The EMBO Journal vol.14 no.17 pp.4196–4203, 1995

The molecular basis for the recognition of misfolded glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase

Marcelo Sousa and Armando J.Parodi[1]

Instituto de Investigaciones Bioquímicas Fundación Campomar, Antonio Machado 151, 1405 Buenos Aires, Argentina

[1]Corresponding author

The UDP-Glc:glycoprotein glucosyltransferase is a soluble enzyme of the endoplasmic reticulum that glucosylates protein-linked $Man_{7-9}GlcNAc_2$ to form the monoglucosylated derivatives. *In vivo* the reaction products are immediately deglucosylated by glucosidase II. The glucosyltransferase has a unique property: it glucosylates misfolded, but not native, glycoproteins. It has been proposed that the glucosyltransferase participates, together with calnexin, in the control mechanism by which only properly folded glycoproteins can exit from the endoplasmic reticulum. In this paper it is demonstrated that the glucosyltransferase recognizes two elements in the acceptor substrates: the innermost *N*-acetylglucosamine unit of the oligosaccharide and protein domains exposed in denatured, but not in native, conformations. Both determinants have to be covalently linked. In many cases the first element is not accessible to macromolecular probes in native conformations. Concerning the protein domains, it is demonstrated here that the glucosyltransferase interacts with hydrophobic amino acids exposed in denatured conformations. More disordered conformations, i.e. those exposing more hydrophobic amino acids, were found to be those having higher glucose acceptor capacity. It is suggested that both accessibility of the innermost *N*-acetylglucosamine unit and binding to hydrophobic patches determine the exclusive glucosylation of misfolded conformations by the glucosyltransferase.

Keywords: endoplasmic reticulum/folding/glucosyltransferase/glycoproteins

Introduction

The lumen of the endoplasmic reticulum is the subcellular site of several modifications in protein structure, such as cleavage of signal peptides, glycosylation of asparagine units and initial oligosaccharide processing reactions, acquisition of secondary, tertiary and in some cases also quaternary structures and anchoring of certain proteins to phosphatidylinositol-containing structures.

Oligosaccharide processing reactions occurring in the endoplasmic reticulum involve removal of three glucose and up to two to three mannose units from the transferred oligosaccharide ($Glc_3Man_9GlcNAc_2$ in most species) (Kornfeld and Kornfeld, 1985) and reglucosylation of the protein-linked $Man_9GlcNAc_2$, $Man_8GlcNAc_2$ and $Man_7GlcNAc_2$ thus formed (Parodi *et al.*, 1984). The newly glucosylated species formed ($Glc_1Man_9GlcNAc_2$, $Glc_1Man_8GlcNAc_2$ and $Glc_1Man_7GlcNAc_2$) are immediately deglucosylated by glucosidase II *in vivo* (Parodi *et al.*, 1983; Trombetta *et al.*, 1989). The glucosylating activity (UDP-Glc:glycoprotein glucosyltransferase) has been detected in mammalian, plant, fungal and protozoan cells (Trombetta *et al.*, 1989; Fernández *et al.*, 1994). The enzyme appeared to be a soluble protein of the lumen of the endoplasmic reticulum and to have a unique property: it glucosylated denatured, but not native, glycoproteins or glycopeptides in cell free assays (Trombetta *et al.*, 1989, 1991; Sousa *et al.*, 1992). The glucosyltransferase behaved, therefore, as a sensor of unfolded, partially folded and misfolded conformations.

Proteins that fail to fold properly are retained in the endoplasmic reticulum, where they are proteolytically degraded. This implies that a very stringent quality control mechanism, able to distinguish between native and misfolded conformations, is operative in the endoplasmic reticulum. A model for such quality control applicable to glycoproteins has recently been proposed (Hammond *et al.*, 1994; Helenius, 1994). According to this model, high mannose-type oligosaccharides in the endoplasmic reticulum shuttle between monoglucosylated and unglucosylated structures, their formation catalyzed by the UDP-Glc:glycoprotein glucosyltransferase and glucosidase II respectively. A membrane-bound chaperone, calnexin, which has a lectin-like activity that recognizes the monoglucosylated oligosaccharides, would bind the monoglucosylated structures and thus retain glycoproteins in the endoplasmic reticulum as long as the protein moieties are not properly folded. On attaining the correct native conformations glycoproteins would become substrates for the glucosidase but not for the glucosyltransferase and thus be liberated from the calnexin anchor. Glycoproteins would then be able to be transported to the Golgi apparatus.

In this paper we establish the molecular basis for the UDP-Glc:glycoprotein glucosyltransferase exclusive recognition of misfolded conformations.

Results

The UDP-Glc:glycoprotein glucosyltransferase interacts with the innermost N-acetylglucosamine residue

It was previously reported that denatured, endo H (endo-β-*N*-acetylglucosaminidase H)-deglycosylated glycoproteins inhibited glucosylation of denatured glycoproteins by UDP-Glc:glycoprotein glucosyltransferase (Sousa *et al.*, 1992; endo H cleaves the bond between the internal *N*-acetylglucosamine units of the oligosaccharides). The inhibition was not glycoprotein specific, as, for instance,

