Protein Folding

DOI: 10.1002/ange.200502723

Synthetic Substrates for an Endoplasmic Reticulum Protein-Folding Sensor, UDP-Glucose: Glycoprotein Glucosyltransferase**

*Kiichiro Totani, Yoshito Ihara, Ichiro Matsuo, Hiroyuki Koshino, and Yukishige Ito**

Proteins destined for secretory pathways are transported to the lumen of the endoplasmic reticulum (ER). The ER lumen is a machine for post-translational modifications that is equipped with a quality-control system.[1] It assists with the folding of proteins and prevents aggregation and transport to the Golgi apparatus of misfolded glycoproteins. Protein glycosylation with an Asn-linked (N-linked) oligosaccharide is a co-translational process that occurs in the ER. N-linked glycan is attached to nascent protein as a tetradecasaccharide (Glc3Man9GlcNAc2; Glc = D-glucose, Man = D-mannose, GlcNAc = *N*-acetyl-D-glucosamine), and is trimmed sequentially by glucosidase I (G-I) and glucosidase II (G-II) to tridecasaccharide (Glc2Man9GlcNAc2), dodecasaccharide (Glc1Man9GlcNAc2), and then to undecasaccharide (Man9GlcNAc2; Figure 1).

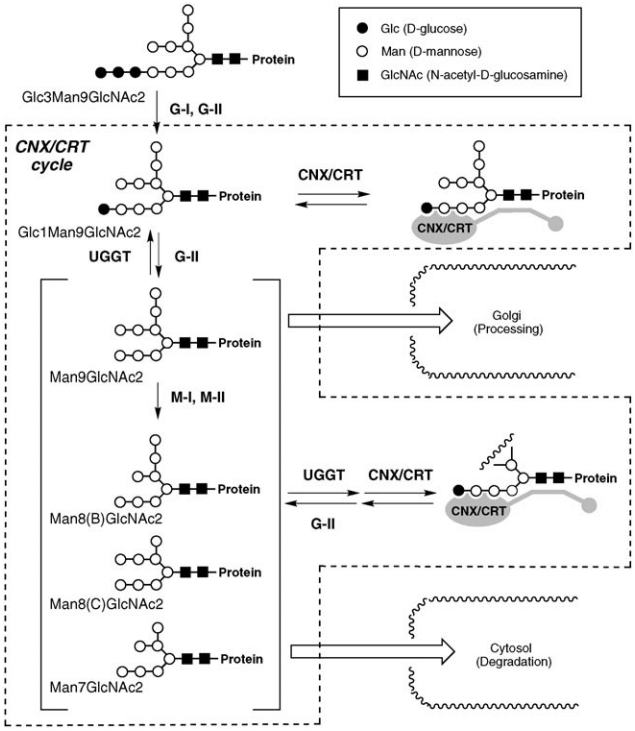


Figure 1. Processing of N-linked oligosaccharides in the ER.

The calnexin (CNX) cycle plays a central role in glycoprotein folding (Figure 1).[1] CNX and its soluble homologue calreticulin (CRT) are lectin chaperones, which have a similar specificity for monoglucosylated high-mannose-type glycans, such as the Glc1Man9GlcNAc2 of N-linked glycoproteins, as primary ligands. These chaperones assist with the folding of glycoproteins by recruiting the disulfide isomerase-like protein ERp57.[2] Clearly, CNX/CRT and G-II compete for the same glycan, and a certain proportion of glycoproteins may reach the nonglucosylated stage (Man9GlcNAc2) without being folded correctly.

