The endoplasmic reticulum UDP-Glc:glycoprotein glucosyltransferase (GT) exclusively glucosylates non-native glycoprotein conformers. GT sequence analysis suggests that it is composed of at least two domains: the N-terminal domain, which composes 80% of the molecule, has no significant similarity to other known proteins and was proposed to be involved in the recognition of non-native conformers and the C-terminal or catalytic domain, which displays a similar size and significant similarity to members of glycosyltransferase family 8. Here, we show that N- and C-terminal domains from *Rattus norvegicus* and *Schizosaccharomyces pombe* GTs remained tightly but not covalently bound upon a mild proteolytic treatment and could not be separated without loss of enzymatic activity. The notion of a two-domain protein was reinforced by the synthesis of an active enzyme upon transfection of *S. pombe* GT null mutants with two expression vectors, each of them encoding one of both domains. Transfection with the C-terminal domain-encoding vector alone yielded an inactive, rapidly degraded protein, thus indicating that the N-terminal domain is required for proper folding of the C-terminal catalytic portion. If, indeed, the N-terminal domain is, as proposed, also involved in glycoprotein conformation recognition, the tight association between N- and C-terminal domains may explain why only N-glycans in close proximity to protein structural perturbations are glucosylated by the enzyme. Although *S. pombe* and *Drosophila melanogaster* GT N-terminal domains display an extremely poor similarity (16.3%), chimeras containing either yeast N-terminal and fly C-terminal domains or the inverse construction were enzymatically and functionally active in *vivo*, thus indicating that the N-terminal domains of both GTs shared three-dimensional features.

Most proteins following the secretory pathway in eukaryotic cells are N-glycosylated in the endoplasmic reticulum (ER).\(^1\) A glycan (Glc\(_3\)-Man\(_9\)-GlcNAc\(_2\)) is transferred to Asn in growing polypeptides. Glucoses are then trimmed by the action of glucosidase I, which removes the external Glc unit, followed by glucosidase II, which excises both Glc residues remaining in the glycan (1). Monoglucosylated N-glycans may be formed by partial deglucosylation of the transferred oligosaccharide or by reglucosylation of Glc-free glycans by the UDP-Glc:glycoprotein glucosyltransferase (GT) (2). This enzyme is a sensor of glycoprotein conformations, as it exclusively glucosylates N-glycans in not properly folded conformers. One or two ER-resident α-mannosidases may degrade Man\(_{5}\)-GlcNAc\(_2\) to Man\(_{5}\)-GlcNAc\(_2\) and Man\(_{9}\)-GlcNAc\(_2\), which may also be reglucosylated by GT. Folding glycoproteins oscillate then between monoglucosylated and unglucosylated forms catalyzed by the opposing activities of GT and glucosidase II. The monoglucosylated forms are recognized by two ER-resident lectins, calnexin and calreticulin. Upon reaching the proper tertiary structures, glycoproteins become substrates of glucosidase II, but not of GT. Properly folded molecules thus liberated from the lectins are then free to pursue their transit to the Golgi. Proteins that fail to properly fold are retained in the ER and are eventually transported to the cytosol, where they are degraded in the proteasomes. Interaction of monoglucosylated glycans with ER lectins not only retains misfolded glycoproteins in the ER, but also facilitates glycoprotein folding by preventing aggregation. GT is therefore the key constituent of the ER quality control of glycoprotein folding, as it is the only element in such a process that discriminates between glycoprotein conformers.

*In vitro* and *in vivo* assays have shown that, under both conditions, GT preferentially glucosylates glycoproteins not in extended conformations, but at more advanced folding, molten globule-like stages, when glycoprotein substrates already display secondary structures and some long-range interactions (3–5). Solvent-accessible hydrophobic amino acid patches have been identified as the structural elements recognized by GT in non-native conformers, as they are the only structural features exclusive of molten globule-like folding intermediates (6). Moreover, *in vitro* assays have shown that GT preferentially glucosylates N-glycans in the close vicinity of protein structural perturbations (6, 7).