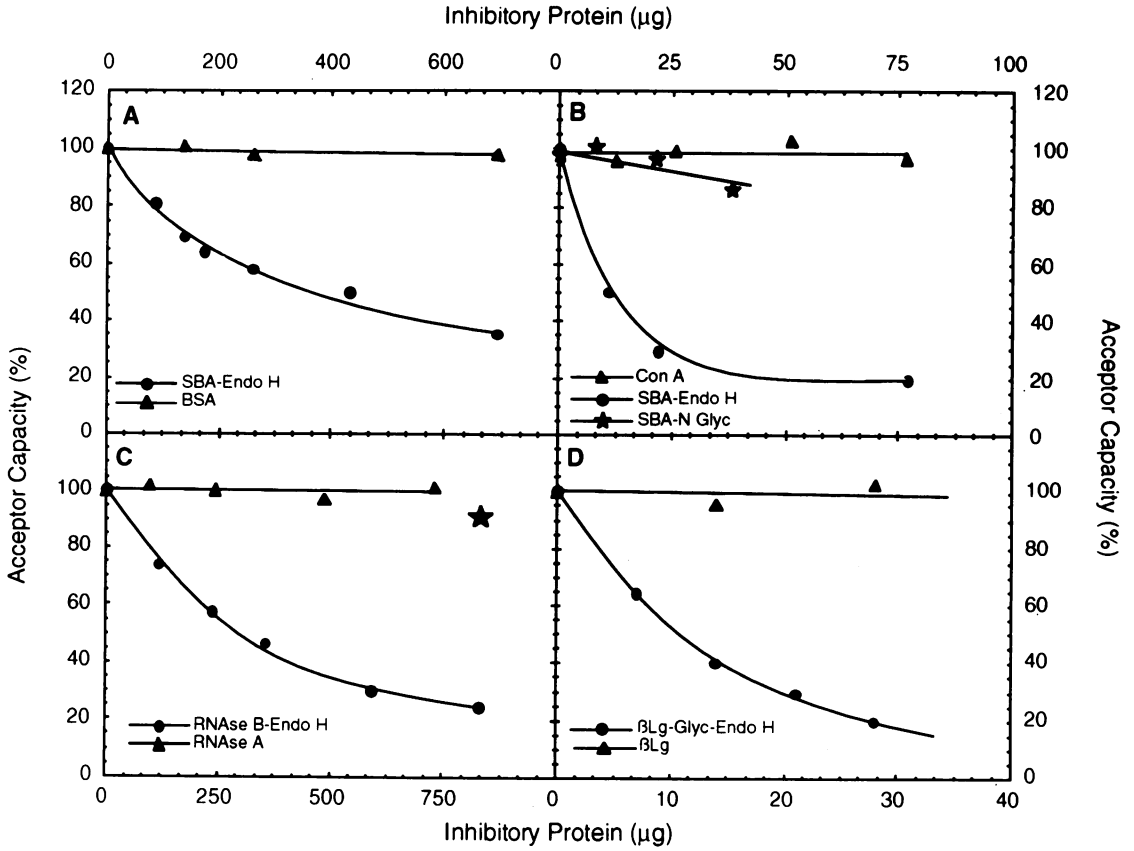

Fig. 1. The interaction of glucosyltransferase with the innermost *N*-acetylglucosamine unit. The following denatured glycoproteins were used as glucose acceptors: (**A**) soybean agglutinin (260 μg/assay); (**B** and **C**) ribonuclease B (80 and 200 μg/assay respectively); (**D**) the neoglycoprotein formed by chemical coupling of β-lactoglobulin to high mannose-type glycopeptides (20 μg/assay). The following proteins were added to the incubation mixtures as potential inhibitors: (A) denatured bovine serum albumin (BSA, ▲—▲) and denatured, endo H-treated soybean agglutinin (SBA-endo H, ●—●); (B) denatured concanavalin A (Con A, ▲—▲), denatured, endo H-treated soybean agglutinin (SBA-endo H, ●—●) and denatured soybean agglutinin treated with endoglycosidase F-N-glycosidase F (SBA-N Glyc, ★—★); (C) denatured ribonuclease A (RNase A, ▲—▲) and denatured, endo H-treated ribonuclease B (RNase B-endo H, ●—●) [the star (★) corresponds to native, endo H-treated ribonuclease B]; (D) denatured β-lactoglobulin (βLg, ▲—▲) and the denatured, endo H-treated neoglycoprotein formed by chemical coupling of β-lactoglobulin to high mannose-type glycopeptides (βLg-Glyc-endo H, ●—●). For further details see Materials and methods.

denatured, endo H-deglycosylated phytohemagglutinin inhibited glucosylation of denatured thyroglobulin. Native bovine serum albumin did not affect the transfer reaction. This was taken as evidence that the enzyme recognized protein domains only exposed in denatured conformations. A rather unexpected result later obtained was the fact that not only native, but also denatured non-glycosylated proteins failed to inhibit glycoprotein glucosylation. A comparison of the effect of denatured, endo H-deglycosylated soybean agglutinin and of denatured bovine serum albumin on the glucosylation of denatured soybean agglutinin is depicted in Figure 1A.

Two possible explanations for the fact that denatured, endo H-deglycosylated glycoproteins inhibited glucosylation, whereas denatured non-glycosylated proteins had no effect, may be envisaged: (i) for glucose transfer the glucosyltransferase had to recognize protein domains exposed in denatured conformations and the innermost *N*-acetylglucosamine unit in the oligosaccharide (i.e. the residue left linked to the protein by endo H) or, alternatively, (ii) the enzyme recognized protein domains exposed in denatured glycoproteins, but not in denatured non-glycosylated proteins.

In order to test the first possibility, the effect of denatured, endo H-treated soybean agglutinin on the glucosylation of bovine pancreas ribonuclease B was compared with that of denatured concanavalin A. The latter is a non-glycosylated protein having a high degree of identity with soybean agglutinin, as both are lectins from leguminous plants (Adar *et al.*, 1989). As shown in Figure 1B, denatured, deglycosylated soybean agglutinin inhibited glucosylation, whereas denatured concanavalin A had no effect. Similarly, denatured soybean agglutinin treated with endoglycosidase F-N-glycosidase F at pH 8.8, i.e. under conditions in which the bonds preferentially cleaved are those between the oligosaccharides and the asparagine units, did not inhibit glucosylation of denatured ribonuclease B (Figure 1B).

A similar result was obtained when the effect of denatured, endo H-deglycosylated ribonuclease B on denatured ribonuclease B glucosylation was compared with that of denatured ribonuclease A. The endo H-treated species inhibited glucosylation, whereas denatured ribonuclease A had no effect (Figure 1C). Both ribonuclease A and B have identical amino acid sequences, the only difference between them being the presence of a single oligosaccharide in the latter species. Furthermore, as will be seen below, glucosylation of a denatured neoglyco-

Table I. Glycosylated β-lactoglobulin as acceptor substrate[a]

Glucose acceptor added	Glucose acceptor capacity (c.p.m.)
β-Lactoglobulin	309
Glycopeptides	139
Glycopeptides + β-lactoglobulin	610
Glycopeptides + β-lactoglobulin (crosslinked)	20 823

[a]For further details see Materials and methods.

protein formed by chemical coupling of β-lactoglobulin and high mannose-type glycopeptides was inhibited by the same species treated with endo H. In contrast, denatured β-lactoglobulin had no effect.

As also depicted in Figure 1C, whereas denatured, deglycosylated ribonuclease B was a good inhibitor, native, endo H-deglycosylated ribonuclease B had no effect. These results show that glucosylation inhibition was the result of an interaction of the glucosyltransferase with both protein domains exposed in denatured, but not in native, conformations and with the *N*-acetylglucosamine unit left linked to glycoproteins by endo H.

Protein domains recognized by the glucosyltransferase are also exposed in denatured proteins

In order to discard the possibility that denatured, non-glycosylated proteins failed to inhibit glycoprotein glucosylation because they lacked the protein domains recognized by the glucosyltransferase, a denatured neoglycoprotein formed by coupling a denatured, non-glycosylated protein to high mannose-type glycopeptides was tested as a glucose acceptor. As shown in Table I, the glycopeptides, obtained by exhaustive thyroglobulin digestion with pronase, were not glucosylated when incubated with the glucosyltransferase either alone or in the presence of denatured β-lactoglobulin. The denatured neoglycoprotein formed by coupling both elements with glutaraldehyde appeared to be, however, a very good glucose acceptor.