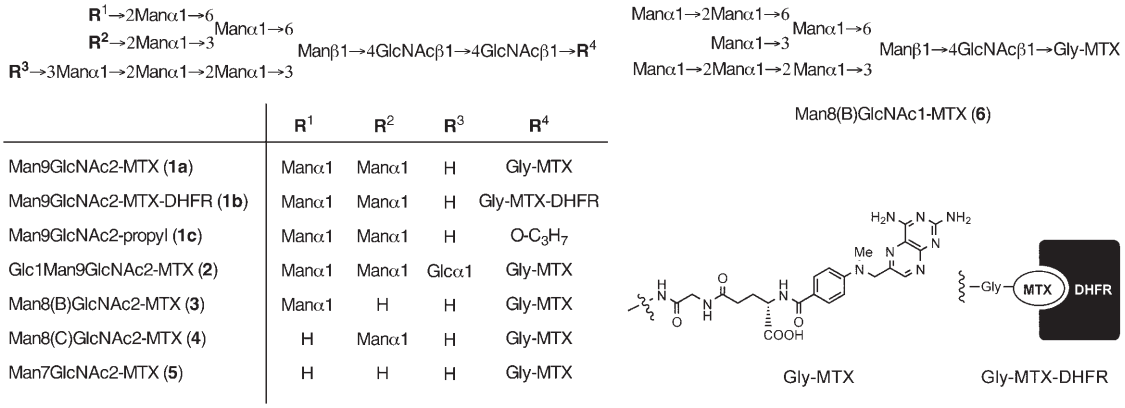
Uridine 5′-diphosphate (UDP)-glucose:glycoprotein glucosyltransferase (UGGT) plays an important role in the CNX/CRT cycle as the "folding sensor".[3] It glucosylates the Man9GlcNAc2 of incompletely folded glycoproteins that arise from immature deglucosylation, and regenerates the CNX/CRT ligand Glc1Man9GlcNAc2. With repeated attempts at folding (CNX/CRT), deglucosylation (G-II), and reglucosylation (UGGT), this system maximizes the chance that newly formed glycoproteins will attain the correct structure. Successfully folded proteins escape from this cycle

[*] Dr. K. Totani, Dr. I. Matsuo, Dr. H. Koshino, Dr. Y. Ito
RIKEN (The Institute of Physical and Chemical Research)
2-1 Hirosawa, Wako, Saitama 351-0198 (Japan)
E-mail: yukito@riken.jp
Dr. K. Totani, Dr. I. Matsuo, Dr. Y. Ito
CREST, Japan Science and Technology Agency (JST)
Kawaguchi, Saitama 332-1102 (Japan)
Fax: (+81) 48-482-4680
Dr. Y. Ihara
Department of Biochemistry & Molecular Biology of Disease
Atomic Bomb Disease Institute
Nagasaki University Graduate School of Biomedical Sciences
1-12-4 Sakamoto, Nagasaki 852-8523 (Japan)

[**] This work was supported by the special postdoctoral researchers program at RIKEN. We thank A. Takahashi for technical assistance. Financial support from the Presidential Fund of RIKEN and the Ministry of Education, Culture, Sports, Science, and Technology (Grant-in-Aid for Young Scientists (B) No. 17750170 and Grant-in-Aid for Creative Scientific Research No. 17GS0420) is acknowledged. UDP = uridine 5′-diphosphate.

