCTGGCAATTATCTTTGCCGTTAGCGATTATCGTCATGGCGATCATATTA
TTCGCAATATTAACGCAGCCGGTGGTGGTGTTGTTGCACCGCTGACCCTGAATGCCGATAAT
CTGCAGCTGTTTTGTCTGAGCTATTATCCGGATTCACAGATTGCACTGCAGCCGGAACCGCT
GTGGCAGTATGATCGTCAGACCGTTGTTCGTCTGTTTTTTGCACTGCTGAGCGGTCGTGCA
CTGCCGACACCGGCAGCACATCAGAAACGTGAACATCTGCTGGCATGGCTGCCGGAACGT
CTGAAAGAAATTGATAGCCTGGAATTTCTGCCTGGTAAAGTTCTGCATGATGTGTATATGCAT
TGTAGCTATGCCGATCTGCCGGAAAAACATCGTATTAAACAAGAAATCAATCGTCTGACCGC
ACGTGCACTGGAACAGACCTATGCAGATTGTCTGCCGGTTCGTGCACCGGAAGCAGCCCG
TCAGAAACCGGTTCTGGCAGTTGTTCTGGAATGGTTTACCTGTCAGCATAGCATTTATCGTA
CCCATAGCACCAGCATGCGTGCCCTGCGTGAACATTTTCATCTGCTGGGTATTGCACAGCC
TGGTGCAACCGATGAAATTACCCGTGAAGTTTTTGATGAATTTCGTGAACTGAGCGCAGAAA
ATGTTGTTGGTGATGCAATTCGTTGCCTGAGCGAAGTTCGTCCGGATGTTATCTATTATCCTA
GCGTTGGTATGTTTCCGCTGACCGTTTATCTGACAGCACTGCGTCTGGCACCACTGCAGCT
GATGGCACTGGGTCATCCGGCAACCACCTGGTCAGAACATATTGATGGTGTTCTGGTTGAA
GAAGATTATCTGGGCGATCCGGCATGTTTTAGCGAAACCGTTTGTGCAGTTCCGAAAGATGC
CATTCCGTATATTCCGCCTGCAAGCACCGAACGTGTTCTGCCTGAACGTACCCCGTTTCGT
GATCGTGCAAAAGCAGCATGGCCTGCAGCACTGCCGGTGCGTGTTGCAGTTTGTGCAAGC
GTTATGAAAATTAACCCTGGTTTTCTGGATACCCTGCGCGAAATTAGCGATCGTAGCCGTGTT
CCGGTTCAGTTTTGTTTTTGGATGGGTTTTGCCCAGGGCCTGACCCTGGATTATCTGCGTC
GTGCCATTCGTCAGGCGCTGCCGACAGCCGAAGTTAATGCACACATGCCGGTTCAGGCATA
TCAGCAGGCACTGAATAGCTGTGAACTGTTTGTTAATCCGTTTCCGTTTGGTAATACCAATG
GTCTGGTTGATACCGTTCGCCAGGGTCTGCCAGGTGTTTGTATGACCGGTCCGGAAGTTCA
TACCCATATCGATGAAGGTCTGTTTCGTCGTCTGGGCCTGCCGGAAGCACTGATTGCACGT

GATCGCGAAGAATATATCACCGCAGTTCTGAGCCTGACCGAAACACCGCGTCTGCGCGAAC
GTCTGCAGAAATATCTGACGGAAAATGATGTTGAGAAAGTGCTGTTTGAAGGCCGTCCGGA
TAAATTTGCGGAACGTGTTTGGCAGCTGTGGGAAGCCCGTAGCCATCGTCAAGAAGAAGGC
GCAGAAGGATCCTGGAGCCATCCGCAGTTCGAAAAATAAgtcgaccggctgctaacaaagcccgaaagg
aagctgagttggctgctgccaccgctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaag

DNA sequence for ApNGT (Q469A) in pJL1 plasmid context:

gaaattaatacgactcactatagggagaccacaacggtttccctctagaaataattttgtttaactttaagaaggagatatacat
ATGGAAAACGAGAATAAACCGAACGTGGCAAATTTTGAAGCAGCAGTTGCAGCCAAAGATTA
TGAAAAAGCATGTAGCGAGCTGCTGCTGATTCTGAGCCAGCTGGATAGCAATTTTGGTGGC
ATTCATGAAATCGAGTTCGAGTACCCAGCTCAGCTGCAGGATCTGGAGCAGGAAAAAATTG
TGTACTTTTGCACCCGTATGGCGACTGCCATCACCACCCTGTTCTCTGACCCGGTTCTGGA
AATCTCCGACCTGGGTGTGCAGCGTTTCCTGGTTTATCAGCGTTGGCTGGCGCTGATCTTC
GCTAGCAGCCCGTTCGTGAACGCAGACCACATCCTGCAGACTTACAACCGTGAACCGAAC
CGCAAAAACTCCCTGGAAATTCACCTGGACTCTAGCAAGTCCTCTCTGATTAAATTTTGCAT
CCTGTACCTGCCGGAATCTAATGTGAACCTGAATCTGGATGTGATGTGGAACATTTCCCCGG
AGCTGTGTGCCTCTCTGTGCTTTGCGCTGCAAAGCCCGCGTTTTGTTGGCACCAGCACCG
CCTTTAACAAACGCGCGACCATTCTGCAGTGGTTCCCGCGTCATCTGGACCAGCTGAAAAA
CCTGAACAACATCCCGTCCGCTATCTCTCATGACGTGTATATGCACTGTTCTTACGACACCA
GCGTTAACAAGCACGATGTTAAGCGCGCGCTGAATCACGTGATTCGTCGCCACATCGAATC
CGAATACGGTTGGAAAGATCGTGATGTGGCTCACATCGGTTATCGCAACAACAAACCGGTTA
TGGTCGTTCTGCTGGAACATTTTCATAGCGCGCACTCTATCTACCGTACTCACTCTACCAGC
ATGATCGCCGCGCGCGAACACTTCTATCTGATCGGCCTGGGTTCCCCGAGCGTTGACCAG
GCCGGTCAGGAGGTTTTCGATGAATTCCACCTGGTAGCGGGTGACAACATGAAGCAAAAAC
TGGAATTCATTCGTTCTGTGTGCGAAAGCAATGGTGCGGCAATTTTCTACATGCCGAGCATC
GGTATGGATATGACCACCATCTTCGCGTCCAATACCCGTCTGGCGCCGATTCAGGCAATCG
CCCTGGGCCACCCGGCGACTACTCACTCCGACTTCATTGAATACGTTATCGTGGAAGACGA
CTACGTCGGCTCTGAGGAATGCTTCTCTGAAACCCTGCTGCGTCTGCCGAAAGACGCTCTG
CCGTATGTTCCGTCCGCCCTGGCTCCGGAGAAAGTTGATTACCTGCTGCGTGAAAACCCTG
AAGTTGTCAACATCGGTATTGCCTCTACCACTATGAAGCTGAACCCGTACTTCCTGGAAGCA
CTGAAGGCCATTCGTGACCGTGCGAAGGTGAAAGTGCACTTCCACTTCGCACTGGGCGCG
TCCAATGGTATCACTCACCCTTACGTTGAACGCTTTATCAAATCTTACCTGGGCGACAGCGC
TACCGCGCACCCGCACTCTCCGTACCACCAGTACCTGCGTATTCTGCACAACTGCGATATG
ATGGTAAACCCTTTTCCGTTTGGTAATACCAATGGTATTATTGACATGGTAACCCTGGGTCTG
GTAGGTGTTTGCAAAACCGGTGCGGAAGTCCACGAACATATCGATGAAGGCCTGTTCAAAC
GTCTGGGCCTGCCGGAATGGCTGATTGCAAACACCGTGGACGAATACGTGGAACGTGCAG
TGCGCCTGGCCGAGAACCATCAGGAACGTCTGGAACTGCGTCGTTACATTATTGAAAACAA
TGGCCTGAACACCCTGTTCACCGGCGACCCACGCCCGATGGGTCAGGTGTTCCTGGAAAA
ACTGAACGCATTCCTGAAGGAAAACTAAgtcgaccggctgctaacaaagcccgaaaggaagctgagttggctgct
gccaccgctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaag

DNA sequence for ApNGT in pJL1 plasmid context:

gaaattaatacgactcactatagggagaccacaacggtttccctctagaaataattttgtttaactttaagaaggagatatacat
ATGGAAAACGAGAATAAACCGAACGTGGCAAATTTTGAAGCAGCAGTTGCAGCCAAAGATTA
TGAAAAAGCATGTAGCGAGCTGCTGCTGATTCTGAGCCAGCTGGATAGCAATTTTGGTGGC
ATTCATGAAATCGAGTTCGAGTACCCAGCTCAGCTGCAGGATCTGGAGCAGGAAAAAATTG
TGTACTTTTGCACCCGTATGGCGACTGCCATCACCACCCTGTTCTCTGACCCGGTTCTGGA
AATCTCCGACCTGGGTGTGCAGCGTTTCCTGGTTTATCAGCGTTGGCTGGCGCTGATCTTC
GCTAGCAGCCCGTTCGTGAACGCAGACCACATCCTGCAGACTTACAACCGTGAACCGAAC
CGCAAAAACTCCCTGGAAATTCACCTGGACTCTAGCAAGTCCTCTCTGATTAAATTTTGCAT
CCTGTACCTGCCGGAATCTAATGTGAACCTGAATCTGGATGTGATGTGGAACATTTCCCCGG
AGCTGTGTGCCTCTCTGTGCTTTGCGCTGCAAAGCCCGCGTTTTGTTGGCACCAGCACCG
CCTTTAACAAACGCGCGACCATTCTGCAGTGGTTCCCGCGTCATCTGGACCAGCTGAAAAA
CCTGAACAACATCCCGTCCGCTATCTCTCATGACGTGTATATGCACTGTTCTTACGACACCA
GCGTTAACAAGCACGATGTTAAGCGCGCGCTGAATCACGTGATTCGTCGCCACATCGAATC
CGAATACGGTTGGAAAGATCGTGATGTGGCTCACATCGGTTATCGCAACAACAAACCGGTTA
TGGTCGTTCTGCTGGAACATTTTCATAGCGCGCACTCTATCTACCGTACTCACTCTACCAGC
ATGATCGCCGCGCGCGAACACTTCTATCTGATCGGCCTGGGTTCCCCGAGCGTTGACCAG
GCCGGTCAGGAGGTTTTCGATGAATTCCACCTGGTAGCGGGTGACAACATGAAGCAAAAAC
TGGAATTCATTCGTTCTGTGTGCGAAAGCAATGGTGCGGCAATTTTCTACATGCCGAGCATC
GGTATGGATATGACCACCATCTTCGCGTCCAATACCCGTCTGGCGCCGATTCAGGCAATCG
CCCTGGGCCACCCGGCGACTACTCACTCCGACTTCATTGAATACGTTATCGTGGAAGACGA
CTACGTCGGCTCTGAGGAATGCTTCTCTGAAACCCTGCTGCGTCTGCCGAAAGACGCTCTG
CCGTATGTTCCGTCCGCCCTGGCTCCGGAGAAAGTTGATTACCTGCTGCGTGAAAACCCTG
AAGTTGTCAACATCGGTATTGCCTCTACCACTATGAAGCTGAACCCGTACTTCCTGGAAGCA
CTGAAGGCCATTCGTGACCGTGCGAAGGTGAAAGTGCACTTCCACTTCGCACTGGGCCAG
TCCAATGGTATCACTCACCCTTACGTTGAACGCTTTATCAAATCTTACCTGGGCGACAGCGC
TACCGCGCACCCGCACTCTCCGTACCACCAGTACCTGCGTATTCTGCACAACTGCGATATG
ATGGTAAACCCTTTTCCGTTTGGTAATACCAATGGTATTATTGACATGGTAACCCTGGGTCTG
GTAGGTGTTTGCAAAACCGGTGCGGAAGTCCACGAACATATCGATGAAGGCCTGTTCAAAC
GTCTGGGCCTGCCGGAATGGCTGATTGCAAACACCGTGGACGAATACGTGGAACGTGCAG
TGCGCCTGGCCGAGAACCATCAGGAACGTCTGGAACTGCGTCGTTACATTATTGAAAACAA
TGGCCTGAACACCCTGTTCACCGGCGACCCACGCCCGATGGGTCAGGTGTTCCTGGAAAA
ACTGAACGCATTCCTGAAGGAAAACTAAgtcgaccggctgctaacaaagcccgaaaggaagctgagttggctgct
gccaccgctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaag

DNA sequence for YpNGT in pJL1 plasmid context:

gaaattaatacgactcactatagggagaccacaacggtttccctctagaaataattttgtttaactttaagaaggagatatacat
ATGGCCGATAAAAGCGTTGAACTGACACCGGTTGTTGAAGCACCGGTGGTTTTTAGCCTGC

CGTATTTTGAATTTCTGGTTTGTACCCGTCGTTATGAAGATGCAGGTCGTCTGCTGATTCTGA
TGCTGGAAAAACTGGATACCCAGTATGGTCGTTGGGATGTGTTTAGCCTGAATAAACAGCCG
ATTCAGCAGCAAGAGTATTATTGTAATCGTCTGGCAGCAGCAATTGGTTGTCTGTTTAGCGAT
CCGGGTTTTGTTATTAGCGAAACCGGCTTTCTGCAGCTGATTAACTTTCATCGTTGGATTGC
ACTGATTTTTGCAGCAAGCACCTTTGGTCATGCAGATCATGTTATTACCAATCTGAATGAAGC
CGGTAATGGTTGTAGCCATCCGCTGCGTTTTGAACGTAATAACTTTCTGAAATTTTGCGTGAT
GTATCTGCCGGAAAGCGGTATTCCGCTGCAGCCGGATATTCTGTGGCAGTTTAATCCGCAG
GCAACCGCAGCACTGTTTCTGGCACTGCTGAGTCCGCGTATTCTGCCGAGCGCAGCAGGT
CATGAAAAACGTGAAACCCTGCTGGCATGGCTGCCTGAAAAACTGCTGACCCTGATTAGCC
TGGAAGGCCTGCCGGAACGCATTCTGCATGATGTTTATATGCATTGTAGCTATGCCGATATG
GCCAAAAAACATACCATTAAACGCAGCATCAATTTTCATCTGCGTAAAACCATGCTGAAAAAT
GGTCTGAGCGATATGAATGAACTGCCTCCGCTGCGCAGTAAACCGCTGATGCTGGTTATTCT
GGAATGGTTTAATAGCGGTCATAGCATTTATCGTACCCATAGCAGCACCCTGCGTGCAGCAC
GTGATCAGTTTAGCACACATGGTGTTGCAATTGCCGAAGCAACCGATGATATTACCCGTAAA
GTGTTTGATGATTTTACCGAAGTTAGCCGTACCGGTGCAGTTGAAACCATTATGGCACTGGC
ACAGCAGCTGCGTCCGGATGTTATCTATTTTCCGAGCGTGGGTATGTTTCCGATGACCGTTG
CACTGACCAATCTGCGTCTGGCACCGCTGCAGGTTATGGCCCTGGGTCATCCGGCAACCA
CCCATAGCGATTATATTGATGCAGTTCTGGTGGAAGAAGATTATCTGGGCGATATTGCATGCT
TTAGCGAAAAAGTTGTGAGCCTGCCGAAAGATTGTCTGCCGTATGTTCCGCCTGCAAATATT
ACCCAGCCGGAACCGATCCAGCAGTTTGTGCAGCGTGAAGCCGTTCATATTGCCGTTTGTG
CAAGCGCAATGAAAATCAATCCGCGTTTTCTGGCAGCCTGTGCAGAAATTGCACTGCGTAG
TCCGCTGCCGATTATCTTTCATTTTCTGGTTGGTTTTTGCTGGGGTATTACCCATCGTGTTAT
GGAAAAAGCCGTTAATGAAATGGTTACCAGCGCCAAAGTTTATGAGCATCTGAACTATCAGA
ACTATCTGCAGGTCATTAATCAGTGCGACCTGTTTATTAACCCGTTTCCGTTTGGTAATACCA
ACGGTATTGTTGATACCGTTCGTCAGGGTCTGCCTGGTGTTTGTCTGAGCGGTGAAGAAGT
TCACGAACATATTGATGAAGGTCTGTTTCGTCGTCTGGGTCTGGCAGAAGAACTGATTACCC
ATAATGTTGAACAGTATATCACCGCAACCGTGCGTCTGATTACCGATACCAATTGGCGTAATG
GTCTGCGTCGTCAGCTGCTGCAGACCCAGCCTGATAATGTTCTGTTTACCGGCAAACCGGA
ACAGTTTGGTCAGATTGTTCGTGCCCTGCTGGATAATGGTCATCAGGATGTTAATGGATCCT
GGAGCCATCCGCAGTTCGAAAAATAAgtcgaccggctgctaacaaagcccgaaaggaagctgagttggctgctgc
caccgctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaag

DNA sequence for 4glm7 in pET.BCS.NS plasmid context:

gaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagagcagaattcggtagatctatataacaca
cataaggaggacatATGCGCGCGACTACC**CTGAATGAAAACGTGACC**CGCGCGGGAGGAGAAC
TGGAAAATAGTATTAGTGATTACACAGAGGCTGAGTTTGTTCAACTTCTTAAGGAAATTGAAA
AAGAGCGCGCGACTACC**TGGTACGCCAACGTCACG**CGCGCGGGAGGAGTGTTAGATGTG
TTACTCGAACACTTTGTAAAAATTACTGAGCATCCAGATGGAACGGATCTGATCTATTATCCT
CGCGCGACTACC**TGGGACTATAATTTAACA**CGCGCGGGAGGACCCGAAGGGATTGTCAAG
GAAATTAAAGAATGGCGAGCTGCTAACGGTAAGCCAGGATTTAAACAGGGCCGCGCGACTA

CC**TACATGGGGAATATTTCG**CGCGCGGGAGGAGGATCCCATCACCATCATCACCATTAAgtcg
acgatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataactagcataaccccttggggcc
tctaaacgggtcttgaggggttttttgctgaaag

DNA sequence for GeneralistP1 in pET.BCS.NS plasmid context:
gaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagagcagaattcggtagatctatataacaca
cataaggaggacatATGCGCGCGACTACC**ATCTACGCGAATGTGACACTT**GCGGGAGGACGCGA
ACTGGAAAATAGTATTAGTGATTACACAGAGGCTGAGTTTGTTCAACTTCTTAAGGAAATTGA
AAAAGAGAATGTTGCTGCAACTGATGATGTGTTAGATGTGTTACTCGAACACTTTGTAAAAAT
TACTGAGCATCCAGATGGAACGGATCTGATCTATTATCCTAGTGATAATAGAGACGATAGCCC
CGAAGGGATTGTCAAGGAAATTAAAGAATGGCGAGCTGCTAACGGTAAGCCAGGATTTAAA
CAGGGCGGATCCCATCACCATCATCACCATTAAAgtcgacgatccggctgctaacaaagcccgaaaggaagc
tgagttggctgctgccaccgctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaag

DNA sequence for GeneralistP2 in pET.BCS.NS plasmid context:
gaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagagcagaattcggtagatctatataacaca
cataaggaggacatATGGAACTGGAAAATAGTATTAGTGATTACACAGAGGCTGAGTTTGTTCAAC
TTCTTAAGGAAATTGAAAAAGAGCGCGCGACTACC**ATCTACGCGAATGTGACACTT**GCGGG
AGGACGCGTGTTAGATGTGTTACTCGAACACTTTGTAAAAATTACTGAGCATCCAGATGGAA
CGGATCTGATCTATTATCCTAGTGATAATAGAGACGATAGCCCCGAAGGGATTGTCAAGGAAA
TTAAAGAATGGCGAGCTGCTAACGGTAAGCCAGGATTTAAACAGGGCGGATCCCATCACCA
TCATCACCATTAAgtcgacgatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaata
actagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaag

DNA sequence for GeneralistP3 in pET.BCS.NS plasmid context:
gaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagagcagaattcggtagatctatataacaca
cataaggaggacatATGGAACTGGAAAATAGTATTAGTGATTACACAGAGGCTGAGTTTGTTCAAC
TTCTTAAGGAAATTGAAAAAGAGAATGTTGCTGCAACTGATGATGTGTTAGATGTGTTACTCG
AACACTTTGTAAAAATTACTGAGCATCCAGATGGAACGGATCTGATCTATTATCCTCGCGCGA
CTACC**ATCTACGCGAATGTGACACTT**GCGGGAGGACGCCCCGAAGGGATTGTCAAGGAAA
TTAAAGAATGGCGAGCTGCTAACGGTAAGCCAGGATTTAAACAGGGCGGATCCCATCACCA
TCATCACCATTAAgtcgacgatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaata
actagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaag

DNA sequence for GeneralistP4 in pET.BCS.NS plasmid context:
gaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagagcagaattcggtagatctatataacaca
cataaggaggacatATGGAACTGGAAAATAGTATTAGTGATTACACAGAGGCTGAGTTTGTTCAAC
TTCTTAAGGAAATTGAAAAAGAGAATGTTGCTGCAACTGATGATGTGTTAGATGTGTTACTCG
AACACTTTGTAAAAATTACTGAGCATCCAGATGGAACGGATCTGATCTATTATCCTAGTGATAA
TAGAGACGATAGCCCCGAAGGGATTGTCAAGGAAATTAAAGAATGGCGAGCTGCTAACGGT

AAGCCAGGATTTAAACAGGGCCGCGCGACTACC**ATCTACGCGAATGTGACACTT**GCGGGA
GGACGCGGATCCCATCACCATCATCACCATTAAgtcgacgatccggctgctaacaaagcccgaaaggaagct
gagttggctgctgccaccgctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaag

DNA sequence for SUMO-4glm7 in pET28a plasmid context:

gaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagaaataattttgtttaactttaagaaggagatatac
cat*ATGGAGAAAAAAATCCGTGGTTCTCACCACCACCACCACCATATGGCTAGCGGATCGGA*
CTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCAC
ATCAATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCG
CTGCGTAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACTCCTTAAGAT
TCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGAT
*AACGATATTATTGAGGCTCACAGAGAACAGATTGGTGGA*GGATCCTCTGGAGGTGGATCGG
GATCGGGCGGCGGCAGCGGTGGCGGCTCGGGCGGGATGCGCGCGACTACC**CTGAATGA**
AAACGTGACCCGCGCGGGAGGAGAACTGGAAAATAGTATTAGTGATTACACAGAGGCTGA
GTTTGTTCAACTTCTTAAGGAAATTGAAAAAGAGCGCGCGACTACC**TGGTACGCCAACGTC**
ACGCGCGCGGGAGGAGTGTTAGATGTGTTACTCGAACACTTTGTAAAAATTACTGAGCATCC
AGATGGAACGGATCTGATCTATTATCCTCGCGCGACTACC**TGGGACTATAATTTAACA**CGCG
CGGGAGGACCCGAAGGGATTGTCAAGGAAATTAAAGAATGGCGAGCTGCTAACGGTAAGC
CAGGATTTAAACAGGGCCGCGCGACTACC**TACATGGGGAATATTTCG**CGCGCGGGAGGAT
AAgatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataactagcataaccccttggggc
ctctaaacgggtcttgaggggttttttgctgaaag

DNA sequence for SUMO-2glm7 in pETBCS.NS plasmid context:

gaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagagcagaattcggtagatctatataacaca
cataaggaggacat*ATGGAGAAAAAAATCCGTGGTTCTCACCACCACCACCACCATATGGCTAGC*
GGATCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCT
GAGACTCACATCAATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAG
ACCACTCCGCTGCGTAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACT
*CCTTAAGATTCTTGTACGAC*GGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACA
TGGAGGATAACGATATTATTGAGGCTCACAGAGAACAGATTGGTGGAGGATCCTCTGGAGG
TGGATCGGGATCGGGCGGCGGCAGCGGTGGCGGCTCGGGCGGGATGCGCGCGACTACC
TGGGACTATAATTTAACACGCGCGGGAGGAGAACTGGAAAATAGTATTAGTGATTACACAGA
GGCTGAGTTTGTTCAACTTCTTAAGGAAATTGAAAAAGAGAATGTTGCTGCAACTGATGATG
TGTTAGATGTGTTACTCGAACACTTTGTAAAAATTACTGAGCATCCAGATGGAACGGATCTGA
TCTATTATCCTAGTGATAATAGAGACGATAGCCCCGAAGGGATTGTCAAGGAAATTAAAGAAT
GGCGAGCTGCTAACGGTAAGCCAGGATTTAAACAGGGCCGCGCGACTACC**TACATGGGGA**
ATATTTCGCGCGCGGGAGGATAAgtcgacgatccggctgctaacaaagcccgaaaggaagctgagttggctgctgc
caccgctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaag

DNA sequence for SUMO-2gFc in pETBCS.NS plasmid context:

gaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagagcagaattcggtagatctatataacaca
cataaggaggacat*ATGGAGAAAAAAATCCGTGGTTCTCACCACCACCACCACCATATGGCTAGC*
GGATCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCT
GAGACTCACATCAATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAG
ACCACTCCGCTGCGTAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACT
*CCTTAAGATTCTTGTACGAC*GGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACA
TGGAGGATAACGATATTATTGAGGCTCACAGAGAACAGATTGGTGGAATGGAACCGAAAAG
CTGTGATAAAACCCATACCTGTCCGCCTTGTCCGGCACCGGAACTGCTGGGTGGTCCGAG
CGTTTTTCTGTTTCCGCCTAAACCGAAAGATACCCTGATGATTAGCCGTACACCGGAAGTTA
CCTGTGTTGTTGTTGATGTTAGCCATGAAGATCCGGAAGTGAAATTTAACTGGTATGTTGATG
GTGTGGAAGTGCATAATGCAAAAACCAAACCGCGTGAAGAACGCGCGACTACC**TGGGACT**
ATAATTTAACACGCGCGGGAGGACGTGTTGTTAGCGTTCTGACCGTTCTGCATCAGGATTG
GCTGAATGGTAAAGAATACAAATGCAAAGTGAGCAACAAAGCACTGCCTGCACCGATTGAA
AAAACCATTAGCAAAGCAAAAGGTCAGCCTCGTGAACCGCAGGTTTATACCCTGCCTCCGA
GCCGTGATGAACTGACCAAAAATCAGGTTAGCCTGACCTGTCTGGTGAAAGGTTTTTATCCG
AGCGATATTGCAGTTGAATGGGAAAGCAATGGTCAGCCGGAAAATAACTATAAAACCACCCC
TCCGGTTCTGGATAGTGATGGTAGCTTTTTTCTGTATAGCAAACTGACCGTTGATAAAAGCC
GTTGGCAGCAGGGTAATGTTTTTAGCTGTAGCGTTATGCATGAAGCCCTGCATAATCATTATA
CCCAGAAAAGCCTGAGCCTGAGTCCGGGTAAAGGTAGCCGCGCGACTACC**TACATGGGGA**
ATATTTCGCGCGCGGGAGGTTAAgtcgacgatccggctgctaacaaagcccgaaaggaagctgagttggctgctgc
caccgctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaag

DNA sequence for PdST6Mut in pJL1 plasmid context:

gaaattaatacgactcactatagggagaccacaacggtttccctctagaaataattttgtttaactttaagaaggagatatacatATGG
AGAAAAAAATCCATCACCATCATCACCATGGCGGATCCGGACTGGTTCCGCGTGGTAGCCA
CATGTGTAATAGCGATAACACCAGCCTGAAAGAAACCGTTAGCAGCAATAGCGCAGATGTTG
TTGAAACCGAAACCTATCAGCTGACCCCGATTGATGCACCGAGCAGCTTTCTGAGCCATAG
CTGGGAACAGACCTGTGGCACCCCGATTCTGAATGAAAGCGATAAACAGGCAATCAGCTTT
GATTTTGTTGCACCGGAACTGAAACAGGATGAGAAATATTGCTTTACCTTCAAAGGCATTACC
GGTGATCATCGTTATATTACCAATACCACCCTGACCGTTGTGGCACCGACCCTGGAAGTTTA
TATTGATCATGCAAGCCTGCCGAGCCTGCAGCAGCTGATTCATATTATTCAGGCCAAAGATG
AATATCCGAGCAATCAGCGTTTTGTTAGCTGGAAACGTGTTACCGTTGATGCAGATAATGCC
AACAAACTGAACATTCATACCTATCCGCTGAAAGGCAATAATACCAGTCCGGAAATGGTTGC
AGCAATTGATGAATATGCACAGAGCAAAAATCGCCTGAACATCGAGTTTTATACCAATACATAT
CACGTGTTTAATAACCTGCCTCCGATTATTCAGCCGCTGTATAATAACGAGAAAGTGAAAATT
AGCCACATCAGCCTGTATGATGACGGTAGCTATGAATATGTTAGCCTGTATCAGTGGAAAGAT
ACCCCGAACAAAATTGAAACACTGGAAGGTGAAGTTAGCCTGCTGGCAAATTATCTGGCAG
GCACCTCACCGGATGCTCCGAAAGGTATGGGTAATCGTTATAATTGGCACAAACTGTATGAC

ACCGACTATTACTTTCTGCGCGAAGATTATCTGGATGTTGAAGCAAATCTGCATGATCTGCGT
GATTACCTGGGTAGCAGTGCAAAACAAATGCCGTGGGATGAATTTGCAAAACTGAGCGATA
GCCAGCAGACCCTGTTTCTGGATATTGTTGGTTTTGATAAAGAACAGCTGCAGCAACAGTAT
AGCCAGAGTCCGCTGCCGAATTTTATCTTTACCGGCACCACCACCTGGGCAGGCGGTGAAA
CCAAAGAATATTATGCCCAGCAGCAGGTTAACGTGATTAACAATGCAATTAATGAAACCAGCC
CGTACTATCTGGGTAAAGATTATGACCTGTTTTTCAAAGGTCATCCTGCCGGTGGTGTGATTA
ATGATATTATTCTGGGTAGCTTCCCGGATATGATTAACATTCCGGCAAAAATTAGCTTCGAGG
TTCTGATGATGACCGATATGCTGCCGGATACCGTTGCAGGTATTGCAAGCAGTCTGTATTTC
ACAATTCCGGCAGATAAAGTGAACTTCATTGTTTTTACCAGCAGCGATACCATTACCGATCGT
GAAGAAGCACTGAAAAGTCCGCTGGTTCAGGTTATGCTGACCCTGGGTATTGTTAAAGAAA
AAGATGTTCTGTTTTGGGCATAAgtcgaccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccg
ctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaag

DNA sequence for NmLgtBTrunc in pET.BCS.NS plasmid context:

gaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagagcagaattcggtagatctatataacaca
cataaggaggacatATGGAGAAAAAAATCTCTGCGTGGAGCCATCCGCAGTTCGAAAAAGGTGGT
GGTTCTGGTGGTGGTTCTGGTGGTTCTGCGTGGAGCCATCCGCAGTTCGAAAAAGGATCC
CAGAACCATGTTATTAGCCTGGCAAGCGCAGCAGAACGTCGTGCACATATTGCAGATACCTT
TGGTCGTCATGGTATTCCGTTTCAGTTTTTTGATGCACTGATGCCGAGCGAACGTCTGGAAC
AGGCAATGGCAGAACTGGTTCCGGGTCTGAGCGCACATCCGTATCTGAGCGGTGTTGAAA
AAGCATGTTTTATGAGCCATGCAGTTCTGTGGAAACAGGCACTGGATGAAGGTCTGCCGTAT
ATTACCGTTTTTGAAGATGATGTTCTGCTGGGTGAAGGTGCAGAAAAATTTCTGGCAGAAGA
TGCCTGGCTGCAAGAACGTTTTGATCCGGATACCGCATTTATTGTTCGTCTGGAAACCATGT
TTATGCATGTTCTGACCAGCCCGAGCGGTGTGGCAGATTATTGTGGTCGTGCATTTCCGCT
GCTGGAAAGCGAACATTGGGGCACCGCAGGTTATATCATTAGCCGTAAAGCAATGCGCTTTT
TTCTGGATCGTTTTGCAGCACTGCCTCCGGAAGGCCTGCATCCGGTTGATCTGATGATGTTT
AGCGATTTTTTTGATCGTGAAGGTATGCCGGTTTGTCAGCTGAATCCGGCACTGTGTGCACA
AGAACTGCACTATGCAAAATTTCATGATCAGAATAGCGCACTGGGTAGCCTGATTGAACATG
ATCGTCTGCTGAATCGTAAACAGCAGCGTCGTGATAGTCCGGCAAATACCTTTAAACATCGT
CTGATTCGTGCCCTGACCAAAATTAGCCGTGAAgtcgacgatccggctgctaacaaagcccgaaaggaagct
gagttggctgctgccaccgctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaag

DNA sequence for HiNGT in pET.BCS.NS plasmid context:

gaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagagcagaattcggtagatctatataacaca
cataaggaggacatATGACCAAAGAGAACCTGCAGAGCGTTCCGCAGAATACCACCGCAAGCCT
GGTTGAAAGCAATAATGATCAGACCAGCCTGCAGATTCTGAAACAGCCTCCGAAACCGAAT
CTGCTGCGTCTGGAACAGCATGTTGCAAAAAAAGATTATGAACTGGCATGCCGTGAACTGAT
GGCAATTCTGGAAAAAATGGATGCCAATTTTGGTGGCGTGCACGATATTGAATTTGATGCAC
CGGCACAGCTGGCATATCTGCCGGAAAAACTGCTGATTCATTTTGCAACCCGTCTGGCAAA
TGCAATTACCACCCTGTTTAGCGATCCGGAACTGGCAATTAGCGAAGAGGGTGCACTGAAA

