The Recognition Motif of the Glycoprotein-Folding Sensor Enzyme UDP-Glc:Glycoprotein Glucosyltransferase†

Kiichiro Totani,[‡,§] Yoshito Ihara,[||] Takashi Tsujimoto,[‡,⊥,#] Ichiro Matsuo,[‡,▽] and Yukishige Ito*[,‡]

RIKEN Advanced Science Institute, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan, Department of Biochemistry, Wakayama Medical University, Wakayama 641-8509, Japan, and RIKEN Research Center for Allergy and Immunology, Yokohama, Kanagawa 230-0045, Japan

Received November 5, 2008; Revised Manuscript Received February 16, 2009

ABSTRACT: The folding of glycoproteins is primarily mediated by a quality control system in the ER, in which UDP-Glc:glycoprotein glucosyltransferase (UGGT) serves as a "folding sensor". In this system, client glycoproteins are delivered to UGGT after the trimming of their innermost glucose residue by glucosidase II, which releases them from the lectin chaperones calnexin (CNX) and calreticulin (CRT). UGGT is inactive against folded proteins, allowing them to proceed to the Golgi apparatus for further processing to complex- or hybrid-type glycoforms. On the other hand, this enzyme efficiently glucosylates incompletely folded glycoproteins to monoglucosylated structures, providing them with an opportunity to interact with CNX/CRT. In order to clarify the mode of this enzyme's substrate recognition, we conducted a structure−activity relationship study using a series of synthetic probes. The inhibitory activities of various glycans suggest that UGGT has a strong affinity for the core pentasaccharide ($Man_3GlcNAc_2$) of high-mannose-type glycans. Our comparison of the reactivity of acceptors that have been modified by various aglycons supports the hypothesis that UGGT recognizes the hydrophobic region of client glycoproteins. Moreover, we discovered fluorescently labeled substrates that will be valuable for highly sensitive detection of UGGT activity.

The endoplasmic reticulum (ER),[1] the upstream compartment of the protein secretory pathway, is a factory for the processing, folding, and oligomeric assembly of newly generated proteins (*1*). The protein maturation process in the ER is regulated by a protein quality control system (*2, 3*), in which oligosaccharides function as tags. A majority of the polypeptides delivered to the ER are N-glycosylated at the asparagine residue of Asn-Xaa-Ser/Thr with a high-mannose-type tetradecasaccharide $Glc_3Man_9GlcNAc_2$ (G3M9) (Figure 1A). This oligosaccharide is transferred as a block from dolichyl pyrophosphate by a multisubunit enzyme, oligosaccharyl transferase (*4*). The processing of glycoproteins in the ER is mediated by a range of glycosidases such as glucosidase I (G-I), glucosidase II (G-II), mannosidase I (M-I), and mannosidase II (M-II), which is also known as Man2C1, a cytosolic mannosidase (*5, 6*). The glycoforms generated by these enzymes are able to interact specifically with molecular chaperones, cargo receptors, mannosidase-like proteins, and enzymes to maximize the production and transport of folded proteins (*2, 3*).

In the glycoprotein quality control system, the calnexin (CNX)/calreticulin (CRT) cycle plays a central role (*2*). It relies upon the interconversion of high-mannose-type glycans between their mono- and nonglucosylated forms (Figure 1B). CNX (*7*) and its soluble homologue CRT (*8*) are lectin-like chaperones, which specifically capture glycoproteins that contain monoglucosylated high-mannose-type glycans, most typically $Glc_1Man_9GlcNAc_2$. Subsequently, the carrier enzyme G-II converts client glycoproteins to their nonglucosylated forms (e.g., $Man_9GlcNAc_2$) (*9*), surrendering them to UDP-glucose:glycoprotein glucosyltransferase (UGGT) (*10*). UGGT functions as a "folding sensor" and is able to reglucosylate the $Man_9GlcNAc_2$ (M9) groups of partially folded proteins to regenerate $Glc_1Man_9GlcNAc_2$ (G1M9), which in turn is captured by CNX/CRT. By contrast, folded proteins are not accepted by this enzyme and are instead transported to the Golgi apparatus for further processing into complex-type and hybrid-type structures. Curiously, evidence suggests that severely misfolded proteins are also ignored by UGGT and are delivered to the ER-associated degradation (ERAD) pathway (*3*). As such, this system maximizes the production of folded glycoproteins, while minimizing the nonproductive entry of terminally misfolded glycoproteins

† This work was partially supported by the Ministry of Education, Culture, Sports, Science, and Technology [Grant-in-Aid for Young Scientists (B), No. 17750170 (K.T.) and Grant-in-Aid for Creative Scientific Research, No. 17GS0420 (Y.I.)], the Human Frontier Science Program (RGP0031/2005-C), and the RIKEN Chemical Genomics Program.

* To whom correspondence should be addressed. Telephone: 81-48-467-9430. Fax: 81-48-462-4680. E-mail: yukito@riken.jp.

‡ RIKEN Advanced Science Institute.

§ Current address: Department of Materials and Life Science, Seikei University, Tokyo 180-8633 Japan.

|| Wakayama Medical University.

⊥ RIKEN Research Center for Allergy and Immunology.

Current address: Department of Chemistry and Biological Science, Aoyama Gakuin University, Sagamihara, Kanagawa 229-8558, Japan.

▽ Current address: Department of Chemistry and Chemical Biology, Gunma University, Gunma 376-8515 Japan.

[1] Abbreviations: ER, endoplasmic reticulum; UGGT, UDP-glucose: glycoprotein glucosyltransferase; CNX, calnexin; CRT, calreticulin; G-II, glucosidase II; MTX, methotrexate; HPLC, high-performance liquid chromatography; PEG, polyethylene glycol.

