THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 289, NO. 35, pp. 24521–24532, August 29, 2014

Substrate Specificity of Cytoplasmic *N*-Glycosyltransferase*[S]

Received for publication, May 22, 2014, and in revised form, June 16, 2014 Published, JBC Papers in Press, June 24, 2014, DOI 10.1074/jbc.M114.579326

Andreas Naegeli[‡1], Gaëlle Michaud[§], Mario Schubert[¶], Chia-Wei Lin[‡], Christian Lizak[¶2], Tamis Darbre[§], Jean-Louis Reymond[§], and Markus Aebi[‡3]

From the [‡]Department of Biology, Institute of Microbiology, ETH Zurich, CH-8093 Zurich, the [§]Department of Chemistry and Biochemistry, University of Berne, 3012 Berne, and the [¶]Department of Biology, Institute of Molecular Biology and Biophysics, ETH Zurich, CH-8093 Zurich, Switzerland

Background: *N*-Glycosyltransferases represent a novel class of *N*-glycosylation-catalyzing enzymes.
Results: A quantitative activity assay allowed us to determine substrate specificities of *N*-glycosyltransferase.
Conclusion: *N*-Glycosyltransferase exhibits a relaxed sugar substrate specificity and a peptide specificity strikingly similar to that of oligosaccharyltransferase.
Significance: This study highlights the convergent evolution of two *N*-glycosylation systems and might lay the basis for a novel route for glycoengineering in bacteria.

***N*-Linked protein glycosylation is a very common post-translational modification that can be found in all kingdoms of life. The classical, highly conserved pathway entails the assembly of a lipid-linked oligosaccharide and its transfer to an asparagine residue in the sequon N*X*(S/T) of a secreted protein by the integral membrane protein oligosaccharyltransferase. A few species in the class of γ-proteobacteria encode a cytoplasmic *N*-glycosylation system mediated by a soluble *N*-glycosyltransferase (NGT). This enzyme uses nucleotide-activated sugars to modify asparagine residues with single monosaccharides. As these enzymes are not related to oligosaccharyltransferase, NGTs constitute a novel class of *N*-glycosylation catalyzing enzymes. To characterize the NGT-catalyzed reaction, we developed a sensitive and quantitative *in vitro* assay based on HPLC separation and quantification of fluorescently labeled substrate peptides. With this assay we were able to directly quantify glycopeptide formation by *Actinobacillus pleuropneumoniae* NGT and determine its substrate specificities: NGT turns over a number of different sugar donor substrates and allows for activation by both UDP and GDP. Quantitative analysis of peptide substrate turnover demonstrated a strikingly similar specificity as the classical, oligosaccharyltransferase-catalyzed *N*-glycosylation, with N*X*(S/T) sequons being the optimal NGT substrates.**

N-Linked protein glycosylation is one of the most frequent post-translational protein modification in eukaryotes, with more than half of all eukaryotic proteins predicted to be *N*-glycoproteins (1). The modification is involved in many biological processes such as, for example, protein folding and quality control in the endoplasmic reticulum (2, 3), cell-cell recognition (4), or immune modulation and inflammation (5, 6). It entails the covalent attachment of a carbohydrate moiety to the side chain amide of asparagine residues in the conserved sequon N*X*(S/T) (where $X \neq$ P) (7). *N*-Glycosylation is essential in all eukaryotes but it is also found in some bacteria (8) and in many archaea (9). The basic layout of the pathway is conserved in all domains of life: assembly of an oligosaccharide on an isoprenoid lipid in the cytoplasm, translocation across the membrane (the endoplasmic reticulum membrane in eukaryotes, the plasma membrane in prokaryotes), and *en bloc* transfer to an asparagine side chain of an acceptor substrate protein. The central enzyme, oligosaccharyltransferase (OST),[4] is a membrane protein complex in most eukaryotes, but single subunit OSTs can be found in prokaryotes and some protozoan species (10). OST catalyzes the formation of an *N*-glycosidic bond between the amide nitrogen of an asparagine side chain and the C1 carbon of the reducing end sugar of the glycan. However, amides are planar and the nitrogen lone-pair is delocalized with the π-system of the carbonyl group, making amides poor nucleophiles and generally very unreactive. The amide group therefore requires activation by the enzyme for *N*-linked protein glycosylation to occur. Several reaction mechanisms have been proposed including tautomerization of the amide into an imidol involving the hydroxyl group of the amino acid at position +2 of the sequon and a basic residue of the enzyme (11), as well as transfer of a proton from the amide nitrogen via the hydroxyl amino acid to a general base (12). However, determination of the three-dimensional x-ray crystal structure of the *Campylobacter lari* OST PglB (13) has shown that the +2 amino acid of the substrate is not involved in catalysis. This structure has led to the formulation of a new mechanistic model. The amide is

* The work was supported in part by the Swiss National Science Foundation Grants CRSII3_12733 and 31003A-127098 (to M. A.) and 200020_140349 (to J. L. R.) and ETH Zürich.
[S] This article contains supplemental Figs. S1 and S2.
[1] Member of the Life Science Zurich Graduate Program for Molecular Life Science.
[2] Present address: Dept. of Biochemistry and Molecular Biology, Centre for Blood Research, University of British Columbia, Vancouver, British Columbia V6T1Z3, Canada.
[3] To whom correspondence should be addressed: ETH Zurich, Institute of Microbiology, Vladimir-Prelog-Weg 4, 8093 Zürich, Switzerland. Tel.: 41-44-632-64-13; Fax: 41-44-632-11-48; E-mail: markus.aebi@micro.biol.ethz.ch.

[4] The abbreviations used are: OST, oligosaccharyltransferase; NGT, *N*-glycosyltransferase; ApNGT *Actinobacillus pleuropneumoniae* NGT; 5CF, 5-carboxyfluorecein; hS, homoserine; DCM, dichloromethane; NMP, *N*-methyl-2-pyrrolidone; Fmoc, fluorenylmethoxy carbonyl; MeOH, methanol; PyBOP, benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate; ACN, acetonitrile; PDB, Protein Data Bank.

activated by the formation of two hydrogen bonds between acidic residues of the enzyme and the amide group of the acceptor asparagine side chain. Due to the orientation of the H bond acceptors, this leads to a twisting of the C-N bond and therefor to the loss of conjugation of the nitrogen lone pair and the π-system of the carbonyl group. The lone pair can participate in a nucleophilic attack on the sugar substrate and form the *N*-glycosidic bond. However, the available three-dimensional structures of OST are either at too low resolution to accurately measure bond lengths and angles proving amide twisting (13), or they lack a bound acceptor peptide (14). Unlike all previously described mechanisms for amide activation, the "twisted amide" model is consistent with biochemical data and supported by the three-dimensional x-ray crystal structure (13, 15, 16).

A few species in the class of γ-proteobacteria encode a very unusual *N*-linked protein glycosylation system. First described in the human pathogen *Haemophilus influenzae* (17–19), this pathway entails the cytoplasmic modification of asparagine side chains of proteins with single hexose residues. The enzyme responsible, HMW1C, is a soluble, cytoplasmic protein that uses nucleotide-activated sugars (UPD-Glc, UPD-Gal) as donor substrates. It is phylogenetically and structurally unrelated to OST and constitutes the founding member of a novel class of *N*-glycosylation catalyzing enzymes: the *N*-glycosyltransferases (NGTs). NGTs can be found in other Pasteurellaceae species as well as in many members of the genus *Yersinia*. The proteins of *Actinobacillus pleuropneumoniae* (20, 21) and *Yersinia enterocolitica* (21) have proven *N*-glycosylation activity. The NGT protein of *A. pleuropneumoniae* is the best studied example on a biochemical level and its three-dimensional x-ray crystal structure has been determined (22). Detailed analysis of its reaction product has shown that it is an inverting glycosyltransferase that glycosylates the same consensus sequon as eukaryotic OST (N*X*(S/T), where $X \neq$ P) (21), although with a more relaxed substrate specificity toward non-consensus sequons (23). Even though NGT is able to modify many different polypeptide substrates under laboratory conditions (21), the only endogenous substrate proteins described so far are adhesin proteins of the outer membrane (17, 23, 24). NGTs are classified in glycosyltransferase family GT41 (Carbohydrate Active Enzymes database (25)) and show similarity to eukaryotic UDP-GlcNAc:peptide *N*-acetylglucosaminyltransferases (*O*-GlcNAc transferase). These soluble protein glycosyltransferases catalyze the transfer of GlcNAc to serine or threonine residues of many cytoplasmic, nucleoplasmic, and mitochondrial proteins (26).

In this study, we set out to perform a detailed biochemical characterization of NGT of the porcine pathogen *A. pleuropneumoniae* (ApNGT) and study its substrate specificity in more detail. We established an *in vitro* assay for NGT activity based on HPLC separation of fluorescently labeled substrate peptides from their glycopeptide products. This assay allowed for direct and sensitive quantification of glycopeptide formation by ApNGT and was applied to quantify sugar and peptide substrate specificities of ApNGT. We found a strikingly diverse sugar donor substrate specificity with flexibility for both the transferred sugar residue as well as the activating nucleotide. Quantification of turnover of different peptides revealed a substrate specificity remarkably similar to that of OST.