Sequence analysis (BLAST search using a BLOSUM62 matrix) of mammalian, insect, and yeast GTs has suggested that they are composed of at least two domains: the N-terminal domain, which composes 80% of the molecule, has no significant similarity to other known proteins, and has been suggested to be involved in non-native conformer recognition; and the C-terminal domain, which binds 5-azido-[β-32P]UDP-glucose and displays a similar size and significant similarity to members of glycosyltransferase family 8 (8–12). Members of this family conserve the anomeric configuration of the monosaccharide transferred from a sugar nucleotide, presumably by forming a monosaccharide-enzyme intermediate. All GT C-terminal domains from different species share a significant
similarity (65–70%), but no such similarity may occur between N-terminal domains. For instance, Rattus norvegicus GT (GenBankTM/EBI accession number AAF67072) and Dro sophila melanogaster GT (accession number Q09332) N-terminal domains share a 32.6% similarity, but they show only 15.5 and 16.3% similarities, respectively, to the same portion of Schizo saccharomyces pombe GT (accession number SP66969) (8–10). Although there is both structural and experimental evidence suggesting the idea that the C-terminal domain is the catalytic portion of the enzyme, the notion that the N-terminal domain is responsible for recognition of non-native conformers has not been experimentally confirmed yet.

The work reported here provides experimental evidence for the two-domain GT structure and shows that both domains are tightly bound. If the N-terminal domain is indeed responsible for conformation recognition, this finding provides a molecular rationale for the close proximity between protein structural perturbations and N-glycans required for glucosylation. Moreover, the S. pombe GT N-terminal domain linked to the D. melanogaster GT C-terminal portion or the inverse construction formed enzymatically and functionally active enzymes in vivo. We therefore concluded that N-terminal domains from different GTs share three-dimensional features despite showing an extremely poor similarity in their primary sequences. In addition to the suggested role in the recognition of misfolded conformers, the results presented here indicate that the N-terminal domains are required for proper folding of the catalytic C-terminal portions, as the latter were enzymatically and functionally inactive and rapidly degraded in vivo when expressed in the absence of the former.

EXPERIMENTAL PROCEDURES

Yeast Strains, Plasmids, Expression Vectors, and Yeast Growth Media—The following S. pombe strains were used: wild-type, Sp95 (h−, ura4-d18, ade6-M216, leu1-32, Sp61G4 (h+, ade6-M210, leu1-32, ura4-d18, gpt::ura4-d1684), and Sp61G4a (h+, ade6-M210, ade1-1, leu1-32, ura4-d18, gpt::ura4-d1684, alqg::ura4 (′)) (13). Plasmid pBluescript SKgt1′33A and expression vector pREP5Xgt1′, which encode S. pombe GT, were generated in our laboratory (13). Plasmid pOT2gt1 (D. melanogaster GT) was a generous gift from Dr. S. E. Trombetta (Department of Cell Biology, Yale University Medical School). Expression vectors pREP2 and pREP3X encode the S. pombe ura4 gene and the Saccharomyces cerevisiae LEU2 gene, respectively. The latter complements the S. pombe leu1 mutation. Rich medium contained 3% glucose (Sigma) and 0.5% yeast extract (Difco). Minimal medium was as described by Alfa et al. (14). The final concentrations of adenine, uracil, and leucine supplements were 75, 75, and 250 μg/ml, respectively.

Materials—[14C]Glucose (300 Ci/mol) was from New England Nuclear. UDP-[14C]Glc (300 Ci/mol) was prepared according to Wright and Robbins (15) with slight modifications. N-Methyloxycinnamoyl and endo-β-N-acetylglucosaminidase II were from Sigma. Anti-c-Myc monoclonal antibodies were a gift from Dr. S. E. Trombetta.