In addition, denatured β-lactoglobulin did not inhibit glucosylation, whereas the denatured, endo H-treated neoglycoprotein was a potent inhibitor (Figure 1D), thus confirming the conclusion drawn above. The results shown in Table I also indicate that both elements recognized by the glucosyltransferase, the protein domains and the oligosaccharide, had to be covalently linked in order to allow glucosylation.

Glycoproteins having similar secondary, but different tertiary, structures have different glucose acceptor capacities

Staphylococcal nuclease is a relatively small (149 amino acids), non-glycosylated protein devoid of cysteine residues. One of these residues was introduced in position 70 by replacing a lysine (K70C) (Ermácora *et al.*, 1994). A high mannose-type glycopeptide was then linked to that position using *N*-succinimidyl 3-(2-pyridyldithio)-propionate, a bifunctional reagent that reacts with amino and sulfhydryl groups. The neoglycoprotein thus formed (K70C-Glyc) had the same specific nuclease activity (4000 U/mg) as the unmodified protein (Shortle and Meeker, 1989; Ermácora *et al.*, 1994), thus suggesting that both species had the same tertiary structure. The neoglycoprotein had a very low glucose acceptor capacity (Table II). Addition of the nuclease inhibitor 3′,5′-diphosphothymidine (pdTp), a stabilizer of the native nuclease conformation and an inducer of proper folding in truncated species (see below), further diminished the glucose acceptor capacity. The neoglycoprotein was, nevertheless, efficiently glucosylated when previously denatured with 8 M urea. That pdTp had no effect on the glucosyltransferase activity was checked using denatured soybean agglutinin as glucose acceptor (not shown).

Table II. The influence of secondary and tertiary structures on the glucose acceptor capacity[a]

Glucose acceptor added	pdTp	Glucose acceptor capacity		Nuclease activity (U/mg)
		c.p.m.	%	
None	–	205	0	
K70C-Glyc (native)	–	717	13	4000
K70C-Glyc (native)	+	275	2	
K70C-Glyc (denatured)	–	4064	100	
1–135 K70C-Glyc	–	3964	98	1000
1–135 K70C-Glyc	+	1728	43	
1–135 H124C-Glyc	–	3163	77	1480
1–135 H124C-Glyc	+	1451	32	

[a]For further details see Materials and methods.

It has been reported that a truncated nuclease lacking the last 14 amino acids at the C-terminal end is *per se* (i.e. without any denaturing treatment) in a compact, but disordered, conformation, but that the enzyme can be induced to properly fold in the presence of Ca^{2+} and pdTp or the substrate (DNA), as judged by far UV circular dichroism (CD) and nuclear magnetic resonance spectra (Shortle and Meeker, 1989; Flanagan *et al.*, 1992). This large fragment showed the same specific nuclease activity as the full-length enzyme, thus indicating that both species had the same tertiary structure.

Two truncated neoglycoproteins were synthesized using this large fragment as a protein backbone, one with the oligosaccharide attached to a cysteine introduced at position 70 (1–135 K70C-Glyc), i.e. in the same position as in the full-length species, and the other at position 124 (1–135 H124C-Glyc). In this case the cysteine was introduced in place of a histidine. Both truncated neoglycoproteins were efficiently glucosylated when pdTp was omitted from the reaction mixture (no denaturing treatment was performed) (Table II). This suggests that the position of the oligosaccharide is not an important factor for glucosylation. Upon addition of pdTp, the glucose acceptor capacity of both truncated neoglycoproteins diminished by ~60%, but without reaching the basal levels of the full-length neoglycoprotein (K70C-Glyc) in the presence of pdTp (it is worth mentioning that Ca^{2+}, the other element required for inducing proper folding, is always present in reaction mixtures, as it is required for glucosyltransferase activity).

In order to check that incubation of the truncated neoglycoproteins with Ca^{2+} and pdTp had actually led to the transition to a native conformation, the far UV CD spectra of the truncated neoglycoprotein 1–135 K70C-

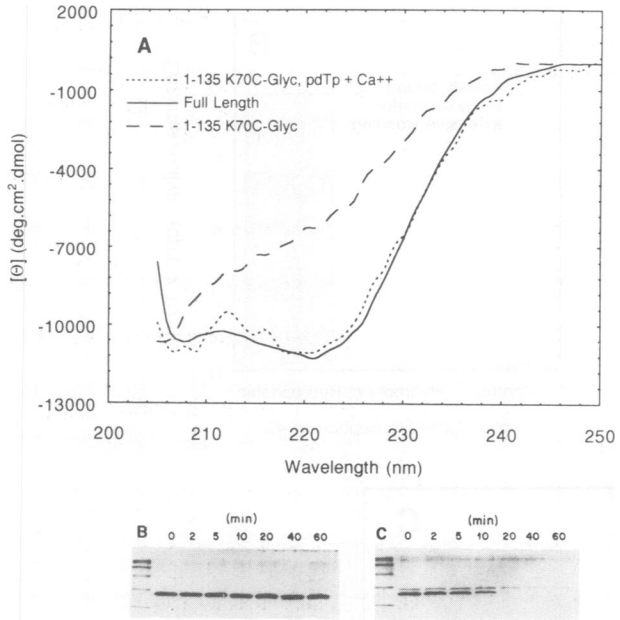

Fig. 2. The secondary and tertiary structures of proteins and neoglycoproteins. (**A**) Circular dichroism spectra of the truncated staphylococcal nuclease neoglycoprotein having the glycopeptide linked to position 70 (1–135 K70C-Glyc) in the absence (— —) or presence (– – –) of Ca^{2+} and pdTp or of the native full-length staphylococcal nuclease (——). (**B**) The full-length neoglycoprotein (K70C-Glyc) and (**C**) the truncated species (1–135 K70C-Glyc) were submitted to controlled proteolysis in the presence of Ca^{2+} and pdTp followed by polyacrylamide gel electrophoresis under denaturing conditions. Lanes at the left of (B) and (C) correspond to the following molecular mass markers (from top to bottom): 66 200, 45 000, 31 000, 21 500 and 14 400 kDa. For further details see Materials and methods.

Glyc and the full-length wild-type enzyme were compared. As depicted in Figure 2A, the truncated neoglycoprotein yielded a spectrum strikingly different to that of the wild-type enzyme, but both spectra became superimposable when Ca^{2+} and pdTp were added to the former species. This indicated that both the truncated neoglycoprotein in the presence of Ca^{2+} and pdTp and the wild-type enzyme had the same secondary structure.