EXPERIMENTAL PROCEDURES

Protein Purification—Purification of NGT proteins was adapted from previously described protocols (20, 21): 2 liters of LB medium, supplemented with 100 μg/ml of ampicillin, was inoculated with *Escherichia coli* DH5α carrying the appropriate expression plasmid. The cells were grown at 37 °C to an A_{600} of 0.5 and protein expression was induced by addition of 0.2% arabinose. After incubation at 25 °C for 4 h, cells were harvested by centrifugation (8,000 × *g*, 10 min, 4 °C). The cell pellet was resuspended in 25 ml of 20 mM Tris-HCl, pH 8.0, 250 mM NaCl, 5 mM β-mercaptoethanol, 20 mM imidazole, supplemented with protease inhibitor mixture (Complete EDTA-free, Roche Applied Science), 1 mM EDTA, and 1 mg/ml of lysozyme. After incubation for 1 h at 4 °C, cells were broken by 3 passes through the French press (1000 p.s.i.), centrifuged twice (4,000 × *g*, 15 min, 4 °C and 40,000 × *g*, 30 min, 4 °C respectively) and the supernatant was loaded onto a 1-ml HisTrap HP column (GE Healthcare). The column was washed with 25 ml of 20 mM Tris-HCl, pH 8.0, 250 mM NaCl, 5 mM β-mercaptoethanol, 50 mM imidazole, and the purified protein was eluted with the same buffer containing 300 mM imidazole. ApNGT in the pooled elution fractions was further purified using a HiLoad 16/60 Superdex 200 size exclusion column (GE Healthcare), equilibrated in 50 mM HEPES, pH 7.0, 200 mM NaCl, 0.1 mM EDTA, 5% glycerol. Finally, the protein was concentrated to approximately 10 mg/ml using an Amicon Ultracel 50-kDa filter unit (Merck Millipore Ltd.) and the protein concentration was determined using the Bio-Rad Protein Assay kit. $AtaC_{1866-2428}$ was purified as described before (23).

UDP Detection Assay—Formation of UDP as a byproduct of the NGT reaction was measured using a prototype version of the UDP-Glo™ Glycosyltransferase Assay (Promega). The reactions were performed in a 10-μl volume containing 100 mM HEPES, pH 8, 500 mM NaCl, 5 μM purified $AtaC_{1866-2428}$ as peptide substrate, 500 μM UDP-Glc, and 1 μM purified ApNGT. After incubation for 3 h at 30 °C, the reactions were stopped by mixing with 10 μl of UDP-Glo detection reagent. The samples were transferred into an opaque white 384-well plate (Greiner) and luminescence was recoded after an incubation time of 60 min using a PerkinElmer Victor^3 spectrometer. For absolute quantification a UDP standard curve was determined. All measurements were performed in triplicates.

NMR Measurements—NMR spectra were recorded on a 500 MHz Avance III Bruker spectrometer equipped with an inverse triple-resonance cryogenetic probe. All NMR spectra were measured in a buffer of 50 mM NaH_2PO_4/Na_2HPO_4, pH 8.0, and 20 mM NaCl in D_2O. For the exchange of H_2O to D_2O, the sample was lyophilized and dissolved in the same volume of D_2O. Standard two-dimensional ^{13}C heteronuclear single-quantum coherence, ^{13}C heteronuclear multiple-bond correlation, and ^1H-^1H total correlation spectroscopy spectra were applied to assign UDP-Glc and its hydrolyzation products. Spectra are referenced to 2,2-dimethyl-2-silapentane sulfonic acid. ^{13}C chemical shifts are indirectly referenced using a scal-

ing factor of 0.251449530 (27). Spectra were processed with Topspin (Bruker) and analyzed with Sparky (46).

Synthesis of Fluorescently Labeled Acceptor Peptides—Peptide COK_1394$_{57-76}$ was synthesized as follows. Synthesis was initiated by the loading of TentaGel S RAM resin (500 mg, loading: 0.24 mmol g^{-1}) in a 10-ml polypropylene syringe fitted with a polypropylene frit, a Teflon stopcock, and a stopper. The resin was swollen in DCM (6 ml, 20 min). After removal of the DCM, the Fmoc protecting group of the resin was removed using a solution of 20% piperidine in NMP. Stirring the reaction mixture was performed by attaching the closed syringes to a rotating axis. The completion of the reaction was checked using the 2,4,6-trinitrobenzenesulfonic acid test. Removal of the Fmoc protecting group was performed by using a solution of 20% piperidine in NMP (6 ml, 20 min). After filtration, the resin was washed with NMP (2 × 6 ml), MeOH (2 × 6 ml), and DCM (2 × 6 ml). Coupling of the Fmoc protected amino acids was performed using Fmoc protected amino acids (3 eq), PyBOP (3 eq), and *N,N*-diisopropylethylamine (6 eq) in NMP (6 ml). The resin was stirred for 2 h. The resin was then washed with NMP (2 × 6 ml), MeOH (2 × 6 ml), and DCM (2 × 6 ml). Acetylation of the resin was performed after each amino acid coupling using a solution of acetic anhydride:DCM (1:1). Coupling of the 5-carboxyfluorescein was performed by using 5-carboxyfluorescein (2 eq), hydroxybenzotriazole (5 eq), and *N,N′*-diisopropylcarbodiimide (5 eq) in NMP (6 ml). The resin was stirred overnight and protected from light using aluminum foil. The resin was washed with NMP (2 × 6 ml), MeOH (2 × 6 ml), and DCM (2 × 6 ml). A solution of 20% piperidine in NMP (6 ml, 6 × 5 min) was then added to the resin to remove excess free 5-carboxyfluorescein. The resin was finally washed with DCM (6 × 6 ml). After coupling of the fluorophore, the compound was protected from light with aluminum foil. TFA cleavage was performed by adding a solution of TFA/triisopropylsilane/H_2O (94:5:1, v/v/v) to the resin for 1 h. The peptide was then precipitated with *tert*-butylmethylether then dissolved in H_2O/ACN with 0.1% TFA mixture. The linear peptide was purified by preparative RP-HPLC and obtained as TFA salts after lyophilization. The fluorescently labeled peptides containing different acceptor sequons were synthesized as described previously (15, 16).

In Vitro Enzyme Assays—Nucleotide-activated sugars were obtained from Sigma. With the exception of UDP-Xyl, which was acquired from CarboSource Services (Complex Carbohydrate Research Center, University of Georgia). For assay development and determination of sugar substrate specificity, reactions were carried out in a 5-μl volume containing 50 μM peptide COK_1394$_{57-76}$, the indicated concentration of sugar donor and 100 nM purified ApNGT in a buffer with 100 mM HEPES, pH 8.0, and 500 mM NaCl (unless stated otherwise). Reactions were stopped by the addition of 20 μl of 0.2% formic acid. The samples were diluted 80-fold in HPLC buffer I (20% ACN, 20 mM potassium phosphate, pH 8.0) and 50-μl samples were analyzed by reverse phase HPLC (Phenomenex Aeris PEPTIDE 3.6 μm XB-C18, 150 × 4.6 mm column, isocratic elution with HPLC buffer I for 15 min). For determination of sequon turnover rates, reactions were carried out in a 5-μl volume containing 200 μM substrate peptide, 2.5 mM UDP-Glc, and the indicated concentration of purified ApNGT in a buffer with 100 mM HEPES, pH 8.0, and 500 mM NaCl. Reactions were stopped by the addition of 20 μl of 0.2% formic acid. The samples were diluted 300-fold in HPLC buffer II (10% ACN, 20 mM potassium phosphate, pH 8.0) and 50-μl samples were analyzed by reverse phase HPLC (Phenomenex Aeris PEPTIDE 3.6 μm XB-C18, 150 × 4.6-mm column, gradient elution from 10–14% ACN over 15 min). The peptides were detected via the 5-carboxyfluorescein fluorophore (excitation wavelength: 494 nm, emission wavelength: 521 nm).

Analysis of Kinetic Data—Peaks in the HPLC profiles were integrated using Chromeleon 6 (Dionex) and formation of glycopeptide was calculated as follows,

$$c_{gp} = \frac{\text{Peak area}_{gp}}{\text{Peak area}_{p} + \text{Peak area}_{gp}} \times c_{p\,\text{total}}$$

c_{gp} refers to the concentration of glycopeptide, peak area$_{p/gp}$ to the measured areas under the curve for the peptide and the glycopeptide, respectively. $c_{p,\text{total}}$ refers to the concentration of peptide originally added to the reaction). Kinetic data were analyzed using GraphPad Prism 6 and the data were either fitted using linear regression (turnover rates) or non-linear regression according to the Michaelis-Menten equation (determination of K_m and k_{cat}).

Mass Spectrometry—In vitro reactions were performed as above and stopped after 8 h. All samples were desalted by C18 Zip-tip (Millipore Corp., Billerica, MA) according to supplier's recommendation, lyophilized, and dissolved in 50% ACN with 0.1% formic acid. After being mixed 1:4 with α-cyano-4-hydroxycinnamic acid matrix (5 mg/ml in 50% ACN in water with 0.1% formic acid), samples were spotted onto a matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) target plate. Data acquisition was performed manually on either MALDI LTQ Orbitrap^{TM} XL (Thermo Scientific) equipped with an N_2 laser or model 4800 Proteomics Analyzer (Applied Biosystems, Framingham, MA) with an Nd:YAG laser and 1,000 shots were accumulated in the reflectron positive ion mode. To confirm glycosylation of peptides 5CF-GSDQQATF and 5CF-GSDQhSATF, theses samples were also analyzed by LC-MS/MS on a calibrated LTQ-Orbitrap Velos mass spectrometer (Thermo Fischer Scientific, Bremen, Germany) coupled to an Eksigent-Nano-HPLC system (Eksigent Technologies, Dublin, CA).

Protein Sequence Alignments—Protein sequences were retrieved from the UniProt database (accession numbers B3H2N2, Q93DC6, A5UCI8, O15294, O18158, Q7KJA9, and Q8PC69) and a structure-aided multiple sequence alignment was performed using Expresso (28). The following crystal structures were used as templates for alignment: human *O*-GlcNAc transferase (PDB code 4N3C (29) for human, *Caenorhabditis elegans* and *Drosophila melanogaster O*-GlcNAc transferase sequences, *Xanthomonas campestris O*-GlcNAc transferase (PDB code 2XGS (30) for *X. campestris O*-GlcNAc transferase) and *A. pleuropneumoniae* NGT (PDB code 3Q3I (22) for *A. pleuropneumoniae* NGT, *Y. enterocolitica* RscC, and *H. influenzae* HMW1C). CLC Workbench 6 was used for visualization.