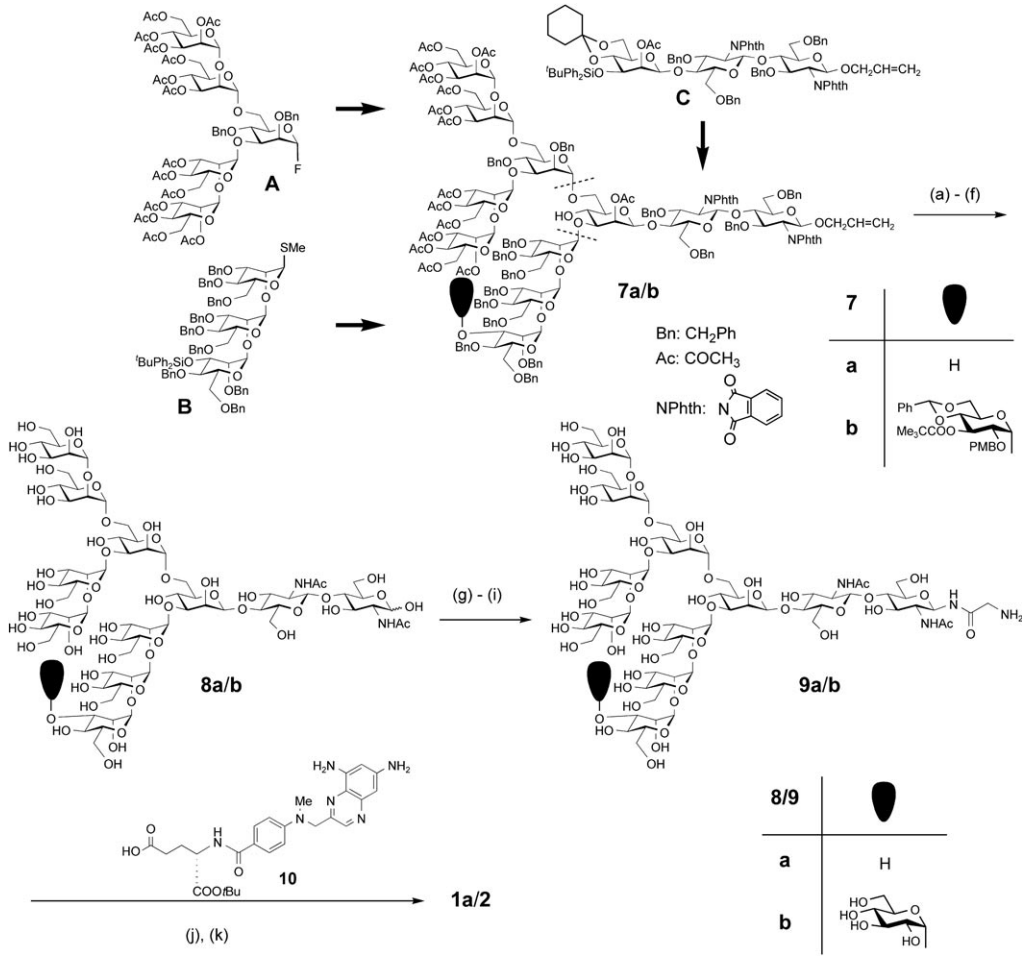
Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

and are transported to the Golgi apparatus for further processing and secretion, while misfolded proteins are retrotranslocated to the cytosol for degradation.[4] ER-resident α-mannosidases may digest Man9GlcNAc2 to Man8GlcNAc2, Man7GlcNAc2, and possibly Man6-5GlcNAc2.[5] UGGT also converts these glycans to monoglucosylated forms that can be CNX/CRT ligands, when protein folding is incomplete, thus constituting an intricate fail-safe system.



	R1	R2	R3	R4
Man9GlcNAc2-MTX (**1a**)	Manα1	Manα1	H	Gly-MTX
Man9GlcNAc2-MTX-DHFR (**1b**)	Manα1	Manα1	H	Gly-MTX-DHFR
Man9GlcNAc2-propyl (**1c**)	Manα1	Manα1	H	O-C_3H_7
Glc1Man9GlcNAc2-MTX (**2**)	Manα1	Manα1	Glcα1	Gly-MTX
Man8(B)GlcNAc2-MTX (**3**)	Manα1	H	H	Gly-MTX
Man8(C)GlcNAc2-MTX (**4**)	H	Manα1	H	Gly-MTX
Man7GlcNAc2-MTX (**5**)	H	H	H	Gly-MTX

Figure 2. Synthetic oligosaccharides used in this study.



