JACS Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

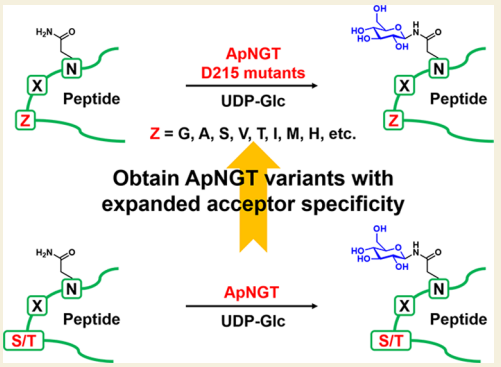
pubs.acs.org/jacsau
Article

Investigation of the Catalytic Mechanism of a Soluble N-glycosyltransferase Allows Synthesis of N-glycans at Noncanonical Sequons

Zhiqiang Hao,[#] Qiang Guo,[#] Yuanyuan Feng, Zihan Zhang, Tiantian Li, Zhixin Tian, Jianting Zheng, Lin-Tai Da,[*] and Wenjie Peng[*]

Cite This: *JACS Au* 2023, 3, 2144−2155
Read Online
ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The soluble N-glycosyltransferase from *Actinobacillus pleuropneumoniae* (ApNGT) can establish an N-glycosidic bond at the asparagine residue in the Asn-Xaa-Ser/Thr consensus sequon and is one of the most promising tools for N-glycoprotein production. Here, by integrating computational and experimental strategies, we revealed the molecular mechanism of the substrate recognition and following catalysis of ApNGT. These findings allowed us to pinpoint a key structural motif (^{215}DVYM^{218}) in ApNGT responsible for the peptide substrate recognition. Moreover, Y222 and H371 of ApNGT were found to participate in activating the acceptor Asn. The constructed models were supported by further crystallographic studies and the functional roles of the identified residues were validated by measuring the glycosylation activity of various mutants against a library of synthetic peptides. Intriguingly, with particular mutants, site-selective N-glycosylation of canonical or noncanonical sequons within natural polypeptides from the SARS-CoV-2 spike protein could be achieved, which were used to investigate the biological roles of the N-glycosylation in membrane fusion during virus entry. Our study thus provides in-depth molecular mechanisms underlying the substrate recognition and catalysis for ApNGT, leading to the synthesis of previously unknown chemically defined N-glycoproteins for exploring the biological importance of the N-glycosylation at a specific site.



KEYWORDS: *computational modeling, N-glycosyltransferase, catalytic mechanisms, noncanonical sequons, enzyme engineering*

INTRODUCTION

Asparagine (N)-linked protein glycosylation, one of the most common post-translational modifications (PTMs) found in all domains of life, plays essential roles in numerous biological processes, including protein folding, host−pathogen recognition, as well as inflammatory and immune responses.[1−3] In the mammalian biosynthesis of N-glycoproteins, an oligomeric membrane-associated oligosaccharyltransferase (OST) catalyzes the *en bloc* transfer of an N-glycan precursor ($Glc_3Man_9GlcNAc_2$) from a dolichol-pyrophosphate donor, also known as a lipid-linked oligosaccharide (LLO), to the side-chain amide nitrogen of an asparagine residue in the consensus sequon Asn-Xaa-Ser/Thr (where Xaa ≠ Pro, N-X-S/T).[4] Various N-glycoproteins, including recombinant enzymes, vaccines, and monoclonal antibodies, have been approved for medical diagnosis and clinic treatments.[5−7] In general, these therapeutic glycoproteins are derived from cell culture. Due to the non-template-driven nature of glycosylation *in vivo*, natural and recombinant glycoproteins are composed of structurally diverse glycans, referred to as glycosylation microheterogeneity.

It is well documented that distinct N-glycoforms significantly impact the bioactivity and pharmacokinetic profile of a given N-glycoprotein.[8−12] Consequently, there have been continuous efforts to develop novel strategies to produce homogeneous N-glycoproteins for detailed structure−function studies.[13−15] Among them, enzymatic synthesis, which can be done in high yield and with high stereo- and regioselectivity, has attracted much attention. Impressive efforts have been made to employ membrane-bound OSTs for preparing homogeneous N-glycoproteins. However, due to the limited availability of the preassembled LLO donors and active OST enzymes, *in vitro* synthesis of structurally defined N-glycoproteins using this approach remains unrealized.[16] Furthermore, bacterial OSTs (e.g., *Cj*PglB) possess more stringent specificity for the (D/E-X)-N-X-S/T sequon, which

Received: May 2, 2023
Revised: July 21, 2023
Accepted: July 24, 2023
Published: August 7, 2023

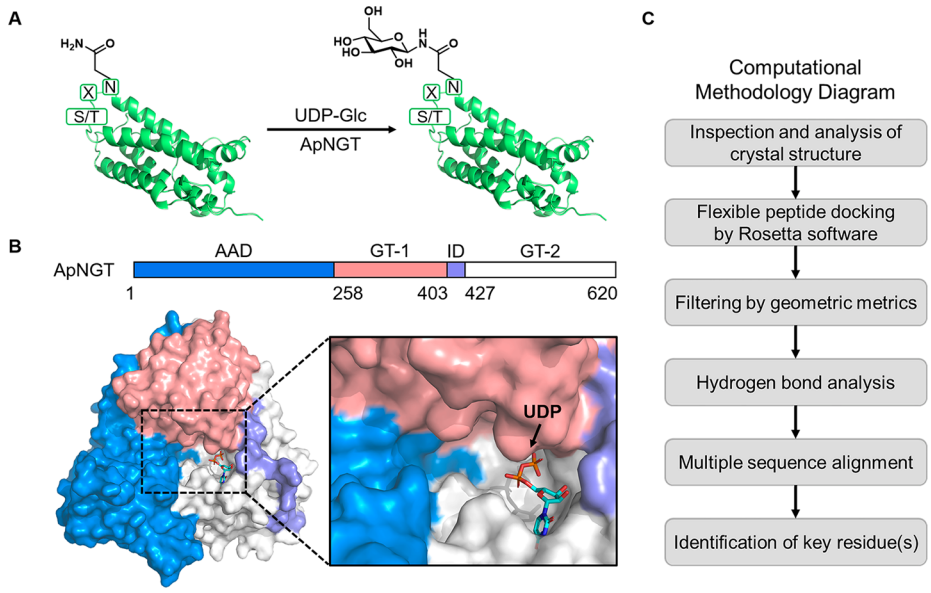


Figure 1. Illustration of the function and structure of ApNGT and the computational workflow applied in this study. (A) Schematic diagram of N-linked glycosylation catalyzed by ApNGT. Crystal structure of human erythropoietin (PDB 1EER) is used as a protein representation. (B) Structural domains of ApNGT and surface representation of previously solved crystal structure of ApNGT in complex with UDP (PDB 3Q3H). UDP is shown in cyan sticks; AAD, GT-1, GT-2, and ID are shown in blue, pink, light gray, and purple cartoons, respectively. The substrate binding groove is highlighted in the zoomed-in image. (C) Computational methodology flowchart for this study.

limits their use for *in vitro* N-glycoprotein synthesis.[17−19] Thus, it is still a challenge to generate an N-glycosidic bond between an oligosaccharide and a protein/peptide using OST enzymes.

Functionally similar to OST is the cytoplasmic soluble N-glycosyltransferase (NGT) produced in some bacteria. This enzyme (GT41 in the CAZy database) can transfer a glucose or less preferred galactose moiety to an asparagine residue in the N-X-S/T sequon (Figure 1A).[20−24] Mrksich and co-workers have screened 41 putative NGTs and characterized the acceptor sequence preference of four variants, among which the NGT from *Actinobacillus pleuropneumoniae* (ApNGT) and its mutant (ApNGT^{Q469A}) display the highest N-glycosylation activity.[25] Additionally, a sequential glycosylation strategy was developed to install distinct pseudo-N-glycoforms at the specific sites within one protein by combination of ApNGT and *endo-β-N*-acetylglucosaminidase (ENGase) mutants.[26] Notably, ApNGT is unable to transfer a GlcNAc moiety and therefore fails to create a natural GlcNAc−Asn linkage.[20,21,27] Despite that, the NGT-mediated N-glycosylation system has attracted growing attention for N-glycoprotein/glycopeptide production *in vitro* and *in vivo*.[28−30]

Although the X-ray crystal structures of ApNGT complexes have been solved, electron density for the glucose moiety of the UDP-glucose (UDP-Glc) donor and the peptide acceptor substrate was not observed (Figure 1B and Table S1).[31] The ApNGT structure consists of three discrete domains, including an all α-helical domain (AAD) at the N terminus (residues 1 to 257) and two Rossmann-like domains that create a GT-B fold at the C terminus (residues 258 to 620).[31] The GT-B fold is composed of the GT-1 domain (residues 258 to 403), the GT-2 domain (residues 427 to 620), and an interdomain (ID) region that connects GT-1 and GT-2.[31] The GT-1 and GT-2 domains form the UDP binding site at their interface (Figure 1B). The absence of comprehensive structural description for ApNGT in complex with its intact substrates (both sugar donor and acceptor peptide) makes it a great challenge to understand its catalytic mechanism and further impedes the rational engineering of the enzyme to accept peptide sequons with broader scope.

In this study, we used a molecular docking strategy to construct ternary complexes for ApNGT^{Q469A} bound with UDP-Glc and different peptide substrates. Structural inspection and statistical analysis allowed us to pinpoint the key residues in ApNGT potentially responsible for recognizing the +2 Ser/Thr in the N-X-S/T sequon. The constructed models were further supported by our crystal structure of an ApNGT binary complex. The functional roles of the identified residues (D215, M218, Y222, and H371) in ApNGT responsible for the substrate recognition and enzymatic catalysis were validated by biochemical assays with site-directed mutants. By combination of an ENGase mutant Endo CC^{N180H}, we further applied certain ApNGT mutants to site-selectively install human-like N-glycans to asparagine residues of natural polypeptides originated from the SARS-CoV-2 spike protein, achieving N-glycosylation at both canonical and noncanonical sequons. These synthesized N-glycopeptides were used to explore the biological roles of N-glycosylation in membrane fusion during virus entry. This work provides deep structural insights into the catalytic mechanism of NGT-mediated N-glycosylation and allows synthesis of previously unknown chemically defined N-glycoproteins. Importantly, our results will facilitate the optimization of ApNGT activity, thereby developing valuable enzymatic tools to generate site-specific N-glycoproteins/glycopeptides for fundamental research and therapeutic applications.