Analysis of ApNGT Mutants—Plasmid pEC415(NGTmyc) was used for NGT expression, pMLBAD(His10-AtaC-1866–2428) was used for expression of $AtaC_{1866-2428}$ (23). Constructs for expression of ApNGT mutants were generated using the Phusion Site-directed Mutagenesis Kit (Thermo Scientific) with pEC415(NGTmyc) as a template. *E. coli* DH5α cells transformed with the appropriate expression constructs were grown in LB medium supplemented with ampicillin (100 μg/ml) and trimethoprim (100 μg/ml) to an A_{600} of 0.5. Protein expression was induced by addition of 0.2% arabinose and cells were harvested after an incubation of 4 h at 37 °C. Whole cell extracts were prepared and analyzed by SDS-PAGE on a 7% acrylamide gel followed by immunoblot analysis.

RESULTS

ApNGT Hydrolyzes Its Sugar Donor Substrates—In previous studies, ApNGT activity was quantified using an indirect assay that measured the formation of UDP as a byproduct of the glycosyltransferase reaction (20, 22). We set up a similar approach using purified ApNGT, the previously described *A. pleuropneumonaie* autotransporter adhesin fragment $AtaC_{1866-2428}$ as substrate protein (23) and UDP-Glc as sugar donor. A bioluminescent UDP detection assay was used for quantification of UDP formation as a proxy for NGT activity. We noted that the acceptor peptide substrate was not required for formation of UDP (Fig. 1*A*, *columns 1* and *5*). Even though UDP formation increased ~2-fold upon addition of the acceptor protein substrate, also in its absence, UDP-Glc was hydrolyzed and an increase in free UDP could be detected. This UDP formation could not be observed when ApNGT, UDP-Glc, or both were missing (Fig. 1*A*, *columns 2–4* and *6–8*). We also detected ApNGT-dependent hydrolysis of UDP-Gal, albeit to a much smaller degree. No significant increase in UDP concentration could be detected when UDP-GlcNAc was incubated with ApNGT (Fig. 1*B*). Further evidence for sugar substrate hydrolysis by ApNGT was obtained by incubating UDP-Glc with purified ApNGT alone and analyzing the reaction product by NMR spectroscopy. The NMR spectrum clearly indicates the formation of free glucose upon incubation of UDP-Glc with ApNGT (approximately 50% of total UDP-Glc was hydrolyzed after 20 h) (Fig. 1*D*). To make sure that the observed UDP-Glc hydrolysis was due to NGT activity and not a contaminating protein in our preparation, we repeated the bioluminescent UDP detection assay with a catalytically inactive mutant of ApNGT (ApNGT-K441A (22)). UDP formation dropped back to background levels (Fig. 1*A*, *columns 9* and *10*), indicating that UDP-Glc hydrolysis is indeed dependent on ApNGT activity. Formation of UDP is therefore not a suitable measure for NGT activity and a more direct activity assay is needed.

A Direct Assay for NGT Activity—As the relevant reaction product is the glycopeptide and not UDP, a more accurate way to measure NGT activity is to quantify the formation of glycopeptide directly. For that, we developed an HPLC-based assay using fluorescently labeled substrate peptides (Fig. 2*A*). Based on the sequence of *Mannheimia hemolytica* autotransporter adhesin protein COK_1394 around Asn^{67}, which was previously shown to be glycosylated efficiently by ApNGT (23), we synthesized the fluorescently labeled peptide COK_1394_{57-76} (5CF-SGAMGDKSVA*NAT*YSLALGS-NH_2) as a new model substrate. This peptide contains a single glycosylation site and is easily detectable via its fluorophore. We incubated this peptide together with purified ApNGT and UDP-Glc and separated the product from educt using RP-HPLC with fluorescence detection (Fig. 2*B*). Upon incubation with ApNGT, a new peak (retention time 9.3 min), corresponding to the glycosylated peptide, appeared. This peak neither appeared when the peptide was incubated with UDP-Glc alone, with ApNGT alone, or with UDP-Glc and inactive ApNGT-K441A mutant. Integration of the fluorescence signal yielded a direct quantification of product and educt of the glycosyltransferase reaction. Therefore, our assay allows a sensitive and accurate characterization of the ApNGT-catalyzed reaction (Fig. 2*C*). We applied this HPLC-based assay to optimize reaction conditions with regard to pH and ionic strength of the buffer. Salt concentration had little influence on reaction speed (with a small maximum around 500 mM NaCl) (Fig. 2*D*). Replacing sodium for potassium or a sodium/potassium mixture had no effect on reaction speeds (data not shown). The pH optimum of ApNGT was between 7 and 8, with sharp drops in activity in both the acidic and basic pH range (Fig. 2*E*). We therefore used buffer with pH 8 containing 500 mM NaCl for all further experiments.

ApNGT Is a Glucosyltransferase but Exhibits a Relaxed Sugar Substrate Specificity—We first set out to perform a detailed analysis of the sugar substrate specificity of ApNGT *in vitro*. For that, we incubated peptide COK_1394_{57-76} with purified ApNGT and different nucleotide-activated monosaccharides and analyzed the reaction products by MALDI mass spectrometry (Fig. 3*A*). As shown previously (20, 21), ApNGT was able to transfer glucose and galactose from UDP onto the peptide. No glycopeptide could be detected when substituted hexoses (GlcA, GlcNAc, or GalNAc) were used as sugar donor but ApNGT was able to transfer a pentose (UDP-Xyl). Surprisingly, ApNGT also turned over sugars activated with GDP rather than UDP. When glucose was activated by GDP it still served as a donor substrate and to a very marginal degree GDP-activated mannose was also turned over. CMP-Neu5Ac, however, could not serve as a substrate under our assay conditions. Having identified which sugar donor substrates could be turned over by ApNGT (UDP-Glc, UDP-Gal, UDP-Xyl, GDP-Glc, GDP-Man), we set out to quantitatively assay the kinetic properties of ApNGT using these donor substrates. We applied our HPLC-based *in vitro* assay to measure reaction velocities at a fixed peptide substrate concentration and various sugar substrate concentrations (Fig. 3*B*). The apparent K_m and k_{cat} values were determined and are depicted in Table 1. In agreement with previous studies (20), our analysis showed that UDP-Glc was the preferred substrate of ApNGT ($k_{cat}/K_m = 502.8\ M^{-1}\ s^{-1}$). UDP-Gal and UDP-Xyl were turned over much slower (1.5 and 0.7% of UDP-Glc turnover, respectively) but showed K_m values in a similar range (3–5 times higher). When glucose was activated by GDP, the K_m was increased more than 12-fold and the turnover dropped to 3% of the value determined for UDP-Glc. The turnover of GDP-Man was too slow to be reliably quantified by this approach (data not shown).