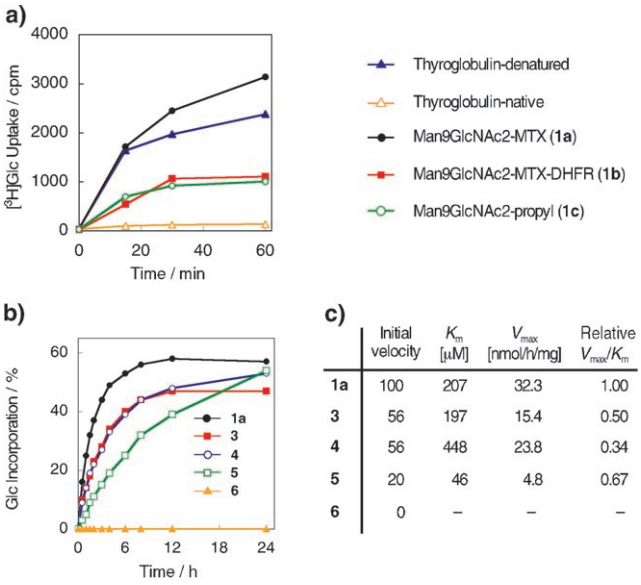
Scheme 1. Synthesis of CHO-MTX conjugates **1a** and **2**. Reagents and conditions: a) ethylenediamine, *n*BuOH, 70°C; b) Ac_2O, pyridine, DMAP; c) $[Ir(cod)(PMePh_2)_2]PF_6$, H_2, THF then I_2, $NaHCO_3$, H_2O; d) Ac_2O, pyridine, DMAP; e) $Pd(OH)_2$, H_2, 80% AcOH, RT; f) NaOMe/MeOH, RT (66% for **1a**, 81% for **2**, six steps); g) saturated aq. NH_4HCO_3, 40°C; h) FmocGlyCl, $NaHCO_3$, dioxane/H_2O (1:1), 0°C; i) piperidine, DMF (81% for **1a**, 79% for **2**, three steps); j) **10**, DMT-MM, MeOH/H_2O (1:1); k) CF_3COOH (62% for **1a**, 65% for **2**, two steps). DMAP = *N*,*N*-dimethylamino-4-pyridine, cod = 1,5-cyclooctadiene, DMT-MM = 4-(4,6-dimethoxy-1,3,5-thiazin-2-yl)-4-methylmorpholinium chloride.

The detailed analysis of UGGT has been problematic because of its unusual specificity; it accepts unglucosylated high-mannose-type glycans only when they are covalently bound to incompletely folded proteins. Remarkably, UGGT rejects folded as well as severely misfolded glycoproteins, and prefers a native-like, compact, molten-globule-type conformation.[3,6,7] Previous studies of UGGT have been conducted using artificially denatured glycoproteins,[6,8] their tryptic digests,[8f] or neoglycopeptides.[7] The development of a simpler and better-defined substrate of UGGT is challenging, because this enzyme requires the coexistence of three separate elements: an acceptor (Manα1→2Man) element, a GlcNAcβ1→Asn element, and a cluster of hydrophobic amino acids exposed on the protein surface. Herein, we describe the first analysis of UGGT-catalyzed glucose transfer, by using chemically synthesized, fully defined, nonpeptidic substrates.

We have recently reported oligosaccharide–methotrexate conjugates (CHO-MTX) as the precursor of artificial glycoproteins having homogeneous carbohydrates.[9] MTX and its derivatives have been investigated as anticancer and antimicrobial drug candidates,[10] because of their inhibitory activity toward dihydrofolate reductase (DHFR), a key enzyme in the biosynthesis of nucleotide precursors of DNA. *Escherichia coli* DHFR is a relatively small ($\approx$18 kDa) protein, which has been used as a tag for induced protein dimerization[11] and the in vitro labeling of proteins.[12] CHO-MTX can be converted to a tight-binding complex with DHFR, by virtue of the strong affinity ($K_A \approx 10^{10}\,M^{-1}$) between MTX and DHFR.[13] In the course of preparing a series of complexes of CHO-MTX and DHFR (Figure 2), we observed that CHO-MTX conjugate **1a** was an excellent acceptor substrate of UGGT.

The synthesis of Man9GlcNAc2-MTX (**1a**) was conducted as shown in Scheme 1. Thus, the protected undecasaccharide **7a**,[14] which was prepared from synthetic blocks **A**, **B**, and **C**, was deprotected to **8a** and converted to glycosylamine under the conditions proposed by Kochetkov et al.[15] Subsequent reaction with FmocGlyCl and removal of Fmoc afforded Man9GlcNAc2-Gly **9a**. Coupling with MTX-(α-*t*Bu)[9,16] was conducted under the conditions reported by Kunishima et al.[17] and cleavage of the *t*Bu ester afforded **1a**. Compound **1a** retained affinity for DHFR, which was confirmed by inhibition experiments[18] and lectin (Con A) binding experiments (see the Supporting Information). The expected UGGT product Glc1Man9GlcNAc2-MTX (**2**) was prepared from the protected dodecasaccharide **7b** in a similar manner.

