CMLS, Cell. Mol. Life Sci. 60 (2003) 229–240
1420-682X/03/020229-12

CMLS Cellular and Molecular Life Sciences

Evolutionary and functional implications of the complex glycosylation of Skp1, a cytoplasmic/nuclear glycoprotein associated with polyubiquitination

C. M. West

Department of Anatomy and Cell Biology, University of Florida, College of Medicine, Gainesville, Florida 32610-0235 (USA), Fax: +1 352 392 3305, e-mail: westcm@ufl.edu

Abstract. Protein degradation is regulatory for the cell cycle, signal transduction and gene transcription. A critical step is the selective marking of the target protein, resulting in polyubiquitination by one of a large number of E3-ubiquitin ligases. Both target marking and E3-ubiquitin ligase activity are associated with common as well as unusual posttranslational modifications. For example, hydroxylation of Pro-residues and modification of Asn-residues by high-mannose sugar chains can target the modified proteins for rapid polyubiquitination in the mammalian cytoplasm. Both prolyl hydroxylation and glycosylation also occur on Skp1, a subunit of the SCF class of E3-ubiquitin ligases, from *Dictyostelium*. In this case, a pentasaccharide containing Gal, Fuc and *N*-acetyl-D-glucosmine (GlcNAc) is attached to the HyPro-residue. The sugars are added sequentially by enzymes that reside in the cytoplasm rather than the secretory pathway. Two of the glycosyltransferases appear to be positioned in ancient evolutionary lineages that bridge prokaryotes and eukaryotes. The first, which attaches GlcNAc to HyPro, is related to enzymes that form α-GalNAc- and α-GlcNAc-Ser/Thr linkages in the Golgi. GlcNAc is extended by a bifunctional glycosyltransferase that mediates the ordered addition of β1,3-linked Gal and α1,2-linked Fuc, using an architecture resembling that of two-domain prokaryotic glycosyltransferases involved in glycosaminoglycan synthesis. Mutational and pharmacological perturbation of glycosylation alters the subcellular localization of Skp1 and growth properties of cells. Prolyl hydroxylation and complex O-glycosylation provide the cell with new options for epigenetic regulation of protein turnover in its cytoplasmic and nuclear compartments.

Key words. Cytoplasmic glycosylation; proline hydroxylation; *Dictyostelium*; ubiquitination; Skp1; E3-SCF ubiquitin ligase.

Skp1 is a subunit of the E3^{SCF}-ubiquitin ligase and other multiprotein complexes

Proteolytic cleavage and subsequent protein degradation are irreversible regulatory events reserved for key cellular processes including the cell cycle, response to nutritional change, supramolecular assembly, some signaling pathways and apoptosis. Proteolytic cleavage of proteins such as cyclin F, cyclin D, p21^{sic1}, IκB, β-catenin, Smad3, Cubitus interruptus, and transcriptional factors involved in nutritional regulation are signaled by polyubiquitination (Ub) [1]. Ubiquitination of target proteins is catalyzed by enzymes known as E2- and E3-Ub ligases.

Skp1 is a small protein of 160–180 amino acids found in all eukaryotes. It has been independently discovered multiple times consistent with emerging evidence for many roles in the cell. Its best-understood function is an adaptor in the SCF class of Zn-Ring finger E3-Ub ligases in yeast and humans [2]. Most Ring finger Ub-ligases consist of a catalytic subunit known as an E2 enzyme and a Ring-H2 protein that recognizes the target for ubiquitination, which together form an isopeptide bond between the C-terminal Gly of Ub and the ε-NH_2 group of a specific Lys-residue of the target.

Additional subunits are present in the more elaborate SCF class of Ub ligases. The Ring finger protein Roc1/Rbx1/Hrt1 recognizes cullin-1, an elongate scaffold-type protein which has Skp1 mounted at its opposite end (fig. 1) [3]. Skp1 in turn recognizes a protein with an F-box, a 40-amino acid motif present on >38 predicted proteins in

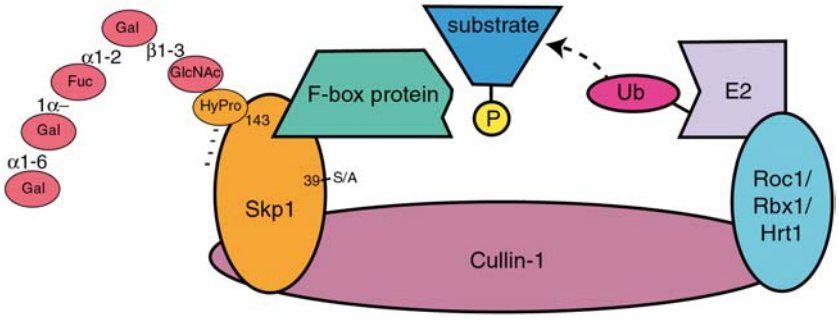


Figure 1. Schematic diagram of $E3^{SCF}$ ubiquitin ligase and the Skp1 glycan. See text for explanation. Drawing after [2, 3].