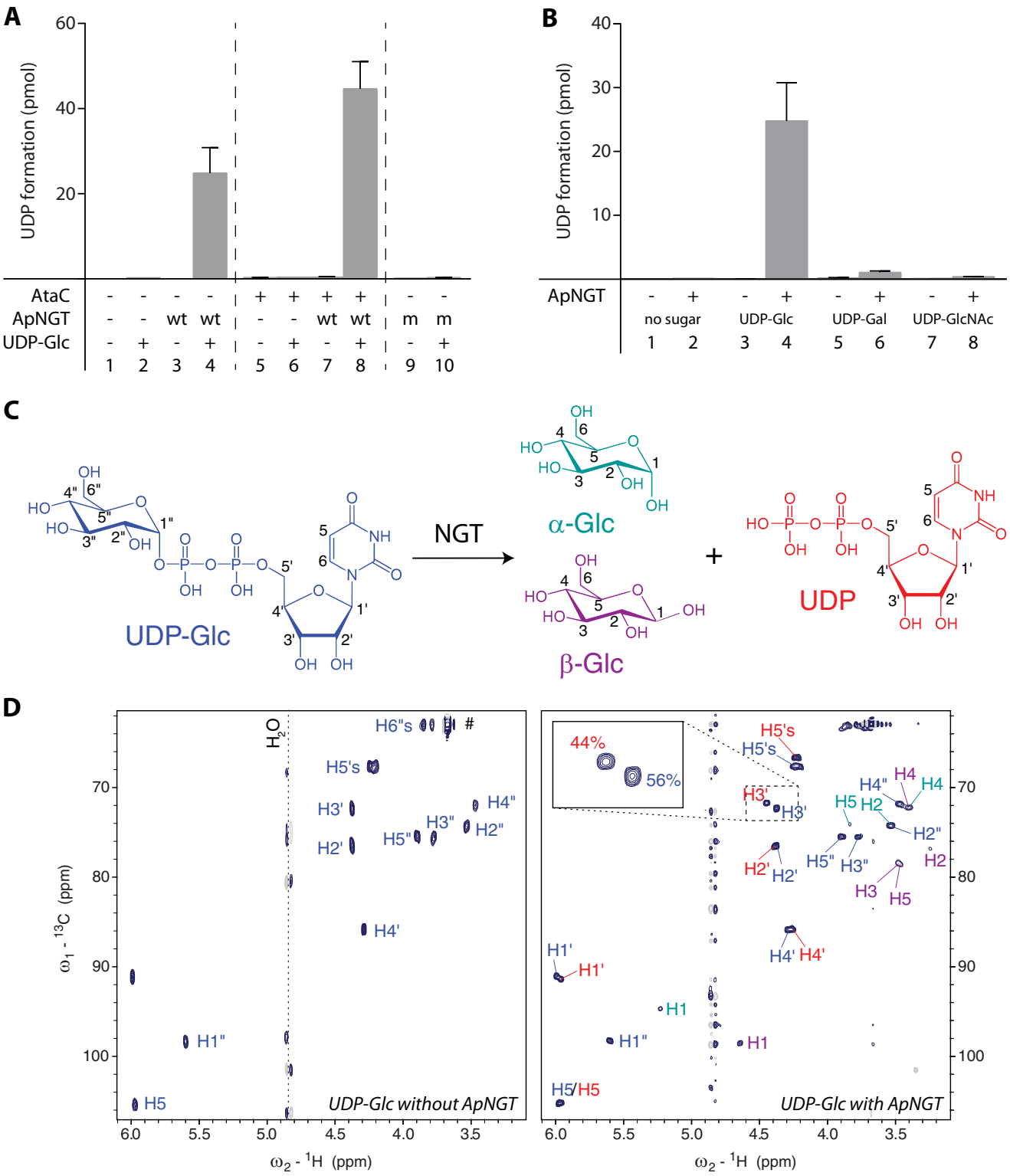

FIGURE 1. **ApNGT hydrolyzes nucleotide-activated sugars.** *A,* substrate protein $AtaC_{1866–2428}$ was incubated with UDP-Glc and purified ApNGT for 3 h at 30 °C and free UDP was determined using the bioluminescent UDP-Glo glycosyltransferase assay. Note that UDP is also formed in the absence of substrate protein but not in the absence of ApNGT or UDP-Glc or by the catalytically inactive ApNGT-K441A mutant. *B,* hydrolysis of different UDP-activated sugars. ApNGT hydrolyzed UDP-Glc and UDP-Gal (~25 times slower) but not UDP-GlcNAc. *C,* schematic representation of UDP-Glc hydrolysis reaction. The color code corresponds to the color label in the NMR spectra. *Numbers* indicate the nomenclature of the protons. *D,* ^{1}H-^{13}C heteronuclear single-quantum coherence spectrum of 1 mM UDP-Glc measured with 2 scans at 293 K (*left*) and ^{1}H-^{13}C heteronuclear single-quantum coherence spectrum of 1 mM UDP-Glc after incubation for 20 h with 10 μM NGT measured with 32 scans at 293 K (*right*). The labels of the different reaction products are color coded corresponding to *panel C*. The assigned chemical shifts of UDP-Glc and UDP match previously reported values (45).

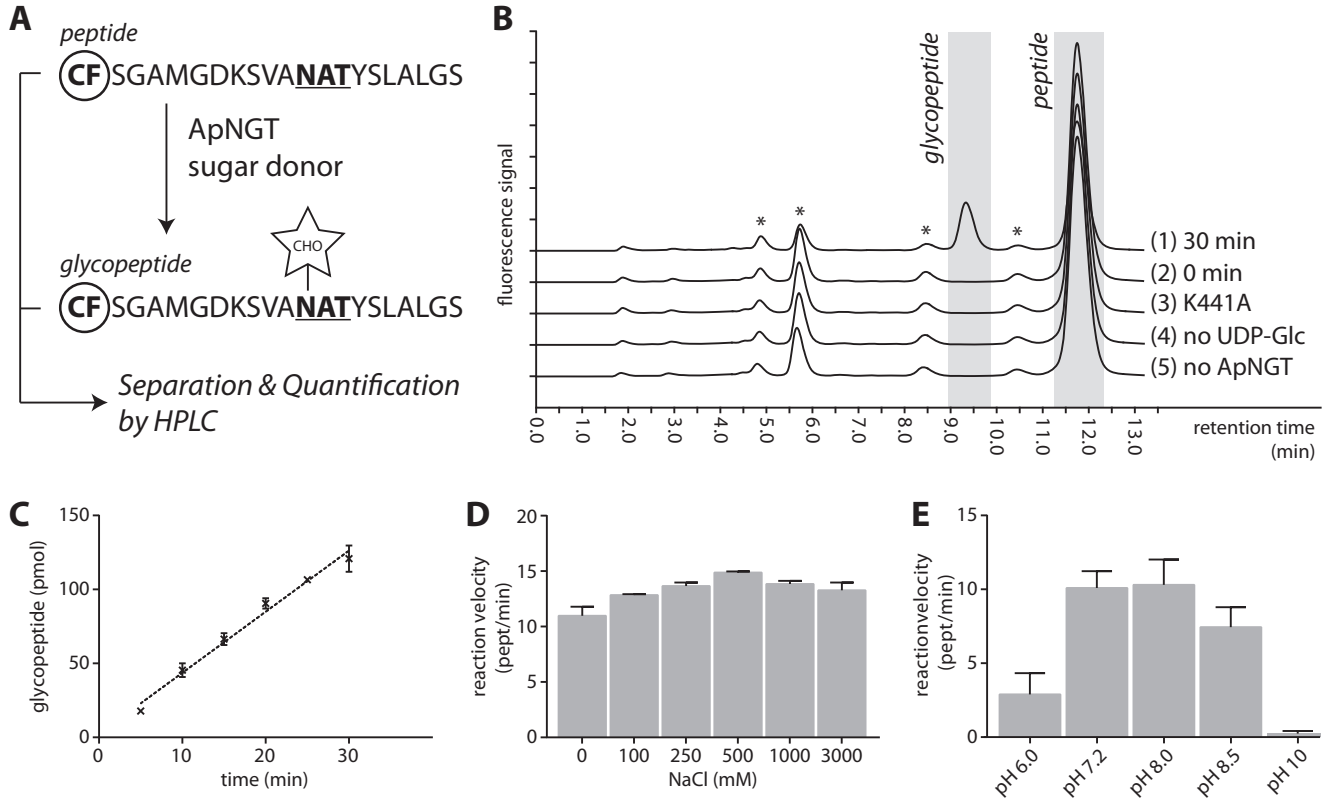

FIGURE 2. **A direct assay for NGT activity.** *A,* schematic overview of assay set up. A fluorescently labeled peptide (*CF*, 5-carboxyfluorescein) is incubated with purified ApNGT and a sugar donor substrate. Product (glycopeptide) is separated from educt (peptide) using RP-HPLC, and both peptides are quantified separately. *B,* representative chromatogram of analysis of glycosylation of peptide COK_1394$_{57–76}$. The peptide was incubated for 30 min with UDP-Glc and purified ApNGT. Samples were analyzed before (*2*) and after incubation time (*1*). The following controls were included: peptide with UDP-Glc and ApNGT-K441A (*3*), peptide with ApNGT alone (*4*), and peptide with UDP-Glc alone (*5*). The unmodified peptide elutes after 11.75 min, glycosylation decreases the retention time to 9.3 min. Glycosylation was further confirmed by MALDI mass spectrometry. Contaminants are marked with an *asterisks*. *C,* determination of the reaction speed of glycosylation. *D* and *E,* determination of the reaction speeds of NGT reaction at different salt concentrations and different pH values.

Quantitative Analysis of the Peptide Substrate Specificity of ApNGT—ApNGT glycosylates substrate proteins at N*X*(S/T) sites (21), but also tolerates Ala, Gly, Val, and Asp at position +2 when the glycosylation system is reconstituted in *E. coli* (23). However, this analysis was based on qualitative mass spectrometry data and only allowed limited conclusions about the sequon preferences of ApNGT. We therefore wanted to quantify the turnover of peptides containing different sequons *in vitro*. For that, we made use of short, fluorescently labeled synthetic peptides containing different glycosylation sequons that were incubated with purified ApNGT and UDP-Glc and glycosylation was quantified using the HPLC-based approach. As small peptides are poor substrates for ApNGT (21), we were unable to reach peptide concentrations that would saturate the enzyme and were therefore unable to determine kinetic parameters according to Michaelis-Menten analysis. We determined initial turnover rates as a quantitative measure of substrate specificity instead (Fig. 4, Table 2). Of the canonical glycosylation sequons (N*X*(S/T)), the NAT sequon was turned over more than 10-fold faster than NAS. Glycosylation of alternative sequons (NAA and NAV), although clearly detectable, was extremely slow (approximately 10,000-fold reduction in turnover rate in relationship to NAT). ApNGT is also able to glycosylate amino acids other than asparagine, such as glutamine and serine (23). To investigate the chemical requirements of the glycosyltransferase toward the acceptor site amino acid and quantify the reaction speed of non-asparagine glycosylation, we also tested synthetic peptides with different natural and unnatural amino acids at the zero position of the glycosylation sequon and assessed glycosylation by mass spectrometry (supplemental Fig. S1). Under these conditions, ApNGT was not able to glycosylate aspartic acid, *N*-methylasparagine, or diaminobutyric acid. Glycopeptides could be detected for the peptides containing glutamine and homoserine (supplemental Fig. S1*C*). Quantification of turnover rates by the HPLC-based approach showed that glutamine was glycosylated at a very low rate (more than 15,000-fold reduction in turnover rate as compared with asparagine) but homoserine was glycosylated only 100 times slower than asparagine (Fig. 4, Table 2).

Catalysis by ApNGT Does Not Rely on Basic Residues—Activation of the acceptor substrate through deprotonation by a catalytic base of the enzyme has been suggested as a reaction mechanism for many inverting glycosyltransferases (31), and this type of catalysis was originally also proposed for the amide nitrogen activation by OST (11). For ApNGT, His^{277} has been proposed as a possible catalytic base (22). To further investigate this hypothesis, we made use of the three-dimensional x-ray crystal structure of ApNGT in complex with UDP (PDB code 3Q3H (22)) to identify potentially basic residues near the active site. As the available crystal structures do not reveal the exact location of the sugar moiety or the acceptor peptide, we decided to investigate all basic residues in an 8-Å radius from the