Methods—GT was assayed using UDP-[14C]Glc as a sugar donor and denatured thyroglobulin as a glucosyl acceptor as described (16). Proteins were microsequenced at the Department of Biochemistry and Molecular Biology of the University of Nebraska (Omaha, NE). In all constructs described below, the accuracy of all DNA junctions generated was checked by DNA sequencing. S. pombe cells were labeled with [35S]methionine and [35S]cysteine (5% of total cell radioactivity) for 4 h and then chased with cold medium containing 50 μg/ml cycloheximide. Cells were harvested and cell pellets were washed three times with cold TNE buffer. Reduced and nonreduced samples were run side by side on the same gel to allow direct comparison of both type of samples. Reduced and nonreduced samples were heated at 95 °C for 3 min, and N-ethylmaleimide (25 mM) was added to the reduced samples to allow direct comparison of both type of samples. Antisense and sense oligonucleotides were employed. Purified GT (2 μg) was resuspended in 10 μl of solution containing 10% glycerol, 2% SDS, and 50 mM Tris-HCl (pH 7.6). N-Ethylmaleimide was added to nonreduced samples at a final concentration of 10 μM. Both reduced and nonreduced samples were heated at 95 °C for 3 min, and N-ethylmaleimide (25 mM) was added to the reduced samples to allow direct comparison of both type of samples. Reduced and nonreduced samples were run side by side on the same gel.

Western Blot Analysis of Proteins Bearing the c-Myc Epitope—Microsomes prepared from yeast cells grown to A600nm = 0.5–1.0 were resuspended in 2% SDS, 5 mM EDTA, 50 mM Tris-HCl, 100 mM dithiothreitol, 1 mM phenylmethylsulfonl fluoride, 1 mM Np-tosyl-l-lysyl chloromethyl ketone, 1 mM Np-tosyl-l-phenylalanoyl chloromethyl ketone, 1 μg E-64, 1 μg pepstatin, and 10 μM leupeptin. Samples were separated by 10% SDS-PAGE, electrotransferred to nitrocellulose filters, and developed with anti-c-Myc antisera according to standard procedures.

Antibodies against the S. pombe GT C-terminal Domain—A 519-bp fragment of S. pombe GT (bases 3658–4176) was amplified using Pfu DNA polymerase (Roche Applied Bioscience), pBluescript SKgt1′33A as template, and the following oligonucleotides as primers: 5′-CGGAATTCGACATTGCATACTGACGACGGCTG-3′ (sense) and 5′-GCGCCGCCTTAATCAGTGGTTTGC-3′ (antisense). The fragment was released by EcoRI and NofI treatment (the primers employed create cleavage sites for these two enzymes) was ligated into the pET22b expression vector (Novagen) previously treated with the same enzymes, and the construct was amplified in Escherichia coli NovaBlue cells (Novagen). A protein comprising amino acids 1220–1382 of S. pombe GT fused to a polyhistidine tail was synthesized by transforming E. coli BL26(DE3) cells (Novagen) with above-described construct. The protein inclusion bodies was washed; dissolved in 6 M urea, 25 mM Tris-HCl (pH 8.0), 0.5 M NaCl, and 5 μM imidazole buffer; and purified by affinity chromatography on Ni2+–immunodiacetic acid-Sepharose (Amersham Biosciences). A single 20-KDa band appeared in 10% SDS-PAGE. Antibodies raised in rabbits reacted with the recombinant protein at least at a 1:1000 serum dilution.

Construct with the D. melanogaster GT N-terminal and S. pombe GT C-terminal Domains—Two synonymous mutations (A3461G and A3464C) were introduced into the gpt1′ sequence encoding the N- and C-terminal domain junction in expression vector pREP3Xgt1′, thus creating a Stul site (cleavage with this enzyme generates a blunt end). For this purpose, the mutagenic oligonucleotide 5′-TTTCAAGCCGTAAGAGGCCCTCTAAAATTTTTCTTGGC-3′ (sense; generates a blunt end) and 5′-GTGACATCCGACGGGGCTCATGAGAAGGCGT-3′ (antisense; generates a BlpEI site), plasmid pOT2gt1 as template, and Pwo DNA polymerase. The PCR product was digested with BlpEI and ligated to pREP3Xgt1Stul previously digested with Stul and BlpEI. The latter cuts S. pombe gpt1′ just before the ER retrieval sequence. The expression vector generated was termed pREP3Xgt1CTDm (Fig. 1A).