Nevertheless, the tertiary structures of K70C-Glyc (full-length neoglycoprotein) and of 1–135 K70C-Glyc (truncated neoglycoprotein) were different: when submitted to limited proteolysis in the presence of 1 mM pdTp and 10 mM $CaCl_2$, K70C-Glyc was barely cleaved, even after a 60 min incubation with trypsin, whereas 1–135 K70C-Glyc was rapidly degraded (Figure 2B and C). Moreover, the specific nuclease activities of both truncated neoglycoproteins (1–135 K70C-Glyc and 1–135 H124C-Glyc) were much lower than that of the full-length neoglycoprotein (K70C-Glyc) (Table II). The tertiary structures of the truncated neoglycoproteins in the presence of Ca^{2+} and pdTp or DNA were, therefore, looser than that of the truncated non-glycosylated species. The latter had, as mentioned above, the same specific nuclease activity as the full-length enzyme with or without the oligosaccharide. The different tertiary structures between full-length and truncated neoglycoproteins may explain why the latter entities had a residual glucose acceptor capacity in the presence of Ca^{2+} and pdTp.

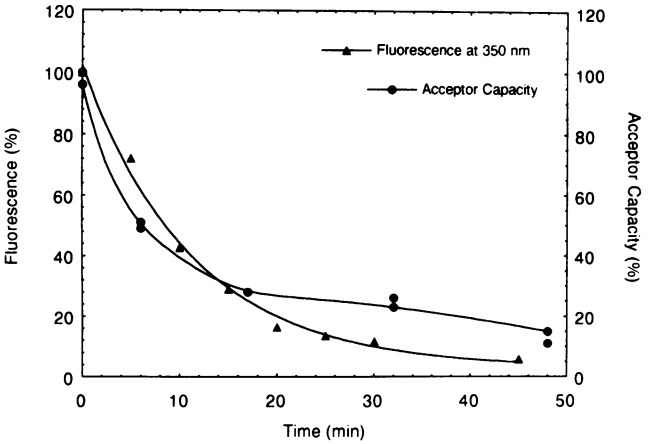

Fig. 3. Glucose acceptor capacity and folding status of glycoproteins. Soybean agglutinin was denatured in 6 M guanidine hydrochloride, diluted and allowed to renature. Fluorescence at 350 nm (▲—▲) and glucose acceptor capacity (●—●) were measured at the indicated times. For further details see Materials and methods.

The glucose acceptor capacity parallels the folding status of glycoproteins

The results presented above suggest that denatured glycoproteins with tertiary structures closely resembling those of native species have less glucose acceptor capacity than those having a more disordered structure. In order to confirm this conclusion, which might provide valuable hints on the nature of the protein domains recognized by the glucosyltransferase, soybean agglutinin was denatured with 6 M guanidine hydrochloride, diluted and allowed to renature under controlled conditions. Renaturation was followed by measuring the fluorescent emission of tryptophan at 350 nm. The maximum wavelengths of tryptophan emission in soybean agglutinin are 328 and 350 nm in the native and random coil conformations respectively. As depicted in Figure 3, the decrease in glucose acceptor capacity closely followed renaturation, thus confirming the conclusion reached above.

Under the standard glucosyltransferase assay conditions used in previous studies the acceptor glycoproteins employed formed a part of large aggregates, as judged by gel filtration. As no such aggregates were formed under the conditions employed in the experiment shown in Figure 3, it may be concluded that the information for glucosylation resides in single glycoprotein species and not necessarily in aggregates.

The UDP-Glc:glycoprotein glucosyltransferase interacts with hydrophobic amino acids

One of the main differences between native and denatured tertiary structures is the exposure of large patches of hydrophobic amino acids in the latter conformations. In order to study whether UDP-Glc:glycoprotein glucosyltransferase interacted with hydrophobic amino acids, two nonapeptides, one containing five hydrophobic aliphatic and four hydrophilic (HISLIMTAN) and the other nine hydrophilic (HTSQHNTQS) amino acids were covalently coupled through their N-termini to Sepharose 4B. The glucosyltransferase was loaded onto resin-containing columns and the enzymatic activity assayed first in eluted fractions and then, after extensive washing, in the resins mechanically extruded from the columns. As shown in