Rat liver UGGT[19] and UDP-[^{3}H]Glc were used for the efficient glucosylation of **1a** (Figure 3a). Under the assay conditions (50 μM substrate), the reactivity of the synthetic substrate **1a** was at least comparable to that of denatured thyroglobulin, which has been used in conventional UGGT assays.[8a,19] Native thyroglobulin was completely inactive, as reported previously.[19b] We speculate that MTX mimics the property of unfolded proteins, and presumably has an exposed hydrophobic patch. In fact, Man9GlcNAc2 linked to a simpler and less hydrophobic aglycon (**1c**) was much less active, as was the complex Man9GlcNAc2-MTX-DHFR (**1b**).



	Initial velocity	K_m [μM]	V_{max} [nmol/h/mg]	Relative V_{max}/K_m
1a	100	207	32.3	1.00
3	56	197	15.4	0.50
4	56	448	23.8	0.34
5	20	46	4.8	0.67
6	0	–	–	–

Figure 3. Reaction of UGGT with the synthetic substrates. a) Time course of [^{3}H]Glc uptake; Man9GlcNAc2-MTX (**1a**), Man9GlcNAc2-MTX-DHFR (**1b**), Man9GlcNAc2-propyl (**1c**), denatured thyroglobulin, and native thyroglobulin. b) Time course of Glc incorporation; Man9-GlcNAc2-MTX (**1a**), Man8(B)GlcNAc2-MTX (**3**), Man8(C)GlcNAc2-MTX (**4**), Man7GlcNAc2-MTX (**5**), and Man8(B)GlcNAc1-MTX (**6**). c) Initial velocity (calculated from the value at 1 h incubation on (b)), K_m, and V_{max} values of **1a**, **3**, **4**, **5**, and **6**.

In the latter case, MTX may be buried in the folded protein, which makes it less accessible to UGGT.

The enzymatic transformation of **1a** to Glc1Man9GlcNAc2-MTX (II) was readily traced by HPLC and the characteristic UV absorption of MTX (Figure 4a). The product II was separated by preparative HPLC to achieve its characterization. Enzymatically produced II was identical to synthetic **2**, judging from MALDI-TOF mass spectrometry and ^{1}H NMR spectroscopy (1D, differential NOE, 2D TOCSY; see Figure 4b–d and the Supporting Information). Furthermore, when measured by isothermal titration calorimetry (ITC),[20] UGGT-generated II had a binding affinity for CRT with a K_A value ($5.17 \times 10^6\,M^{-1}$) essentially identical to that of the authentic **2** and previously reported Glc1Man9GlcNAc2-propyl[14,21] ($5.26 \times 10^6\,M^{-1}$). These results provided proof of the glycan structure of the UGGT product Glc1Man9GlcNAc2.

We then prepared a series of high-mannose glycan–MTX conjugates, Man8(B)GlcNAc2-MTX (**3**), Man8(C)GlcNA2-MTX (**4**), Man7GlcNAc2-MTX (**5**), and Man8(B)GlcNAc1-MTX (**6**; Figure 2), to examine the specificity of UGGT. As shown in Figure 3b and c, the order of reactivity was Man9 > Man8 > Man7, with relative values of initial velocity of 100, 56, and 20, respectively. These values were in good agreement with previously reported values that were estimated by using protease-digested thyroglobulin.[8a] The B and C isomers of Man8GlcNAc2-MTX had a similar reactivity. The K_m and V_{max} values for **1a**, **3**, **4**, and **5** (calculated from Lineweaver–Burk plots, see the Supporting Information) are provided in Figure 3c, and show that **1a** was the most reactive substrate. Interestingly, compound **6**, which had a single GlcNAc instead