RESULTS AND DISCUSSION

Computational Construction of ApNGT Ternary Complexes for Different Peptide Substrates

To identify key residues in ApNGT potentially responsible for recognizing the N-X-S/T consensus sequon, we constructed ternary complexes of ApNGT bound with both sugar donor

https://doi.org/10.1021/jacsau.3c00214
JACS Au 2023, 3, 2144−2155

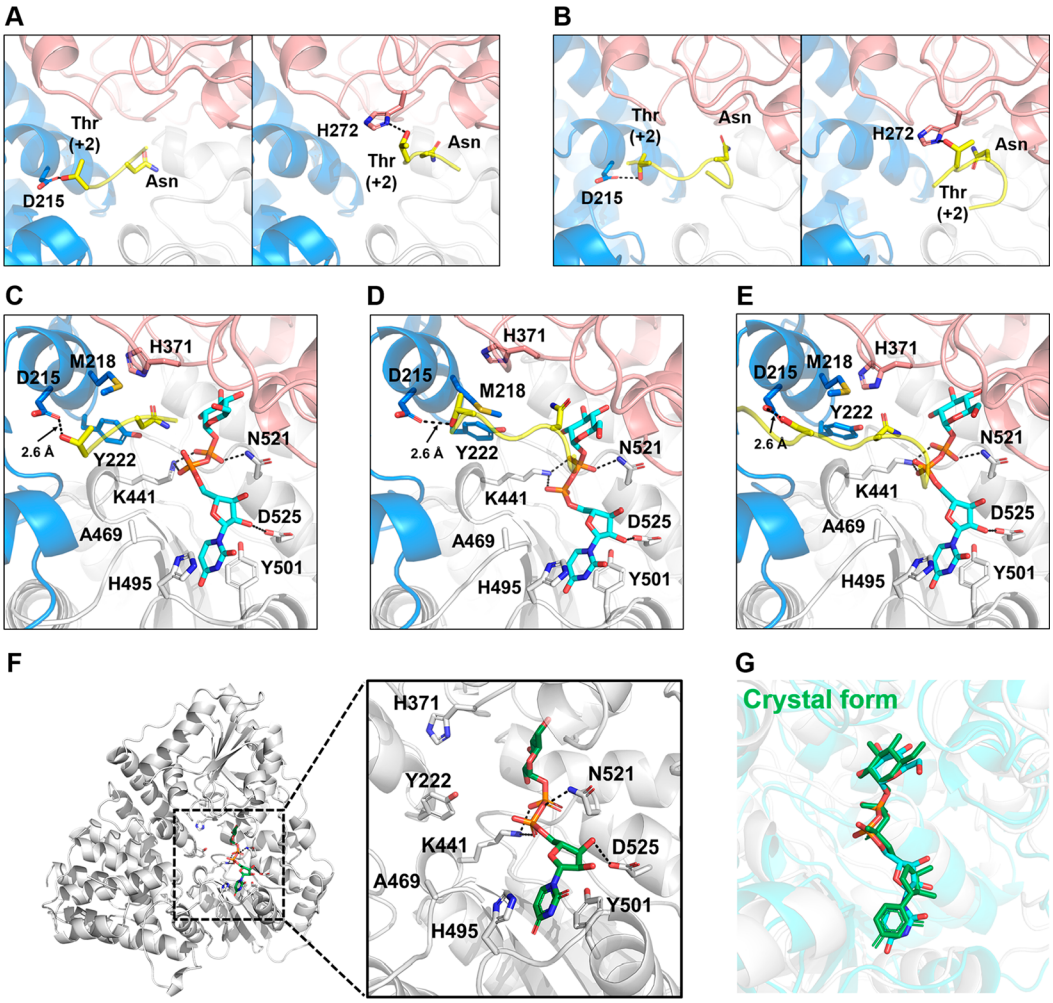


Figure 2. Structural features of the constructed ApNGT complexes. Comparisons of representative docked complexes for peptide GNWT (A) and GGNWTT (B) when D215 or H272 serves as the anchor residue to stabilize the +2 Thr. Zoomed-in views of the binding sites in the constructed ternary complexes with GNWT (C), GGNWTT (D), and KNLNESLID (E) when D215 serves as the anchor residue. The peptide substrates are shown in yellow cartoons; UDP-Glc is displayed in cyan sticks. Critical ApNGT residues are labeled and shown in stick models. HB interactions are highlighted with black dashed lines. (F) Crystal structure of ApNGT$^{Q469A/M218A}$ with an intact UDP-Glc (PDB 8J30). The UDP-Glc binding pocket is highlighted on the right with critical residues shown in stick models. (G) Superposition of modeled UDP-Glc (cyan) and crystal UDP-Glc (green) with RMSD of 0.899 Å.

and acceptor peptide. To achieve this, we first performed molecular docking to explore all possible binding modes for a given peptide toward ApNGT. We selected three peptide sequences with varied lengths as substrates, i.e., GNWT and GGNWTT, which have been previously identified as the ideal N-glycosylation acceptors for ApNGT, and ^{1191}KNLNESLID^{1199} derived from the SARS-CoV-2 spike protein.[32,33] The ApNGT structure was adopted from one crystal form of ApNGT by removing the bound UDP ligand (PDB 3Q3H, see Figure 1B).[31] In addition, we introduced a Q469A mutation in ApNGT (termed hereafter as wild-type ApNGT or wtApNGT) since this substitution was found to profoundly enhance N-glycosylation activity.[25,34,35] Using this structure, we performed flexible peptide docking for each of the three peptides with Rosetta.[36] A total of 50,000 docking models were generated for each peptide substrate. Finally, we added the UDP-Glc structure into each peptide−ApNGT docking complex to obtain the corresponding ternary structure (see Figure 1C and Experimental Section in Supporting Information).

Next, for each peptide substrate, we filtered the 50,000 ternary models based on two criteria: (1) To ensure a distance close enough for the formation of an N-glycosidic bond, we only allowed a distance less than 5 Å between the amide nitrogen (nucleophile) of the acceptor Asn in the N-X-S/T sequon and the anomeric carbon of UDP-Glc (denoted as Distance 1). (2) To eliminate conformations with steric clashes between sugar donor and acceptor peptide, we set the minimum distance of the heavy atoms (non-hydrogen) between UDP-Glc and peptide substrate (denoted as Distance 2) larger than 1 Å. Notably, the angle of the incoming nucleophile is not taken into account. Thus, according to the above two geometric metrics, we obtained 1192, 1066, and 517 hits for peptides GNWT, GGNWTT, and KNLNESLID, respectively (Figure S1).

Identification of Key Residues in ApNGT that Recognize the +2 Ser/Thr in the N-X-S/T Sequon

With the filtered docking models in hand, we sought to pinpoint key residues in ApNGT responsible for recognizing the +2 Ser/Thr in the N-X-S/T sequon. Considering that both

https://doi.org/10.1021/jacsau.3c00214
JACS Au 2023, 3, 2144−2155

Ser and Thr contain a hydroxyl group, we conducted hydrogen bond (HB) analysis to identify which residue(s) in ApNGT might anchor the hydroxyl group of the +2 Ser/Thr to the protein. The results indicate that both D215 and H272 are most likely to form HB interactions with the hydroxyl group of the +2 Ser/Thr for the chosen peptides (Figure S2). Sequence alignment analysis further indicates that D215 is highly conserved among ten active NGTs.[20,21,25,26,31,32,34,35] On the other hand, H272 is not conserved and could be replaced by Trp in some NGTs (Figures S2 and S3).

Detailed structural inspection demonstrates that D215 resides within the acceptor binding groove formed by the AAD and GT-1/2 domains, resulting in an ideal peptide loading mode in the cleft for both GNWT and GGNWTT (Figure 2A,B). In contrast, H272 lies outside the cavity, and the loaded peptide is not embedded in the binding groove. This finding is further supported by the more stable packing interactions (Packstat) between the peptide substrate and ApNGT when D215 serves as an anchor residue, accompanied by a larger buried solvent-accessible surface area (SASA) upon binding (Figures S4E,F and S5E,F). Altogether, we conclude that D215 in ApNGT is likely responsible for recognizing the +2 Ser/Thr in the N-X-S/T sequon.

Structural Features of the ApNGT Ternary Complex

To provide structural insights into ApNGT in complex with both sugar donor and acceptor peptide, we constructed the ternary complex for each peptide. As described above, after docking the peptide into ApNGT, UDP-Glc was aligned to the active site and the complexes were filtered via two geometric criteria. Then, we selected one representative ternary model from the filtered structures in which the HB between D215 and the +2 Ser/Thr was formed, and subjected it to energy minimization (Figure 2C−E). The resulting complexes show that the UDP moiety can form polar or nonpolar contacts with the ApNGT residues, including K441, H495, Y501, N521, and D525, complying well with the interaction network observed in the crystal structure (Figure S6A).

Since the glucose moiety of UDP-Glc was not observed in the electron density map of former crystal structure, we thus attempted to solve the binary complex of ApNGT with an intact UDP-Glc. To slow down the glycosylation rate, we introduced an additional M218A mutation to wtApNGT (Figure S11A). We were able to obtain a binary crystal structure of ApNGT$^{Q469A/M218A}$ at a resolution of 2.8 Å (Table S8), and the ApNGT structure demonstrates a very similar fold to the previously solved structure (PDB 3Q3H), with a backbone C_α RMSD value of ∼0.450 Å (Figure S6D). Fortunately, we observed clear electron densities of the glucose moiety in our binary complex (PDB 8J30), which matches well with the UDP-Glc structure from our modeled ternary complexes. Thus, this crystallographic structure further validates our theoretical models (Figure 2F,G).