10.1021/bi8020586 CCC: $40.75
Published on Web 02/17/2009

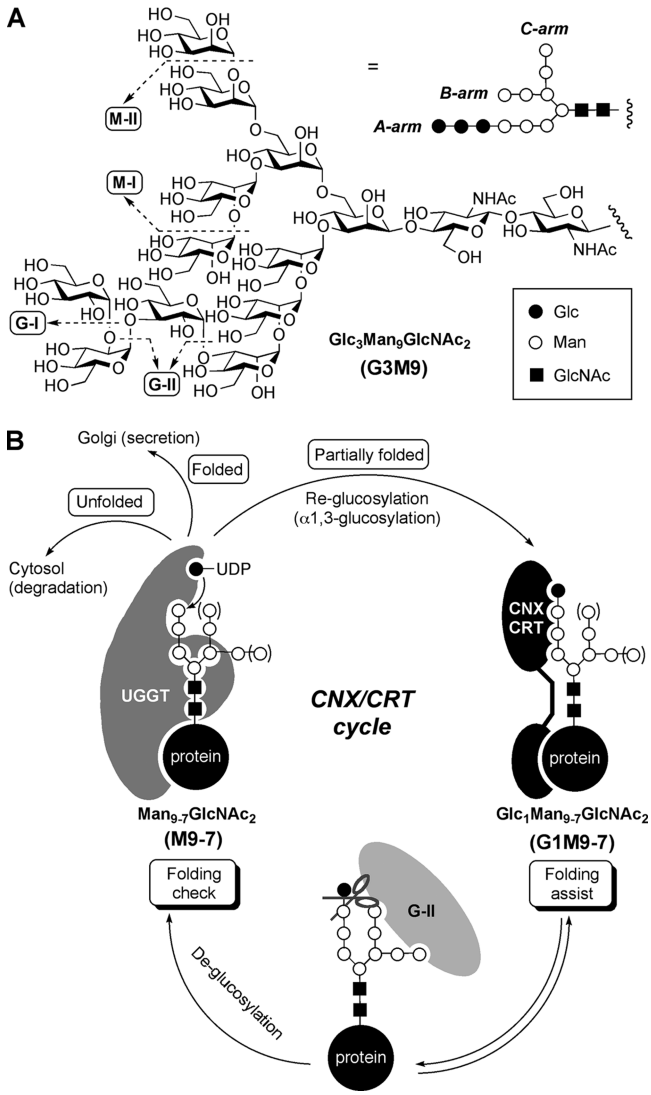

FIGURE 1: Glycan-controlled ER glycoprotein quality control. (A) The structure of the glycan tag $Glc_3Man_9GlcNAc_2$ and its processing in the ER. (B) Schematic diagram of the calnexin/calreticulin cycle, the central mechanism of ER glycoprotein quality control.

into the CNX/CRT cycle. To achieve a precise understanding of this process, we conducted quantitative analyses of the oligosaccharide specificities of CRT (*11, 12*), G-II (*13*), and UGGT (*14*), using synthetic oligosaccharides and their derivatives (*15, 16*), and have provided some insights into the mechanisms of CRT (*17*) and G-II (*18*). On the other hand, the detailed analysis of UGGT has been problematic because of its unusual specificity. In this study, we describe the quantitative structure−activity relationship study of UGGT utilizing chemically synthesized molecular probes and inhibitors.

UGGT is a soluble glycoprotein with a MW of 170 kDa that is ubiquitously expressed in the ER lumen of most eukaryotic cells (*19*−*22*). It consists of a large (80%) N-terminal folding sensor domain and a smaller (20%) catalytic C-terminal domain (*23*). Previous studies using glycoproteins or neoglycoproteins have suggested that UGGT senses exposed hydrophobic patches of unfolded polypeptides (*24*−*28*). In fact, glycopeptides obtained by trypsin digestion (>12 amino acids) that contained hydrophobic amino acid clusters close to their glycosylation site were reactive (*29*), while free glycans and shorter glycopeptides were poor substrates (*30*). Our results show that UGGT strongly recognizes the core pentasaccharide $Man_3GlcNAc_2$ (M3) region of high-mannose-type glycoprotein oligosaccharides, suggesting its capacity to accept a wide range of *N*-glycan structures. Our study also revealed a correlation between the relative reactivity of substrates and the proximity of the recognition elements (the reaction site, the core pentasaccharide M3, and the aglycon hydrophobic patch). In addition, highly reactive substrates of UGGT were found, which will be useful as effective molecular probes for detecting minute UGGT activity.

EXPERIMENTAL PROCEDURES

Synthetic Glycans. The preparation of the oligosaccharide derivatives that were screened as substrates or inhibitors of UGGT was conducted as described in Supporting Information. All of these compounds except **M1-G-MTX** and **M5A-G-MTX** were chemically synthesized based on our convergent strategy (*15*). Compounds **M1-G-MTX** and **M5A-G-MTX** were prepared by enzymatic digestion of **M3-G-MTX** and **M7A-G-MTX**, respectively. All of them were rigorously characterized by ^{1}H NMR and matrix-assisted laser desorption ionization time-of-flight mass spectroscopy.

Inhibition Assay of UGGT. Glycan-G-MTX (**GN1-**, **GN2-**, **M1-**, **M3-**, **M5A-**, or **M7A-G-MTX**) (0, 50, 250, or 500 μM) were included in reaction mixtures containing, in a total volume of 100 μL, 5 mM UDP-Glc, 2 μg of soluble UGGT purified from a rat's liver (*32*), 0.5 mM deoxynojirimycin, 0.5 mM deoxymannojirimycin, 10 mM $CaCl_2$, 0.6% Triton X-100, and 4 mM Tris·HCl (pH 8.0). After a 30 min preincubation period at 37 °C, the acceptor substrate **M9-G-MTX** (50 μM) was added. After being incubated at 37 °C for 1 h, the mixtures were diluted with 100 μL of H_2O and 200 μL of CH_3CN and then heated at 100 °C for 1 min to inactivate the enzyme. The percentage of glucose transfer in each reaction was analyzed by a HPLC column (a TSK-GEL Amide-80 (4.6 mm i.d. × 25 cm) with 3% AcOH−Et_3N (pH 7.3)/CH_3CN mixed solvent (35:65 to 50:50, linear gradient for 50 min) at 40 °C and a 1 mL/min flow rate. The glycan-G-MTX were detected by absorption at 304 nm. The percentage inhibition was calculated by comparing the glucose transfer between the reactions with and without the inhibitor.