Conduct with the D. melanogaster GT N-terminal and S. pombe GT C-terminal Domains—Four mutations (C56G, A57C, A58G, and G58R) were introduced into the gpt1′ sequence encoding the signal recognition domain between the S. pombe GT signal sequence and the N-terminal domain to create a BsiHI site. These mutations resulted in the point mutation K20R at the amino acid level. Mutations were generated by conducting three PCRs. The first one used primers 5′-CGGCGATTCGCCGCGGTCAGG-3′ (sense; hybridizes with a noncoding region ~ 100 bp upstream of the BsiHI site). The second one used primers 5′-CATCTACATACATTTTCTCTTGGC-3′ (sense; generates a blunt end) and 5′-TGCTACACATCCGACGGGGCTCATGAGAAGGCGT-3′ (antisense; generates a BspEI site), plasmid pOT2gt1′ as template, and Pwo DNA polymerase. The PCR product was digested with BspEI and ligated to pREP3Xgt1Stul previously digested with Stul and BspEI. The latter cuts S. pombe gpt1′ just before the ER retrieval sequence. The expression vector generated was termed pREP3Xgt1CTDm (Fig. 1A).
primers were those employed in the first and second reactions, respectively. The 606-bp fragment obtained was digested with BamHI and BstXI and ligated to expression vector pREP3Xgpt1StuI (Fig. 1A) previously treated with the same restriction enzymes, thus generating expression vector pREP3Xgpt1StuI-BssHII (Fig. 1B). The fragment encoding the N-terminal domain of D. melanogaster GT was synthesized using primers 5'-TGCCGCGCAATCCAGTCAGAGCTATCC-3' (sense) and 5'-CCGTATCCTCATCAGAGGC-3' (antisense). The first primer encodes the retrieval plus stop sequences, and the second one encodes the N- and C-terminal domain junction (centered in base 3444 from the ATG codon). The fragment was purified, phosphorylated, reannealed and ligated to the above-described constructs previously digested with BstHI and ligated to pREP3Xgpt1StuI-BssHII previously digested with StuI and BssHII. The expression vector generated was termed pREP3Xgpt1StuIBssHII (Fig. 1B).

Truncated Protein Constructs—The expression vector encoding the S. pombe GT signal peptide followed by the N-terminal domain and the ER retrieval sequence was synthesized using first performing inverse PCR using pBluescript SKgpt1'33A (which encodes full-length GT) as template and oligonucleotide primers 5'-CCGGACGAACCTTGAACCTG-3' (sense) and 5'-AAAGAAGCTGAGAAGACTT-3' (antisense). The first primer encodes the retrieval plus stop sequences, and the second one encodes the N- and C-terminal domain junction (centered in base 3444 from the ATG codon). The fragment was purified, phosphorylated, reannealed, ligated and amplified in E. coli DH5α cells (Invitrogen). The resulting plasmid (pBluescript SKgpt1INT) was digested with SnaBI and BspEI. The 355-bp fragment generated was ligated to expression vector pREP3Xgpt1 previously treated with the same restriction enzymes. The resulting expression vector was pREP3xgpt1INT (Fig. 1C). The expression vector encoding the GT signal peptide followed by the C-terminal domain and the ER retrieval sequence was constructed in a similar way, but with the following primers: 5'-AATTTCAACGCTGAAGAAGG-3' (sense) and 5'-GCGGCGCAATCCAAGCTTGAACCTG-3' (antisense). The first primer encodes the retrieval plus stop sequences, and the second one starts at the sequence coding for the junction between the GT signal sequence and the mature protein sequence. The resulting plasmid (pBluescript SKgpt1CINT) was digested with BamHI and NruI. The resulting 590-bp fragment was ligated to the expression vector pREP3Xgpt1 previously treated with the same restriction enzymes. The resulting expression vector was pREP3xgpt1CINT (Fig. 1C).