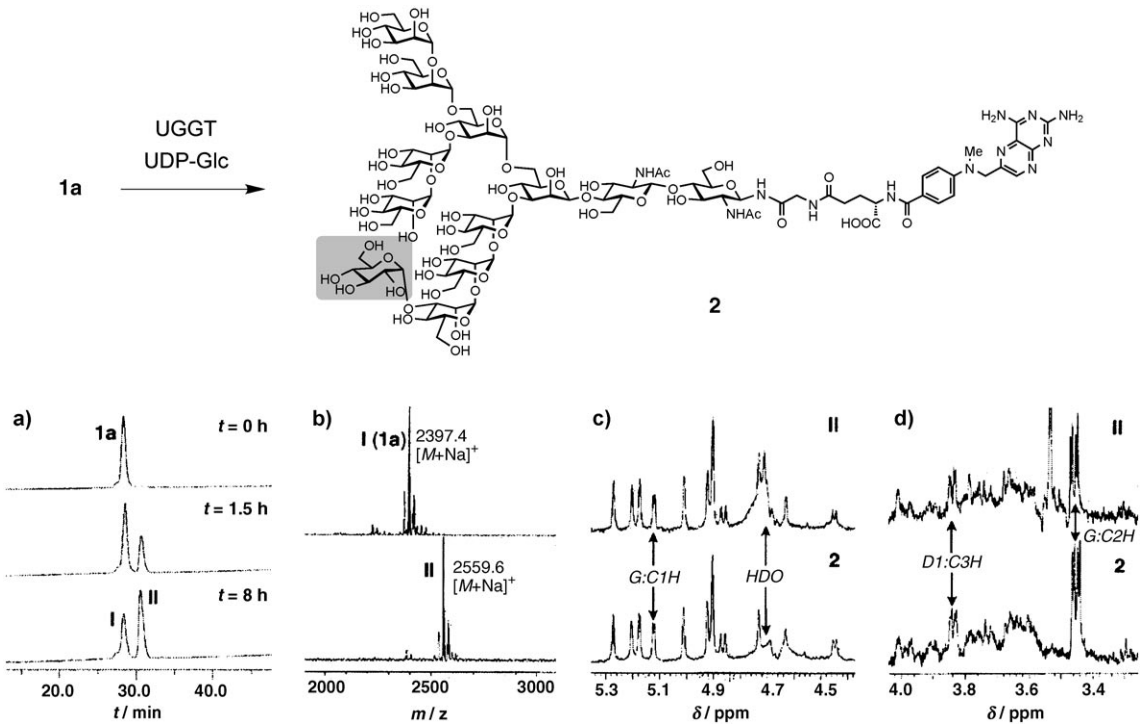


Figure 4. Characterization of the UGGT product. a) HPLC analysis of the glucosylation of Man9GlcNAc2-MTX (**1a**): TSK-GEL Amide-80 column (4.6 mmφ × 25 cm), mobile phase CH_3CN/3% AcOH-Et_3N, pH 7.3, linear gradient from 65:35 to 50:50 in 50 min, flow rate 1.0 mL min^{-1} at 40 °C. b) MALDI-TOF mass spectra of HPLC peaks of compounds **I** (**1a**) and **II** (corresponds to Glc1Man9GlcNAc2-MTX (**2**)). c) ^{1}H NMR spectra of synthetic **2** and the resulting UGGT product **II** in D_2O. The HDO signal was suppressed by DANTE (delays alternating with nutations for tailored excitation) presaturation. d) Differential NOE spectra (trace through the G:C1H resonance) of **2** and **II**.

of chitobiose ($GlcNAc_2$), was completely devoid of activity. UGGT was reported to recognize the innermost GlcNAc moiety,[8b] so the fact that **6** was not accepted might reflect an insufficient distance between the acceptor site Man and innermost GlcNAc.

Although Man7-9GlcNAc2-MTX were all accepted as substrates of UGGT, the glucosylation stopped at 50–60% conversion in every case. Addition of alkaline phosphatase had no effect, which suggests that the result was not caused by accumulated UDP. Interestingly, when recombinant CRT[22] (25 mol%) was added to the incubation mixture well after the reaction reached its plateau (150 h), the glucosylation was boosted (Figure 5a). We speculate that there was inhibition by glucosylated products; CRT trapped the products, segregated them from UGGT, and alleviated the inhibition. In line with this conjecture, a lectin-deficient CRT mutant (D135A)[23] had no effect (Figure 5b). For all the substrates examined (**1a**, **3**, and **4**), enhancement of the conversion (up to >70%) was observed in a dose-dependent manner (Figure 5c).