Glucose Transfer Assay (Method A for Figure 4A and Method B for Figure 4B−*E).* Method A. Each assay was carried out in a 50 μL reaction mixture [4 mM Tris·HCl (pH 8.0), 10 mM $CaCl_2$, 50 μM deoxynojirimycin, and 0.6% Triton-X100] containing UGGT (4 μg), UDP-[^{3}H]Glc (0.4 μCi), and 50 μM acceptor substrate (**M9-G-MTX**, **M9-G-E**, **M9-G-E-Fmoc**, **M9-G**, **M9-G-Fmoc**, or **M9**). After incubation at 37 °C for 1 h, the reactions were stopped by adding 700 μL of cold Con A-binding buffer [20 mM Tris·HCl (pH 8.0), 150 mM NaCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, 1 mM $MnCl_2$, and 0.05% NP40]. The reaction mixtures were treated with 30 μL of Con A−Sepharose beads (Amersham Biosciences, 50% slurry) at 4 °C for 1 h with agitation in an orbital shaker and centrifuged for 2 min at 1500*g*, before the supernatants were discarded. The Con A beads were rinsed three times with the binding buffer and centrifuged as above. The beads were boiled for 5 min in 100 μL of elution buffer (6 M guanidine hydrochloride), and

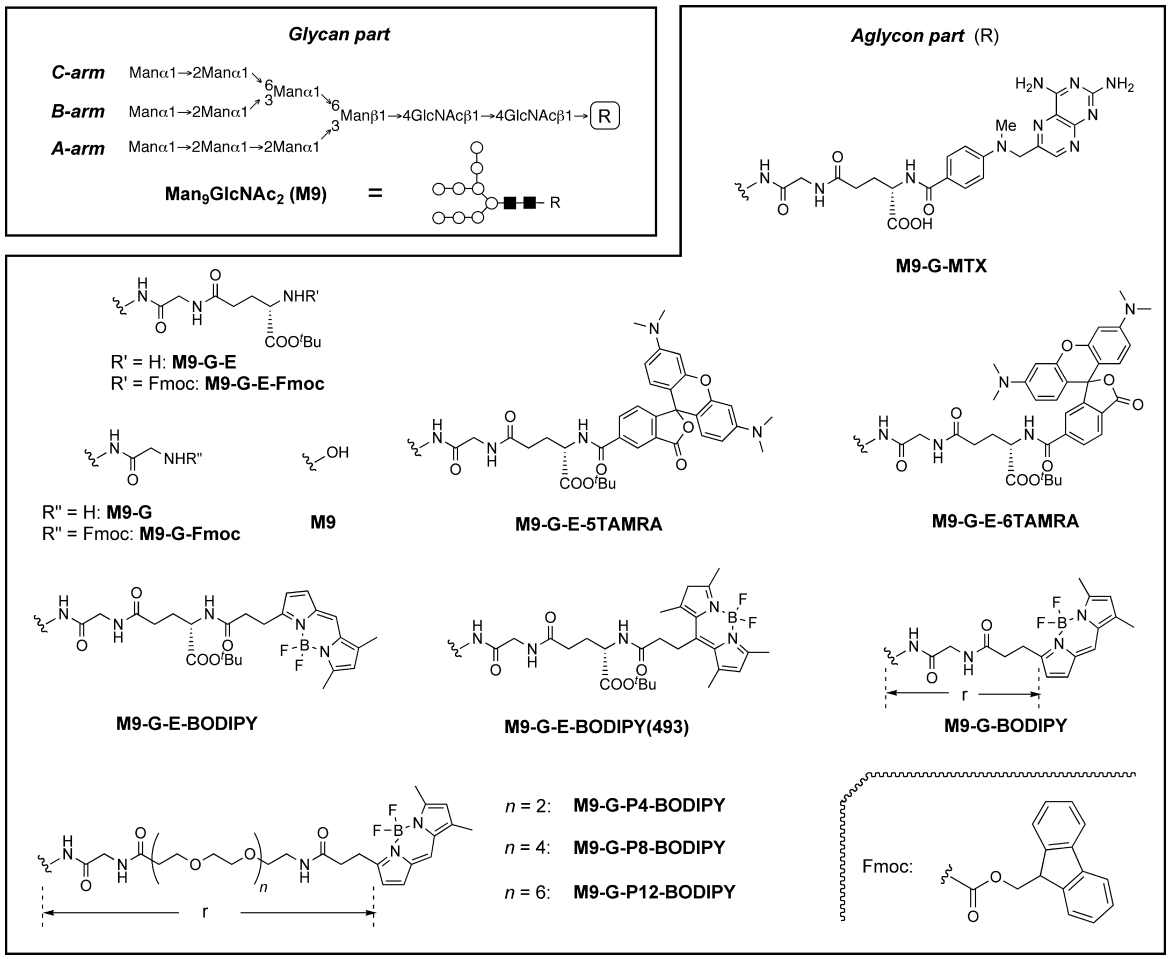

FIGURE 2: The synthetic UGGT substrate used in this study.

the bead-associated radioactivity was quantified by liquid scintillation counting. Assays without an acceptor were used as the control. The initial velocity of each substrate was calculated from the value of [^{3}H]Glc uptake at 1 h of incubation.

Method B. The reaction mixtures contained, in a total volume of 40 μL, 25 μM acceptor substrate (**M9-G-MTX**, **M9-G-E-5TAMRA**, **M9-G-E-6TAMRA**, **M9-G-E-BODIPY**, **M9-G-E-BODIPY(493)**, **M9-G-BODIPY**, **M9-G-P4-BODIPY**, **M9-G-P8-BODIPY**, or **M9-G-P12-BODIPY**), 2.5 mM UDP-Glc, 0.2 mg of soluble rat liver UGGT, 250 μM deoxynojirimycin, 250 μM deoxymannojirimycin, 10 mM $CaCl_2$, 0.6% Triton X-100, and 4 mM Tris·HCl (pH 8.0). After 1 or 4 h at 37 °C, 20 μL of the mixtures was removed by a micropipet and diluted with 20 μL of CH_3CN to stop the enzymatic reaction. The percentage of glucose transfer in each reaction was analyzed by HPLC as described in the Inhibition Assay of UGGT section with slight modifications [mobile phase, 3% AcOH−Et_3N (pH 7.3)/CH_3CN for **M9-G-MTX**, H_2O/CH_3CN for fluorescent substrates; detection, **M9-G-MTX** (absorption at 304 nm), **M9-G-E-5TAMRA**, **M9-G-E-6TAMRA** (fluorescence at 580 nm, excitation 555 nm), **M9-G-E-BODIPY**, **M9-G-BODIPY**, **M9-G-P4-BODIPY**, **M9-G-P8-BODIPY**, **M9-G-P12-BODIPY** (fluorescence at 513 nm, excitation 504 nm), **M9-G-E-BODIPY(493)** (fluorescence at 503 nm, excitation 493 nm)]. The initial velocity of each substrate was calculated from the value of Glc transfer at 1 h of incubation.