In the constructed ternary complexes, the bound peptide fits into the acceptor binding groove formed by the AAD and GT-1/2 domains. The hydroxyl group of the +2 Ser/Thr establishes a HB with the D215 side chain. In addition, the side-chain methyl group of the +2 Thr forms nonpolar contacts with M218, a highly conserved residue among NGTs (Figure 2C,D). This stabilizing interaction is absent when Ser is at the +2 position (Figure 2E), which might explain why ApNGT has higher N-glycosylation activity toward the N−A−T sequon compared with N−A−S.[27] In addition, the acceptor Asn fits into a pocket surrounded by two conserved residues, Y222 and H371, poised for nucleophilic attack on the anomeric carbon of UDP-Glc. Notably, to fulfill the above critical substrate recognition, the core -NXT/S- region of the incoming peptides with varied sequences and lengths should adopt similar backbone conformations, thereby leading to a broad substrate specificity for ApNGT. The flanking regions, however, are presumably flexible. Former computational work performed by Walvoort and co-workers adopted a much longer peptide substrate (decapeptide) and also observed a highly flexible nature of the binding peptides.[37] Our findings thus provide deep structural insights into the catalysis for ApNGT, guiding further enzyme engineering.

Screening Acceptor Specificity of ApNGT D215 Mutants

To investigate the importance of D215 in recognizing the +2 Ser/Thr in the N-X-S/T sequon, site-saturated mutagenesis was performed at this site. Thus, a total of 19 mutants were successfully constructed and expressed in *E. coli*. The acceptor specificity of D215 mutants was screened by an HPLC-based method against an array of synthetic peptides with variations at the +2 position of the sequence GGNWXT (where X ≠ Cys). In brief, each peptide substrate (1 mM), UDP-Glc (10 mM), and the enzyme (1.25 μM) were mixed in Tris buffer (10 μL, 50 mM, pH 8.0). After incubation at 37 °C for 30 min, reversed-phase high-performance-liquid chromatography (RP-HPLC) was applied to determine the product formation. Accordingly, the synthetic peptide library contained 18 analogues in addition to the optimal GGNWTT sequence.[32] As expected, wtApNGT had strict specificity toward the N-W-S/T sequon, which showed over 4-fold higher activity for Thr than Ser at the +2 position (Figure S7). However, for nearly all D215 mutants, the glycosylation activity drastically decreased. The exception was the D215N mutant, which displayed similar glycosylation activity to wtApNGT against the N-W-T sequon (in 87% and 93% yields, respectively).

Strikingly, around half of D215 mutants exhibited weak or moderate activity toward noncanonical sequons, in which the +2 position was neither Ser nor Thr (Figure S7). For example, the D215R mutant could glycosylate the N-W-A or N-W-I sequon in 19% yield, and the D215G mutant enabled the glycosylation of the N-W-M sequon in 29% yield. Notably, Yeo and co-workers found that the D215A mutant completely lost glycosylation activity toward its natural acceptor substrate, the HMW1 protein.[31] However, in our assays, the D215A mutant still retained 17% activity with the peptide containing the N-W-T sequon. We reason that the optimal peptide GGNWTT is much smaller than the HMW1 protein and the N-glycosylation site can be more easily accessed by ApNGT.

Encouraged by above results, we carried out more detailed investigations on the acceptor specificity of D215 mutants at a higher enzyme concentration. Thus, we performed another round of enzymatic assays with the 20-fold higher enzyme concentration (25 μM). As expected, under the new condition, wtApNGT still maintained strict specificity toward the N-W-S/T sequon (Figures 3A and S8). Both the N-W-S and N-W-T sequons were fully glycosylated by wtApNGT. Notably, the N-W-T sequon was tolerated by most of D215 mutants. The mutation of D215 to an amino acid residue with a relatively small side chain (e.g., Gly, Ala, Ser, Cys, Thr, and Asn) reduced the glycosylation activity but attenuated the stringency of ApNGT toward the +2 position of the N-X-S/T sequon. On the other hand, mutation to a residue with a bulky side chain,

https://doi.org/10.1021/jacsau.3c00214
JACS Au 2023, 3, 2144−2155

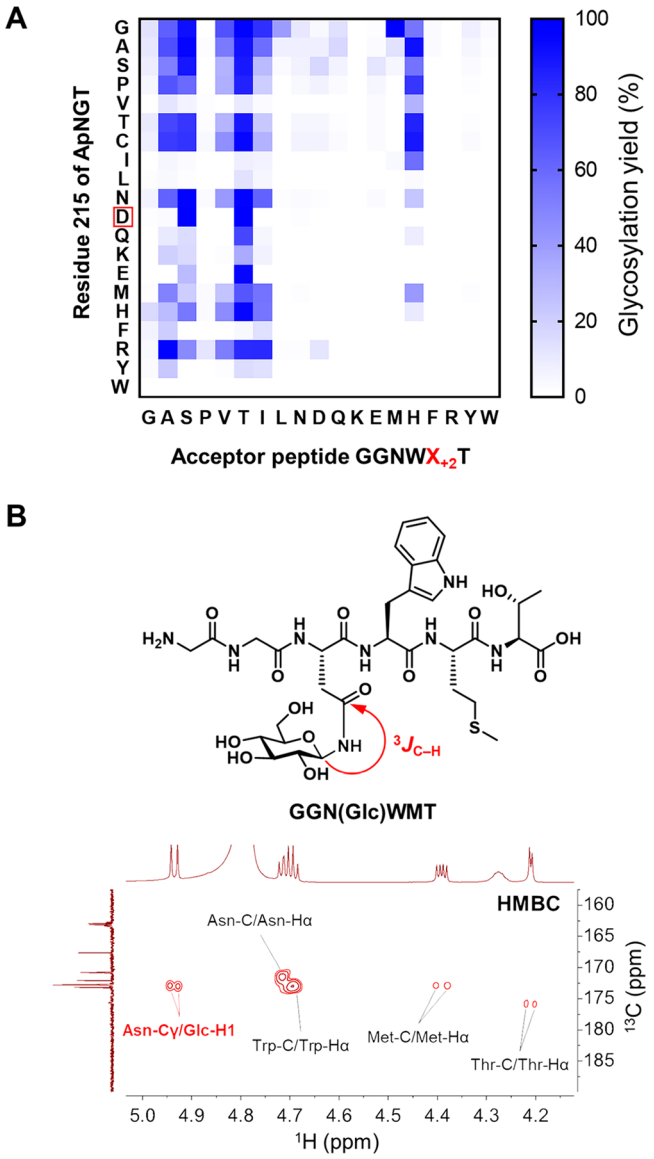


Figure 3. Peptide preferences of wtApNGT and its D215 mutants. (A) Relative N-glycosylation activities were determined at high enzyme concentration (25 μM) against a panel of synthetic peptides with variation at the +2 position of GGNWX$_{+2}$T (also see Figure S8 with reaction yields). (B) ^{1}H−^{13}C correlations in HMBC spectrum of glycopeptide GGN(Glc)WMT. The anomeric proton of glucose (Glc-H1) is correlated with the γ-carbon of Asn (Asn-Cγ), which is indicated in red and highlighted in the chemical structure.

such as Val, Leu, Ile, Phe, Tyr, and Trp, generally gave rise to a more severe reduction in glycosylation activity and reduced the stringency of peptide sequons as well. However, the D215R mutant glycosylated the N-W-A, N-W-T, and N-W-I sequons in excellent yields. Although Asn is similar to Asp in size, the D215N mutant exhibited broader acceptor specificity than wtApNGT.

N-glycosylation products formed by the mutants were further characterized by MS and NMR spectroscopic analysis. Each N-glycosylation product is 162 m/z bigger than its corresponding peptide substrate (Figures S20−36). To directly prove glycosylation of the peptide, we characterized one representative glycosylation product GGN(Glc)WMT using 2D NMR spectroscopy. Using a combination of ^{1}H−^{1}H COSY, ^{1}H−^{13}C HMBC, and ^{1}H−^{13}C HSQC experiments, we successfully confirmed the formation of the product, GGN-(Glc)WMT (see Experimental Section in Supporting Information). In particular, the anomeric proton of the glucose (Glc-H1) in the ^{1}H NMR spectrum appeared at 4.94 ppm with a coupling constant of 9.2 Hz and was correlated with the γ-carbon of Asn (at 172.8 ppm) in the HMBC spectrum (Figure 3B, indicated in red). This data supports the formation of an N-glycosidic bond with the beta stereochemistry.

To further examine whether H272 is involved in substrate recognition, the glycosylation activities of the H272A variant were measured against the synthetic peptides (Figure S11B). Clearly, the H272A mutant maintained good activity toward N-W-S and N-W-T. Although the D215 mutants can recognize peptides with Ala, Val, Ile, or His at the +2 position, the H272A mutant did not show obvious activity toward these noncanonical sequons at either low (1.25 μM) or high (25 μM) enzyme concentration. This finding further suggests that H272 is not involved in the peptide recognition.

Kinetic Analysis of ApNGT D215 Mutants

To further characterize enzymatic properties of the enzyme, kinetic parameters of wtApNGT and certain D215 mutants were measured at a saturated concentration of UDP-Glc by a HPLC-based method (Tables 1 and S6, Figure S9). As expected, the catalytic efficiency of wtApNGT against the N-W-T sequon was 18-fold greater than that against N-W-S (k_{cat}/K_m = 5.84 and 0.32 s^{-1} mM^{-1}, respectively). For D215 mutants, the N-W-T sequon was favored over N-W-S as well, although with much lower k_{cat}/K_m values. Among D215 mutants, the D215N mutant had the highest k_{cat}/K_m value (1.71 s^{-1} mM^{-1}) toward the N-W-T sequon, but was still 3.4-fold lower than that of wtApNGT. The results are consistent with the fact that high concentration of the D215 mutant is required to fully glycosylate the N-W-S/T sequon.