Insertion of the c-Myc Epitope at the C Terminus—For inserting the c-Myc epitope (EQKLISEEDLN) into chimeric and truncated proteins before the ER retrieval sequence, oligonucleotide primers 5'-CCGGACGCAAACACTCATCTGACAGGATCTGAAATGCT-3' (sense) and 5'-CCG-GATTCAGATCCTCTTCTGAGATGAGTTTTTGTT-3' (antisense) were annealed and ligated to the above-described constructs previously digested with BspEI. The sequence coded for the epitope and left, at both ends, sequences able to bind DNA previously digested with the restriction enzyme. All constructs had an additional Pro residue before the c-Myc epitope. The resulting expression vector designations include “c-myc” to indicate insertion of the epitope-encoding sequence.

RESULTS

Linkage between the N- and C-Terminal Domains Is Extremely Sensitive to Endoproteolysis—Fig. 2A shows that all active S. pombe GT preparations yielded three bands when subjected to 10% SDS-PAGE. The same result was obtained also with some R. norvegicus enzyme preparations (Fig. 2B). The sizes of the fragments were 161, 130, and 35 kDa for yeast GT and 172, 135, and 38 kDa for the mammalian enzyme. The results obtained from microsequencing the N termini are depicted in Fig. 2 (A and B). They reveal that, in both cases, the smaller fragments had been produced by endoproteolytic breakage of the full-length species at the junction between the N- and C-terminal domains (amino acids 1149 and 1220 from the initial Met residue for the yeast and mammalian GTs, respectively) (Figs. 2A, B, and C). As expected, the largest and smallest (but not the middle) fragments of S. pombe GT reacted with polyclonal antibodies directed against a C-terminal 172-amino acid fragment of the molecule (Figs. 1A and 2C).

Noncovalent Linkages Tightly Bind Both Domains—The N- and C-terminal domains of R. norvegicus and S. pombe GTs proved to be tightly bound, as the small fragment yielded by endoproteolytic degradation could not be separated from the middle one by gel filtration on a Bio-Sil SEC 250 column (Fig. 4, A and C), by chromatography on a phenyl-Superose column (Fig. 4, B and D), by ion-exchange chromatography on a Mono Q HR 5/5 column, and by filtration through a Centricon-100 filter (data not shown). Moreover, the linkage between both...
domains was not mediated by disulfide bonding as revealed by 10% SDS-PAGE performed under reducing and nonreducing conditions (Fig. 4, E and F). The small and middle fragments of both R. norvegicus and yeast GTs were retained on a concanavalin A-Sepharose column despite the fact that the mammalian enzyme C-terminal domain lacks N-glycosylation consensus sequences and that the only one present at such a location in the yeast enzyme was unoccupied as revealed by 10% SDS-PAGE following endo-\(\text{H}9252\)-N-acetylglucosaminidase \(\text{H}\) treatment (data not shown). This result shows that the N- and C-terminal domains could not be separated also by lectin affinity chromatography.

Noncovalently Bound N- and C-terminal Domains Are Enzymatically Active and Able to Discriminate between Native and Non-native Conformers—We observed that certain S. pombe preparations were enzymatically active despite almost totally lacking the full-length protein component. A slight modification of the purification procedure was then introduced to fully eliminate it: the interaction period of the applied material with the concanavalin A-Sepharose column before elution with \(\text{H}9251\)-methylmannoside was increased from 1 to 12 h. As depicted in Fig. 5 (A and C), although the resulting preparation completely lacked the full-length enzyme, it displayed enzymatic activity. As expected, the largest fragment in the extensively degraded preparation did not react with antibodies directed against the last portion of the C terminus (Fig. 5 B). As the same procedure (longer interaction of the applied material with concanavalin A-Sepharose) did not yield an R. norvegicus preparation devoid of the full-length enzyme, the purified GT was incubated with protease V8. As shown in Fig. 5 D, the full-length R. norvegicus enzyme disappeared after a 5-min incubation, yielding 135- plus 38-kDa fragments that displayed almost full enzymatic activity. Microsequencing of the latter yielded the sequence SFKWG, indicating that protease V8 cleavage had occurred at amino acid 1211 from the initial Met residue, i.e. 9 amino acids ahead of that produced by endoproteolytic cleavage (see above). The 38-kDa protein was gradually further degraded to smaller fragments, and the enzymatic activity was concomitantly lost (Fig. 5 D). Both R. norvegicus and S. pombe enzyme preparations devoid of full-length compo-
Domains Glucosyltransferase