Molecular Modeling of BODIPY-Linked M9. All calculations were performed on a Redhat Linux server using MacroModel version 8.1 software. The initial structure ($Man_3GlcNAc_2$-Gly) built within Maestro version 5.0 software was subjected to conjugate gradient energy minimization with the AMBER* force field and the GB/SA water model, and the dielectric constant was set at 4.0. The Monte Carlo (MC) approach was used for the global conformational search. Six-membered rings were fixed in their most stable $^4C_1$ conformation, and all other torsion angles were randomly modified at each MC step. MC steps (10000) were carried out for all compounds, and after each MC step, the resultant geometry was minimized using 1000 gradient conjugate steps. All conformers found within 50 kJ/mol of the global minimum were stored. The conformer with the lowest potential energy was then converted into high-mannose substrates ($Man_9GlcNAc_2$). From this initial structure, modification to appropriate substrates was performed with Maestro version 5.0 software. After the construction of each structure, the MC approach was also used for the global conformation search. In this search, the moiety of the initial structure ($Man_3GlcNAc_2$-Gly) was fully restricted. The other six-membered rings remained as above, and all other torsion angles were randomly modified at each MC step. MC steps (10000) were carried out for all compounds, and after each MC step, the resultant geometry was minimized using 1000 gradient conjugate steps. All conformers found within 50

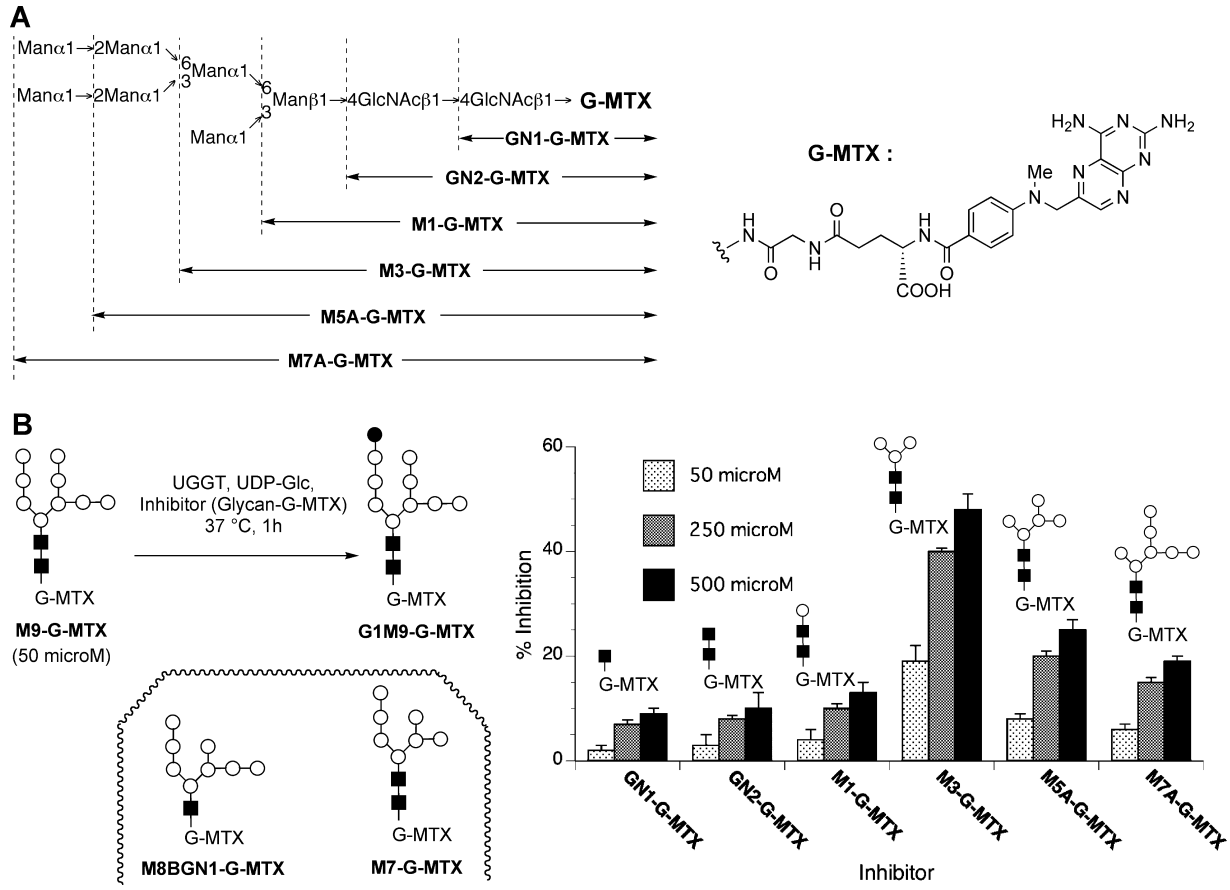

FIGURE 3: An analysis of how glycan is recognized by UGGT. (A) The structure of the synthetic inhibitors of UGGT. (B) Inhibition of UGGT by various methotrexate-conjugated glycans that lack the reaction site. The enzyme assays were carried out in a 100 μL reaction mixture containing 5 mM UDP-Glc, 2 μg of UGGT, 0.5 mM deoxynojirimycin, 0.5 mM deoxymannojirimycin, 50−500 μM inhibitor, 0.6% Triton X-100, 10 mM $CaCl_2$, and 4 mM Tris·HCl (pH 8.0). After a 30 min preincubation at 37 °C, the acceptor substrate **M9-G-MTX** (50 μM) was added. After being incubated at 37 °C for 1 h, the mixtures were diluted with 100 mL of H_2O and 200 μL of CH_3CN and then heated at 100 °C for 1 min to inactivate the enzyme. The reaction mixtures were analyzed by HPLC under the following conditions: a TSK-GEL Amide-80 column (4.6 mm i.d. × 25 cm) and a mobile phase of 3% AcOH−Et_3N (pH 7.3)/CH_3CN mixed solvent (35:65 to 50:50, linear gradient for 50 min) at 40 °C with a flow rate of 1.0 mL/min.

kJ/mol of the global minimum were stored. The conformers of each substrate with the lowest potential energy were compared.

RESULTS AND DISCUSSION

Our previous study revealed that a fully synthetic nonpeptidic substrate, $Man_9GlcNAc_2$-Gly-MTX (**M9-G-MTX**; MTX, methotrexate), is an excellent substrate of UGGT for producing $Glc_1Man_9GlcNAc_2$-Gly-MTX (**G1M9-G-MTX**) (Figure 2). Its activity is at least comparable to that of denatured thyroglobulin, a benchmark glycoprotein substrate of UGGT (*10*). In addition, the specificity of UGGT for glycan was analyzed using a series of MTX derivatives that contain ER-type oligosaccharides, showing that (1) their activity toward UGGT is diminished as the numbers of mannose residues in their B- and C-arms decrease, (2) the presence of a chitobiose ($GlcNAc_2$) is an absolute requirement, and (3) the glucosylation is accelerated in the presence of CRT, which recognizes the product of UGGT ($Glc_1Man_{9-7}GlcNAc_2$).