In all cases, mutations of D215 led to a significant decrease in enzyme turnover (smaller k_{cat} value) and substrate binding affinity (larger K_m value). For example, compared to wtApNGT, the K_m value of the D215A mutant toward the N-W-T sequon increased from 1.9 to 13.28 mM, and its k_{cat} value sharply dropped from 11.07 to 0.65 s^{-1}. The results indicate that D215 is essential for recognizing the N-X-S/T consensus sequon by forming HB interactions with the +2 Ser/Thr, as proposed in our computational study. Notably, all tested D215 mutants show weak catalytic efficiency against noncanonical sequons.

Since the above analyses are based on the optimized substrate GGNWTT, we then sought to determine the kinetic parameters of wtApNGT and its D215 mutants against three natural peptides derived from the SARS-CoV-2 spike protein, i.e., VVNIQK, LNEVAK, and KNLNESLID. As expected, wtApNGT only exhibited enzymatic activity toward KNLNESLID that contains a canonical sequon, although with a limited efficiency (1.13 s^{-1} M^{-1}, Tables 1 and S7, Figure S10). In comparison, all three tested D215 mutants had relatively weak activities toward LNEVAK (enzyme efficiency <0.1 s^{-1} M^{-1}, Table S7), and only D215A and D215G were active against VVNIQK (k_{cat}/K_m = 0.26 and 0.33 s^{-1} M^{-1}, respectively; Table 1). These results further confirmed that the D215 mutants could indeed be applied to modify the noncanonical sequons within natural peptides.

https://doi.org/10.1021/jacsau.3c00214
JACS Au 2023, 3, 2144−2155

Table 1. Kinetic Parameters of wtApNGT and Representative D215 Mutants[a]

ApNGT variants	peptide substrates	K_m [mM]	k_{cat} [s^{-1}]	k_{cat}/K_m [s^{-1} mM^{-1}]
WT	GGNWTT	1.90 ± 0.18	11.07 ± 0.38	5.84
WT	GGNWST	16.49 ± 1.04	5.19 ± 0.19	0.32
WT	KNLNESLID			0.00113
D215A	GGNWTT	13.28 ± 0.85	0.65 ± 0.03	0.05
D215A	GGNWST	23.33 ± 2.40	0.70 ± 0.05	0.03
D215A	GGNWVT	11.83 ± 1.47	0.25 ± 0.02	0.02
D215A	VVNIQK			0.00026
D215G	GGNWTT	9.32 ± 0.61	1.44 ± 0.05	0.15
D215G	GGNWST	26.39 ± 2.85	0.67 ± 0.05	0.03
D215G	GGNWMT	11.81 ± 0.95	1.48 ± 0.08	0.13
D215G	VVNIQK			0.00033
D215N	GGNWTT	4.12 ± 0.19	6.95 ± 0.17	1.71
D215N	GGNWST	12.83 ± 0.93	0.76 ± 0.03	0.06
D215N	GGNWAT	12.21 ± 1.93	0.23 ± 0.02	0.02
D215N	KNLNESLID			0.00013
D215R	GGNWTT	8.45 ± 0.99	0.27 ± 0.02	0.03
D215R	GGNWST	17.55 ± 2.50	0.07 ± 0.01	0.004
D215R	GGNWAT	7.50 ± 1.03	0.76 ± 0.05	0.10

[a]The apparent values are determined by varying the concentration of peptide substrate with fixed concentration of UDP-Glc (100 mM). For peptides KNLNESLID and VVNIQK, linear fitting based on $v_0 = c[\text{NGT}] \times k_{cat}/K_m \times [\text{peptide}]$ was used to obtain k_{cat}/K_m values.[32]

Biological Role of the ^{215}DVYM^{218} Motif in ApNGT

In our constructed ternary complex, M218 makes additional hydrophobic contacts with the side-chain methyl group of the +2 Thr. To investigate the functional role of M218 in recognizing the +2 Ser/Thr, M218 was mutated to a small nonpolar residue, Ala, or a polar residue, His. The glycosylation activity toward GGNWTT and GGNWST was quantified by RP-HPLC. The M218A and M218H mutants retained only 4% and 0.9% activity with GGNWTT, respectively, while neither of the mutants showed obvious activity with GGNWST (Figure S11A). These results support the notion that M218 provides additional interactions with the +2 Thr.

Though they are phylogenetically and structurally unrelated to each other, the fact that NGT and OST both recognize the N-X-S/T consensus sequon prompts us to make comparisons about their substrate recognition and catalytic mechanism. OST enzymes have highly conserved WWD and MXXI/DK motifs (the MXXI motif in bacteria, the DK motif in archaea and eukaryotes) responsible for recognizing the N-X-S/T sequon (Figures S6E−G).[17,38−42] In NGT, we wonder if the highly conserved ^{215}DVYM^{218} motif fulfills a similar role (Figure S6H).

Structurally, the side chains of V216 and Y217 are not facing toward the active pocket (Figure S6C), thus, ^{216}VY^{217} may serve as a structural motif rather than a functional motif. We tested the glycosylation activity of mutants V216A and Y217A, respectively, and both mutants retained considerable activity toward GGNWTT and GGNWST (Figure S11A). These results indicate that V216 and Y217 do not participate in recognizing the +2 Ser/Thr, consistent with our computational models.

Determining Functional Roles of Y222 and H371

For OST-catalyzed N-glycosylation, several catalytic mechanisms have been proposed. Imperiali and co-workers proposed the formation of an imidate tautomer intermediate, that is, the carboxamido oxygen of the asparagine side chain could form HBs with the side-chain hydroxyl group and the backbone amide of the +2 Ser/Thr, followed by deprotonation of the amide nitrogen by a basic amino acid residue.[43] Our computational modeling does not support this tautomerization mechanism for NGT, as no such general base exists in the active site of ApNGT. Another proposed mechanism for yeast OST involves the carboxamide twisting to activate the amide nitrogen by D47 and E360 (corresponding to D56 and E319 in PglB, respectively), which are bridged by a metal ion.[38−40,44] However, no such acidic residue pair or metal ion can be found in the ApNGT active site, and ApNGT is known as a metal-independent glycosyltransferase.[21,45,46]

Notably, in our constructed models, we observe that two conserved residues, Y222 and H371, are positioned to establish HBs with the acceptor Asn, implying their potential involvement in the catalytic reaction, e.g., activating the nucleophilic ability of the acceptor Asn. To further examine the functional roles of Y222 and H371, we substituted these two residues with Ala or Phe to abolish the HBs. The Y222A, Y222F, or H371A mutation greatly reduced the glycosylation activity against GGNWTT, and the double mutant of Y222A/H371A was completely inactive (Figure S11A). For GGNWST, the Y222A, Y222F, H371A, and Y222A/H371A mutants all led to loss of activity. These results suggest that Y222 and H371 are likely involved in activating the nucleophilic ability of the acceptor Asn, e.g., via amide twisting by HB formation, to attack the anomeric carbon of UDP-Glc from the opposite direction (S_N2 mechanism).

It is noteworthy that the O-GlcNAc transferase (OGT), an O-glycosyltransferase transferring a GlcNAc moiety from UDP-GlcNAc to the hydroxyl group of Ser/Thr (S/T-OH) in an acceptor peptide, belongs to the same CAZy family as NGT (GT41).[24,47−50] They share a similar GT-B structural fold but with distinctive N-terminal tetratricopeptide repeats (TPR) (Figures S6A−C).[47−50] Crystal structures of human OGT (hOGT) in complex with peptide substrate and UDP-GlcNAc or its analog UDP-5S-GlcNAc have been reported in a number of investigations (Figure S6B).[51−54] The lack of a clearly defined consensus sequence for OGT is consistent with the fact that the peptide forms direct interactions with the

https://doi.org/10.1021/jacsau.3c00214
JACS Au 2023, 3, 2144−2155

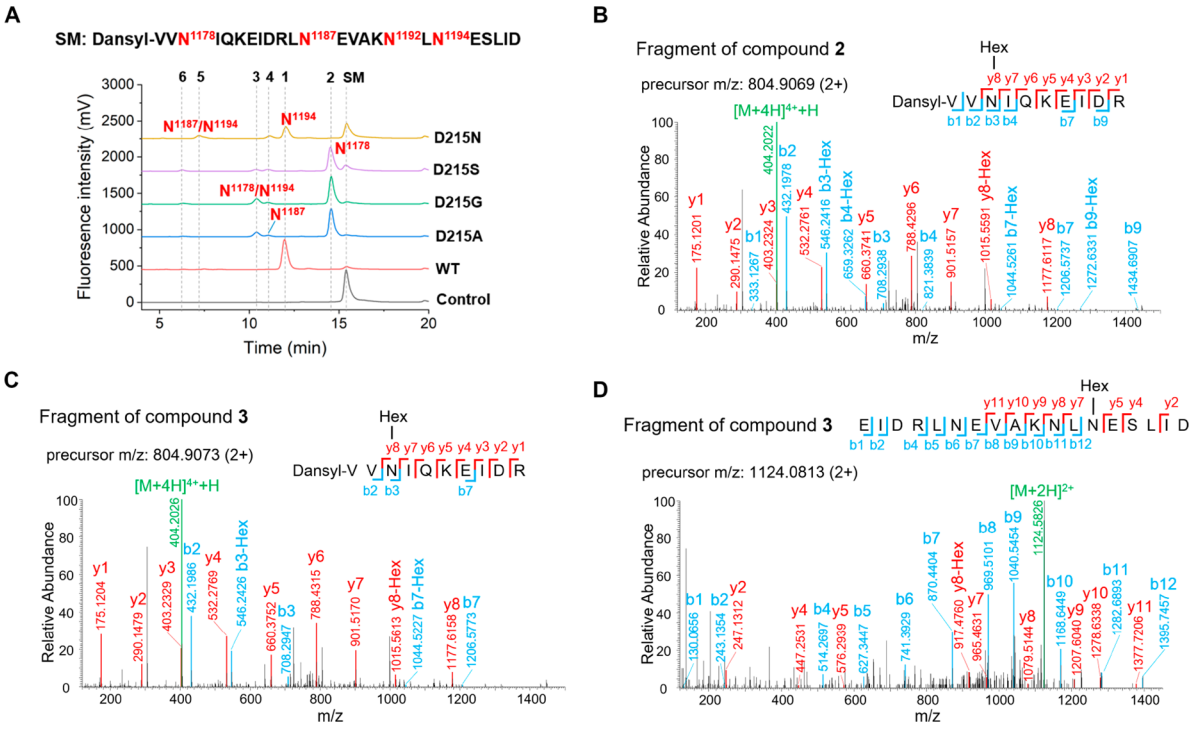


Figure 4. N-Glycosylation of Dansyl-HR2P with wtApNGT or its D215 mutants. (A) N-glycosylation reactions were monitored and quantified by HPLC. Compound **1**: 88% (wtApNGT) and 34% (D215N mutant). Compound **2**: 73% (D215A mutant), 67% (D215G mutant), and 62% (D215S mutant). Compound **3**: 14% (D215A mutant), 16% (D215G mutant), and 5% (D215S mutant). (B–D) Identification of glycosylation sites for compounds **2** and **3** by ESI-MS/MS analysis after trypsin digestion.