humans. Typically, a Leu-rich or WD40 domain in the F-box protein in turn mediates specific interaction with a target protein [4], which is recognized based on a specific posttranslational modification. This ususally involves a phosphate moiety, but a recent report implicates a high-mannose N-linked glycan [5] as another recognition determinant, Pro-hydroxylation is recognized by the related $E3^{VHL}$-Ub ligase [6], and a subfamily of plant F-box proteins possess a lectin-like domain [7]. Formation of this complex leads to transfer of Ub from the E2 intermediate to either the F-box protein [8] or the candidate target if present. This process repeats itself at least four times, resulting in a linear polyubiquitin chain associated with a Lys-residue on the target protein. Biochemical studies suggest that this SCF complex then docks on the 19S cap of the 26S proteasome in an ATP-dependent fashion [9], the Ubs are removed and recycled, and the target protein is degraded within the core 20S subunit. There is evidence that this process can extract the target protein from a multiprotein complex, allowing the rest of the complex to exit the proteasome undamaged [10].

Genetic studies, yeast two-hybrid screens and co-pull-down studies have identified additional candidate interacting partners for Skp1 [11], suggesting other functions. For example, the Skp1-$p58^{ctf13}$ heterodimer is a subunit of the Cbf3 DNA-binding protein complex associated with kinetochore formation [12, 13]. Skp1 is transiently associated with RAV1 and RAV2 in assembling the V_0:V_1 heterodimer of the vacuolar-type H^+-ATPase. Skp1 is also associated with factors connected with recycling of the v-SNARE Snc1p in yeast endocytosis [14]. In plants there is evidence that Skp1 interacts with Snf1-related protein kinases and the α4/PAD1 20S-proteasomal subunit in a multiprotein complex that may involve the rest of the SCF complex [15]. Skp1 seems to be associated with an unusually large number of protein complexes for such a small protein.

The use of one protein for multiple functions might result in cross-regulation if Skp1 is in rate-limiting supply, as has been suggested for its role in E3-$SCF^{\beta TRCF}$- and E3-SCF^{cdc4}-Ub ligase activity [16]. Alternatively, there may be multiple isoforms for discrete Skp1 complexes. In support of this model, there are 17 Skp1-like proteins in *Caenorhabditis elegans*, and only six of these interact in a two-hybrid assay with Cul-1, the cullin-type found in $E3^{SCF}$-Ub ligases [17, 18]. In contrast, only a single similar Skp1 is found in yeast. Humans and other mammals may also express only a single Skp1, though there is evidence for 3 Skp1-like genes and alternatively-spliced isoform sequences in EST-databases. Plants contain at least 20 Skp1-like genes [15], and *Dictyostelium* expresses 2, which differ by only an S/A-substitution at position 39 [19, 20]. Interestingly, this polymorphism also exists in *Drosophila*, which contains at least 6 Skp1-like genes. In addition, there is evidence for phosphorylation of Skp1 in insects [12] and budding yeast [13] in the region that differs beween the 2 *Dictyostelium* isoforms, and modification by an unusual pentasaccharide that is attached to a hydroxylated Pro-residue at position 143 in the lower eukaryote *Dictyostelium* [20], which is reviewed here.

Information about the biochemical pathway for assembling the SCF complex would help to understand its regulation. A recent study in *Dictyostelium* suggested that the complex is not assembled until the F-box protein interacts with its target [21], consistent with evidence that F-box proteins compete for available Skp1 docking sites [1, 15, 16]. The Cul-1 scaffold is modified at its C-terminal end by the Ub-like protein NEDD8, which might control Cul-1 assembly with Roc1 and the E2 enzyme [3, 22]. Skp1, whose purpose has never been adequately explained, might regulate assembly at the N-terminal end of the Cul-1 scaffold. Supporting the concept of regulation, Skp1 interaction with the F-box protein Grr1 is nutritionally regulated in yeast [23], mammalian Skp1 can dimerize in vitro [24], which might compete with complex formation, and some Skp1 functions depend on its glycosylation in *Dictyostelium* (see below). In addition, there is evidence for other protein factors that regulate the E3-SCF^{Skp2}-Ub ligase [25].