In their pioneering work, Sousa et al. (*30, 31*) revealed the role of the innermost GlcNAc residue as a recognition element of UGGT. Namely, a denatured endo-H-treated glycoprotein, which retains the innermost GlcNAc residue, but not a nonglycosylated denatured protein, inhibits UGGT. With the prototypical substrate **M9-G-MTX** in hand, we set out to identify the glycan motif that is recognized by UGGT by inhibition assays using the truncated structures **GN1-G-MTX**, **GN2-G-MTX**, **M1-G-MTX**, **M3-G-MTX**, **M5A-G-MTX**, and **M7A-G-MTX** (Figure 3A) (see also the Supporting Information) (Figure 3B). While all of them exhibited inhibition in a dose-dependent manner, their potency was clearly dependent upon the structure of the glycans. When a 10-fold concentration of **GN1-G-MTX** was included, the glucosylation was slightly (ca. 10%) retarded, which is consistent with the study of Sousa et al. (*30, 31*). As the disaccharide **GN2-G-MTX** and trisaccharide **M1-G-MTX** homologues exhibit similar activities to **GN1-G-MTX**, the contribution of the second GlcNAc and the core mannose (Manβ) seems to be marginal. The observation that **M8BGN1-G-MTX**, which lacks the innermost GlcNAc, is devoid of both glucosylation activity as an acceptor substrate (*14*) and inhibitory activity toward UGGT (data not shown) suggests that the distance between the innermost GlcNAc and the reaction site is important.

By contrast, a markedly higher degree of inhibition (~50%) was observed with **M3-G-MTX**, which corresponds to the core pentasaccharide of the high-mannose-type oligosaccharides. To our surprise, **M5A-G-MTX** and **M7A-G-MTX**, which have additional mannose residues on their