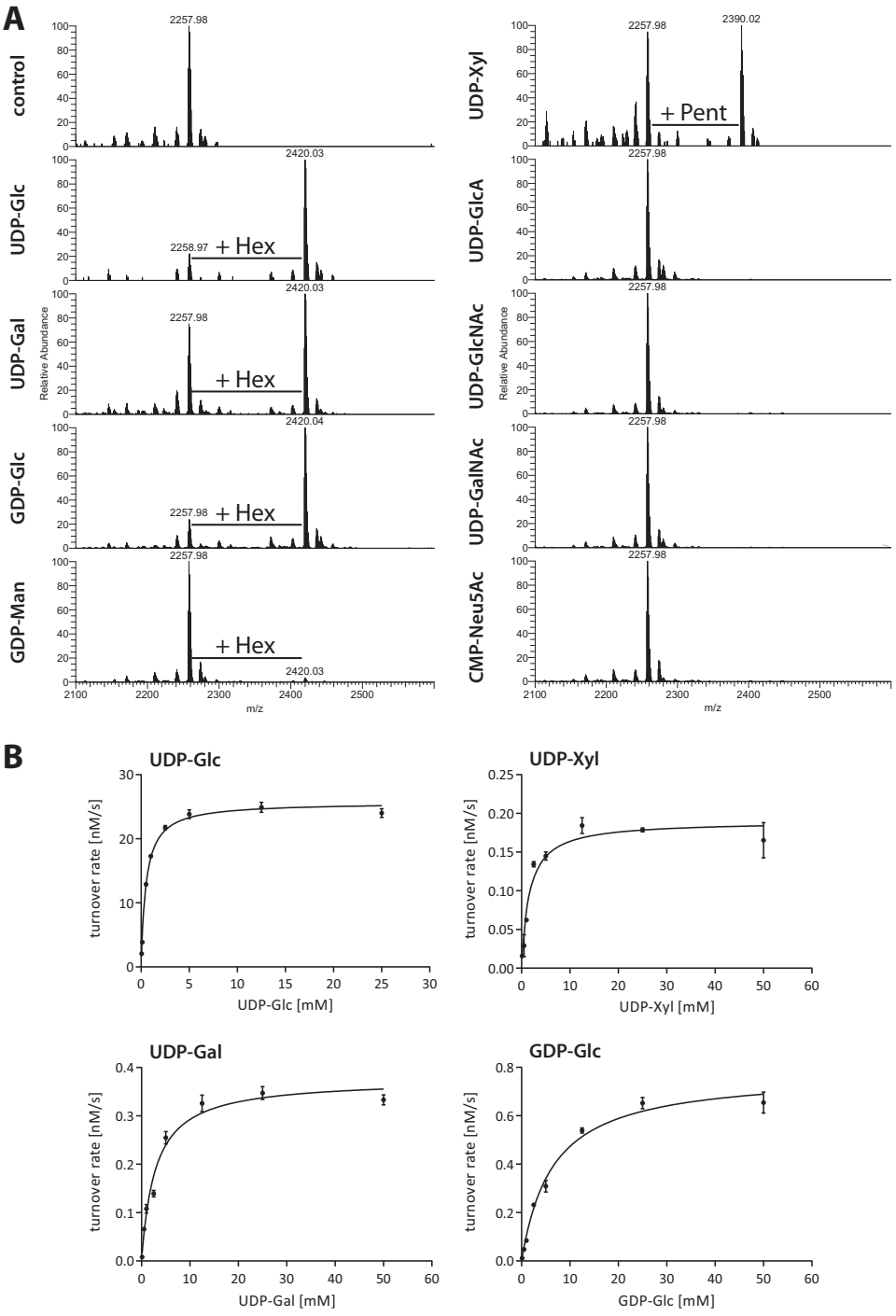

FIGURE 3. **ApNGT is able to utilize different nucleotide-activated sugar donor substrates.** *A,* 50 μM peptide COK_1394$_{57-76}$ was incubated with 5 μM purified ApNGT and 2 mM of the different nucleotide-activated sugars and the reaction products were analyzed by MALDI mass spectrometry. The unmodified peptide was detected at $[M + H]^+ = 2257.98$. Upon incubation with UDP-Glc, UDP-Gal, GDP-Glc, and GDP-Man (*left panel*), an additional peak corresponding to the hexosylated peptide ($[M + H]^+ = 2420.03$) was detected. Incubation with UDP-Xyl resulted in a peak corresponding to peptide modified with a pentose ($[M + H]^+ = 2390.02$). No modification of the peptide with GlcA ($[M + H]^+ = 2434.01$), GlcNAc or GalNAc ($[M + H]^+ = 2461.06$), or Neu5Ac ($[M + H]^+ = 2549.08$) could be detected. *B,* determination of Michaelis-Menten kinetics of turnover of different sugar donor substrates (UDP-Glc, UDP-Gal, UDP-Xyl, and GDP-Glc). Data were fitted using non-linear regression according to the Michaelis-Menten formula ($R^2 = 0.9884$, 0.9637, 0.9345, and 0.9769, respectively). The kinetic parameters are summarized in Table 1.

β-phosphate of UDP (His^{272}, His^{277}, Arg^{281}, His^{371}, Lys^{441}, and His^{495}) (Fig. 5*A*). An alignment of the ApNGT sequence with other known NGTs (*H. influenzae* HMW1C and *Y. enterocolitica* RscC) and other, more distantly related GT41 family glycosyltransferases (such as eukaryotic *O*-GlcNAc transferases as well as a bacterial homolog from *Xanthomonas campestris*), showed that of these residues, only His^{277} and Lys^{441} were absolutely conserved. Arg^{281} is conserved among NGTs but not OGTs, whereas the other residues were not conserved at all. Alanine point mutants of all these residues were assayed for glycosylation activity of the mutant enzymes by co-expression with the substrate protein AtaC$_{1866-2428}$ in *E. coli* followed by

immunoblot analysis with Asn-Glc specific serum (23, 32). Residue Lys[441] is an unlikely catalytic base as it interacts with the pyrophosphate moiety of UDP and stabilizes the negative charge. Mutation of this residue has previously been shown to inactivate the enzyme (22) and the corresponding mutant was therefore used as a control. None of the mutations influenced ApNGT expression (Fig. 5*C*, *top panel*) and, based on shift in electrophoretic mobility of the substrate protein and reactivity to serum MS14, no significant reduction in activity was detected for any other mutant than K441A (Fig. 5*C*, *middle* and

TABLE 1

Kinetic parameters of ApNGT in turnover of different sugar substrates

The values depicted are apparent values that were determined by varying the concentration of the sugar donor substrate while keeping the peptide substrate concentration constant (50 μM). The values were determined from the kinetic data represented in Fig. 3*B*. K_m and k_{cat} are given as absolute values (top) and relative to UDP-Glc parameters (bottom, depicted as fold-increase and percentage respectively).

Sugar substrate	K_m and fold-increase	k_{cat} and relative k_{cat}	k_{cat}/K_m
	m_M	s^{-1}	$_M{}^{-1}\,s^{-1}$
UDP-Glc	0.5 ± 0.04	0.3 ± 3.70 × 10^{-3}	502.8
	1-Fold	100%	
UDP-Gal	2.8 ± 0.4	3.8 × 10^{-3} ± 1.2 × 10^{-4}	1.3
	5.6-Fold	1.5%	
UDP-Xyl	1.6 ± 0.3	1.9 × 10^{-3} ± 8.9 × 10^{-5}	1.2
	3.2-Fold	0.7%	
GDP-Glc	6.9 ± 0.8	8.0 × 10^{-3} ± 2.9 × 10^{-4}	1.2
	13.4-Fold	3.1%	

TABLE 2

Initial turnover rates of the glycosylation reaction of peptides with different sequons

Turnover rates are depicted as number of peptide that each molecule of ApNGT is turning over per minute. For determination of relative turnover, the fastest turnover rate (5CF-GSDQNATF) was defined as 100% and the other turnover rates are given in relationship to that rate. Parameters were determined from the kinetic data represented in Fig. 4. Unnatural amino acids: hS, homoserine; Dab, diaminobutyric acid; N(CH_3), *N*-Methyl-asparagine (side chain structures are depicted in supplemental Fig. S1*B*).

Peptide sequence	Glycopeptide detected	Initial turnover rate	Relative turnover
		peptide/min	%
Sequon numbering		$5CF\text{-}G_{-4}S_{-3}D_{-2}Q_{-1}X_0A_{+1}Y_{+2}F_{+3}$	
Variation at position +2			
5*CF*-GSDQNATF	Yes	0.04 ± 0.001	100
5*CF*-GSDQNASF	Yes	2.9 × 10^{-3} ± 2.7 × 10^{-4}	7.9
5*CF*-GSDQNAAF	Yes	3.6 × 10^{-6} ± 5.7 × 10^{-7}	0.01
5*CF*-GSDQNAVF	Yes	4.5 × 10^{-6} ± 5.6 × 10^{-7}	0.01
Variation at position 0			
5*CF*-GSDQQATF	Yes	2.2 × 10^{-6} ± 3.8 × 10^{-7}	0.006
5*CF*-GSDQhSATF	Yes	3.3 × 10^{-4} ± 1.5 × 10^{-5}	0.9
5*CF*-GSDQDATF	No	ND[a]	ND
5*CF*-GSDQDabATF	No	ND	ND
5*CF*-GSDQN(CH_3)ATF	No	ND	ND

[a] ND, not determined.