Skp1 is modified by a novel pentasaccharide in *Dictyostelium*

Structural studies show that Skp1 is modified by a core trisaccharide equivalent to the blood group H type 1 antigen, Fucα1,2Galβ1,3GlcNAc, to which are attached two α-linked Gal residues, possibly as a Galα1,6Gal disaccharide in α-linkage to Fuc (fig. 1) [26]. The complete pentasaccharide is attached to a hydroxylated Pro residue at position 143 of the protein. Western blot and mass-spectrometry (MS) studies suggest that $\geq 90\%$ of both isoforms of Skp1 are fully glycosylated at steady state [27].

This complex glycan is distinct from a frequent modification in the cytoplasm and nucleus consisting of O-linked β-GlcNAc attached to Ser or Thr residues of proteins [28]. Although they each contain a GlcNAc linked to a hydroxyamino acid, the amino acid type and linkage anomericity differ, as Skp1 GlcNAc is predicted to be α-linked (see below). A related modification in the cytoplasm is the attachment of O-GlcNAc in α-linkage to Ser and Thr residues of small GTP-binding proteins by the α-toxin of *Clostridium novyi* [29]. However, this modification is, like O-β-GlcNAc, not thought to be extended. Although there are numerous suggestions from the literature for complex (two or more sugars) glycosylation of other cytoplasmic proteins, the evidence needs to be confirmed by direct structural studies [20]. *Dictyostelium* Skp1 is the best-documented example of a glycoprotein modified by a complex glycan consisting of multiple kinds of monosaccharides in the cytoplasm or nucleus.

Skp1 is glycosylated in the cytoplasm

In principle, Skp1 might be transiently translocated into the secretory pathway to be modified by conventional glycosyltransferases, or there might be a novel glycosylation pathway in the cytoplasm dedicated to the selective modification of cytoplasmic proteins. Skp1 does not possess a conventional signal peptide for directing it to the rough endoplasmic reticulum (rER), and a potential N-glycosylation sequon is not utilized [30], together suggesting that Skp1 might be glycosylated in the cytoplasm. Since enzymes of this type had not been previously described in this compartment, in vitro assays were developed using recombinantly expressed Skp1 structural mutants that are incompletely glycosylated in vivo or Skp1 from glycosylation mutant strains. These assays were used to search for and characterize the Skp1-modification enzymes, and to guide their purification for a proteomics approach to clone two of the genes.

Prolyl hydroxylation

The Skp1 prolyl hydroxylase (PH) probably modifies the 4-position of Pro143, as a synthetic peptide containing 4-HyPro is a good substrate for the Skp1 GlcNAcTase (see below). Treatment of cells with agents that inhibit animal and algal 4-prolyl hydroxylase (4-PH) causes accumulation of a slightly lower M_r (~1 kDa) isoform of Skp1 that is not glycosylated [27]. This is consistent with the in vitro data that Skp1 is modified by a conventional 4-PH.

To biochemically characterize the activity, an assay used for 4-PH from the rER [31], based on replacement of a proton on the 4-carbon atom with a hydroxyl group, yielding ^3H_2O from [3,4-^{3}H]Pro, was performed using recombinant *Dictyostelium* Skp1A-His_{10} [32] from *Escherichia coli* cells metabolically labeled with [3,4-^{3}H]Pro. A cytosolic fraction (S100) of *Dictyostelium* exhibited substantial activity that was not seen using the vesicle (P100) fraction even when the membrane barrier was disrupted using non-ionic detergents [West et al., unpublished data]. Activity was dependent on substrates and cofactors of known prolyl hydroxylases. Thus, the Skp1-dependent activity has properties expected of a 4-PH except that it is not associated with the vesicular (rER) fraction of the cell.

Recently, 4-PHs have also been found in the cytoplasm of animal cells, where they modify the transcriptional factor HIF-1α (hypoxia inducing factor) in an O_2-dependent manner [33]. These enzymes are evolutionarily related to the rER 4-PHs, and a *Dictyostelium* homolog encoded on chromosome 2 is a candidate for modifying Skp1. However, the *Dictyostelium* genome also harbors genes more closely related in sequence to the rER 4-PHs, except that they lack signal peptides [20]. These are also candidates for encoding the Skp1 PH that might differ from the HIF-1α type in their O_2-dependence threshold.