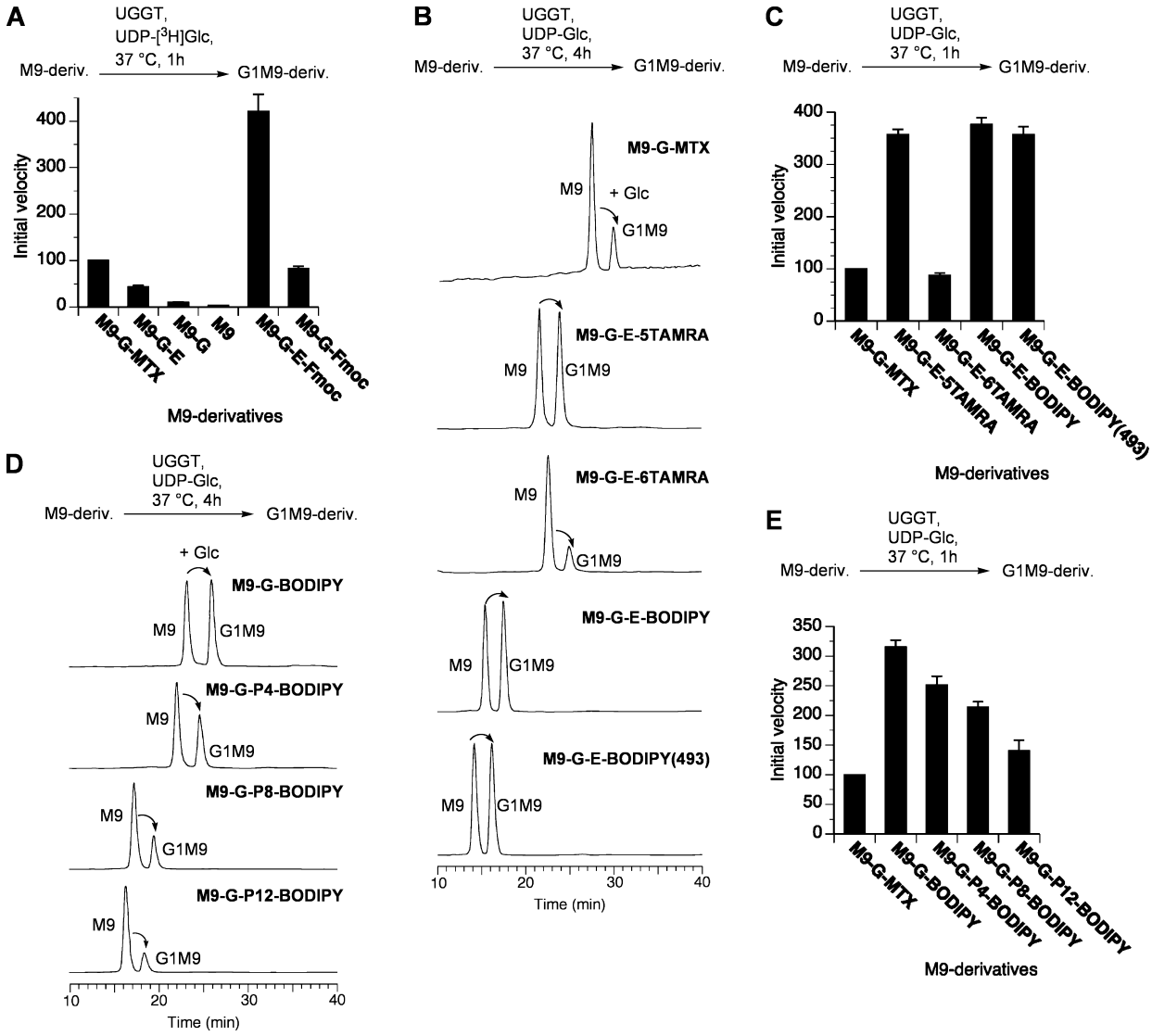

FIGURE 4: An analysis of how aglycon is recognized by UGGT. (A) UGGT mediates the uptake of [^{3}H]Glc by synthetic $Man_9GlcNAc_2$-Gly-MTX and its aglycon-trimmed analogues. Each assay was carried out in a 50 μL reaction mixture [4 mM Tris·HCl (pH 8.0), 10 mM $CaCl_2$, 50 μM deoxynojirimycin, 0.6% Triton X-100] containing UGGT (4 μg), UDP-[^{3}H]Glc (0.4 μCi), and 50 μM acceptor substrate. After incubation at 37 °C for 1 h, the reactions were stopped by adding 700 μL of cold Con A-binding buffer [20 mM Tris·HCl (pH 8.0), 150 mM NaCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, 1 mM $MnCl_2$, and 0.05% NP40]. The reaction mixtures were treated with 30 μL of Con A−Sepharose beads at 4 °C for 1 h. After being rinsed with binding buffer, the beads were boiled for 5 min in 100 μL of elution buffer (6 M guanidine hydrochloride), and the bead-associated radioactivity was quantified by liquid scintillation counting. Assays without an acceptor were used as the background. The initial velocity of each substrate was calculated from the value of [^{3}H]Glc uptake at 1 h of incubation. (B) HPLC profiles of the glucosylation of various fluorescent synthetic substrates: Exploring second-generation substrates for UGGT. Conditions: TSK-GEL Amide-80 column (4.6 mm i.d. × 25 cm); mobile phase 3% AcOH−Et_3N (pH 7.3)/CH_3CN mixed solvent (for **M9-G-MTX**) or H_2O/CH_3CN (for fluorescent substrates) (35:65 to 50:50, linear gradient for 50 min) at 40 °C with a flow rate of 1.0 mL/min. (C) The initial velocity of the fluorescent UGGT substrates. Assay conditions: The reaction mixtures contained, in a total volume of 40 μL, 25 μM acceptor substrate, 2.5 mM UDP-Glc, 0.2 mg of UGGT, 250 μM deoxynojirimycin, 250 μM deoxymannojirimycin, 10 mM $CaCl_2$, 0.6% Triton X-100, and 4 mM Tris·HCl (pH 8.0). After incubation for 1 h at 37 °C, 20 μL of each mixture was removed by a micropipet and diluted with 20 μL of CH_3CN to stop the enzymatic reaction. The percentage of glucose transfer in each reaction was analyzed by HPLC as described in part B. (D) The HPLC profiles of the glucosylation of PEG-linked fluorescent substrates with a systematic elongated linker. The conditions are described in part B. (E) The initial velocity of the PEG-linked fluorescent UGGT substrates. The assay conditions are described in part C. The HPLC analysis conditions are described in part B.

B- and C-arms, showed lower levels of inhibition. These results imply that a bulky B- or C- arm has a negative effect on the binding of UGGT. In fact, our previous study (*14*) revealed that **M9-G-MTX** exhibits a 4.5-fold higher K_m value than **M7-G-MTX**, which lacks the terminal Man residue in both its B- and C-arms. However, the activity of the latter as an acceptor is markedly lower than that of **M9-G-MTX**, implying that Man residues at the B- and C-arms contribute to the binding of the A-arm to the active site. Interestingly, we recently found that M7-glycoforms exhibit a strong inhibition on glucosidase II activity (*13*). Therefore, Man trimming at the B- and C-arms seems to function as a regulation system, in which accumulated M7 attenuates the glycoprotein's entry into the CNX/CRT cycle, reduces the efficiency of the CNX/CRT cycle, and consequently delivers glycoproteins to the ERAD process (*3*).

Our effort was then extended to examine the contribution of the aglycon structure. The aglycon portion of **M9-G-MTX** consists of glycine, glutamic acid, and an aromatic substituent. To define the structural requirements for recognition by

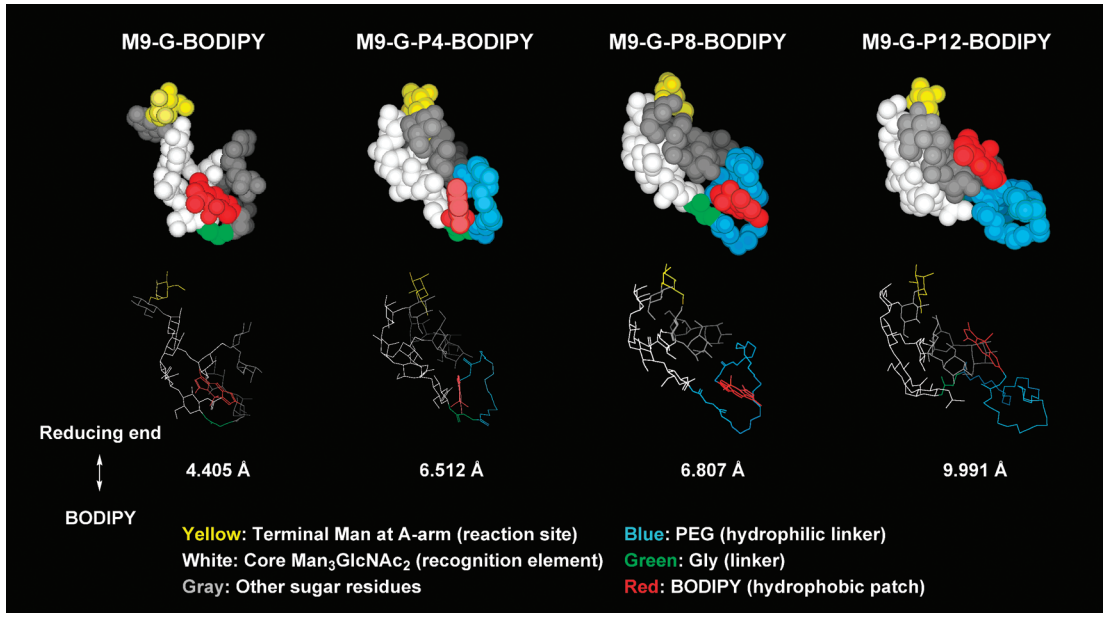

FIGURE 5: Molecular modeling of the PEG-linked fluorescent substrates.

UGGT, we prepared the M9 derivatives **M9-G-E** (see also the Supporting Information) and **M9-G**, in which the aglycon is partially deleted from **M9-G-MTX** (Figure 2). Formation of the G1M9 products was quantified by glucose uptake after incubation with UGGT and the donor substrate UDP-[^{3}H]Glc at 37 °C for 1 h (Figure 4A). The activity of **M9-G-E**, which lacks the aromatic ring system, was reduced to ca. 40% of the activity of **M9-G-MTX**. Further reduction of the activity was observed (ca. 10%) when the aglycon was simplified to a glycine such as **M9-G**, and barely any glucosylation was observed by reducing oligosaccharide **M9** (*15*). These results indicate that UGGT has an affinity for the aromatic part of **M9-G-MTX**, which is consistent with the hypothesis that UGGT recognizes a hydrophobic amino acid cluster that is proximal to the glycosylation site (*29*).