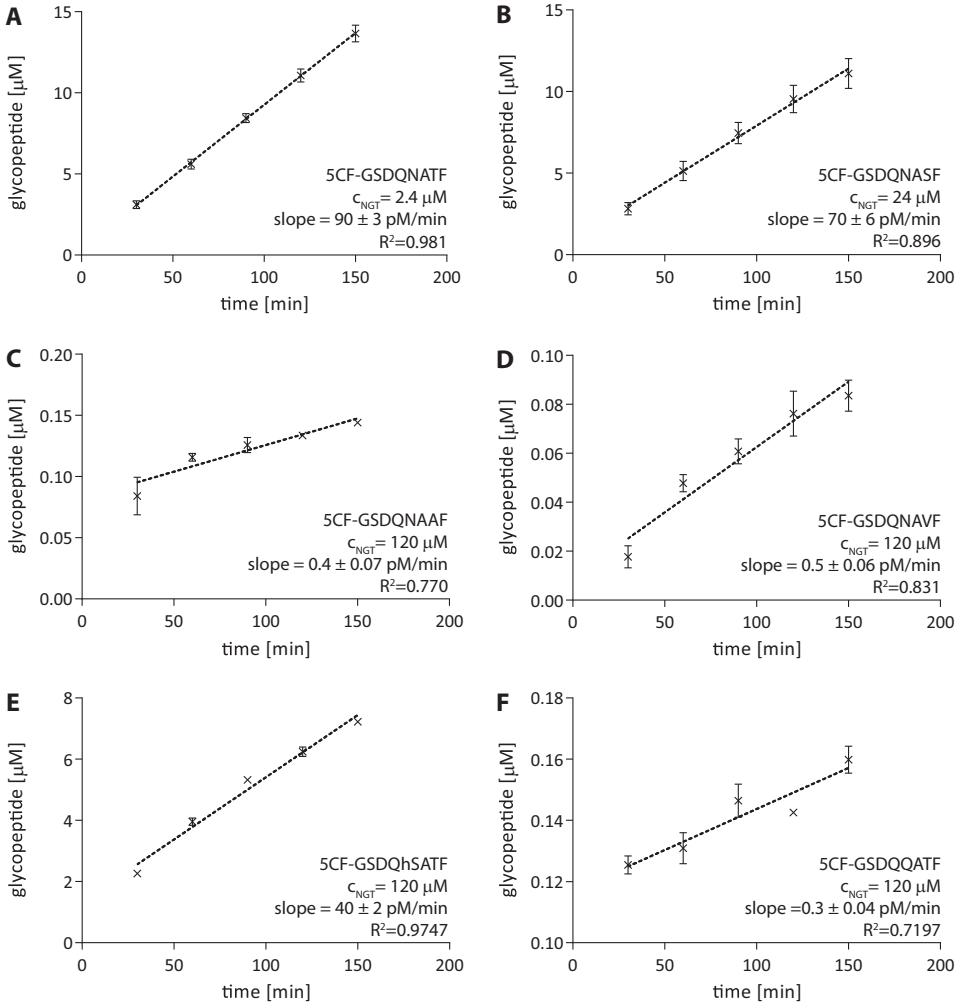

FIGURE 4. **Quantification of turnover rates of different substrate peptides.** Determination of turnover rates of glycosylation reaction with different synthetic peptides. The amount of glycopeptide at each time point was determined based on HPLC separation and fluorescence detection. Each measurement was performed in triplicate. Parameters derived from this analysis are listed in Table 2.

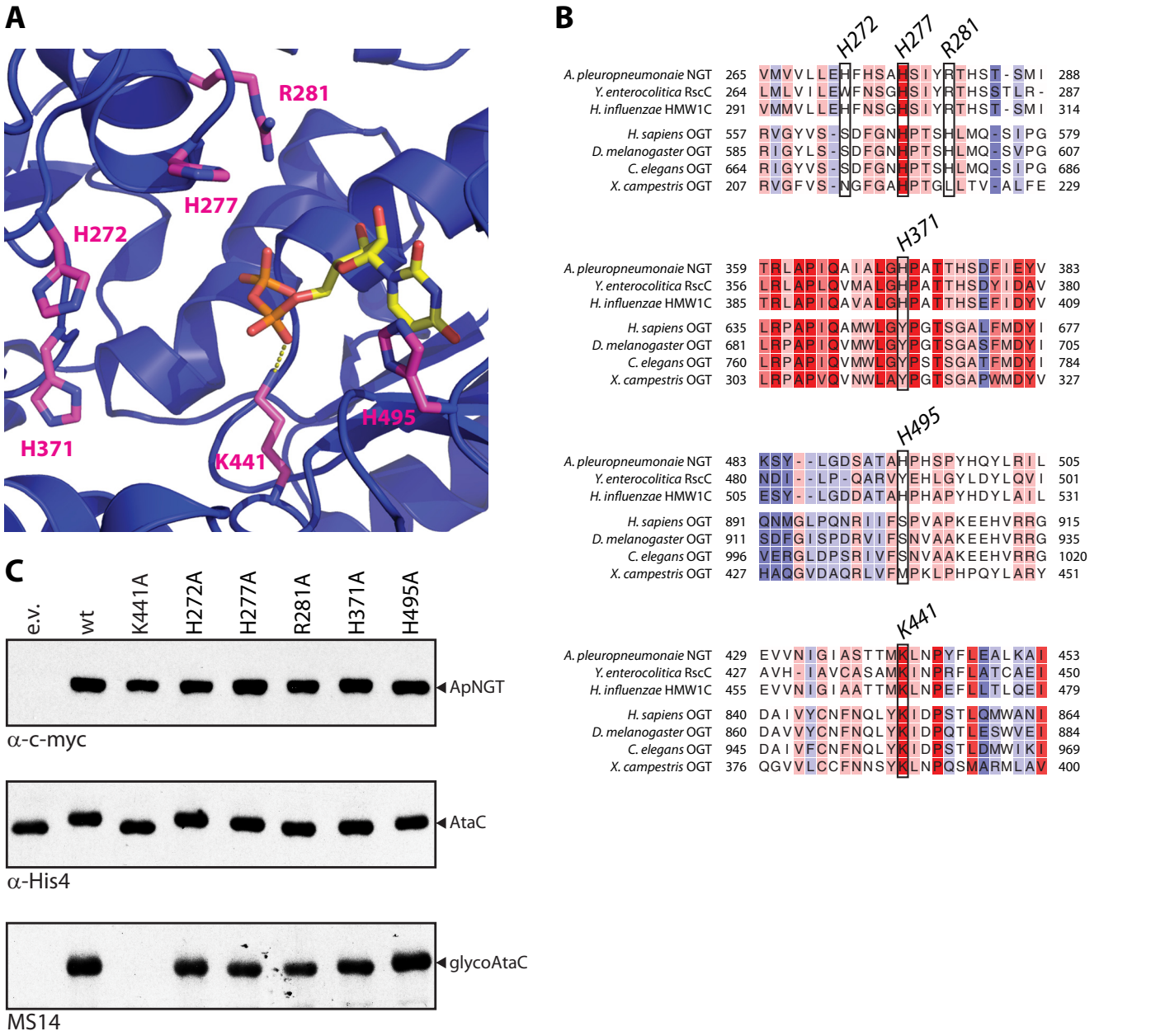

FIGURE 5. **Basic amino acids around the active site of APNGT are not involved in catalysis.** *A,* crystal structure of ApNGT in complex with UDP (PBD code 3Q3H) (22). All basic residues in an 8-Å radius from the β-phosphate of the UDP moiety (in *yellow*) are marked in *magenta*. *B,* structure-aided sequence alignments of several GT41 glycosyltransferase sequences. The region around the residues marked in *A* is shown. For full alignment see supplemental Fig. S2. Alignment was performed using Expresso (28). Coloring indicates conservation of residues (from *blue,* not conserved, to *red*, fully conserved). *C, in vivo* activity assay of ApNGT mutants. The different mutants of ApNGT were co-expressed in *E. coli* with acceptor substrate $AtaC_{1866-2428}$ and whole cell protein extracts were analyzed by immunoblot. NGT proteins were detected via the Myc epitope (*top panel*), $AtaC_{1866-2428}$ via the His_{10} tag (*middle panel*), and glycosylation was detected using Asn-Glc specific serum MS14 (*bottom panel*) (23, 32).

bottom panels). The mutation of His^{277} to alanine has previously been shown to greatly reduce enzymatic activity of ApNGT *in vitro* (22). As we could not detect such a reduction with our *in vivo* approach we sought to further investigate this mutant and clarify the apparent discrepancy between our results and previous studies. We therefore purified the ApNGT mutants H277A and K441A and analyzed the enzymatic activity using the HPLC-based approach (Fig. 6). ApNGT-K441A was catalytically inactive also under *in vitro* conditions. However, this analysis revealed a reduction in activity of ApNGT-H277A (~14% of wild-type), which had not been detected by the *in vivo* approach. Nevertheless, mutation of none of the basic residues in the vicinity of the active site led to a reduction of activity big enough to support a direct role of these residues in catalysis.

DISCUSSION

Biochemical characterization of the ApNGT catalyzed reaction has so far been relying on altered electrophoretic mobility of substrate peptides upon glycosylation (20, 21), resulting in limited separation and quantification power. Alternatively, formation of UDP as a byproduct of sugar transfer was used as an

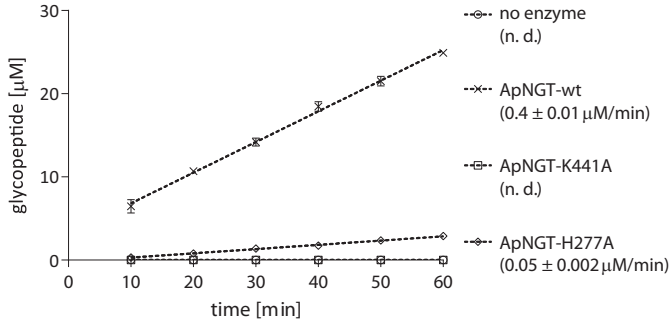

FIGURE 6. ***In vitro* analysis of ApNGT mutants H227A and K441A.** 50 μM peptide COK_1394$_{57-76}$ was incubated with 10 mM UDP-Glc and 0.04 μM purified ApNGT (or the indicated mutant). Samples were taken every 10 min and analyzed by RP-HPLC. Turnover rates were determined by analyzing the data using linear regression and are depicted in *brackets* below the legend. For the control without any NGT enzyme and ApNGT-K441A no activity could be detected (*n.d.*)