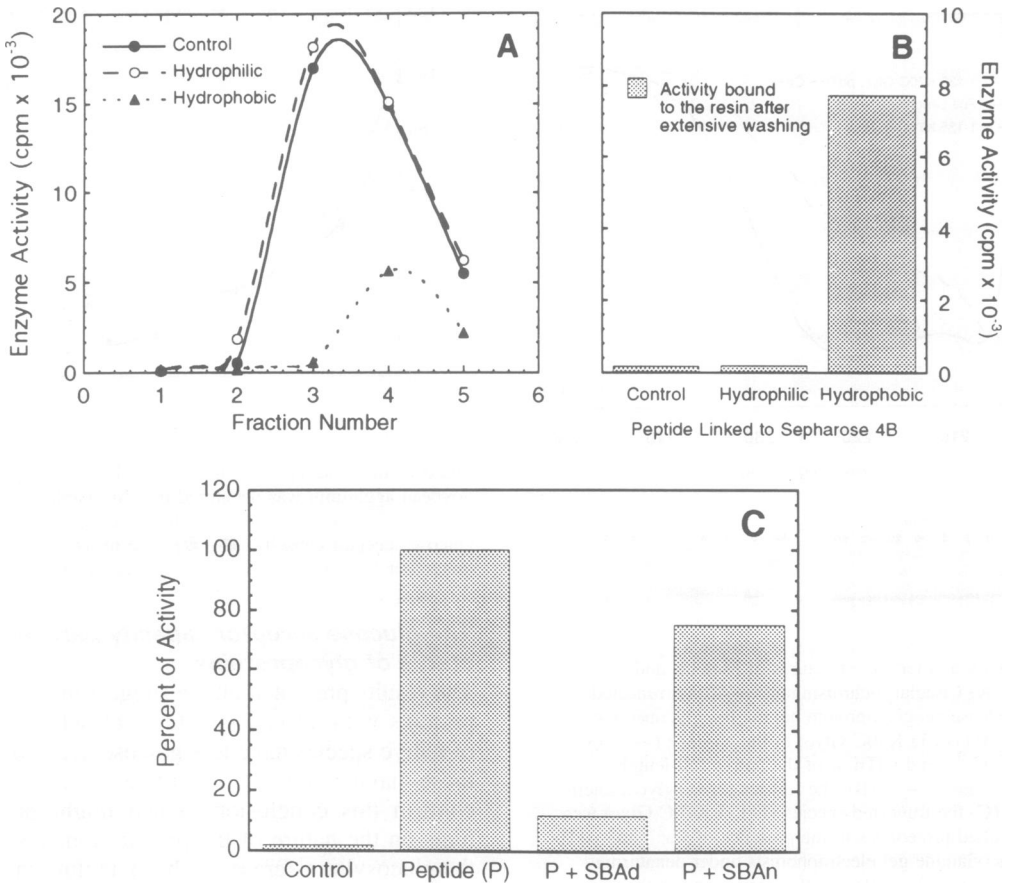

Fig. 4. The interaction of the glucosyltransferase with hydrophobic amino acids. (**A**) The glucosyltransferase was loaded onto columns containing Sepharose 4B coupled to a hydrophilic peptide (○—○) or to a hydrophobic aliphatic peptide (▲—▲) or to no peptide (●—●). Activity was assayed in eluted fractions. (**B**) The resins used in (A) were mechanically extruded from the columns after extensive washing and the glucosyltransferase activity assayed in them. (**C**) The glucosyltransferase was loaded into tubes containing Sepharose 4B coupled to a hydrophobic aromatic peptide (P) or to the same resin, but the tubes contained, in addition, denatured soybean agglutinin (P+SBAd) or the native lectin (P+SBAn). The control refers to Sepharose 4B without a coupled peptide. The glucosyltransferase activity was assayed in the resins after extensive washing. For further details see Materials and methods.

Figure 4A, the control (i.e. the resin without coupled peptides) and the column containing the hydrophilic peptide gave similar results, with all the activity in the eluted fractions. On the other hand, a sizable portion of the activity appeared bound to the hydrophobic peptide-containing resin (Figure 4B). It is worth mentioning that, as described under Materials and methods, the activities in the eluted fractions cannot be quantitatively compared with those in the resins, as different assay conditions were employed.

In order to check if the same elements recognized by the glucosyltransferase in hydrophobic peptides were exposed in denatured, but not in native, conformations, a nonapeptide composed of five hydrophobic aromatic and four hydrophilic amino acids (HFSFYFTFS) was coupled to Sepharose 4B and the binding of glucosyltransferase to the resin was assayed in the presence of native or denatured soybean agglutinin or in the absence of the lectin. The experimental conditions were similar to those of the experiment shown in Figure 4A and B, but the resins were in test tubes, not in columns. Only the activity bound to the resins after extensive washing was assayed. As depicted in Figure 4C, the glucosyltransferase was retained by the resin containing the hydrophobic aromatic peptide, but not by the peptide-free resin. Moreover, the addition of denatured soybean agglutinin produced a severe decrease in glucosyltransferase binding, whereas the same amount of native lectin produced a several fold lower effect.

As will be extensively discussed below, there is evidence suggesting that all glycoproteins are transiently glucosylated *in vivo* (Gañán *et al.*, 1991). Moreover, already published experiments have shown that both mammalian and yeast glucosyltransferases may glucosylate glycoproteins isolated from mammals or plants, provided that they are in denatured conformations (Sousa *et al.*, 1992; Fernández *et al.*, 1994). These facts suggest that the glucosyltransferase might recognize hydrophobic domains composed of a broad spectrum of hydrophobic amino acids. The peptides employed above were designed, therefore, without following any particular pattern and their only specific features were the presence of either aromatic or aliphatic or aliphatic hydrophobic or only hydrophilic amino acids. Nonapeptides were chosen as it has been reported that binding protein (BiP) recognizes peptides containing seven or more amino acids (Flynn *et al.*, 1991; Blond-Elguindi *et al.*, 1993). On the other hand, these two reports indicated that the chaperone preferentially binds peptides containing either aliphatic or aromatic hydrophobic amino acids respectively. The glucosyl-

transferase seems to bind nonapeptides containing both types of residues, although it cannot be discounted that it has a higher affinity for one of them or for peptides having a different number of residues.

Discussion

UDP-Glc:glycoprotein glucosyltransferase appears to be an extremely subtle sensor of misfolded conformations: not only did the diminution in glucose acceptor capacity of denatured soybean agglutinin closely follow renaturation, but the enzyme was also able to distinguish between two neoglycoproteins having similar secondary, but different tertiary, structures, such as the glycosylated truncated and full-length staphylococcal nucleases in the presence of Ca^{2+} and the inhibitor pdTp. Even the native full-length neoglycoprotein showed a different glucose acceptor capacity when the assay was performed in the presence or absence of the diphosphonucleoside, a substance that not only induces proper folding of the truncated nuclease, but also stabilizes the native conformation of the full-length enzyme (Shortle and Meeker, 1989; Flanagan *et al.*, 1992).

What are the molecular bases of the recognition of misfolded conformations by the glucosyltransferase? Evidence presented here indicates that one of the elements in the recognition system is the accessibility of the innermost *N*-acetylglucosamine unit of the oligosaccharide to the enzyme. In many native glycoproteins the innermost *N*-acetylglucosamine is not accessible to a macromolecular probe such as endo H (bovine pancreas ribonuclease B is an exception) (Trimble and Maley, 1984). In those cases, cleavage of oligosaccharides by the endoglycosidase requires denaturation of the glycoprotein. It may be speculated that proper folding of most glycoproteins would hinder recognition of the innermost *N*-acetylglucosamine unit by glucosyltransferase and thus prevent glucosylation.

On the other hand, interaction of protein domains exposed in denatured, but not in native, conformations with the glucosyltransferase also appeared to be required for glucosylation. The necessary interaction of the glucosyltransferase not only with the innermost *N*-acetylglucosamine unit, but also with protein domains, was indicated by both the lack of effect of native, endo H-treated ribonuclease B on glucosylation (Figure 1C) and by the fact that pronase-derived glycopeptides were very poorly glucosylated (Sousa *et al.*, 1992), notwithstanding the full accessibility of the entire oligosaccharides to the glucosyltransferase. It is worth mentioning that native ribonuclease B was not a substrate for the glucosyltransferase, although its oligosaccharide was fully accessible to endo H (Trombetta and Parodi, 1992).

The protein domains and the oligosaccharides had to be covalently linked. This is an important restriction. As there are numerous unfolded, partially folded and misfolded non-glycosylated proteins and glycoproteins in the lumen of the endoplasmic reticulum, if the protein domains recognized by the glucosyltransferase and the oligosaccharides were not required to be covalently linked it might be speculated that domains exposed in improperly folded species would induce glucosylation of glycoproteins already in their native conformations, provided that the innermost *N*-acetylglucosamine units are accessible to the glucosyltransferase in the latter species.

At least 50% of all glycoproteins (perhaps all of them) are transiently glucosylated in intact cells (Gañán *et al.*, 1991). Therefore, no amino acid primary sequence or a particular shape or the clustering upon denaturation of specific amino acids distant in the primary sequence may be envisaged to be common to all denatured glycoproteins. Even the consensus glycosylation sequence NxS/T (where x may be any amino acid except P) is not required for glucosylation, as it is absent from β-lactoglobulin. The only structural element up to now known to be common to denatured conformations of all glycoproteins and proteins is the exposure of large patches of hydrophobic amino acids.