We then compared the activities of derivatives that possess other hydrophobic aglycons, in order to exclude the possibility that UGGT has an affinity for the functionalized ring system of MTX. To this end, the Fmoc-carrying compounds **M9-G-E-Fmoc** and **M9-G-Fmoc** were examined (Figure 2) (see also the Supporting Information). Interestingly, the reactivity of **M9-G-E-Fmoc** was strikingly high, being >4 fold higher than **M9-G-MTX**. On the other hand, the reactivity of compound **M9-G-Fmoc**, which has a shorter distance between its Fmoc site and the glycan, is significantly lower and is similar to that of **M9-G-MTX**. In any event, it is clear that the reactivity of **M9-G-MTX** is not derived from the characteristic structure of MTX.

A subsequent study explored the fluorescently labeled M9 derivatives **M9-G-E-5TAMRA**, **M9-G-E-6TAMRA**, **M9-G-E-BODIPY**, and **M9-G-E-BODIPY(493)** (Figure 2). Gratifyingly, the reactivity of **M9-G-E-5TAMRA**, **M9-G-E-BODIPY**, and **M9-G-E-BODIPY(493)** far exceeded that of **M9-G-MTX** (Figure 4B,C), presumably due to the hydrophobicity of these fluorescent tags compared to MTX. In fact, their retention times in reverse-phase HPLC were longer than that of **M9-G-MTX** (data not shown). Rather unexpectedly, **M9-G-E-6TAMRA** exhibited only one-third of the activity of its 5TAMRA congener **M9-G-E-5TAMRA**. These results suggest that the orientation of the aglycon affects its affinity for UGGT.

Having identified the BODIPY-labeled compound **M9-G-E-BODIPY** as an excellent substrate of UGGT, we synthesized **M9-G-P4-BODIPY**, **M9-G-P8-BODIPY**, and **M9-G-P12-BODIPY** (Figure 2), which have PEG-type linkers of various lengths (see also the Supporting Information) to look into the effects of the distance between the glycan and BODIPY. As shown in Figure 4D,E, the activity dropped gradually as the length of the linker increased, indicating that UGGT favors the presence of a hydrophobic substituent in proximity to the reducing end of high-mannose glycans. A modeling study using the AMBER* force field implemented with the MacroModel program (version 8.1) (Figure 5) gave us estimates of the distances between the reducing end and BODIPY (r; Figure 2) of 4.41, 6.51, 6.81, and 9.99 Å for **M9-G-BODIPY**, **M9-G-P4-BODIPY**, **M9-G-P8-BODIPY**, and **M9-G-P12-BODIPY**, respectively. The stable conformations of these compounds depicted in Figure 5 show that the most efficient substrate, **M9-G-BODIPY**, has a compact conformation whose hydrophobic patch (BODIPY: red) is close to the reaction site (a terminal mannose residue at the A-arm: yellow) and the recognition motif of the glycan part (the core pentasaccharide $Man_3GlcNAc_2$: white). However, as the length of the hydrophilic PEG linker (Blue) increases from **M9-G-BODIPY** to **M9-G-P12-BODIPY**, the BODIPY (red) seems to be located further apart from the recognized glycan motif (white). The glucose transfer activity and the proximity of the three recognition elements (yellow, white, and red) are strongly correlated, suggesting that UGGT preferentially recognizes a hydrophobic patch close to the core pentasaccharide $Man_3GlcNAc_2$.

Although it is clear that the most reactive substrate **M9-G-BODIPY** has the shortest distance, compounds **M9-G-P4-BODIPY**~**M9-G-P12-BODIPY** still have excellent reactivity and are all more reactive than our first-generation substrate **M9-G-MTX**. These results are in agreement with the proposal of Taylor et al. (*27*), who hypothesized that the

presence of a flexible linker domain spans a wide range of distances. It is likely that PEG-type linkers with greater lengths compromise the hydrophobicity of BODIPY, resulting in a reduction of its reactivity. The reactivity of compounds that have alkyl-type linkers is less sensitive to the distance between M9 and BODIPY (see the Supporting Information).

CONCLUSIONS

Our study first identified the UGGT-recognizing motif of the glycan part of acceptor substrates by inhibition experiments with glycan-Gly-MTX derivatives. Inhibition experiments revealed that UGGT recognizes the core pentasaccharide $Man_3GlcNAc_2$ structure, in addition to the reaction site. Although the inhibitory activity of $Man_3GlcNAc_2$ *per se* is not strong, this observation would provide a framework to create a specific inhibitor of UGGT, because this glycoform is unlikely to be present in the ER.

It also provides support for the notion that UGGT recognizes hydrophobic aglycons in addition to the glycan part of acceptor substrates. A comparison of various acceptor substrates revealed that their activity is governed by the hydrophobicity and orientation of the aglycon. We also discovered fluorescent substrates that are highly reactive toward UGGT. These compounds will be useful as molecular probes for the detection of UGGT activity. For instance, a comparison of its expression levels in various cell types would be possible. Further studies are in progress along this line and will be reported in due course.

ACKNOWLEDGMENT

We thank A. Takahashi for technical assistance.