OGT residues primarily via its backbone atoms.[51−54] Notably, although the N-Cat ^{495}VHPH^{498} motif in hOGT is structurally equivalent to the ^{215}DVYM^{218} motif of ApNGT (Figure S6C), the former is not involved in peptide recognition.[27,31,51−54] Particularly, the hOGT H498 has been found to either serve as a catalytic base to activate the S/T-OH group or purely participate in stabilizing the sugar donor, while substitution of H498 with Ala could significantly decrease the activity of hOGT.[50−52] Combining our observation that the ^{215}DVYM^{218} motif in ApNGT is responsible for anchoring the +2 Ser/Thr in the N-X-S/T sequon, we thus conclude that the above particular structural motifs in NGT and OGT might dictate their unique peptide acceptor specificities.

Multiglycosylation of Natural Polypeptides by ApNGT D215 Mutants

Although the GGNWTT peptide is an ideal substrate for ApNGT, it is an unnatural substrate. We therefore investigated the glycosylation activity of wtApNGT and all D215 mutants against a natural polypeptide (HR2P), ^{1176}VVNIQKEIDRL-NEVAKNLNESLID^{1199}. This sequence is derived from the heptad repeat 2 (HR2) domain of the SARS-CoV-2 spike protein and contains one canonical N-glycosylation site (N^{1194}ES) and three additional asparagine residues (at sites 1178, 1187, and 1192).[33] The chosen HR2 peptide is responsible for driving the membrane fusion by forming a six-helix bundle (6-HB) structure with HR1 and, therefore, is considered as one of the potential antiviral targets.[55] To facilitate detection by HPLC, we labeled the polypeptide with a dansyl group at the N-terminus (Figure S12).

As expected, the reaction with wtApNGT and UDP-Glc produced a single product bearing only one glucose moiety (compound **1**), as confirmed by HPLC and ESI-MS analysis (Figures 4A, S13, and S37). However, the D215A, D215G, D215N, and D215S mutants formed multiple products as shown in the HPLC spectra (Figures 4A, S13, and S37–42). The products were further characterized by MS/MS analysis to identify the precise N-glycosylation sites (Figures 4B–D and S54–56). The results showed that the D215A, D215G, and D215S mutants glycosylated the N^{1194}ES sequon as well as the N^{1178}IQ sequon, resulting in **2** (monoglycosylation at N^{1178}) and **3** (diglycosylation at N^{1178} and N^{1194}). These three mutants showed higher activity against the noncanonical sequon NIQ than the canonical sequon NES, as **2** was the major product in all reactions. Notably, we also observed a small amount of **4**, resulting from monoglycosylation at N^{1187}. On the other hand, the D215N mutant displayed weak activity against both the N^{1194}ES and N^{1187}EV sequons to produce one diglycosylation product **5**, and two monoglycosylation products **1** and **4**. In this case, product **1**, the canonical N-glycosylation product, was a dominant product and most of starting material was unreacted.

To explore whether the dansyl group at the N-terminus might influence the acceptor specificity of wtApNGT and D215 mutants, we carried out enzymatic assays with the tag-free HR2P. The results were consistent with those of the labeled polypeptide (Figure S14).

Moreover, we further applied wtApNGT and the D215A mutant to the natural 37aa-polypeptide ^{1171}GINASVVNIQ-KEIDRLNEVAKNLNESLIDLQELGKYE^{1207}, an extended version of the above 24aa-polypeptide. As expected, wtApNGT only glycosylated the two canonical sequons N^{1173}AS and N^{1194}ES, resulting in one major diglycosylation product **P3** (Figures S15, S50, S60, and S61) and one minor monoglycosylation product **P4** (Figures S15, S51, and S62). Again, the D215A mutant showed a preference for the

https://doi.org/10.1021/jacsau.3c00214
JACS Au 2023, 3, 2144−2155

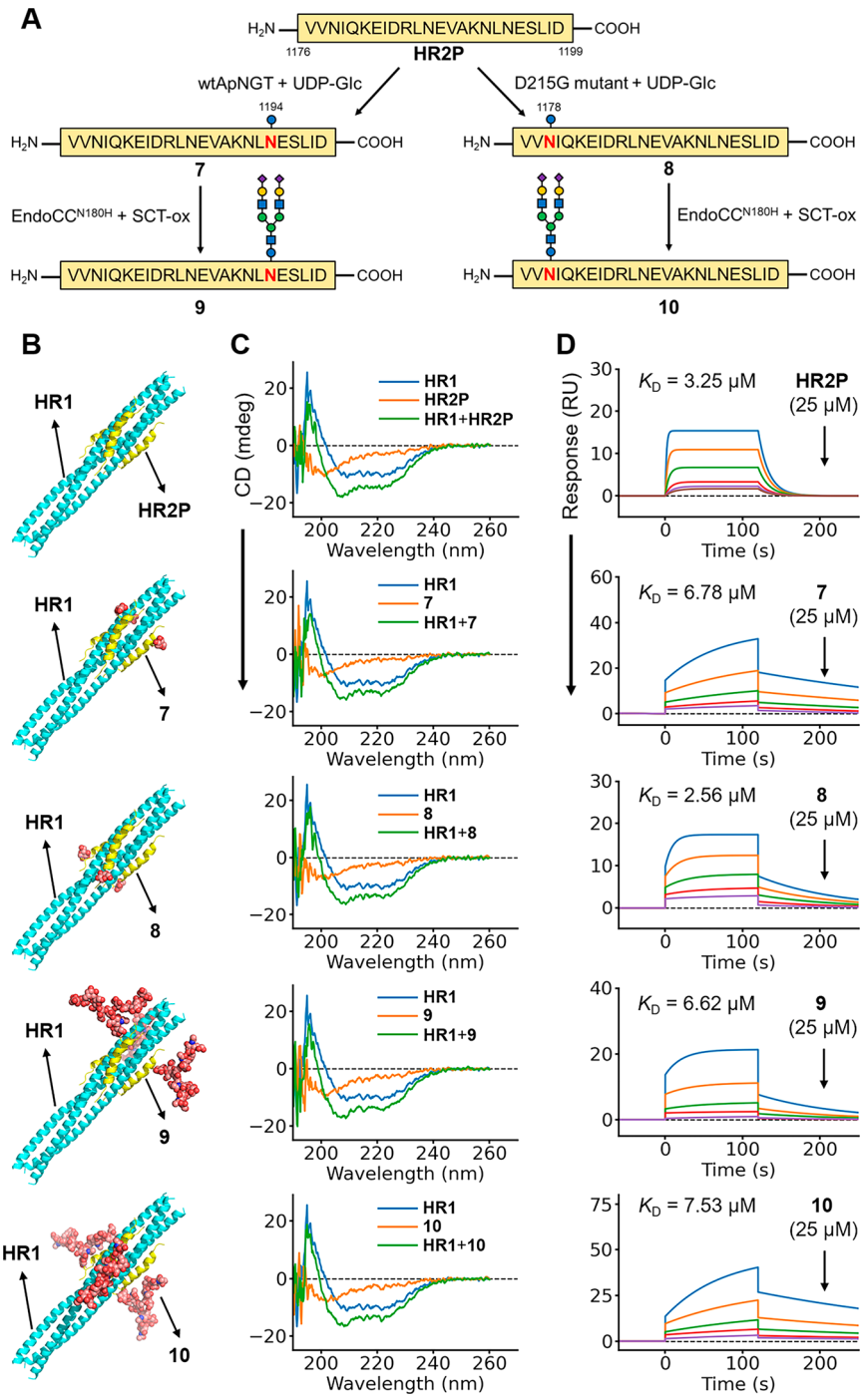


Figure 5. Investigation of the biological importance of N-glycosylation on HR2 interacting with HR1. (A) Synthetic route of compounds **7**−**10**. (B) Molecular modeling of 6-HB structures with N-glycosylated peptides. (C) CD spectra of HR1 and HR2P/HR2P N-glycopeptides (HR2P, **7**−**10** from top to bottom panel). (D) SPR sensorgrams of the interactions between the immobilized HR1 and HR2P/HR2P N-glycopeptides (HR2P, **7**−**10** from top to bottom panel) at 2-fold serial dilutions.

noncanonical sequon NIQ as it did in the 24aa-polypeptide, producing one major product **P2** (monoglycosylation at N^{1178}, see Figures S15, S49, and S59) and one minor product **P1** (diglycosylation at N^{1173} and N^{1178}, see Figures S15, S48, and S58). Notably, the D215A mutant glycosylated the $N^{1173}AS$ sequon instead of $N^{1194}ES$.