activity measure (20). In this study we show that ApNGT hydrolyzes UDP-Glc in the absence of an acceptor peptide substrate. Such sugar substrate hydrolysis is rather common for retaining glycosyltransferases such as, for example, human blood group glycosyltransferases (33) or UDP-GalNAc:polypeptide α-*N*-acetylgalactosaminyltransferases (34, 35) but it has also been observed for glycosyltransferases employing an inverting reaction mechanism including human OGT (36–38). It explains why solely the electron density for UDP and not the glucose moiety was observed in the crystal structure of the ApNGT::UDP-Glc complex (22) because of sugar substrate hydrolysis. The very same effect was also observed when the crystal structure of human OGT was first determined (37). Sugar substrate hydrolysis complicates the quantification of glycosyltransferase activity and limits the general applicability of UDP detection-based glycosyltransferase assays due to the high intrinsic background. To circumvent these problems and minimize experimental artifacts, we developed an activity assay for NGT based on HPLC separation and quantification of fluorescently labeled peptide substrates. This assay allowed for the sensitive and direct quantification of glycopeptide formation by NGT. Although an influence of the fluorescent tag on peptide turnover is very well possible (especially with short peptides where the tag is very close to the glycosylation site), such influences would affect all peptides in the same way and therefore still allow the comparative analysis of different substrates and reaction conditions. With this assay we were able to quantitatively investigate the substrate specificity of ApNGT with respect to its sugar donor and peptide acceptor substrates. ApNGT was able to use a number of different UDP-activated sugars as donor substrates: glucose, galactose, and xylose were all transferred. ApNGT also accepts different activating nucleotides because glucose could be transferred from either UDP or GDP, and we also observed turnover of GDP-activated mannose, albeit at much lower rates. However, we cannot exclude that this very minor turnover of GDP-Man is due to trace contamination of the compound with other nucleotide-activated sugars. Nervertheless, this relaxed specificity for the activating nucleotide is a very unusual feature for a glycosyltransferase. Substituted hexoses (GlcA, GlcNAc, GalNAc, and Neu5Ac) were not accepted, indicating a selectivity filter excluding bulky substituents from the binding pocket. However, of all the different sugar donor substrates, only UDP-Glc is turned over efficiently. The analysis of the reaction kinetics of ApNGT with different nucleotide-activated sugars showed that the main difference between the sugar substrates was not based on affinity. We could reach saturation of the enzyme with all studied nucleotide-activated sugars, resulting in similar K_m values, particularly for the UDP-activated sugars (UDP-Glc, UDP-Gal, and UDP-Xyl). The main difference was observed in reaction speed at saturation (k_{cat}): when incubated with any sugar other the UDP-Glc, ApNGT only reached 0.7–3% of its maximal turnover rate. This suggests that ApNGT contains a rather unspecific nucleotide-binding pocket that is able to accommodate many different nucleotide-activated sugars, as demonstrated by enzyme saturation with different sugar donors. UDP is even able to bind on its own as can be seen in the crystal structure of ApNGT in complex with hydrolyzed UDP-Glc (PDB code 3Q3H (22)). However, only UDP-Glc seemed to bind in an orientation that allows for efficient catalysis of the glycosylation reaction, resulting in much faster turnover than any other sugar donor. This matches with the observed sugar substrate specificity of ApNGT *ex vivo* expressed in *E. coli*, where glucose was the only sugar that was transferred to protein (23). The other sugar donor substrates we assayed, most likely bind in a slightly distorted conformation; either due to the different monosaccharides attached to UDP or, in the case of GDP-Glc, the bigger purine nucleotide that influences the orientation of the glucose moiety. This likely causes the reduction in catalytic efficiency we observed. Nevertheless, ApNGT is able to transfer a variety of different sugars to protein, which opens up new pathways for enzymatic synthesis of novel protein glycoconjugates.

Quantitative analysis of peptide substrate specificity of ApNGT proved to be more difficult. Readily synthetically accessible short peptide substrates were turned over slowly by ApNGT, which made it impossible to determine Michaelis-Menten kinetic parameters. We therefore assayed initial turnover rates as a proxy for enzyme activity. The analysis revealed that although ApNGT is able to glycosylate non-consensus sequons, N*X*(S/T) sites were turned over much faster than any other sequon we tested. ApNGT also exhibited a more than 10-fold preference toward threonine over serine at position +2. Sequons lacking a hydroxyl amino acid at position +2 can also be glycosylated, but at very slow rates. Similar substrate specificities have also been observed for the classical *N*-glycosylation pathway. Both the statistical analysis of occupied sites (39, 40) as well as biochemical data (15, 41) point to a preference of OST toward threonine at position +2. Quantitative analysis of the *C. lari* OST PglB, however, revealed that the difference is much smaller with only a 1.2-fold reduction in turnover (15). Glycosylation of non-consensus sequons by OST has also been observed (15, 40, 42). Interestingly, the reported turnover rates show the same activity pattern as the turnover rates we observed for ApNGT. As noted before (21, 23), there seems to be a convergent evolution of NGT and the eukaryotic *N*-glycosylation system. Even though NGT shares none of the structural features or conserved motifs of OSTs (such as the WWDYG motif that is responsible for binding of the hydroxyl amino acid at position +2 of the sequon in OST(13)), it exhibits a remark-

ably similar peptide substrate specificity. A "molecular mimicry" might be the evolutionary driving force that shaped the peptide substrate specificity of NGTs in pathogenic bacteria.

Although for OST, the catalytic mechanism of asparagine side chain activation gets clearer and clearer, the reaction mechanism of NGT remains elusive. Apart from *N*-glycosylation of asparagine residues, ApNGT also accepts glutamine and catalyzes the *O*-linked glycosylation of homoserine *in vitro* and serine *in vivo* (23). It is unable to glycosylate a carboxyl function (aspartic acid), a secondary amide (*N*-methyl asparagine), or an amine function (diaminobutyric acid). Although these substrate specificities are very similar to what has been reported for the *C. lari* OST PglB (16), a similar reaction mechanism based on a twisted amide transition state is unlikely as none of the structural features necessary, immobilization of asparagine carbonyl, appropriately oriented hydrogen bond acceptors for amide twisting, can be found in NGT. Although twisting the amide leads to a more nucleophilic nitrogen, it also leads to higher reactivity of the C = O group, which behaves essentially as a ketone (43) and would lead to an unwanted side reaction if not properly protected. In OST, the asparagine side chain is stuck in a tunnel formed by the enzyme and only the nitrogen is solvent accessible (13). No such feature is evident from the crystal structure of ApNGT, nor are there any appropriately located hydrogen bond acceptors (22). Catalysis by a basic amino acid side chain of the enzyme would be consistent with the observed substrate specificities. However, such a mechanism seems highly unlikely because all basic amino acids in the vicinity of the active site could be mutated to alanine without loss of enzymatic activity, as demonstrated by the immunoblot, analysis of the glycosylation of a model substrate protein in *E. coli* His^{277} has previously been proposed as a possible catalytic base and its mutation decreased enzyme activity as assayed by measuring UDP formation *in vitro* (22). Our *in vitro* analysis of this mutant also demonstrated reduced activity (~14% of wild-type), indicating that the *in vivo* approach we applied is not sensitive enough to reliably detect reductions of enzymatic activity in this order of magnitude. However, even taking into account the inferior sensitivity of the *in vivo* assay, the observed enzymatic activities are not reduced substantially enough to support a direct involvement of these residues in catalysis. The reduced activity upon mutation of His^{277} is therefore rather due to a structural role of the residue or an indirect involvement in enzymatic function (*e.g.* in substrate binding), explaining its conservation among GT41 family glycosyltransferases.

The closely related *O*-GlcNAc transferase is also an inverting glycosyltransferase with a GT-B-fold that functions without an apparent catalytic base residue. Based on the three-dimensional structure of the tertiary enzyme/sugar donor/peptide acceptor complex, Schimpl *et al.* (44) proposed that not an amino acid side chain but rather the pro-*R* oxygen of the α-phosphate of the UDP moiety acts as the catalytic base. They present biochemical evidence in favor of this proposal as only the S_p-αS thiophosphate analog of UDP-GlcNAc but not its R_p diastereomer could be turned over by OGT. Furthermore, ^{31}P NMR titration experiments indicated an unusually high pK_a of 6.4 for the α-phosphate, making a role of this group as a proton acceptor plausible. As OGT and NGT are closely related and structurally similar, it is therefore tempting to speculate that a similar reaction mechanism applies also for NGT. Alternatively, the enzyme could undergo major conformational changes upon peptide binding that would bring other residues into the vicinity of the active site. The structure of ApNGT that was determined from a crystallization condition containing a peptide substrate did not show any electron density for that peptide. Because this structure is virtually identical to the apo structure (22), it remains unclear, if peptide substrate binding would trigger any conformational changes. Therefore, further biochemical and structural investigation is needed to conclusively demonstrate how NGT catalyzes *N*-linked protein glycosylation.

Acknowledgments—We thank Dr. H. Zegzouti (Promega) for supplying a prototype kit of the UDP-Glo glycosyltransferase assay. We also thank Dr. R. Gauss (ETH Zurich) for help with HPLC, the Functional Genomics Center Zurich for help with mass spectrometry, and Dr. Göran Widmalm (Stockholm University) for sharing ^{1}H chemical shift assignments of UDP and UDP-Glc.

REFERENCES

1. Apweiler, R., Hermjakob, H., and Sharon, N. (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta* **1473,** 4–8
2. Aebi, M., Bernasconi, R., Clerc, S., and Molinari, M. (2010) *N*-Glycan structures: recognition and processing in the ER. *Trends Biochem. Sci.* **35,** 74–82
3. Helenius, A., and Aebi, M. (2004) Roles of *N*-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem.* **73,** 1019–1049
4. Sharon, N., and Lis, H. (1995) Lectins–proteins with a sweet tooth: functions in cell recognition. *Essays Biochem.* **30,** 59–75
5. Kaneko, Y., Nimmerjahn, F., and Ravetch, J. V. (2006) Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation. *Science* **313,** 670–673
6. Rudd, P. M., Elliott, T., Cresswell, P., Wilson, I. A., and Dwek, R. A. (2001) Glycosylation and the immune system. *Science* **291,** 2370–2376
7. Marshall, R. D. (1972) *Glycoproteins. Annu. Rev. Biochem.* **41,** 673–702
8. Nothaft, H., and Szymanski, C. M. (2010) Protein glycosylation in bacteria: sweeter than ever. *Nat. Rev. Microbiol.* **8,** 765–778
9. Calo, D., Kaminski, L., and Eichler, J. (2010) Protein glycosylation in Archaea: sweet and extreme. *Glycobiology* **20,** 1065–1076
10. Kelleher, D. J., and Gilmore, R. (2006) An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* **16,** 47R–62R
11. Imperiali, B., Shannon, K. L., Unno, M., and Rickert, K. W. (1992) A mechanistic proposal for asparagine-linked glycosylation. *J. Am. Chem. Soc.* **114,** 7944–7945
12. Bause, E., Breuer, W., and Peters, S. (1995) Investigation of the active site of oligosaccharyltransferase from pig liver using synthetic tripeptides as tools. *Biochem. J.* **312,** 979–985
13. Lizak, C., Gerber, S., Numao, S., Aebi, M., and Locher, K. P. (2011) X-ray structure of a bacterial oligosaccharyltransferase. *Nature* **474,** 350–355
14. Matsumoto, S., Shimada, A., Nyirenda, J., Igura, M., Kawano, Y., and Kohda, D. (2013) Crystal structures of an archaeal oligosaccharyltransferase provide insights into the catalytic cycle of *N*-linked protein glycosylation. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 17868–17873
15. Gerber, S., Lizak, C., Michaud, G., Bucher, M., Darbre, T., Aebi, M., Reymond, J. L., and Locher, K. P. (2013) Mechanism of bacterial oligosaccharyltransferase: *in vitro* quantification of sequon binding and catalysis. *J. Biol. Chem.* **288,** 8849–8861
16. Lizak, C., Gerber, S., Michaud, G., Schubert, M., Fan, Y. Y., Bucher, M., Darbre, T., Aebi, M., Reymond, J. L., and Locher, K. P. (2013) Unexpected reactivity and mechanism of carboxamide activation in bacterial *N*-linked protein glycosylation. *Nat. Commun.* **4,** 2627