SUPPORTING INFORMATION AVAILABLE

Experimental details for the preparation of substrates and inhibitors and assay data for the UGGT activity of M9-BODIPY conjugates that contain alkyl-type linkers. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Gething, M.-J., and Sambrook, J. (1992) Protein folding in the cell. *Nature 355*, 33–45.
2. Helenius, A., and Aebi, M. (2004) Roles of N-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem. 73*, 1019–1049.
3. Molinari, M. (2007) N-Glycan structure dictates extension of protein folding or onset of disposal. *Nature Chem. Biol. 3*, 313–320.
4. Burda, P., and Aebi, M. (1999) The dolichol pathway of N-linked glycosylation. *Biochim. Biophys. Acta 1426*, 239–257.
5. Suzuki, T., Hara, I., Nakano, M., Shigeta, M., Nakagawa, T., Kondo, A., Funakoshi, Y., and Taniguchi, N. (2006) Man2C1, an α-mannosidase is involved in the trimming of free oligosaccharides in the cytosol. *Biochem. J. 400*, 33–41.
6. Chantret, I., and Moore, S. E. H. (2008) Free oligosaccharide regulation during mammalian protein *N*-glycosylation. *Glycobiology 18*, 210–224.
7. Hammond, C., Braakman, I., and Helenius, A. (1994) Role of N-linked oligosaccharides, glucose trimming and calnexin during glycoprotein folding in the endoplasmic reticulum. *Proc. Natl. Acad. Sci. U.S.A. 91*, 913–917.
8. Peterson, J. R., Ora, A., Van, P. N., and Helenius, A. (1995) Transient lectin-like association of carleticulin with folding intermediates of cellular and viral glycoprotein. *Mol. Biol. Cell 6*, 1173–1184.
9. Grinna, L. S., and Robbins, P. W. (1979) Glycoprotein biosynthesis. Rat liver microsomal glucosidases which process oligosaccharides. *J. Biol. Chem. 254*, 8814–8818.
10. Trombetta, E. S., and Parodi, A. J. (2005) Glycoprotein reglucosylation. *Methods 35*, 328–337.
11. Matsuo, I., Wada, M., Manabe, S., Yamaguchi, Y., Otake, K., Kato, K., and Ito, Y. (2003) Synthesis of monoglucosylated high-mannose-type dodecasaccharide, a putative ligand for molecular chaperone, calnexin and calreticulin. *J. Am. Chem. Soc. 125*, 3402–3403.
12. Ito, Y., Hagihara, S., Matsuo, I., and Totani, K. (2005) Structural approaches to the study of oligosaccharide in glycoprotein quality cotrol. *Curr. Opin. Struct. Biol. 15*, 481–489.
13. Totani, K., Ihara, Y., Matsuo, I., and Ito, Y. (2006) Substrate specificity analysis of endoplasmic reticulum glucosidase II using synthetic high mannose-type glycans. *J. Biol. Chem. 281*, 31502–31508.
14. Totani, K., Ihara, Y., Matsuo, I., Koshino, H., and Ito, Y. (2005) Synthetic substrates for an endoplasmic reticulum protein-folding sensor, UDP-glucose: glycoprotein glucosyltransferase. *Angew. Chem., Int. Ed. 44*, 7950–7954.
15. Matsuo, I., Totani, K., Tatami, A., and Ito, Y. (2006) Comprehensive synthesis of ER related high-mannose-type sugar chain by convergent strategy. *Tetrahedron 62*, 8262–8277.
16. Totani, K., Ihara, Y., Matsuo, I., and Ito, Y. (2006) High-mannose-type glycan modifications of dihydrofolate reductase using glycan-methotrexate conjugates. *Bioorg. Med. Chem. 14*, 5220–5229.
17. Tatami, A., Hon, Y. S., Matsuo, I., Takatani, M., Koshino, H., and Ito, Y. (2007) Analysis of carbohydrate binding property of lectin-chaperone calreticulin. *Biochem. Biophys. Res. Commun. 364*, 332–337.
18. Totani, K., Ihara, Y., Matsuo, I., and Ito, Y. (2008) Effects of macromolecular crowding on glycoprotein processing enzymes. *J. Am. Chem. Soc. 103*, 2101–2107.
19. Trombetta, S., Bosch, M., and Parodi, A. J. (1989) Glucosylation of glycoproteins by mammalian, plant fungal and trypanosomatid protozoa microsomal membranes. *Biochemistry 28*, 8108–8116.
20. Fernandez, F. S., Trombetta, S. E., Hellman, U., and Parodi, A. J. (1994) Purification to homogeneity of UDP-glucose:glycoprotein glucosyltransferase from *Schizosaccharimyces cerevisae*. *J. Biol. Chem. 269*, 30701–30706.
21. Perker, C. G., Fessler, L. I., Nelson, R. E., and Fessler, J. H. (1995) Drosophila UDP-glucose:glycoprotein glucosyltransferase: sequence and characterization of an enzyme that distinguishes between denatured and native proteins. *EMBO J. 14*, 1294–1303.
22. Arnold, S. M., and Kaufman, R. J. (2003) The noncatalytic portion of human UDP-glucose:glycoprotein glucosyltransferase I confers UDP-glucose binding and transferase function to the catalytic domain. *J. Biol. Chem. 278*, 43320–43328.
23. Arnold, S. M., Fessler, L. I., Fessler, J. H., and Kaufman, R. J. (2000) Two homologues encoding human UDP-glucose:glycoprotein glucosyltransferase differ in mRNA expression and enzmatic activity. *Biochemistry 39*, 2149–2163.
24. Ritter, C., and Helenius, A. (2000) Recognition of local glycoprotein misfolding by the ER folding sensor UDP-glucose:glycoprotein glucosyltransferase. *Nat. Struct. Biol. 7*, 278–280.
25. Caramelo, J. J., Castro, O. A., Alonso, L. G., Prat-Gay, G., and Parodi, A. J. (2003) UDP-Glc:glycoprotein glucosyltransferase recognizes structured and solvent accessible hydrophobic patches on molten globule-like folding intermediates. *Proc. Natl. Acad. Sci. U.S.A. 100*, 86–91.
26. Caramelo, J. J., Castro, O. A., Prat-Gay, G., and Parodi, A. J. (2004) The endoplasmic reticulum glucosyltransferase recognizes nearly native glycoprotein folding intermediates. *J. Biol. Chem. 279*, 46280–46285.
27. Taylor, S. C., Ferguson, A. D., Bergeron, J. J. M., and Thomas, D. Y. (2004) The ER protein folding sensor UDP-glucose glycoprotein-glucosyltransferase modifies substrates distant to local changes in glycoprotein conformation. *Nat. Struct. Mol. Biol. 11*, 128–134.
28. Ritter, C., Quirin, K., Kowarik, M., and Helenius, A. (2005) Minor folding defects trigger local modification of glycoproteins by ER folding sensor GT. *EMBO J. 24*, 1730–1738.
29. Taylor, S. C., Thibault, P., Tessier, D. C., Bergeron, J. J. M., and Thomas, D. Y. (2003) Glycopeptide specificity of the secretory protein folding sensor UDP-glucose glycoprotein:glucosyltransferase. *EMBO Rep. 4*, 405–411.

30. Sousa, M. C., Ferrero-Garcia, M. A., and Parodi, A. J. (1992) Recognition of the oligosaccharide and protein moieties of glycoproteins by UDP-Glc:glycoprotein glucosyltransferase. *Biochemistry 31*, 97–105.
31. Sousa, M., and Parodi, A. J. (1995) The molecular basis for the recognition of misfolded glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase. *EMBO J. 14*, 4196–4203.
32. Trombetta, S. E., and Parodi, A. J. (1992) Purification to apparent homogeneity and partial characterization of rat liver UDP-glucose:glycoprotein glucosyltransferase. *J. Biol. Chem. 267*, 9236–9240.

BI8020586