ATGATTAGCCTGCAGCGTTGGCTGACCCTGATTTTTGCAAGCAGCCCGTATGTTAATGCAGA
TCATATCCTGAACAAATACAACATCAATCCGGATAGCGAAGGTGGTTTTCATCTGGCAACCG
ATAATAGCAGCATTGCCAAATTTTGCATCTTTTATCTGCCTGAGAGCAACGTTAATATGAGCCT
GGATGCACTGTGGGCAGGTAATCAGCAGCTGTGTGCAAGCCTGTGTTTTGCACTGCAGAG
CAGCCGTTTTATTGGCACCGCCAGCGCATTTCATAAACGTGCAGTTGTTCTGCAGTGGTTTC
CGAAAAAACTGGCAGAAATTGCAAATCTGGATGAACTGCCTGCAAACATTCTGCATGATGTT
TATATGCATTGCAGCTATGATCTGGCCAAAAACAAACATGATGTTAAACGTCCGCTGAATGAA
CTGGTGCGTAAACATATTCTGACCCAGGGTTGGCAGGATCGTTATCTGTATACCCTGGGTAA
AAAAGATGGTAAACCGGTTATGATGGTTCTGCTGGAACATTTTAACAGCGGTCATAGCATTTA
TCGTACCCATAGCACCAGCATGATTGCAGCACGTGAAAAATTCTATCTGGTTGGTCTGGGTC
ATGAAGGCGTTGATAATATTGGTCGTGAAGTGTTTGATGAGTTCTTTGAGATTAGCAGCAACA
ACATTATGGAACGCCTGTTTTTTATCCGCAAACAGTGTGAAACCTTTCAGCCTGCCGTGTTT
TATATGCCGAGCATTGGTATGGATATCACCACCATTTTTGTTAGCAATACACGTCTGGCACCG
ATTCAGGCAGTTGCACTGGGTCATCCGGCAACCACCCATAGCGAATTTATTGATTATGTGAT
CGTCGAGGATGATTACGTTGGTAGCGAAGATTGTTTTAGCGAAACCCTGCTGCGCCTGCCG
AAAGATGCACTGCCGTATGTTCCGAGCGCACTGGCACCGCAGAAAGTGGATTATGTTCTGC
GTGAAAATCCGGAAGTTGTGAATATTGGTATTGCAGCAACCACCATGAAACTGAATCCGGAA
TTTCTGCTGACACTGCAAGAAATTCGTGATAAAGCCAAAGTGAAAATTCACTTCCATTTTGCG
CTGGGTCAGTCAACCGGTCTGACCCATCCGTATGTGAAATGGTTTATTGAAAGCTATCTGGG
TGATGATGCAACCGCACATCCGCATGCACCGTATCATGATTATCTGGCCATTCTGCGCGATT
GTGATATGCTGCTGAATCCGTTTCCGTTTGGTAATACCAATGGCATTATTGATATGGTTACCCT
GGGCCTGGTTGGTGTTTGTAAAACCGGTGATGAAGTGCATGAACATATTGATGAAGGTCTGT
TTAAACGTCTGGGCCTGCCGGAATGGCTGATTGCAGATACCCGTGAAACCTATATTGAATGT
GCACTGCGCCTGGCGGAAAATCATCAAGAACGCCTGGAACTGCGTCGTTATATCATTGAAA
ATAATGGCCTGCAGAAACTGTTTACCGGTGATCCGCGTCCGCTGGGCAAAATTCTGCTGAA
AAAAACCAATGAGTGGAAACGCAAACATCTGAGCAAAAAAGGATCCTCTGCGTGGAGCCAT
CCGCAGTTCGAAAAAGGTGGTGGTTCTGGTGGTGGTTCTGGTGGTTCTGCGTGGAGCCAT
CCGCAGTTCGAAAAATAAgtcgacgatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctg
agcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaag

DNA sequence for EcNGT in pET.BCS.NS plasmid context:

gaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagagcagaattcggtagatctatataacaca
cataaggaggacatATGAGCCACAAAACCGATACCGCACCGGTTCAAGAACAGGCAGGCCTGAC
CTTTCGTCTGGAAACCTTTGAATGGCAGGTTCATCAGGGTCTGAATGAAGAAGCAGCACGT
AGCCTGATTAGCCTGCTGCAGCTGCTGGATCGTCATTATGCACAGTGGGGTGAAAGCTTTA
GCGCATGGGCACCGGGTATGACCGCAGAAGAAATTAATCCGCATCTGTGTACCCGTATTGC
CGGTGCAATTACCGCACTGTTTAGCCGTCCGGGTTTTCGTGTTAGTGATGGTGGTTTTGCA
GAACTGATGGATTATCATCGTTGGCTGGCAATTATCTTTGCCGTTAGCGATTATCGTCATGGC
GATCATATTATTCGCAATATTAACGCAGCCGGTGGTGGTGTTGTTGCACCGCTGACCCTGAA
TGCCGATAATCTGCAGCTGTTTTGTCTGAGCTATTATCCGGATTCACAGATTGCACTGCAGC

CGGAACCGCTGTGGCAGTATGATCGTCAGACCGTTGTTCGTCTGTTTTTTGCACTGCTGAG
CGGTCGTGCACTGCCGACACCGGCAGCACATCAGAAACGTGAACATCTGCTGGCATGGCT
GCCGGAACGTCTGAAAGAAATTGATAGCCTGGAATTTCTGCCTGGTAAAGTTCTGCATGATG
TGTATATGCATTGTAGCTATGCCGATCTGCCGGAAAAACATCGTATTAAACAAGAAATCAATC
GTCTGACCGCACGTGCACTGGAACAGACCTATGCAGATTGTCTGCCGGTTCGTGCACCGG
AAGCAGCCCGTCAGAAACCGGTTCTGGCAGTTGTTCTGGAATGGTTTACCTGTCAGCATAG
CATTTATCGTACCCATAGCACCAGCATGCGTGCCCTGCGTGAACATTTTCATCTGCTGGGTA
TTGCACAGCCTGGTGCAACCGATGAAATTACCCGTGAAGTTTTTGATGAATTTCGTGAACTG
AGCGCAGAAAATGTTGTTGGTGATGCAATTCGTTGCCTGAGCGAAGTTCGTCCGGATGTTAT
CTATTATCCTAGCGTTGGTATGTTTCCGCTGACCGTTTATCTGACAGCACTGCGTCTGGCAC
CACTGCAGCTGATGGCACTGGGTCATCCGGCAACCACCTGGTCAGAACATATTGATGGTGT
TCTGGTTGAAGAAGATTATCTGGGCGATCCGGCATGTTTTAGCGAAACCGTTTGTGCAGTTC
CGAAAGATGCCATTCCGTATATTCCGCCTGCAAGCACCGAACGTGTTCTGCCTGAACGTAC
CCCGTTTCGTGATCGTGCAAAAGCAGCATGGCCTGCAGCACTGCCGGTGCGTGTTGCAGT
TTGTGCAAGCGTTATGAAAATTAACCCTGGTTTTCTGGATACCCTGCGCGAAATTAGCGATC
GTAGCCGTGTTCCGGTTCAGTTTTGTTTTTGGATGGGTTTTGCCCAGGGCCTGACCCTGGA
TTATCTGCGTCGTGCCATTCGTCAGGCGCTGCCGACAGCCGAAGTTAATGCACACATGCCG
GTTCAGGCATATCAGCAGGCACTGAATAGCTGTGAACTGTTTGTTAATCCGTTTCCGTTTGG
TAATACCAATGGTCTGGTTGATACCGTTCGCCAGGGTCTGCCAGGTGTTTGTATGACCGGTC
CGGAAGTTCATACCCATATCGATGAAGGTCTGTTTCGTCGTCTGGGCCTGCCGGAAGCACT
GATTGCACGTGATCGCGAAGAATATATCACCGCAGTTCTGAGCCTGACCGAAACACCGCGT
CTGCGCGAACGTCTGCAGAAATATCTGACGGAAAATGATGTTGAGAAAGTGCTGTTTGAAG
GCCGTCCGGATAAATTTGCGGAACGTGTTTGGCAGCTGTGGGAAGCCCGTAGCCATCGTC
AAGAAGAAGGCGCAGAAGGATCCTCTGCGTGGAGCCATCCGCAGTTCGAAAAAGGTGGTG
GTTCTGGTGGTGGTTCTGGTGGTTCTGCGTGGAGCCATCCGCAGTTCGAAAAATAAgtcgacg
atccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataactagcataaccccttggggcctcta
aacgggtcttgaggggtttttgctgaaag