In summary, we have developed efficient substrates of UGGT that are completely synthetic. An unambiguous characterization of the UGGT product and the quantitative analysis of the specificity of UGGT for glycan were achieved with these substrates. Studies are in progress to systematically prepare various CHO-MTX conjugates. These compounds are expected to be valuable tools with which to investigate the

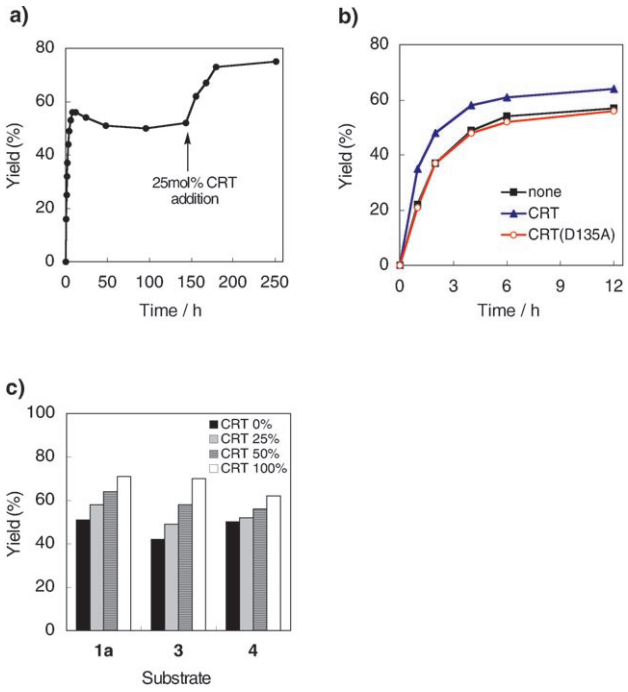


Figure 5. Effect of calreticulin (CRT). a) Effect of 25 mol% CRT addition on **1a**. b) Time course of Glc incorporation for **1a** in the presence or absence of 25 mol% CRT or CRT(D135A) mutant. c) CRT dose-dependent Glc incorporation for **1a**, **3**, and **4** after 48 h incubation at 37 °C.

properties of UGGT and other proteins involved in the ER quality-control system.

Received: August 3, 2005
Published online: November 11, 2005

Keywords: glucose · glycoproteins · glycosylation · protein folding · quality control