N- and C-terminal Domains May Be Expressed Separately, Yielding an Active Enzyme—The notion of a two-domain GT structure able to display enzymatic activity even if not covalently bound was supported by the synthesis of an active enzyme upon transfection of a GT null mutant (Sp61G4) with two plasmids separately encoding the N and C termini (Table I). Nevertheless, microsomes prepared from cells transformed with both expression vectors displayed a lower specific enzymatic activity than those isolated from cells transfected with a plasmid coding for the full-length enzyme. Constructs coding for either one of both domains had segments encoding the GT signal peptide before the enzymatic domains and a c-Myc epitope (EQKLISEEDLN) before them (Fig. 1A, lanes 1–4) and Sp61G4 (lanes 5 and 6) cells expressing full-length Sp. pombe GT (expression vector pREP3Xgp1T-c-myc) and pREP2gpt1NT-c-myc domains and full-length Sp. pombe GT (pREP3Xgp1T-c-myc), and pREP2gpt1NT-c-myc + pREP3Xgp1T-c-myc 1071.5 5840.4 28 42 42 1071 5840.4 2

Expression vectors (S. pombe) and N-terminal (pREP2-gpt1T-c-myc) and C-terminal (pREP3Xgp1T-c-myc) domains and full-length Sp. pombe GT (pREP3Xgp1T-c-myc).

Values are expressed as counts/min/mg of microsomal protein.

nents glucosylated denatured (but not native) thyroglobulin (data not shown).

N- and C-terminal Domains Are Mutually Interchangeable—As shown above, in both yeast and mammalian GTs, the N- and C-terminal domains are tightly bound and show enzymatic activity even in the absence of any covalent linkage between them. This was a rather unexpected result because whereas the C-terminal domains of R. norvegicus, D. melanogaster, and S. pombe share a high similarity (62.3–74.0%), much lower values occur for the N-terminal portions (32.6% for the R. norvegicus and D. melanogaster GTs and 15.5–16.3% for the latter enzymes and S. pombe GT). In other words, to bind highly similar C-terminal domains, the tertiary structures of the N-terminal portions of different GTs should be expected to share three-dimensional features. To confirm that the N-terminal domains from different GTs share elements of tertiary structure, two types of chimeras were constructed in expression vector pREP2X. The first one encoded the D. melanogaster GT N-terminal and S. pombe GT C-terminal domains (Fig. 1B). The second construct coded for the yeast GT N-terminal and fly GT C-terminal domains (Fig. 1A). These constructs, together with those encoding either the full-length S. pombe enzyme or only its C-terminal portion, were transfected into Sp. pombe Sp61G4A mutant cells. All constructs coded for the S. pombe GT signal peptide at the N termini and for the above-mentioned c-Myc epitope before the S. pombe GT ER retrieval sequence (PDEL).

Sp61G4A mutants lack GT and transfer to Manα-GlcNAc₂ instead of the complete glycan, as they lack the dolichol-P-Glc-dependent glucosyltransferase responsible for Glc₁-Man₉-GlcNAc₂-P-dolichol formation. As they synthesize underglycosylated glycoproteins and lack the folding facilitation process mediated by glycoprotein-calnexin interaction, these double mutant cells grow poorly with a round morphology at 28°C and do not grow at 37°C. Their transformation with a GT-encoding expression vector rescued the wild-type phenotype, but no rescue was observed when two point mutations that abolished GT activity were introduced at the C-terminal domain (13). The C-terminal domain was effectively expressed as revealed by Western blot analysis (Table I). On the contrary, the Glc₁-Man₉-GlcNAc₂ shoulder was absent in cells transformed with full-length S. pombe GT or with both chimeras coding for mixed yeast-fly enzymes (Fig. 7A, C, D, and E).

Experimental Procedures.