DNA sequence for ApNGT in pET.BCS.NS plasmid context:

gaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagagcagaattcggtagatctatataacaca
cataaggaggacatATGGAAAACGAGAATAAACCGAACGTGGCAAATTTTGAAGCAGCAGTTGCA
GCCAAAGATTATGAAAAAGCATGTAGCGAGCTGCTGCTGATTCTGAGCCAGCTGGATAGCA
ATTTTGGTGGCATTCATGAAATCGAGTTCGAGTACCCAGCTCAGCTGCAGGATCTGGAGCA
GGAAAAAATTGTGTACTTTTGCACCCGTATGGCGACTGCCATCACCACCCTGTTCTCTGACC
CGGTTCTGGAAATCTCCGACCTGGGTGTGCAGCGTTTCCTGGTTTATCAGCGTTGGCTGGC
GCTGATCTTCGCTAGCAGCCCGTTCGTGAACGCAGACCACATCCTGCAGACTTACAACCGT
GAACCGAACCGCAAAAACTCCCTGGAAATTCACCTGGACTCTAGCAAGTCCTCTCTGATTAA
ATTTTGCATCCTGTACCTGCCGGAATCTAATGTGAACCTGAATCTGGATGTGATGTGGAACAT
TTCCCCGGAGCTGTGTGCCTCTCTGTGCTTTGCGCTGCAAAGCCCGCGTTTTGTTGGCAC
CAGCACCGCCTTTAACAAACGCGCGACCATTCTGCAGTGGTTCCCGCGTCATCTGGACCA

GCTGAAAAACCTGAACAACATCCCGTCCGCTATCTCTCATGACGTGTATATGCACTGTTCTTA
CGACACCAGCGTTAACAAGCACGATGTTAAGCGCGCGCTGAATCACGTGATTCGTCGCCAC
ATCGAATCCGAATACGGTTGGAAAGATCGTGATGTGGCTCACATCGGTTATCGCAACAACAA
ACCGGTTATGGTCGTTCTGCTGGAACATTTTCATAGCGCGCACTCTATCTACCGTACTCACT
CTACCAGCATGATCGCCGCGCGCGAACACTTCTATCTGATCGGCCTGGGTTCCCCGAGCGT
TGACCAGGCCGGTCAGGAGGTTTTCGATGAATTCCACCTGGTAGCGGGTGACAACATGAA
GCAAAAACTGGAATTCATTCGTTCTGTGTGCGAAAGCAATGGTGCGGCAATTTTCTACATGC
CGAGCATCGGTATGGATATGACCACCATCTTCGCGTCCAATACCCGTCTGGCGCCGATTCA
GGCAATCGCCCTGGGCCACCCGGCGACTACTCACTCCGACTTCATTGAATACGTTATCGTG
GAAGACGACTACGTCGGCTCTGAGGAATGCTTCTCTGAAACCCTGCTGCGTCTGCCGAAA
GACGCTCTGCCGTATGTTCCGTCCGCCCTGGCTCCGGAGAAAGTTGATTACCTGCTGCGTG
AAAACCCTGAAGTTGTCAACATCGGTATTGCCTCTACCACTATGAAGCTGAACCCGTACTTC
CTGGAAGCACTGAAGGCCATTCGTGACCGTGCGAAGGTGAAAGTGCACTTCCACTTCGCA
CTGGGCCAGTCCAATGGTATCACTCACCCTTACGTTGAACGCTTTATCAAATCTTACCTGGG
CGACAGCGCTACCGCGCACCCGCACTCTCCGTACCACCAGTACCTGCGTATTCTGCACAAC
TGCGATATGATGGTAAACCCTTTTCCGTTTGGTAATACCAATGGTATTATTGACATGGTAACCC
TGGGTCTGGTAGGTGTTTGCAAAACCGGTGCGGAAGTCCACGAACATATCGATGAAGGCCT
GTTCAAACGTCTGGGCCTGCCGGAATGGCTGATTGCAAACACCGTGGACGAATACGTGGA
ACGTGCAGTGCGCCTGGCCGAGAACCATCAGGAACGTCTGGAACTGCGTCGTTACATTATT
GAAAACAATGGCCTGAACACCCTGTTCACCGGCGACCCACGCCCGATGGGTCAGGTGTTC
CTGGAAAAACTGAACGCATTCCTGAAGGAAAACGGATCCTCTGCGTGGAGCCATCCGCAGT
TCGAAAAAGGTGGTGGTTCTGGTGGTGGTTCTGGTGGTTCTGCGTGGAGCCATCCGCAGT
TCGAAAAATAAgtcgacgatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataact
agcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaag

DNA sequence for ApNGT Q469A in pET.BCS.NS plasmid context:

gaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagagcagaattcggtagatctatataacaca
cataaggaggacatATGGAAAACGAGAATAAACCGAACGTGGCAAATTTTGAAGCAGCAGTTGCA
GCCAAAGATTATGAAAAAGCATGTAGCGAGCTGCTGCTGATTCTGAGCCAGCTGGATAGCA
ATTTTGGTGGCATTCATGAAATCGAGTTCGAGTACCCAGCTCAGCTGCAGGATCTGGAGCA
GGAAAAAATTGTGTACTTTTGCACCCGTATGGCGACTGCCATCACCACCCTGTTCTCTGACC
CGGTTCTGGAAATCTCCGACCTGGGTGTGCAGCGTTTCCTGGTTTATCAGCGTTGGCTGGC
GCTGATCTTCGCTAGCAGCCCGTTCGTGAACGCAGACCACATCCTGCAGACTTACAACCGT
GAACCGAACCGCAAAAACTCCCTGGAAATTCACCTGGACTCTAGCAAGTCCTCTCTGATTAA
ATTTTGCATCCTGTACCTGCCGGAATCTAATGTGAACCTGAATCTGGATGTGATGTGGAACAT
TTCCCCGGAGCTGTGTGCCTCTCTGTGCTTTGCGCTGCAAAGCCCGCGTTTTGTTGGCAC
CAGCACCGCCTTTAACAAACGCGCGACCATTCTGCAGTGGTTCCCGCGTCATCTGGACCA
GCTGAAAAACCTGAACAACATCCCGTCCGCTATCTCTCATGACGTGTATATGCACTGTTCTTA
CGACACCAGCGTTAACAAGCACGATGTTAAGCGCGCGCTGAATCACGTGATTCGTCGCCAC
ATCGAATCCGAATACGGTTGGAAAGATCGTGATGTGGCTCACATCGGTTATCGCAACAACAA

ACCGGTTATGGTCGTTCTGCTGGAACATTTTCATAGCGCGCACTCTATCTACCGTACTCACT
CTACCAGCATGATCGCCGCGCGCGAACACTTCTATCTGATCGGCCTGGGTTCCCCGAGCGT
TGACCAGGCCGGTCAGGAGGTTTTCGATGAATTCCACCTGGTAGCGGGTGACAACATGAA
GCAAAAACTGGAATTCATTCGTTCTGTGTGCGAAAGCAATGGTGCGGCAATTTTCTACATGC
CGAGCATCGGTATGGATATGACCACCATCTTCGCGTCCAATACCCGTCTGGCGCCGATTCA
GGCAATCGCCCTGGGCCACCCGGCGACTACTCACTCCGACTTCATTGAATACGTTATCGTG
GAAGACGACTACGTCGGCTCTGAGGAATGCTTCTCTGAAACCCTGCTGCGTCTGCCGAAA
GACGCTCTGCCGTATGTTCCGTCCGCCCTGGCTCCGGAGAAAGTTGATTACCTGCTGCGTG
AAAACCCTGAAGTTGTCAACATCGGTATTGCCTCTACCACTATGAAGCTGAACCCGTACTTC
CTGGAAGCACTGAAGGCCATTCGTGACCGTGCGAAGGTGAAAGTGCACTTCCACTTCGCA
CTGGGCGCGTCCAATGGTATCACTCACCCTTACGTTGAACGCTTTATCAAATCTTACCTGGG
CGACAGCGCTACCGCGCACCCGCACTCTCCGTACCACCAGTACCTGCGTATTCTGCACAAC
TGCGATATGATGGTAAACCCTTTTCCGTTTGGTAATACCAATGGTATTATTGACATGGTAACCC
TGGGTCTGGTAGGTGTTTGCAAAACCGGTGCGGAAGTCCACGAACATATCGATGAAGGCCT
GTTCAAACGTCTGGGCCTGCCGGAATGGCTGATTGCAAACACCGTGGACGAATACGTGGA
ACGTGCAGTGCGCCTGGCCGAGAACCATCAGGAACGTCTGGAACTGCGTCGTTACATTATT
GAAAACAATGGCCTGAACACCCTGTTCACCGGCGACCCACGCCCGATGGGTCAGGTGTTC
CTGGAAAAACTGAACGCATTCCTGAAGGAAAACGGATCCTCTGCGTGGAGCCATCCGCAGT
TCGAAAAATAAgtcgacgatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataact
agcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaag

DNA sequence for EndoA WT in pGEX/T2 plasmid context: (*ITALICS* indicates GST-Tag)
gttgacaattaatcatcggctcgtataatgtgtggaattgtgagcggataacaatttcacacaggaaacagtattc*ATGTCCCCTA*
TACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATATCTTG
AAGAAAAATATGAAGAGCATTTGTATGAGCGCGATGAAGGTGATAAATGGCGAAACAAAAAG
TTTGAATTGGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGATGGTGATGTTAAATTAACAC
AGTCTATGGCCATCATACGTTATATAGCTGACAAGCACAACATGTTGGGTGGTTGTCCAAAA
GAGCGTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAGATACGGTGTTTCGAG
AATTGCATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTACCTGAAATG
CTGAAAATGTTCGAAGATCGTTTATGTCATAAAACATATTTAAATGGTGATCATGTAACCCATC
CTGACTTCATGTTGTATGACGCTCTTGATGTTGTTTTATACATGGACCCAATGTGCCTGGATG
CGTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTGAAGCTATCCCACAAATTGATAAGTACTT
GAAATCCAGCAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGC
*GACCATCCTCCAAAATCGGATCTGGTTCCGCGTGGATCC*TCTACGTACAACGGCCCGCTGT
CCTCCCATTGGTTTCCAGAGGAACTTGCCCAATGGGAACCAGACAGTGATCCAGACGCACC
CTTTAACAGAAGCCATGTTCCGCTGGAACCAGGCCGCGTTGCGAATAGGGTAAATGCTAAT
GCAGACAAGGACGCACACCTTGTTTCGTTGTCCGCGCTAAACAGGCATACATCAGGTGTTC
CATCGCAAGGAGCGCCAGTTTTCTATGAAAATACGTTCAGCTATTGGCATTATACAGATTTGA
TGGTTTATTGGGCTGGTTCAGCTGGCGAAGGCATTATCGTTCCGCCAAGTGCCGATGTCATT
GATGCATCGCACCGAAATGGGGTGCCGATTTTAGGAAATGTGTTCTTCCCGCCGACGGTTT

ATGGAGGGCAGCTAGAGTGGCTAGAACAAATGTTAGAGCAAGAGGAGGACGGTTCATTCCC
CCTTGCTGACAAATTGCTAGAAGTCGCAGACTATTATGGGTTTGACGGCTGGTTTATTAACC
AAGAAACAGAAGGGGCAGACGAAGGAACAGCCGAAGCCATGCAAGCTTTTCTCGTTTATTT
GCAGGAACAAAAGCCAGAAGGCATGCACATCATGTGGTATGACTCGATGATTGATACAGGG
GCGATCGCCTGGCAAAACCATTTAACGGATCGAAATAAAATGTACTTGCAAAATGGCTCGAC
CCGCGTCGCTGACAGCATGTTTTTGAACTTTTGGTGGCGTGACCAGCGCCAATCGAACGAA
TTGGCACAAGCACTTGGCAGGTCTCCGTATGACCTCTATGCCGGAGTGGATGTGGAAGCAC
GAGGGACAAGTACCCCTGTTCAGTGGGAAGGCCTGTTTCCTGAAGGAGAAAAGGCGCATA
CATCACTCGGGTTATACCGTCCAGATTGGGCATTTCAGTCAAGTGAAACAATGGAAGCGTTT
TATGAAAAAGAACTACAATTTTGGGTTGGCTCGACAGGAAATCCAGCCGAAACAGACGGCC
AGTCAAATTGGCCTGGCATGGCGCACTGGTTTCCCGCGAAAAGCACCGCTACTTCGGTACC
CTTTGTGACTCACTTTAATACGGGCAGCGGCGCTCAGTTTTCGGCAGAAGGCAAAACTGTG
TCGGAACAGGAATGGAATAACCGCAGCCTTCAAGATGTGCTGCCGACATGGCGCTGGATTC
AGCATGGCGGCGATTTAGAGGCAACATTTTCTTGGGAAGAAGCGTTTGAAGGGGGAAGCTC
GTTACAATGGCATGGCTCATTAGCGGAAGGAGAACACGCCCAAATCGAGCTCTATCAAACA
GAGTTGCCGATAAGCGAAGGCACTTCGCTAACGTGGACATTTAAAAGCGAGCACGGCAACG
ATTTAAATGTGGGCTTCCGTTTAGATGGGGAAGAGGACTTCCGTTATGTGGAAGGAGAACA
GCGTGAATCGATAAATGGTTGGACGCAGTGGACGTTGCCGCTGGATGCGTTTGCTGGTCA
GACGATAACAGGGCTGGCATTTGCAGCGGAAGGGAATGAGACTGGGCTGGCAGAATTCTAT
ATTGGACAACTGGCCGTAGGTGCTGATAGCGAAAAGCCTGCCGCTCCAAACGTGAACGTAC
GCCAGTACGACCCAGACCCGAGTGGCATTCAGCTCGTATGGGAAAAACAAAGCAACGTCC
ACCATTACCGCGTTTATAAAGAAACAAAGCACGGCAAAGAGCTAATTGGCACATCTGCTGGA
GATCGAATTTACCTAGAAGGCCTAGTCGAGGAAAGCAAACAAAACGACGTGCGTCTGCATAT
AGAAGCACTAAGTGAAACATTTGTGCCAAGTGATGCTCGCATGATCGACATAAAAAGCGGCT
CGTTTTAGgggaattcatcgtgactgactgacgatctgcctcgcgcgtttcggtgatgacggtgaaaacctctgacacatgcagct
cccggagacggtcacagcttgtctgt

DNA sequence for EndoCC1 N180H in pET41b plasmid context:

gaaattaatacgactcactatagggaattgtgagcggataacaattcccctctagaaataattttgtttaactttaagaaggagatatac
atATGCACCATCACCACCACCATATGCCTATCGCTGGGAAGAAGTTCCACCCCCGGGCTCTG
CCTGAGTTCTGGAGGACGTTCCGGGAGATGGACGAATGGCGGGCCACTCAGACTGGACCT
CAGGCTCGTCCGGCTGAAGGCATTCTCAAGTACGTTCCACGGAAGATTCGGCCTGCTGATA
TCGCAGGGAAAGGTCGACTGTTGGTTTCTCATGACTACAAGGGAGGCTACGTCGAAGATCC
CTTTTCCAAGTCGTATAGCTTCAATTGGTGGTTCTCGACGGATAGCTTCAACTACTTCGCTCA
CCACCGGATAACCATTCCCCCTCCGGAATGGATAAACGCTGCTCATCGCCAGGGTGTACCT
ATTCTCGGCACCATCATCTTCGAAGGTGGAAGCGACGAAGACATCCTCCGGATGGTGATCG
GGAAAACACCAGGAAGCACCAGCAACTTCCACGCCGAACGAAACGCGGAGTACACCGTAC
CAGTTTCGTCGTACTACGCAGAACTCTTCGCAGACCTGGCTGTCGAGCGTGGATTCGATGG
TTGGCTGCTCCACGTTGAGATTGGTCTGCAGGGAGGATCAGAGCAAGCACGAGGTCTGGC
TGCTTGGGTTGCGCTACTGCAGCAGGAGGTTTTGAAGAAGGTTGGCCCACATGGCTTGGT