[1] A. Helenius, M. Aebi, *Annu. Rev. Biochem.* **2004**, *73*, 1019–1049.
[2] E.-M. Frickel, P. Frei, M. Bouvier, W. F. Stafford, A. Helenius, *J. Biol. Chem.* **2004**, *279*, 18227–18287.
[3] E. S. Trombetta, A. J. Parodi, *Methods* **2005**, *35*, 328–337.
[4] a) Y. Yoshida, T. Chiba, F. Tokunaga, H. Kawasaki, K. Iwai, T. Suzuki, Y. Ito, K. Matsuoka, M. Yoshida, K. Tanaka, T. Tai, *Nature* **2002**, *418*, 438–442; b) Y. Oda, N. Hosokawa, I. Wada, K. Nagata, *Science* **2003**, *299*, 1394–1397; c) M. Molinari, K. K. Eriksson, V. Calanca, C. Galli, P. Cresswell, M. Michalak, A. Helenius, *Mol. Cell* **2004**, *13*, 125–135.
[5] Z. Frenkel, W. Gregory, S. Kornfeld, G. Z. Lederkremer, *J. Biol. Chem.* **2003**, *278*, 34119–34124.
[6] J. J. Caramelo, O. A. Castro, G. de Prat-Gay, A. Parodi, *J. Biol. Chem.* **2004**, *279*, 46280–46285.
[7] J. J. Caramelo, O. A. Castro, L. A. Alonso, G. de Prat-Gay, A. J. Parodi, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 86–91.
[8] a) M. C. Sousa, M. A. Ferrero-Garcia, A. J. Parodi, *Biochemistry* **1992**, *31*, 97–105; b) M. Sousa, A. Parodi, *EMBO J.* **1995**, *14*, 4196–4203; c) E. S. Trombetta, A. Helenius, *J. Cell Biol.* **2000**, *148*, 1123–1129; d) C. Ritter, A. Helenius, *Nat. Struct. Biol.* **2000**, *7*, 278–280; e) D. C. Tessier, D. Dignard, A. Zapun, A. Radominska-Pandya, A. J. Parodi, J. J. M. Bergeron, D. Y. Thomas, *Glycobiology* **2000**, *10*, 403–412; f) S. C. Taylor, A. D. Ferguson, J. J. M. Bergeron, D. Y. Thomas, *Nat. Struct. Mol. Biol.* **2004**, *11*, 128–134; g) S. M. Arnold, R. J. Kaufman, *J. Biol. Chem.* **2003**, *278*, 43320–43328.
[9] K. Totani, I. Matsuo, Y. Ito, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 2285–2289.
[10] a) S. P. Ackland, R. L. Schlisky, *J. Clin. Oncol.* **1987**, *5*, 2017–2031; b) H. Zeng, Z.-S. Chen, M. G. Belinsky, P. A. Rea, G. D. Kruth, *Cancer Res.* **2001**, *61*, 7225–7232.
[11] a) S. Gendreizig, M. Kindermann, K. Johnsson, *J. Am. Chem. Soc.* **2003**, *125*, 14970–14971; b) H. Lin, W. M. Abida, R. T. Sauer, V. W. Cornish, *J. Am. Chem. Soc.* **2000**, *122*, 4247–4248.
[12] L. Miller, J. Sable, P. Goelet, M. P. Sheetz, V. W. Cornish, *Angew. Chem.* **2004**, *116*, 1704–1707; *Angew. Chem. Int. Ed.* **2004**, *43*, 1672–1675.
[13] a) J. R. Appleman, E. E. Howell, J. Kraut, M. Kuhl, R. L. Blakley, *J. Biol. Chem.* **1988**, *263*, 9187–9198; b) S. P. Sasso, R. M. Gilli, J. C. Sari, O. S. Rimet, C. M. Briand, *Biochim. Biophys. Acta* **1994**, *1207*, 74–79.
[14] I. Matsuo, M. Wada, S. Manabe, Y. Yamaguchi, K. Otake, K. Kato, Y. Ito, *J. Am. Chem. Soc.* **2003**, *125*, 3402–3403.
[15] a) L. M. Lihkosherstov, O. S. Novikoya, V. A. Dereviskaja, N. K. Kochetkov, *Carbohydr. Res.* **1986**, *146*, C1–C5; b) I. D. Manger, T. W. Rademacher, R. A. Dwek, *Biochemistry* **1992**, *31*, 10724–10732; c) H. J. Stubbs, M. A. Shia, K. G. Rice, *Anal. Biochem.* **1997**, *247*, 357–365.
[16] A. Rosowski, R. Forsch, J. Uren, M. Wick, *J. Med. Chem.* **1981**, *24*, 1450–1455.
[17] M. Kunishima, C. Kawachi, K. Hioki, K. Terao, S. Tani, *Tetrahedron* **2001**, *57*, 1551–1558.
[18] B. C. Widemann, F. M. Balis, P. C. Adamson, *Clin. Chem.* **1999**, *45*, 223–228.
[19] a) S. E. Trombetta, A. J. Parodi, *J. Biol. Chem.* **1992**, *267*, 9236–9240; b) S. E. Trombetta, M. Bosch, A. Parodi, *Biochemistry* **1989**, *28*, 8108–8116.
[20] T. K. Dam, C. F. Brewer, *Chem. Rev.* **2002**, *102*, 387–429.
[21] Y. Ito, S. Hagihara, M. A. Arai, I. Matsuo, M. Takatani, *Glycoconjugate J.* **2004**, *21*, 257–266.
[22] S. Baksh, M. Michalak, *J. Biol. Chem.* **1991**, *266*, 21458–21465.
[23] M. Kapoor, L. Ellgaard, J. Gopalakrishnapai, C. Schirra, E. Gemma, S. Oscarson, A. Helenius, A. Surolia, *Biochemistry* **2004**, *43*, 97–106.