Skp1 P143-hydroxylation does not occur in Skp1 expressed in prespore cells under the control of the prespore cell-specific *cotB* promoter, resulting in accumulation of a low M_r isoform that is not glycosylated [27, West et al., unpublished data]. Hence prolyl hydroxylation appears to be developmentally regulated and is rate limiting in this cell type, but the physiological significance of this is unclear.

GlcNAc-Tase

GlcNAc is transferred to Skp1 HyPro143 from UDP-GlcNAc by an enzyme activity that is associated with a protein, GnT51, after a high degree of purification from *Dictyostelium* [34]. GnT51 is encoded by a two-exon gene and the predicted complementary DNA (cDNA), when expressed in *E. coli*, induced the expression of Skp1 GlcNAcTase activity [35]. The predicted sequence of GnT51 lacks a leader peptide for targeting the protein to the secretory pathway (fig. 2A.i), consistent with its presence in the cytosolic fraction of lysed cells and its soluble nature during

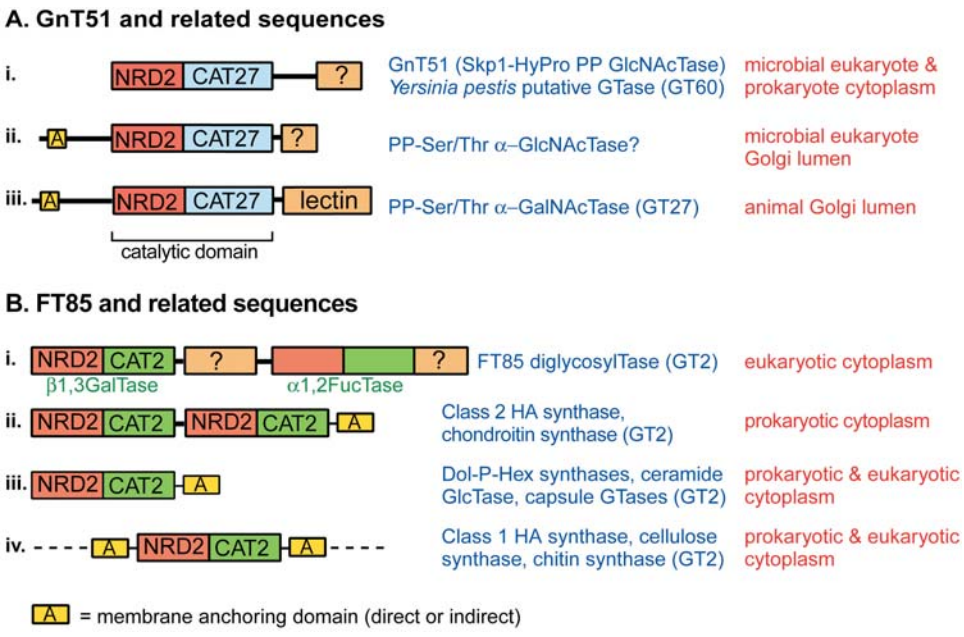


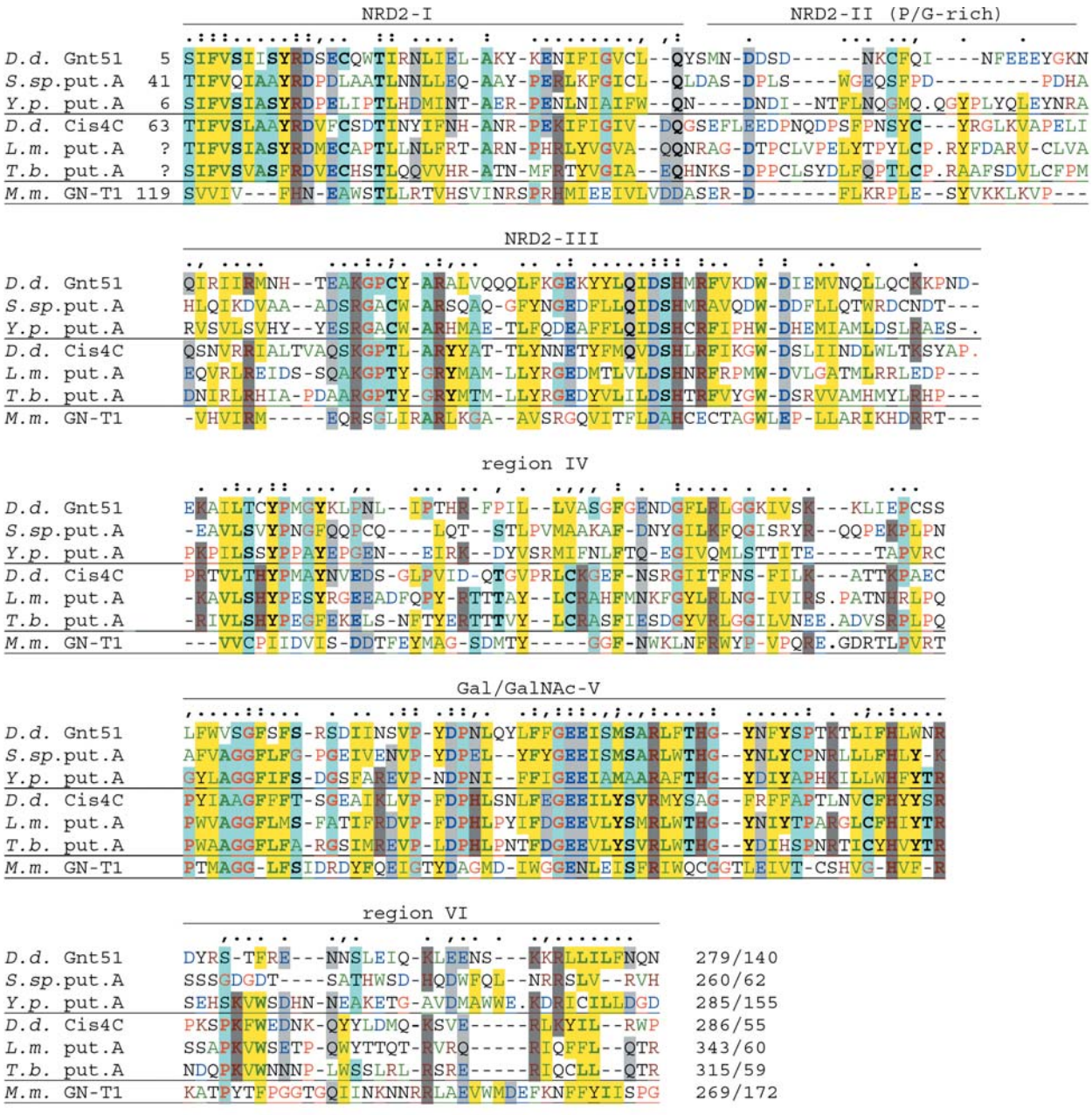
Figure 2. Domain diagrams of Skp1 GTases and related sequences. (*A*) The Skp1 GlcNAc-Tase (GnT51) (i) is compared with the Golgi-associated family GT27 animal PP-Ser/Thr α-GalNAc-Tases (iii) and related sequences predicted to reside in bacterial cytoplasm (i) and eukaryotic microbial Golgi (ii). (*B*) The Skp1 β-Gal-Tase/α-Fuc-Tase (FT85) (i) is compared with two-domain GTases involved in the synthesis of bacterial glycosaminoglycans (ii), and single-domain GTases contributing to the synthesis of either polysaccharides (iv), some with repeating disaccharides, or glycolipids (iii). See text for explanation.