CATTTGGTACGACAGTGTCACTGTCCGTGGGGACCTGTGGTGGCAAGACAGGCTGAACGC
TTTCAACTTGCCGTTCTTCTTGAATTCTTCGGGAATTTTCACAAACTATTGGTGGTACAACGA
TGCACCTCAGAAACAAATCGACTTCCTTTCGAGGGTTGACCCGAATCTCACCGGGCAAACC
GCTGAGCCGCATCAATACAACCTGCAGAAGACGATTCAAGATATCTATATCGGTGTGGATGT
CTGGGGACGCGGTTCGCATGGTGGAGGAGGATTTGGTGCCTACAAAGCTATTGAGCACGC
AGACCCGAAAGGACTCGGGTTTAGCGTTGCCCTCTTCGCTCAAGGATGGACCTGGGAAAC
CGAGGAGGAGAAACCAGGCTGGAACTGGGCACAGTTCTGGGACTACGACTCTAAACTCTG
GGTTGGACCTCCCGGAGTTGTCGAGGCACCTGACCATACCGTCAAACCTGGCGAATACCC
CTGCGTTCACGGACCCTTCCAACCCATCTCCAGCTTCTTCCTGACATATCCACCTCCCGAC
CCGCTAGACCTGCCATTCTACACCAACTTCTGCCCCGGTATCGGAGATGCGTGGTTCGTTG
AGGGCAAGGAGGTCTTCCGCTCCGAGACGGGTTGGACAGACATGGACAAGCAGACCACG
GTTGGCGACTTGGTCTGGCCCCGACCCAAGATTTATGATCTTCCATCTCAAAATGCCAGTCA
GGCTACGTTAAATGCGGCATTCAACTTTAACGATGCGTGGAATGGAGGAAACTCGCTTCAGA
TCAACCTCACCGTCCCTGGAGGAGCGACCACGTATGGAGCCTACTGGGTTCCCATCCAGA
CATTCACATTCTCCAGTCGGCGCCAGTACGAAGCTTCGATCGTCTACAAGCCGGGGTTGAG
TGGGAAGACCCGCTTCGATGCCAAGTACGAGGTGGGTATCCGAACCATCACAGGGGAAGA
CCAAGGCAAGATCATCTCCAACACGACGACGGAAGTTGGAAACGGTTGGCGCAAGGTGCA
CATTTTGTTCGAGATCGAGACGCCAGTTGAAGGAGGGTCCATCATCGTACCTTCATCCATCG
GCCTGGTCATTGCAGTTTCGAACGTATCAACTACCGAGCAATTCGAGTTCCCCTTCCTGGTC
GGCCAGATTACCATCCACCCCCACCTCCCCGATCGTTACAAGGAGTTCAAGCCGGCCCTCC
TGTGGCTTTTGTTCACACCTTCCGCTGGAACTAATAGCCTCGATGGCACCCTCACTTGGGA
CGTCGTCGCAGCCATTGAACGCCCTCCACCAGTCGAAATTAACAACCCCGATGACGCACAA
ATCCCCTGGAACCTGCAGCCGACCAAACAAGAATGGTTCCCCGACTTCCTCTACTTCAATG
TGTACGTGCTGGAGCTCTTGGATGGTGGTGGACAAGGTCCTCCACAGTGGATTGGCACGA
CTGGATACGATGGGGAGAAAAAGAGGTTCTTCATCTATGACGAAAGCTTGCCACCGACGTC
CGGTTTAAGGAGGTTCACGTTCCAGATCGAGGGCGTCCTGGAGACGGGAGAGTCAACGCA
CTGGTACGATGCACCGGCTGCGCCGTCGGCAACGGCGGGAGGAGAGCAGAAGCGGACAC
GACGCACCTCTCTCAAGTCCGTGCTCAGTCCGTTGCGGAGGAAGAAGTCGAAGGGCGATA
TCTCCGTCGCCAAGTGAgaattctgtacaggccttggcgcgcctgcaggcgagctccgtcgacaagcttgcggccgcact
cgagcaccaccaccaccaccaccaccactaattgattaatacctaggctgctaaacaaagcccgaaaggaagctgagttggctgct
gccaccgctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaaggaggaactatatccgg
at

Supplementary Information References:

1. Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P. M.; Henrissat, B., The carbohydrate-active enzymes database (cazy) in 2013. *Nucleic Acids Res.* **2014,** *42* (D1), D490-495.
2. Katoh, K.; Standley, D. M., Mafft multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013,** *30* (4), 772-780.
3. Price, M. N.; Dehal, P. S.; Arkin, A. P., Fasttree 2--approximately maximum-likelihood trees for large alignments. *PLOS ONE* **2010,** *5* (3), e9490.
4. Stamatakis, A., Raxml version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014,** *30* (9), 1312-1313.
5. Letunic, I.; Bork, P., Interactive tree of life (itol) v3: An online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **2016,** *44* (W1), W242-245.
6. Kightlinger, W.; Lin, L.; Rosztoczy, M.; Li, W.; DeLisa, M. P.; Mrksich, M.; Jewett, M. C., Design of glycosylation sites by rapid synthesis and analysis of glycosyltransferases. *Nat. Chem. Biol.* **2018,** *14* (6), 627-635.
7. Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchison, C. A.; Smith, H. O., Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **2009,** *6* (5), 343-345.
8. Jewett, M. C.; Swartz, J. R., Mimicking the escherichia coli cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnol. Bioeng.* **2004,** *86* (1), 19-26.
9. Martin, R. W.; Des Soye, B. J.; Kwon, Y.-C.; Kay, J.; Davis, R. G.; Thomas, P. M.; Majewska, N. I.; Chen, C. X.; Marcum, R. D.; Weiss, M. G.; Stoddart, A. E.; Amiram, M.; Ranji Charna, A. K.; Patel, J. R.; Isaacs, F. J.; Kelleher, N. L.; Hong, S. H.; Jewett, M. C., Cell-free protein synthesis from genomically recoded bacteria enables multisite incorporation of noncanonical amino acids. *Nat. Commun.* **2018,** *9* (1), 1203.
10. Kwon, Y.-C.; Jewett, M. C., High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Sci. Rep.* **2015,** *5*, 8663.
11. Tong, X.; Li, T.; Orwenyo, J.; Toonstra, C.; Wang, L.-X., One-pot enzymatic glycan remodeling of a therapeutic monoclonal antibody by endoglycosidase s (endo-s) from streptococcus pyogenes. *Bioorg. Med. Chem.* **2018,** *26* (7), 1347-1355.
12. Higuchi, Y.; Eshima, Y.; Huang, Y.; Kinoshita, T.; Sumiyoshi, W.; Nakakita, S. I.; Takegawa, K., Highly efficient transglycosylation of sialo-complex-type oligosaccharide using coprinopsis cinerea endoglycosidase and sugar oxazoline. *Biotechnol. Lett.* **2017,** *39* (1), 157-162.
13. Huang, W.; Li, C.; Li, B.; Umekawa, M.; Yamamoto, K.; Zhang, X.; Wang, L. X., Glycosynthases enable a highly efficient chemoenzymatic synthesis of n-glycoproteins carrying intact natural n-glycans. *J. Am. Chem. Soc.* **2009,** *131* (6), 2214-2223.
14. Noguchi, M.; Tanaka, T.; Gyakushi, H.; Kobayashi, A.; Shoda, S., Efficient synthesis of sugar oxazolines from unprotected n-acetyl-2-amino sugars by using chloroformamidinium reagent in water. *J. Org. Chem.* **2009,** *74* (5), 2210-2212.
15. Ochiai, H.; Huang, W.; Wang, L. X., Expeditious chemoenzymatic synthesis of homogeneous n-glycoproteins carrying defined oligosaccharide ligands. *J. Am. Chem. Soc.* **2008,** *130* (41), 13790-13803.
16. Baranova, N.; Loose, M., Single-molecule measurements to study polymerization dynamics of ftsz-

ftsa copolymers. *Methods Cell Biol.* **2017,** *137*, 355-370.
17. Eshima, Y.; Higuchi, Y.; Kinoshita, T.; Nakakita, S.-i.; Takegawa, K., Transglycosylation activity of glycosynthase mutants of endo-β-n-acetylglucosaminidase from coprinopsis cinerea. *PLOS ONE* **2015,** *10* (7), e0132859.
18. Lin, C.-W.; Tsai, M.-H.; Li, S.-T.; Tsai, T.-I.; Chu, K.-C.; Liu, Y.-C.; Lai, M.-Y.; Wu, C.-Y.; Tseng, Y.-C.; Shivatare, S. S.; Wang, C.-H.; Chao, P.; Wang, S.-Y.; Shih, H.-W.; Zeng, Y.-F.; You, T.-H.; Liao, J.-Y.; Tu, Y.-C.; Lin, Y.-S.; Chuang, H.-Y.; Chen, C.-L.; Tsai, C.-S.; Huang, C.-C.; Lin, N.-H.; Ma, C.; Wu, C.-Y.; Wong, C.-H., A common glycan structure on immunoglobulin g for enhancement of effector functions. *Proc. Natl. Acad. Sci. U. S. A.* **2015,** *112* (34), 10611-10616.
19. Jian, W.; Edom, R. W.; Wang, D.; Weng, N.; Zhang, S., Relative quantitation of glycoisoforms of intact apolipoprotein c3 in human plasma by liquid chromatography–high-resolution mass spectrometry. *Anal. Chem.* **2013,** *85* (5), 2867-2874.