To elucidate the sequon selectivity of D215 mutants observed in Figures 4A and S15, we built a three-dimensional structure for the above 37aa-polypeptide using AlphaFold2 (Figure S6I).[56] The results show that the noncanonical sequon $N^{1178}IQ$ is located in a highly flexible region, allowing it easier access into the active site of D215 mutants. Likewise, the canonical sequon $N^{1173}AS$ could also be well recognized by both wtApNGT and D215A owing to its flexible nature. Moreover, $N^{1194}ES$ lies at the junction of one α-helix and random-coil regions, so it can still be catalyzed by wtApNGT or the D215N variant. In contrast, both noncanonical sequons $N^{1187}EV$ and $N^{1192}LN$ are located in the middle of one α-helix region; therefore, the rigid helical structure likely hampers the access of the sequons into the ApNGT active site. These

https://doi.org/10.1021/jacsau.3c00214
JACS Au 2023, 3, 2144−2155

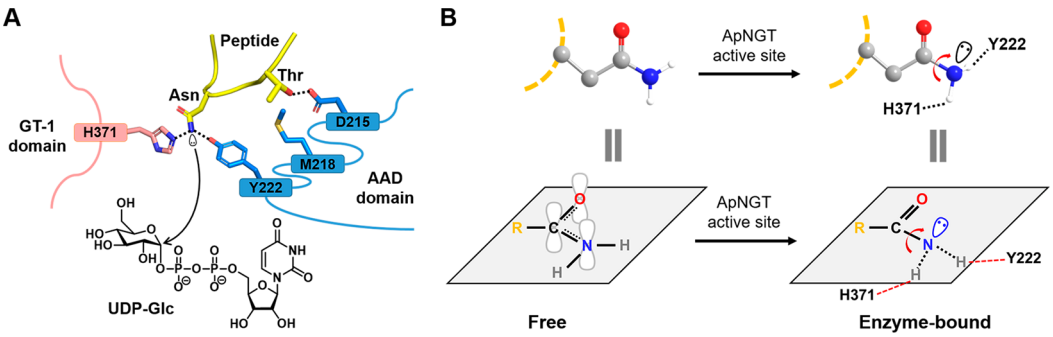


Figure 6. (A) Proposed N-glycosylation mechanism of ApNGT. D215 and M218 from the conserved ^{215}DVYM^{218} motif are proposed to be responsible for the recognition of the consensus sequon N-X-S/T. (B) Y222 and H371 residues are presumably involved in activating the nucleophilic ability of the acceptor Asn to attack the anomeric carbon of UDP-Glc. While bound to ApNGT, the amido group of Asn might form hydrogen bonds with Y222 and H371, leading to a twisting of the C−N bond and therefore to loss of conjugation of the nitrogen lone pair with the carbonyl group, which increases the nucleophilicity of nitrogen.

results suggest that the secondary structure of a glycosylation site could impose significant impact on the ApNGT activity.

Biological Roles of N-Glycosylation on HR2 in HR1 Recognition

It is well-known that glycosylation greatly influences the size, charge, solubility, and stability of a given protein to prevent rapid clearance and degradation from circulation. Meanwhile, certain glycans serve as ligands for cell-surface lectins (carbohydrate-binding receptor), mediating protein trafficking, pathogen recognition and cell signaling.[57] Thus, the glyco-engineering approach has been widely applied to reduce side effects and improve biophysical/biological properties of therapeutic proteins.[58]

In the postfusion conformation of the SARS-CoV-2 S protein, the HR2 domain forms a stable 6-HB structure with HR1, thereby bringing viral and cellular membranes close together for fusion.[59] Notably, there is a canonical N-glycosylation site (N^{1194}) located at the fusion core region of HR2 (aa1179−1197) and a noncanonical N-glycan site at N^{1178}. To investigate the biological importance of the N-glycosylation at N^{1194} and N^{1178} in forming the 6-HB structure between HR2 and HR1, we prepared the corresponding mono-N-glycosylated HR2P by treating HR2P with wtApNGT or D215G mutant and UDP-Glc, leading to products **7** (at N^{1194}) and **8** (at N^{1178}), respectively. Then, the glycosynthase Endo CC^{N180H} was used to transfer a human-like biantennary glycan to **7** and **8**, resulting in a canonical N-glycosylation product **9** and a noncanonical N-glycosylation product **10**, respectively, in good yields (Figure 5A). These results indicated that combination of ApNGT variants and ENGase mutants could site-selectively install complex N-glycans within a target polypeptide bearing multiple canonical and noncanonical N-glycosylation sites.[25]

With the N-glycopeptides **7**−**10** in hand, we set up experiments to characterize their biophysical/biological properties and explore their capability of 6-HB formation with HR1.[60] Upon mixing the recombinant HR1 with HR2P or **7**−**10**, an additional peak was observed in all size exclusion chromatography (SEC) analyses (Figure S16), suggesting that all tested compounds were able to form 6-HBs with HR1. These results were further supported by circular dichroism (CD) spectroscopic analyses, which clearly showed that recombinant HR1 had a relatively low content of α-helicity (with minima at 208 and 222 nm) and HR2P or **7**−**10** had random coil structures as expected (Figure 5C). When HR1 was treated with HR2P or **7**−**10** in equimolar concentrations, the α-helicity was increased. These findings suggested that the interactions between HR1 and HR2P N-glycopeptides could induce and stabilize α-helical conformation.

To quantitatively profile the interactions between HR1 and the glycosylated HR2P, we performed surface plasmon resonance (SPR) studies. Because the glycans were installed at the end of HR2P molecule, His-tagged HR1 was immobilized on an anti-His sensor chip and HR2P glycopeptides served as the analytes. Our SPR results showed that HR2P was quickly bound to HR1 and the N-glycosylation at N^{1194} (**7** and **9**) slightly decreased the binding affinity, leading to the increase in the equilibrium dissociation constant (K_D) values from 3.25 μM to 6.78 and 6.62 μM, respectively (Figure 5D). Due to the location at the edge of the fusion core (Figure 5B), the installation of a glucose at the N^{1178} of HR2P (**8**) had an inconsiderable impact on the binding affinity to HR1 (K_D = 2.56 μM). However, further introduction of bulky sialylated complex type glycan (SCT) at N^{1178} (**10**) resulted in weaker interaction with HR1 (Figure 5D, bottom panel). Our molecular modeling suggests that both N^{1178} and N^{1194} N-glycosylation sites are fully exposed in the solvents but not buried within the HR1−HR2 interface. Therefore, the N-glycosylation likely contributes to positioning HR2 in a suitable orientation for the HR1 recognition, and both the site and the size of N-glycosylation can influence the 6-HB formation (Figure 5B). Future attempts may extend to examining the impact of a broader range of N-glycans on the HR1−HR2 interaction. This work sheds light on the biological roles of the N-glycosylation at HR2 and guides the design of novel antiviral inhibitors.

Altogether, we have shown these NGT mutants are useful tools to provide novel glycopeptides for investigating the biological importance of a specific N-glycosites, facilitating in-depth understanding of N-glycosylation and its myriad of functions.[61] Moreover, the newly installed natural N-glycan can mask the protease cleavage site, while the synthetic glycan with a reactive group (e.g., azide, alkyne) provides a handle for further functionalization, such as conjugation with poly(ethylene glycol) (PEG) chains to prolong the half-life of a protein in the bloodstream or attachment with toxins for drug delivery.[5,62] Finally, as wtApNGT and its variants examined so far fail to transfer a GlcNAc moiety to a canonical N-glycosite, our current work also guides further attempts to gain the

https://doi.org/10.1021/jacsau.3c00214
JACS Au 2023, 3, 2144−2155

GlcNAc-transfer ability for ApNGT via bioengineering, e.g., bump-and-hole technique.[22,34]

CONCLUSION

In summary, by combining computational modeling and biochemical assays, we propose the overall N-glycosylation mechanism for ApNGT and pinpoint the key structural motif in ApNGT responsible for regulating the acceptor substrate specificity. The constructed models were further supported by our crystal structure of the ApNGT binary complex. One highly conserved motif in ApNGT, ^{215}DVYM^{218}, was found to recognize the +2 Ser/Thr site in the N-X-S/T sequon (Figure 6). In addition, the acceptor Asn is situated between Y222 and H371 which might take part in activating its nucleophilic ability, e.g., via amide twisting (Figure 6B). Importantly, the D215 mutants exhibit good N-glycosylation activity for Asn residues in noncanonical sequons, which allowed us to install N-glycans to a polypeptide bearing multiple Asn residues. Furthermore, we explored the biological role of the N-glycosylation on HR2 in the 6-HB formation during SARS-CoV-2 entry. Our results show the N-glycosylation can stabilize the 6-HB structure although the site and size of the N-glycan can affect the binding affinity of HR2 to HR1. Our findings provide new insights into the substrate recognition and activating mechanism of NGT, deepen the understanding of NGT-mediated N-glycosylation, and present a novel tool for chemo-enzymatic synthesis of well-defined N-glycoproteins/glycopeptides for various glycobiology studies.

MATERIALS AND METHODS

All conventional chemical reagents and solvents were purchased from commercial suppliers and used without further purification unless otherwise stated. The genes were ordered from General Biosystems (Anhui, China). NMR experiments were performed with a Bruker AVANCE NEO (700 MHz) instrument. HPLC experiments were performed by Shimadzu LC-2010A with a PDA or fluorescence detector. ESI-MS/MS analyses were performed on a Q Exactive mass spectrometer (Thermo Fisher Scientific). Computations were run on the π 2.0 cluster supported by the Center for High Performance Computing at Shanghai Jiao Tong University.

Methods including cloning, expression, and purification of ApNGT variants and HR1, molecular modeling, multiple sequence alignment, crystallization and structure determination, enzymatic property characterization, glycopeptide synthesis and characterization, ESI-MS and MS/MS analyses, and 6-HB formation determination (SEC-HPLC, CD spectroscopy and SPR experiments) are described in full detail in the Supporting Information.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacsau.3c00214.

Complete experimental procedures and methods, including supplementary tables and figures, primer sequences, reaction conditions, kinetic analysis curves, mass spectra, HPLC chromatograms, and NMR spectra (PDF)

Accession Codes

Atomic coordinates of the crystal structure have been deposited in the Protein Data Bank (PDB) with the entry ID 8J30.