We have demonstrated here that UDP-Glc:glycoprotein glucosyltransferase interacts with hydrophobic peptides. No interaction with peptides containing only hydrophilic amino acids was found. As the interaction with hydrophobic peptides could be prevented by denatured, but not by native, soybean agglutinin, it may be concluded that the enzyme recognized patches of hydrophobic amino acids exposed in denatured conformations. It should be stressed, however, that it has not been definitively demonstrated that hydrophobic amino acids are the elements exposed in denatured conformations that allow glucosylation upon recognition by the glucosyltransferase. Nevertheless, in agreement with the suggestion that hydrophobic amino acids form part of the recognition system is the fact that more disordered conformations, i.e. those exposing more hydrophobic amino acids, were found to be those having higher glucose acceptor capacity.

It is known that the conformation of oligosaccharides in glycoproteins may be influenced by the interaction of certain monosaccharide units with specific amino acids (Sutton and Phillips, 1983; Savvidou *et al.*, 1984). It may be speculated, therefore, that oligosaccharides in denatured glycoproteins might have particular conformations determined by the hydrophobic environment and that what the glucosyltransferase actually recognizes are oligosaccharides having those conformations, and not protein domains. This possibility seems highly unlikely, however, as endo H-treated glycoproteins interacted with the glucosyltransferase (Figure 1) and it is not reasonably expected that the conformation of a single *N*-acetylglucosamine residue in a denatured glycoprotein would mimic the conformation of entire oligosaccharides having compositions $Man_7GlcNAc_2$, $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$.

As expected for an enzyme that glucosylates a variety of glycoproteins provided they are in non-native conformations, the glucosyltransferase appeared to be a very versatile catalyst, not requiring the oligosaccharide to be in a particular position in the amino acid sequence and even glucosylating structures not found in nature, such as the neoglycoproteins synthesized here. On the other hand, glucosylation required accessibility by the glucosyltransferase to elements exposed in improperly folded glycoproteins, namely the innermost *N*-acetylglucosamine residue and large patches of hydrophobic amino acids. The unique property of UDP-Glc:glycoprotein glucosyltransferase, that of covalently tagging denatured conformations, is well suited for its proposed role in the quality control of glycoprotein folding.

Materials and methods

Materials

Bovine thyroglobulin, bovine pancreas ribonucleases A and B, endo H, endoglycosidase F-*N*-glycosidase F and β-lactoglobulin were from Sigma (St Louis, MO). pdTp was from United States Biochemicals (Cleveland, OH). *N*-Succinimidyl 3-(2-pyridyldithio)propionate was from Pierce (Rockford, IL). The following nonapeptides, hydrophilic (HTSQHNTQS), hydrophobic aliphatic (HISLIMTAN) and hydrophobic aromatic (HFSFYFTFS) were kindly synthesized by Dr Gunnar Lindeberg (Department of Immunology, University of Uppsala Biomedical Center, Uppsala, Sweden). Staphylococcal nuclease mutants were a generous gift from Dr Robert O.Fox (Department of Molecular Biophysics, Yale University, New Haven, CT). NAP-5 columns, cyanogen bromide-activated Sepharose 4B and concanavalin A–Sepharose were from Pharmacia (Uppsala, Sweden).

Preparations

Soybean agglutinin and UDP-[^{14}C]Glc (320 Ci/mol) were prepared as indicated previously (Wright and Robbins, 1965; Sousa *et al.*, 1992). Pure rat liver UDP-Glc:glycoprotein glucosyltransferase was employed. It was prepared as described previously (Trombetta and Parodi, 1992). High mannose-type glycopeptides from thyroglobulin were prepared as described (Trombetta *et al.*, 1989).

Coupling of glycopeptides to β-lactoglobulin

Pronase-derived, high mannose-type glycopeptides (1 mg) were dissolved in 0.3 ml water and mixed with 0.1 ml 0.4 M sodium phosphate buffer, pH 7.4, containing 1.5 mg bovine β-lactoglobulin. Glutaraldehyde (0.2 ml of a 20 mM solution) was added dropwise and the mixture was left for 30 min at room temperature. $NaCNBH_3$ (9 μl of a 0.275 M solution in 0.1 M NaOH) was then added and the solution left for another 30 min at room temperature. The solution was then desalted in a fast desalter equilibrated with 50 mM Tris–HCl buffer, pH 7.6, 0.15 M NaCl. The sample was then supplemented with $CaCl_2$, $MnCl_2$ and $MgCl_2$ to a final 2 mM concentration and loaded onto a 3 ml concanavalin A–Sepharose column equilibrated with the same solution. The column was washed and the derivatized β-lactoglobulin was finally eluted at 37°C with 0.2 M α-methylmannoside in the same solution, dialyzed and freeze-dried.

Glycosylated β-lactoglobulin as acceptor substrate

The following substances were tested for their glucose acceptor capacity: assay 1, 350 μg denatured β-lactoglobulin; assay 2, 15 μg pronase-derived high mannose-type glycopeptides; assay 3, 15 μg glycopeptides plus 350 μg β-lactoglobulin; assay 4, 42 μg β-lactoglobulin cross-linked to glycopeptides. The reaction mixture contained, in a total volume of 50 μl, 20 mM Tris–HCl, pH 8.0, 5 μM UDP-[^{14}C]Glc, 10 mM $CaCl_2$ and 0.3 μM glucosyltransferase. Glucose incorporation in assays 1 and 4 was quantitated using the trichloroacetic acid precipitation procedure, whereas incorporation into free glycopeptides was quantitated as previously described (Trombetta *et al.*, 1989).

Coupling of glycopeptides to staphylococcal nuclease

Pronase-derived high mannose-type glycopeptides (1 mg) were dissolved in 0.45 ml 0.2 M sodium borate buffer, pH 7.4. A 30-fold molar excess of *N*-succinimidyl 3-(2-pyridyldithio)propionate dissolved in 50 μl dimethylsulfoxide was then added. The mixture was then left at room temperature for 30 min and desalted in a 57×1 cm Sephadex G-10 column equilibrated with 7% 2-propanol. The oligosaccharide-containing fractions were pooled and dried. The cysteine-containing full-length or truncated nucleases (5 mg) were dissolved in 0.5 ml 50 mM Tris–HCl buffer, pH 8.8, 50 mM dithiothreitol and incubated at room temperature for 20 min. The samples were then desalted in a NAP-5 column equilibrated with 50 mM Tris–HCl buffer, pH 8.8, 1 mM EDTA. The desalted proteins were added to the derivatized glycopeptides and the mixtures incubated overnight at 4°C. The mixtures were diluted with 5 vol of 50 mM Tris–HCl buffer, pH 7.6, 0.15 M NaCl, 2 mM $CaCl_2$, $MnCl_2$ and $MgCl_2$ and loaded onto 3 ml concanavalin A–Sepharose columns equilibrated with the same solution. The columns were washed and the neoglycoproteins eluted with 0.2 M α-methylmannoside at 37°C. The eluted substances were dialyzed against 20 mM Tris–HCl buffer, pH 8.0.