17. Grass, S., Buscher, A. Z., Swords, W. E., Apicella, M. A., Barenkamp, S. J., Ozchlewski, N., and St. Geme, J. W., 3rd. (2003) The *Haemophilus influenzae* HMW1 adhesin is glycosylated in a process that requires HMW1C and phosphoglucomutase, an enzyme involved in lipooligosaccharide biosynthesis. *Mol. Microbiol.* **48,** 737–751
18. Grass, S., Lichti, C. F., Townsend, R. R., Gross, J., and St Geme, J. W., 3rd. (2010) The *Haemophilus influenzae* HMW1C protein is a glycosyltransferase that transfers hexose residues to asparagine sites in the HMW1 adhesin. *PLoS Pathog.* **6,** e1000919
19. Gross, J., Grass, S., Davis, A. E., Gilmore-Erdmann, P., Townsend, R. R., and St Geme, J. W., 3rd. (2008) The *Haemophilus influenzae* HMW1 adhesin is a glycoprotein with an unusual *N*-linked carbohydrate modification. *J. Biol. Chem.* **283,** 26010–26015
20. Choi, K. J., Grass, S., Paek, S., St. Geme, J. W., 3rd, and Yeo, H. J. (2010) The *Actinobacillus pleuropneumoniae* HMW1C-like glycosyltransferase mediates *N*-linked glycosylation of the *Haemophilus influenzae* HMW1 adhesin. *PLoS One* **5,** e15888
21. Schwarz, F., Fan, Y. Y., Schubert, M., and Aebi, M. (2011) Cytoplasmic *N*-glycosyltransferase of *Actinobacillus pleuropneumoniae* is an inverting enzyme and recognizes the NX(S/T) consensus sequence. *J. Biol. Chem.* **286,** 35267–35274
22. Kawai, F., Grass, S., Kim, Y., Choi, K. J., St Geme, J. W., 3rd, and Yeo, H. J. (2011) Structural insights into the glycosyltransferase activity of the *Actinobacillus pleuropneumoniae* HMW1C-like protein. *J. Biol. Chem.* **286,** 38546–38557
23. Naegeli, A., Neupert, C., Fan, Y. Y., Lin, C. W., Poljak, K., Papini, A. M., Schwarz, F., and Aebi, M. (2014) Molecular analysis of an alternative *N*-glycosylation machinery by functional transfer from *Actinobacillus pleuropneumoniae* to *Escherichia coli*. *J. Biol. Chem.* **289,** 2170–2179
24. Gawthorne, J. A., Tan, N. Y., Bailey, U. M., Davis, M. R., Wong, L. W., Naidu, R., Fox, K. L., Jennings, M. P., and Schulz, B. L. (2014) Selection against glycosylation sites in potential target proteins of the general HMWC *N*-glycosyltransferase in *Haemophilus influenzae*. *Biochem. Biophys. Res. Commun.* **445,** 633–638
25. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res.* **37,** D233–D238
26. Hart, G. W., Housley, M. P., and Slawson, C. (2007) Cycling of *O*-linked β-*N*-acetylglucosamine on nucleocytoplasmic proteins. *Nature* **446,** 1017–1022
27. Markley, J. L., Bax, A., Arata, Y., Hilbers, C. W., Kaptein, R., Sykes, B. D., Wright, P. E., and Wüthrich, K. (1998) Recommendations for the presentation of NMR structures of proteins and nucleic acids: IUPAC-IUBMB-IUPAB Inter-Union Task Group on the standardization of data bases of protein and nucleic acid structures determined by NMR spectroscopy. *J. Biomol. NMR* **12,** 1–23
28. Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., Keduas, V., and Notredame, C. (2006) Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Res.* **34,** W604–W608
29. Lazarus, M. B., Jiang, J., Kapuria, V., Bhuiyan, T., Janetzko, J., Zandberg, W. F., Vocadlo, D. J., Herr, W., and Walker, S. (2013) HCF-1 is cleaved in the active site of *O*-GlcNAc transferase. *Science* **342,** 1235–1239
30. Dorfmueller, H. C., Borodkin, V. S., Blair, D. E., Pathak, S., Navratilova, I., and van Aalten, D. M. (2011) Substrate and product analogues as human *O*-GlcNAc transferase inhibitors. *Amino Acids* **40,** 781–792
31. Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G. (2008) Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* **77,** 521–555
32. Lolli, F., Mazzanti, B., Pazzagli, M., Peroni, E., Alcaro, M. C., Sabatino, G., Lanzillo, R., Brescia Morra, V., Santoro, L., Gasperini, C., Galgani, S., D'Elios, M. M., Zipoli, V., Sotgiu, S., Pugliatti, M., Rovero, P., Chelli, M., and Papini, A. M. (2005) The glycopeptide CSF114(Glc) detects serum antibodies in multiple sclerosis. *J. Neuroimmunol.* **167,** 131–137
33. Soya, N., Shoemaker, G. K., Palcic, M. M., and Klassen, J. S. (2009) Comparative study of substrate and product binding to the human ABO(H) blood group glycosyltransferases. *Glycobiology* **19,** 1224–1234
34. Kubota, T., Shiba, T., Sugioka, S., Furukawa, S., Sawaki, H., Kato, R., Wakatsuki, S., and Narimatsu, H. (2006) Structural basis of carbohydrate transfer activity by human UDP-GalNAc: polypeptide α-*N*-acetylgalactosaminyltransferase (pp-GalNAc-T10). *J. Mol. Biol.* **359,** 708–727
35. Raman, J., Fritz, T. A., Gerken, T. A., Jamison, O., Live, D., Liu, M., and Tabak, L. A. (2008) The catalytic and lectin domains of UDP-GalNAc: polypeptide α-*N*-acetylgalactosaminyltransferase function in concert to direct glycosylation site selection. *J. Biol. Chem.* **283,** 22942–22951
36. Moréra, S., Imberty, A., Aschke-Sonnenborn, U., Rüger, W., and Freemont, P. S. (1999) T4 phage β-glucosyltransferase: substrate binding and proposed catalytic mechanism. *J. Mol. Biol.* **292,** 717–730
37. Lazarus, M. B., Nam, Y., Jiang, J., Sliz, P., and Walker, S. (2011) Structure of human *O*-GlcNAc transferase and its complex with a peptide substrate. *Nature* **469,** 564–567
38. Sun, H. Y., Lin, S. W., Ko, T. P., Pan, J. F., Liu, C. L., Lin, C. N., Wang, A. H., and Lin, C. H. (2007) Structure and mechanism of *Helicobacter pylori* fucosyltransferase: a basis for lipopolysaccharide variation and inhibitor design. *J. Biol. Chem.* **282,** 9973–9982
39. Petrescu, A. J., Milac, A. L., Petrescu, S. M., Dwek, R. A., and Wormald, M. R. (2004) Statistical analysis of the protein environment of *N*-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology* **14,** 103–114
40. Zielinska, D. F., Gnad, F., Wiśniewski, J. R., and Mann, M. (2010) Precision mapping of an *in vivo N*-glycoproteome reveals rigid topological and sequence constraints. *Cell* **141,** 897–907
41. Kasturi, L., Eshleman, J. R., Wunner, W. H., and Shakin-Eshleman, S. H. (1995) The hydroxy amino acid in an Asn-*X*-Ser/Thr sequon can influence *N*-linked core glycosylation efficiency and the level of expression of a cell surface glycoprotein. *J. Biol. Chem.* **270,** 14756–14761
42. Valliere-Douglass, J. F., Eakin, C. M., Wallace, A., Ketchem, R. R., Wang, W., Treuheit, M. J., and Balland, A. (2010) Glutamine-linked and non-consensus asparagine-linked oligosaccharides present in human recombinant antibodies define novel protein glycosylation motifs. *J. Biol. Chem.* **285,** 16012–16022
43. Kirby, A. J., Komarov, I. V., and Feeder, N. (2001) Synthesis, structure and reactions of the most twisted amide. *J. Chem. Soc. Perkin Trans.* **2,** 522–529
44. 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., and van Aalten, D. M. (2012) *O*-GlcNAc transferase invokes nucleotide sugar pyrophosphate participation in catalysis. *Nat. Chem. Biol.* **8,** 969–974
45. Landström, J., Persson, K., Rademacher, C., Lundborg, M., Wakarchuk, W., Peters, T., and Widmalm, G. (2012) Small molecules containing hetero-bicyclic ring systems compete with UDP-Glc for binding to WaaG glycosyltransferase. *Glycoconj. J.* **29,** 491–502
46. Goddard, T. D. and Kneller, D. G. (2008) *Sparky*, Version 3, University of California, San Francisco, CA