Transformed cells were incubated with 5 mM [¹⁴C]Glc in the presence of N-methyldeloxynoririmycin, a glucosidase II inhibitor. Analysis of whole cell N-glycosidochelin released by endo-β-N-acetylglucosaminidase H showed that, in addition to peaks migrating as a Man₉-GlcNAc standard, shoulders in the position expected for Glc₁-Man₉-GlcNAc₂ were formed in cells transfected with full-length S. pombe GT or with both chimeras (Fig. 7A, C, D, and E). Labeled glucose residues were present only in substances derived from cells transformed by the expression vector encoding full-length GTs (Fig. 7A, E–H).

The C-terminal domain was effectively expressed as revealed by Western blot analysis, but the protein was extremely unstable, as no expressed protein was detected after a 2-h incubation of intact cells with cycloheximide, thus strongly suggesting that it had been unable to properly fold in the absence of the N-terminal domain (Fig. 6B, lanes 3 and 4). On the contrary, full-length S. pombe GT continued to be detected after a similar cycloheximide treatment (Fig. 6B, lanes 1 and 2).

To confirm that the chimeric enzymes were not only enzymatically but also functionally active in vivo, we investigated whether they rescued the non-growth phenotype of Sp61G4A mutants at 37°C. As shown in Fig. 8, A–C, in all cases, the mutant cells transformed with full-length GT constructs, but not with either the C-terminal domain-encoding expression
Cells were incubated for 30 min with [14C]Glc in the presence of 2.5 mM 5-azido-[32P]UDP-glucose (10). These observations indicate that the C-terminal domain is the catalytic portion of the molecule where formation of the putative glycosyl-enzyme intermediate takes place. Expression of the R. norvegicus GT C-terminal domain in insect cells yielded a protein displaying ~5% of the specific full-length enzymatic activity (10). It was not reported, however, whether the recombinant truncated enzyme discriminated between properly and not properly folded glycoprotein substrates. As shown above and will be further discussed below, we were unable to obtain an enzyme active either in vivo or in vitro upon expressing the S. pombe C-terminal domain in a yeast GT null mutant. It has been suggested that the GT N-terminal domain is responsible for sensing the folding status of glycoprotein substrates. An additional or alternative role (or perhaps the sole one) is, as demonstrated here, to contribute to proper folding of the catalytic or C-terminal domain within the ER luminal environment.

To confirm the presence of common tertiary structural elements in the N-terminal domains of different GTs, as suggested by the tight association between highly similar C-terminal and highly divergent N-terminal domains, chimeras containing the D. melanogaster GT N-terminal and S. pombe GT C-terminal domains or the yeast GT N-terminal and fly GT C-terminal domains were expressed in S. pombe GT null mutants. Chimeric proteins were enzymatically active in vivo and were able to rescue the wild-type phenotype (growth at 37 °C) of gpt1/alg6 double mutant cells (Sp61G4A). These mutants lack GT and synthesize underglycosylated glycoproteins, as they transfer (inefficiently) Man9-GlcNAc instead of Glc3-Man9-GlcNAc (13). The severe ER stress to which they are submitted pre-
domain structure. The results presented here, however, show that, despite the extremely poor primary sequence similarity shown by GT N-terminal domains from different species, such portions probably share substantial elements of tertiary structure, as yeast and fly GT N-terminal domains were able to facilitate folding and stabilization of the GT C-terminal portions of both species alike.

A very unique feature of GT is that it glucosylates only N-glycans in the very near proximity of protein structural perturbations. It was recently reported that only the N-glycan attached to the misfolded subunit of an artificial dimer formed by properly folded and misfolded RNase B monomers was glucosylated by GT in vitro assays (6). It was further demonstrated that the N-glycan was glucosylated only when present in the immediate vicinity of the localized protein structural perturbation introduced into a mutant RNase B monomer by abolishing one of the four disulfide bonds present in the wild-type molecule (7). These recent results agree with previous ones reporting that N-glycans and misfolded protein subunits have to be covalently linked to be glucosylated (19). The close proximity between both structural elements required for efficient glucosylation (N-glycans and structural perturbations) may be best explained by the results reported herein, i.e. by the intimate association between the catalytic (C-terminal) and putative misfolded recognition (N-terminal) domains by noncovalent bonding.