AUTHOR INFORMATION

Corresponding Authors

Lin-Tai Da − *Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai 200240, China;* orcid.org/0000-0002-1589-5754; Email: darlt@sjtu.edu.cn

Wenjie Peng − *Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai 200240, China;* orcid.org/0000-0001-5093-7115; Email: wenjiep@sjtu.edu.cn

Authors

Zhiqiang Hao − *Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai 200240, China*

Qiang Guo − *Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai 200240, China*

Yuanyuan Feng − *State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China*

Zihan Zhang − *Shanghai Key Laboratory of Chemical Assessment and Sustainability, School of Chemical Science & Engineering, Tongji University, Shanghai 200092, China*

Tiantian Li − *Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai 200240, China;* orcid.org/0000-0003-0414-8064

Zhixin Tian − *Shanghai Key Laboratory of Chemical Assessment and Sustainability, School of Chemical Science & Engineering, Tongji University, Shanghai 200092, China;* orcid.org/0000-0002-2877-8282

Jianting Zheng − *State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China;* orcid.org/0000-0003-1250-3556

Complete contact information is available at: https://pubs.acs.org/10.1021/jacsau.3c00214

Author Contributions

#Z.H. and Q.G. contributed equally. Z.H., Q.G., L.D., and W.P. designed the project. Z.H. performed the computational study. Q.G. performed the site-directed mutagenesis, assayed the glycosylation activity, and synthesized the glycopeptides. Y.F. and J.Z. carried out crystallization and structure determination. Z.Z. and Z.T. performed MS/MS analysis. T.L. performed NMR experiments. Z.H., Q.G., L.D., and W.P. wrote the manuscript, and all authors reviewed and edited the manuscript. CRediT: **Zhiqiang Hao** investigation, methodology, writing-original draft; **Qiang Guo** investigation, methodology, writing-original draft; **Yuanyuan Feng** investigation, methodology, writing-review & editing; **Zihan Zhang** investigation, methodology, writing-review & editing; **Tiantian Li** methodology; **Zhixin Tian** supervision, writing-review & editing; **Jianting Zheng** supervision, writing-review & editing; **Lin-Tai Da** conceptualization, project administration, writing-original draft, writing-review & editing; **Wenjie Peng** conceptualiza-

https://doi.org/10.1021/jacsau.3c00214
JACS Au 2023, 3, 2144−2155

tion, project administration, supervision, writing-original draft, writing-review & editing.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The work was supported by National Natural Science Foundation of China (91853121, 21977066, and 22177069 to W.P.; 22177072 to L.D.), Shanghai Pilot Program for Basic Research-Shanghai Jiao Tong University (21TQ1400210 to W.P.), Medical-Engineering Interdisciplinary Research Foundation of Shanghai Jiao Tong University (YG2022ZD001 to W.P.), and Natural Science Foundation of Shanghai (23ZR1432500 to W.P.).

■ REFERENCES

(1) *Essentials of Glycobiology*; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2022.

(2) Moremen, K. W.; Tiemeyer, M.; Nairn, A. V. Vertebrate protein glycosylation: diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol.* **2012**, *13* (7), 448−462.

(3) Schjoldager, K. T.; Narimatsu, Y.; Joshi, H. J.; Clausen, H. Global view of human protein glycosylation pathways and functions. *Nat. Rev. Mol. Cell Biol.* **2020**, *21* (12), 729−749.

(4) Schwarz, F.; Aebi, M. Mechanisms and principles of N-linked protein glycosylation. *Curr. Opin. Struct. Biol.* **2011**, *21* (5), 576−582.

(5) Zeng, Y.; Tang, F.; Shi, W.; Dong, Q.; Huang, W. Recent advances in synthetic glycoengineering for biological applications. *Curr. Opin. Biotechnol.* **2022**, *74*, 247−255.

(6) Leader, B.; Baca, Q. J.; Golan, D. E. Protein therapeutics: a summary and pharmacological classification. *Nat. Rev. Drug Discovery* **2008**, *7* (1), 21−39.

(7) Dimitrov, D. S. Therapeutic Proteins. In *Therapeutic Proteins: Methods and Protocols*; Voynov, V., Caravella, J. A., Eds.; Humana Press: Totowa, NJ, 2012; pp 1−26.

(8) Boune, S.; Hu, P. S.; Epstein, A. L.; Khawli, L. A. Principles of N-linked glycosylation variations of IgG-based therapeutics: pharmacokinetic and functional considerations. *Antibodies* **2020**, *9* (2), 22.

(9) Jefferis, R. Glycosylation as a strategy to improve antibody-based therapeutics. *Nat. Rev. Drug Discov* **2009**, *8* (3), 226−234.

(10) Van Landuyt, L.; Lonigro, C.; Meuris, L.; Callewaert, N. Customized protein glycosylation to improve biopharmaceutical function and targeting. *Curr. Opin. Biotechnol.* **2019**, *60*, 17−28.

(11) *Total Chemical Synthesis of Proteins*; Brik, A., Dawson, P., Liu, L., Eds.; Wiley: 2021.

(12) Delobel, A. Glycosylation of Therapeutic Proteins: A Critical Quality Attribute. In *Mass Spectrometry of Glycoproteins: Methods and Protocols*; Delobel, A., Ed.; Springer US: New York, NY, 2021; pp 1−21.

(13) Fairbanks, A. J. Chemoenzymatic synthesis of glycoproteins. *Curr. Opin. Chem. Biol.* **2019**, *53*, 9−15.

(14) Li, C.; Wang, L. X. Chemoenzymatic methods for the synthesis of glycoproteins. *Chem. Rev.* **2018**, *118* (17), 8359−8413.

(15) Carlo, U.; Yasuhiro, K. Recent advances in the chemical synthesis of N-linked glycoproteins. *Curr. Opin. Chem. Biol.* **2018**, *46*, 130−137.

(16) Kightlinger, W.; Warfel, K. F.; DeLisa, M. P.; Jewett, M. C. Synthetic glycobiology: parts, systems, and applications. *ACS Synth. Biol.* **2020**, *9* (7), 1534−1562.

(17) Gerber, S.; Lizak, C.; Michaud, G.; Bucher, M.; Darbre, T.; Aebi, M.; Reymond, J. L.; Locher, K. P. Mechanism of bacterial oligosaccharyltransferase: in vitro quantification of sequon binding and catalysis. *J. Biol. Chem.* **2013**, *288* (13), 8849−8861.

(18) Ollis, A. A.; Zhang, S.; Fisher, A. C.; DeLisa, M. P. Engineered oligosaccharyltransferases with greatly relaxed acceptor-site specificity. *Nat. Chem. Biol.* **2014**, *10* (10), 816−822.

(19) Jaroentomeechai, T.; Stark, J. C.; Natarajan, A.; Glasscock, C. J.; Yates, L. E.; Hsu, K. J.; Mrksich, M.; Jewett, M. C.; DeLisa, M. P. Single-pot glycoprotein biosynthesis using a cell-free transcription-translation system enriched with glycosylation machinery. *Nat. Commun.* **2018**, *9* (1), 2686.

(20) Schwarz, F.; Fan, Y. Y.; Schubert, M.; Aebi, M. Cytoplasmic N-glycosyltransferase of Actinobacillus pleuropneumoniae is an inverting enzyme and recognizes the NX(S/T) consensus sequence. *J. Biol. Chem.* **2011**, *286* (40), 35267−35274.

(21) Choi, K. J.; Grass, S.; Paek, S.; St. Geme, J. W.; Yeo, H. J. The Actinobacillus pleuropneumoniae HMW1C-Like glycosyltransferase mediates N-linked glycosylation of the Haemophilus influenzae HMW1 adhesin. *PLoS One* **2010**, *5* (12), No. e15888.

(22) Liu, Z.; Li, K.; Liu, X.; Zhao, J.; Yu, Y.; Wang, L.; Kong, Y.; Chen, M. Production of microhomogeneous glycopeptide by a mutated NGT according FuncLib with unique sugar as substrate. *Enzyme Microb. Technol.* **2022**, *154*, No. 109949.

(23) Naegeli, A.; Neupert, C.; Fan, Y.-Y.; Lin, C.-W.; Poljak, K.; Papini, A. M.; Schwarz, F.; Aebi, M. Molecular analysis of an alternative N-glycosylation machinery by functional transfer from Actinobacillus pleuropneumoniae to Escherichia coli. *J. Biol. Chem.* **2014**, *289* (4), 2170−2179.

(24) Cantarel, B. L.; Coutinho, P. M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res.* **2009**, *37*, D233−D238.

(25) Lin, L.; Kightlinger, W.; Prabhu, S. K.; Hockenberry, A. J.; Li, C.; Wang, L. X.; Jewett, M. C.; Mrksich, M. Sequential glycosylation of proteins with substrate-specific N-glycosyltransferases. *ACS Central Sci.* **2020**, *6* (2), 144−154.

(26) Lomino, J. V.; Naegeli, A.; Orwenyo, J.; Amin, M. N.; Aebi, M.; Wang, L. X. A two-step enzymatic glycosylation of polypeptides with complex N-glycans. *Bioorg. Med. Chem.* **2013**, *21* (8), 2262−2270.

(27) Naegeli, A.; Michaud, G.; Schubert, M.; Lin, C. W.; Lizak, C.; Darbre, T.; Reymond, J. L.; Aebi, M. Substrate specificity of cytoplasmic N-glycosyltransferase. *J. Biol. Chem.* **2014**, *289* (35), 24521−24532.

(28) Kightlinger, W.; Duncker, K. E.; Ramesh, A.; Thames, A. H.; Natarajan, A.; Stark, J. C.; Yang, A.; Lin, L.; Mrksich, M.; DeLisa, M. P.; Jewett, M. C. A cell-free biosynthesis platform for modular construction of protein glycosylation pathways. *Nat. Commun.* **2019**, *10* (1), 5404.

(29) Ahangama Liyanage, L.; Harris, M. S.; Cook, G. A. In vitro glycosylation of membrane proteins using N-glycosyltransferase. *ACS Omega* **2021**, *6* (18), 12133−12142.