Glycosylated nucleases as acceptor substrates

The indicated potential substrates (45 μg) were incubated with 0.3 μM glucosyltransferase in a medium containing 20 mM Tris–HCl buffer, pH 8.0, 4.5 μM UDP-[^{14}C]Glc and 10 mM $CaCl_2$ and, where indicated, 1 mM pdTp.

Far UV circular dichroism

The far UV CD spectra of native full-length staphylococcal nuclease (65 μM), 1–135 K70C-Glyc fragment (18 μM) and 1–135 K70C-Glyc (16 μM) plus Ca^{2+} (10 mM) and pdTp (1 mM) were recorded in a Jasco spectropolarimeter. An average of five spectra minus the spectra of the corresponding buffer (20 mM Tris–HCl, pH 8.0, with or without Ca^{2+} and pdTp) is shown for each protein.

Limited proteolysis

The full-length (K70C-Glyc, 25 μg) or the truncated (1–135 K70C-Glyc, 25 μg) neoglycoproteins were incubated in a solution containing 20 mM Tris–HCl buffer, pH 8.0, 10 mM $CaCl_2$ and 1 mM pdTp, with 50 ng trypsin at 37°C. Samples were withdrawn at the indicated times, mixed with cracking buffer and heated at 100°C for 5 min. The samples were analyzed by SDS–PAGE as described before (Schägger and von Jagow, 1987).

Binding of the glucosyltransferase to hydrophobic amino acids

The hydrophilic (HTSQHNTQS) and hydrophobic peptides (HISLIMTAN and HFSFYFTFS) in 5% dimethylsulfoxide (1.5 mM concentration) were coupled to cyanogen bromide-activated Sepharose 4B according to the instructions of the manufacturer. For the experiments depicted in Figure 4A and B the glucosyltransferase (50 μl) was mixed with an equal volume of 40 mM HEPES, pH 7.0, 0.3 M NaCl and poured into columns containing 0.5 ml of the resin equilibrated with 20 mM HEPES, pH 7.0, 0.15 M NaCl (binding buffer). In these figures the peptides coupled to the resin were hydrophilic (HTSQHNTQS) or aliphatic hydrophobic (HISLIMTAN). Elution was performed with binding buffer and fractions of 0.1 ml were collected. After the fifth fraction the columns were washed three times with 1 ml binding buffer and the resins mechanically extruded. Glucosyltransferase was assayed in the fractions (Figure 4A) in incubation mixtures containing 110 μg denatured soybean agglutinin, 10 mM $CaCl_2$, 10 mM Tris–HCl buffer, pH 8.0, 5 μM UDP-[^{14}C]Glc and 10 μl each fraction in a total volume of 50 μl. Incubations lasted for 5 min at 37°C. Activity in the resins (Figure 4B) was assayed in incubation mixtures containing 480 μg denatured soybean agglutinin, 10 mM $CaCl_2$, 10 mM Tris–HCl buffer, pH 8.0, 10 μM UDP-[^{14}C]Glc and 50 μl each resin in a total volume of 150 μl. Incubations lasted for 10 min at 37°C. The activities in the eluted fractions cannot be quantitatively compared, therefore, with those in the resins, as different assay conditions were employed. For the experiment depicted in Figure 4C the glucosyltransferase (1.5 μM) in 20 μl binding buffer and, where indicated, 3 μg native or denatured soybean agglutinin, were loaded into tubes containing 20 μl Sepharose 4B coupled to the hydrophobic aromatic peptide (HFSFYFTFS) equilibrated with binding buffer. After 10 min at room temperature in a shaker the resins were thrice washed with 0.2 ml binding buffer. The activity in the resins was then assayed in incubation mixtures containing 150 μg denatured soybean agglutinin, 10 mM $CaCl_2$, 10 mM Tris–HCl buffer, pH 8.0, 5 μM UDP-[^{14}C]Glc and 20 μl each resin in a total volume of 50 μl. Incubations lasted for 10 min at 37°C.

Treatment of denatured glycoproteins with endo H and endoglycosidase F-N-glycosidase F

Treatment with endo H was performed as described previously (Sousa *et al.*, 1992) and that with endoglycosidase F-N-glycosidase F was performed in 0.6 ml 75 mM sodium borate, pH 8.8, containing 2.7 mg denatured soybean agglutinin and 1 U enzyme. The incubations lasted for 18 h at 30°C. After degradation with any one of the endoglycosidases the incubation mixtures were dried and material therein resuspended in 8 M urea in 0.1 M Tris–HCl, pH 8.0, and heated at 65°C for 6 h. This was followed by dialysis against 10 mM Tris–HCl, pH 8.0.

Enzymatic assays

The UDP-Glc:glycoprotein glucosyltransferase and staphylococcal nuclease were assayed as described before (Cuatrecasas *et al.*, 1967; Trombetta *et al.*, 1989).

Denaturation and renaturation of proteins and glycoproteins

Non-glycosylated proteins and glycoproteins were denatured as indicated above. In the case of the experiment shown in Figure 3, denaturation and renaturation of soybean agglutinin was performed as described (Nagai and Yamaguchi, 1993), with some modifications. Pure soybean

agglutinin was denatured in 6 M guanidine hydrochloride, 50 mM HEPES, pH 7.0. The sample was 100-fold diluted in 100 mM HEPES, pH 7.0, 0.1 mM $CaCl_2$, 0.1 mM $MnCl_2$ to a final protein concentration of 40 μg/ml and allowed to renature at 23°C. At the indicated times the intrinsic fluorescence at 350 nm was recorded (excitation wavelength 280 nm). Aliquots of the renaturation mixture (40 μl) were withdrawn at the indicated times and their glucose acceptor capacity measured in a final volume of 50 μl with 0.3 μM glucosyltransferase. Final concentrations in the reaction mixture were 40 mM HEPES, pH 7.0, 10 mM $CaCl_2$, 80 μM $MnCl_2$, 42 μM UDP-[^{14}C]Glc and 32 μg/ml soybean agglutinin. Incubations lasted for 5 min at 37°C.