It should be stressed that ascription of the location of the conformation recognition site to the N-terminal domain is merely speculative. It could be that its sole role is, as demonstrated here, to facilitate folding and stabilization of the catalytic C-terminal portion and that the conformation-sensing site resides also in this last part of the molecule. The sole evidence against this possibility is that enzymes belonging to glycosyltransferase family 8, with which GT C-terminal domains share a similar size and significant sequence similarity, probably do not have the capacity to discriminate between glycoprotein conformers, as most of them are not involved in glycoprotein glucosylation. On the other hand, the close proximity between the N-glycan and the protein structural perturbation required for glucosylation could also be explained if both the catalytic and conformation recognition sites reside in the relatively small (∼35 kDa) C-terminal domains.

REFERENCES
1. Kornfeld, R., and Kornfeld, S. (1985) Annu. Rev. Biochem. 54, 631–664
2. Parodi, A. J. (2000) Annu. Rev. Biochem. 69, 69–95
3. Trombetta, S. E., and Helenius, A. (2000) J. Cell Biol. 148, 1123–1129
4. Caramelo, J. J., Castro, O. A., Alonso, L. G., de Prat Gay, G., and Parodi, A. J. (2000) Proc. Natl. Acad. Sci. U. S. A. 106, 86–91
5. Labriola, C., Cazzulo, J. J., and Parodi, A. J. (1999) Mol. Biol. Cell 10, 1381–1394
6. Ritter, C., and Helenius, A. (2000) Nat. Struct. Biol. 7, 278–280
7. Quinn, K. (2002) Substrate Specificity and Enzymatic Properties of Bovine UDP-Glucoprotein Glucosyltransferase, an ER Folding Sensor. Diploma thesis, Swiss Federal Institute of Technology, Zurich.
8. Parker, C. G., Fessler, L. I., Nelson, R. E., and Fessler, J. H. (1995) EMBO J. 14, 1294–1303
9. Fernández, F., Jannatipour, M., Heliman, U., Rokeach, L., and Parodi, A. J. (1996) EMBO J. 15, 705–713
10. Tessier, D. C., Dignard, D., Zapun, A., Radominska-Pandya, A., Parodi, A. J., Begeron, J. J. M., and Thomas, D. Y. (2000) Glycobiochemistry 10, 403–412
11. Arnold, S. M., Fessler, L. I., Fessler, J. H., and Kaufman, R. J. (2000) Biochemistry 39, 2149–2163
12. Trombetta, S. E., and Parodi, A. J. (2002) Adv. Protein Chem. 59, 303–344
13. Fanchiotti, S., Fernández, F., D’Alessio, C., and Parodi, A. J. (1998) J. Cell Biol. 143, 625–635
14. Alfa, C., Fantes, P., Hyams, J., McLeod, M., and Warbrick, E. (1993) Experiments with Fission Yeast: A Laboratory Course Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
15. Wright, A., and Robbins, P. W. (1965) Biochim. Biophys. Acta 104, 594–595
16. Trombetta, S. E., Bosco, M., and Parodi, A. J. (1989) Biochemistry 28, 8108–8116
17. Fernández, F. S., Trombetta, S. E., Heliman, U., and Parodi, A. J. (1994) J. Biol. Chem. 269, 30701–30706
18. Trombetta, S. E., and Parodi, A. J. (1992) J. Biol. Chem. 267, 9236–9240
19. Sousa, M., and Parodi, A. J. (1995) EMBO J. 14, 4196–4203
The UDP-glucose:Glycoprotein Glucosyltransferase Is Organized in at Least Two Tightly Bound Domains from Yeast to Mammals
Marcelo Guerin and Armando J. Parodi

J. Biol. Chem. 2003, 278:20540-20546.
doi: 10.1074/jbc.M300891200 originally published online March 20, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M300891200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 17 references, 6 of which can be accessed free at http://www.jbc.org/content/278/23/20540.full.html#ref-list-1