(30) Prabhu, S. K.; Yang, Q.; Tong, X.; Wang, L.-X. Exploring a combined Escherichia coli-based glycosylation and in vitro trans-glycosylation approach for expression of glycosylated interferon alpha. *Bioorg. Med. Chem.* **2021**, *33*, No. 116037.

(31) Kawai, F.; Grass, S.; Kim, Y.; Choi, K.-J.; St. Geme, J. W.; Yeo, H.-J. Structural insights into the glycosyltransferase activity of the Actinobacillus pleuropneumoniae HMW1C-like protein. *J. Biol. Chem.* **2011**, *286* (44), 38546−38557.

(32) Kightlinger, W.; Lin, L.; Rosztoczy, M.; Li, W. H.; 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.

(33) Watanabe, Y.; Allen, J. D.; Wrapp, D.; McLellan, J. S.; Crispin, M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* **2020**, *369* (6501), 330−333.

(34) Xu, Y. Y.; Wu, Z. G.; Zhang, P. R.; Zhu, H.; Zhu, H. L.; Song, Q. T.; Wang, L.; Wang, F. X.; Wang, P. G.; Cheng, J. S. A novel enzymatic method for synthesis of glycopeptides carrying natural eukaryotic N-glycans. *Chem. Commun.* **2017**, *53* (65), 9075−9077.

(35) Song, Q. T.; Wu, Z. G.; Fan, Y. Y.; Song, W. R.; Zhang, P. R.; Wang, L.; Wang, F. X.; Xu, Y. Y.; Wang, P. G.; Cheng, J. S. Production

https://doi.org/10.1021/jacsau.3c00214
JACS Au 2023, 3, 2144−2155

of homogeneous glycoprotein with multisite modifications by an engineered N-glycosyltransferase mutant. *J. Biol. Chem.* **2017**, *292* (21), 8856−8863.
(36) Raveh, B.; London, N.; Zimmerman, L.; Schueler-Furman, O. Rosetta FlexPepDock ab-initio: simultaneous folding, docking and refinement of peptides onto their receptors. *PLoS One* **2011**, *6* (4), No. e18934.
(37) Yakovlieva, L.; Ramírez-Palacios, C.; Marrink, S. J.; Walvoort, M. T. C. Semiprocessive hyperglycosylation of adhesin by bacterial protein N-glycosyltransferases. *ACS Chem. Biol.* **2021**, *16* (1), 165−175.
(38) Wild, R.; Kowal, J.; Eyring, J.; Ngwa, E. M.; Aebi, M.; Locher, K. P. Structure of the yeast oligosaccharyltransferase complex gives insight into eukaryotic N-glycosylation. *Science* **2018**, *359* (6375), 545−550.
(39) Lizak, C.; Gerber, S.; Numao, S.; Aebi, M.; Locher, K. P. X-ray structure of a bacterial oligosaccharyltransferase. *Nature* **2011**, *474* (7351), 350−355.
(40) Matsumoto, S.; Shimada, A.; Nyirenda, J.; Igura, M.; Kawano, Y.; Kohda, D. Crystal structures of an archaeal oligosaccharyltransferase provide insights into the catalytic cycle of N-linked protein glycosylation. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (44), 17868−17873.
(41) Matsumoto, S.; Taguchi, Y.; Shimada, A.; Igura, M.; Kohda, D. Tethering an N-glycosylation sequon-containing peptide creates a catalytically competent oligosaccharyltransferase complex. *Biochemistry* **2017**, *56* (4), 602−611.
(42) Ramirez, A. S.; de Capitani, M.; Pesciullesi, G.; Kowal, J.; Bloch, J. S.; Irobalieva, R. N.; Reymond, J. L.; Aebi, M.; Locher, K. P. Molecular basis for glycan recognition and reaction priming of eukaryotic oligosaccharyltransferase. *Nat. Commun.* **2022**, *13* (1), 7296.
(43) Imperiali, B.; Shannon, K.; Unno, M.; Rickert, K. Mechanistic proposal for asparagine-linked glycosylation. *J. Am. Chem. Soc.* **1992**, *114* (20), 7944−7945.
(44) Bai, L.; Wang, T.; Zhao, G. P.; Kovach, A.; Li, H. L. The atomic structure of a eukaryotic oligosaccharyltransferase complex. *Nature* **2018**, *555* (7696), 328−333.
(45) Kong, Y.; Li, J.; Hu, X. Y.; Wang, Y. G.; Meng, Q. Y.; Gu, G. F.; Wang, P. G.; Chen, M. N-Glycosyltransferase from Aggregatibacter aphrophilus synthesizes glycopeptides with relaxed nucleotide-activated sugar donor selectivity. *Carbohydr. Res.* **2018**, *462*, 7−12.
(46) Meng, Q. Y.; Li, K.; Rong, Y. H.; Wu, Q. Z.; Zhang, X. L.; Kong, Y.; Chen, M. Probing peptide substrate specificities of N-glycosyltranferase isoforms from different bacterial species. *Carbohydr. Res.* **2019**, *473*, 82−87.
(47) Clarke, A. J.; Hurtado-Guerrero, R.; Pathak, S.; Schüttelkopf, A. W.; Borodkin, V.; Shepherd, S. M.; Ibrahim, A. F. M.; van Aalten, D. M. F. Structural insights into mechanism and specificity of O-GlcNAc transferase. *EMBO J.* **2008**, *27* (20), 2780−2788.
(48) Jínek, M.; Rehwinkel, J.; Lazarus, B. D.; Izaurralde, E.; Hanover, J. A.; Conti, E. The superhelical TPR-repeat domain of O-linked GlcNAc transferase exhibits structural similarities to importin α. *Nat. Struct. Mol. Biol.* **2004**, *11* (10), 1001−1007.
(49) Martinez-Fleites, C.; Macauley, M. S.; He, Y.; Shen, D. L.; Vocadlo, D. J.; Davies, G. J. Structure of an O-GlcNAc transferase homolog provides insight into intracellular glycosylation. *Nat. Struct. Mol. Biol.* **2008**, *15* (7), 764−765.
(50) Lazarus, M. B.; Nam, Y.; Jiang, J.; Sliz, P.; Walker, S. Structure of human O-GlcNAc transferase and its complex with a peptide substrate. *Nature* **2011**, *469* (7331), 564−567.
(51) Schimpl, M.; Zheng, X.; Borodkin, V. S.; Blair, D. E.; Ferenbach, A. T.; Schüttelkopf, A. W.; Navratilova, I.; Aristotelous, T.; Albarbarawi, O.; Robinson, D. A.; Macnaughtan, M. A.; van Aalten, D. M. F. O-GlcNAc transferase invokes nucleotide sugar pyrophosphate participation in catalysis. *Nat. Chem. Biol.* **2012**, *8* (12), 969−974.
(52) Lazarus, M. B.; Jiang, J. Y.; Gloster, T. M.; Zandberg, W. F.; Whitworth, G. E.; Vocadlo, D. J.; Walker, S. Structural snapshots of the reaction coordinate for O-GlcNAc transferase. *Nat. Chem. Biol.* **2012**, *8* (12), 966−968.
(53) Lazarus, M. B.; Jiang, J.; Kapuria, V.; Bhuiyan, T.; Janetzko, J.; Zandberg, W. F.; Vocadlo, D. J.; Herr, W.; Walker, S. HCF-1 is cleaved in the active site of O-GlcNAc transferase. *Science* **2013**, *342* (6163), 1235−1239.
(54) Pathak, S.; Alonso, J.; Schimpl, M.; Rafie, K.; Blair, D. E.; Borodkin, V. S.; Schüttelkopf, A. W.; Albarbarawi, O.; van Aalten, D. M. F. The active site of O-GlcNAc transferase imposes constraints on substrate sequence. *Nat. Struct. Mol. Biol.* **2015**, *22* (9), 744−750.
(55) de Vries, R. D.; Schmitz, K. S.; Bovier, F. T.; Predella, C.; Khao, J.; Noack, D.; Haagmans, B. L.; Herfst, S.; Stearns, K. N.; Drew-Bear, J.; Biswas, S.; Rockx, B.; McGill, G.; Dorrello, N. V.; Gellman, S. H.; Alabi, C. A.; de Swart, R. L.; Moscona, A.; Porotto, M. Intranasal fusion inhibitory lipopeptide prevents direct-contact SARS-CoV-2 transmission in ferrets. *Science* **2021**, *371* (6536), 1379−1382.
(56) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583−589.
(57) Varki, A. Biological roles of glycans. *Glycobiology* **2017**, *27* (1), 3−49.
(58) Ma, B.; Guan, X. Y.; Li, Y. H.; Shang, S. Y.; Li, J.; Tan, Z. P. Protein glycoengineering: an approach for improving protein properties. *Front. Chem.* **2020**, *8*, 622.
(59) Shi, W.; Cai, Y.; Zhu, H.; Peng, H.; Voyer, J.; Rits-Volloch, S.; Cao, H.; Mayer, M. L.; Song, K.; Xu, C.; Lu, J.; Zhang, J.; Chen, B. Cryo-EM structure of SARS-CoV-2 postfusion spike in membrane. *bioRxiv* 2022, https://www.biorxiv.org/content/10.1101/2022.12.05.519151v1.
(60) Xia, S.; Zhu, Y.; Liu, M.; Lan, Q.; Xu, W.; Wu, Y.; Ying, T.; Liu, S.; Shi, Z.; Jiang, S.; Lu, L. Fusion mechanism of 2019-nCoV and fusion inhibitors targeting HR1 domain in spike protein. *Cell. Mol. Immunol.* **2020**, *17* (7), 765−767.
(61) Griffin, M. E.; Hsieh-Wilson, L. C. Tools for mammalian glycoscience research. *Cell* **2022**, *185* (15), 2657−2677.
(62) Pelegri-O'Day, E. M.; Lin, E. W.; Maynard, H. D. Therapeutic protein-polymer conjugates: advancing beyond PEGylation. *J. Am. Chem. Soc.* **2014**, *136* (41), 14323−14332.

https://doi.org/10.1021/jacsau.3c00214
JACS Au 2023, 3, 2144−2155