Analytical methods

Proteins and saccharides were quantitated as described previously (Lowry *et al.*, 1951; Dubois *et al.*, 1956)

Acknowledgements

We are grateful to Dr R.O.Fox and co-workers (Department of Molecular Biophysics, Yale University, New Haven, CT) for providing the nuclease mutants and to Dr Gunnar Lindeberg (Department of Immunology, University of Uppsala Biomedical Center, Uppsala, Sweden) for synthesizing the peptides. The invaluable help and advice of Drs Mario Ermácora and José M.Delfino (Department of Biochemistry, School of Pharmacy and Biochemistry, University of Buenos Aires, Buenos Aires, Argentina) are gratefully acknowledged. Labeled UDP-Glc was prepared by Ms Susana Raffo. This work was supported by United States Public Health Service grant 44500, by the Swedish Agency for Research Cooperation with Developing Countries (SAREC) and by the University of Buenos Aires. M.C.S. is a Doctoral Fellow of the University of Buenos Aires and A.J.P. is a Career Investigator of the National Research Council (Argentina).

References

Adar,R., Richardson,M., Lis,H. and Sharon,N. (1989) The amino acid sequence of *Erytrina corallodendron* lectin and its homology with other legume lectins. *FEBS Lett.*, **257**, 81–85.

Blond-Elguindi,S., Cwiria,S.E., Dower,W.J., Lipshutz,R.J., Sprang,S.R., Sambrook,J.F. and Gething,M.-J.H. (1993) Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BiP. *Cell*, **75**, 717–728.

Cuatrecasas,P., Fuchs,S. and Anfinsen,C.B. (1967) Catalytic properties and specificity of the extracellular nuclease of *Staphylococcus aureus*. *J. Biol. Chem.*, **242**, 1541–1547.

Dubois,M., Gilles,K.A., Hamilton,J.K., Rebers,P.A. and Smith,F. (1956) Colorimetric method for determination of sugars and related substances. *Anal. Chem.*, **28**, 350–356.

Ermácora,M.R., Ledman,D.W., Hellinga,H.W., Hsu,G.W. and Fox,R.O. (1994) Mapping staphylococcal nuclease conformation using an EDTA–Fe derivative attached to genetically engineered cysteine residues. *Biochemistry*, **33**, 13625–13641.

Fernández,F.S., Trombetta,S.E., Hellman,U. and Parodi,A.J. (1994) Purification to homogeneity of UDP-Glc:glycoprotein glucosyltransferase from *Schizosaccharomyces pombe* and apparent absence of the enzyme from *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **269**, 30701–30706.

Flanagan,J.M., Kataoka,M., Shortle,D. and Engenman,D.M. (1992) Truncated staphylococcal nuclease is compact but disordered. *Proc. Natl Acad. Sci. USA*, **89**, 748–752.

Flynn,G.C., Pohl,J., Flocco,M.T., and Rothman,J.E. (1991) Peptide-binding specificity of the molecular chaperone BiP. *Nature*, **353**, 726–730.

Gañán,S., Cazzulo,J.J. and Parodi,A.J. (1991) A major proportion of *N*-glycoproteins are transiently glucosylated in the endoplasmic reticulum. *Biochemistry*, **30**, 3098–3104.

Hammond,C., Braakman,I. and Helenius,A. (1994) Role of *N*-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control. *Proc. Natl Acad. Sci. USA*, **91**, 913–917.

Helenius,A. (1994) How *N*-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. *Mol. Biol. Cell*, **5**, 253–265.

Kornfeld,R. and Kornfeld,S. (1985) Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.*, **54**, 631–664.

Lowry,O.H., Rosenbrough,N.J., Farr,A.L. and Randall,R.J. (1951) Protein measurement with the Folin phenol reagent. *J. Biol. Chem.*, **193**, 265–275.

Nagai,K. and Yamaguchi,H. (1993) Direct demonstration of the essential role of the intramolecular high-mannose oligosaccharide chains in the folding and assembly of soybean (*Glycine max*) lectin polypeptides. *J. Biochem.*, **113**, 123–125.

Parodi,A.J., Mendelzon,D.H. and Lederkremer,G.H. (1983) Transient glucosylation of protein bound $Man_9GlcNAc_2$, $Man_8GlcNAc_2$ and $Man_7GlcNAc_2$ in calf thyroid cells. A possible recognition signal in the processing of glycoproteins. *J. Biol. Chem.*, **258**, 8260–8265.

Parodi,A.J., Mendelzon,D.H., Lederkremer,G.H. and Martín-Barrientos,J. (1984) Evidence that transient glucosylation of protein-linked $Man_9GlcNAc_2$, $Man_8GlcNAc_2$ and $Man_7GlcNAc_2$ occurs in rat liver and *Phaseolus vulgaris* cells. *J. Biol. Chem.*, **259**, 6351–6357.

Savvidou,G., Klein,M., Grey,A.A., Dorrington,A.J. and Carver,J.P. (1984) Possible role for peptide–oligosaccharide interactions in differential oligosaccharide processing at asparagine-107 of the light chain and asparagine-297 of the heavy chain in a monoclonal IgGk. *Biochemistry*, **23**, 3736–3740.

Schägger,H. and von Jagow,G. (1987) Tricine–sodium dodecylsulfate–polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.*, **166**, 368–379.

Shortle,D.A. and Meeker,A. (1989) Residual structure in large fragments of staphylococcal nuclease: effects of amino acid substitutions. *Biochemistry*, **28**, 936–944.

Sousa,M., Ferrero-García,M.A. and Parodi,A.J. (1992) Recognition of the oligosaccharide and protein moieties of glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase. *Biochemistry*, **31**, 97–105.

Sutton,B.J. and Phillips,D.C. (1983) The three dimensional structure of the carbohydrate within the Fc fragment of immunoglobulin G. *Biochem. Soc. Trans.*, **11**, 130–132.

Trimble,R.B. and Maley,F. (1984) Optimizing hydrolysis of *N*-linked high-mannose oligosaccharides by endo-β-*N*-acetylglucosaminidase H. *Anal. Biochem.*, **141**, 515–522.

Trombetta,S.E. and Parodi,A.J. (1992) Purification to apparent homogeneity of rat liver UDP-Glc:glycoprotein glucosyltransferase. *J. Biol. Chem.*, **267**, 9236–9240.

Trombetta,S.E., Bosch,M. and Parodi,A.J. (1989) Glucosylation of glycoproteins by mammalian, plant, fungal and trypanosomatid protozoa microsomal proteins. *Biochemistry*, **28**, 8108–8116.

Trombetta,S.E., Gañán,S. and Parodi,A.J. (1991) The UDP-Glc:glycoprotein glucosyltransferase is a soluble protein of the endoplasmic reticulum. *Glycobiology*, **1**, 155–161.

Wright,A. and Robbins,P.W. (1965) Enzymatic synthesis of uridine diphosphate [^{14}C]glucose. *Biochim. Biophys. Acta*, **104**, 594–595.

Received on April 6, 1995; revised on June 9, 1995