purification. Thus, the Skp1 GlcNAc-Tase differs from conventional protein-modifying GTases, which are type II membrane proteins with their catalytic domains projecting into the lumen of the Golgi apparatus [36].

The sequence of GnT51 shows significant similarity to the ~275-amino-acid-long catalytic domain of family GT27 Golgi mucin-type polypeptide GalNAc-Tases (fig. 2A.iii) [20]. The catalytic domains of these enzymes contain a 125-amino acid NRD2 (nucleotide recognition domain-2)-like motif followed by a 150-amino acid cat27 motif containing the Gal/GalNAc-like motif [37]. Though GnT51 is only 14% identical to murine PP GalNAc-T1, its similarity is 45% based on the alignment in figure 3. Family GT27 GTases have a highly conserved DxH motif that is related to the metal-binding DxD-tripeptide motif that characterizes GTases of the SpsA superfamily [36, 38]. Site-directed mutagenesis of D or H of the DxH motif of GnT51 abolishes its catalytic activity in vitro [35], as shown previously for family GT27 GalNAc-Tases [37]. C-terminal to the catalytic domain is a second, 70-amino acid domain separated by a 90-amino acid region rich in Asn and Ser residues (fig. 2A.i). Polypeptide GalNAc-Tases also have a C-terminal domain (fig. 2A.iii), which has been proposed to be involved in acceptor substrate recognition [39]. Though there is no apparent sequence homology between their C-terminal domains, it is required for activity of recombinant GnT51 [35]. Both GnT51 and family GT27 enzymes catalyze the linkage of HexNAc-moieties to hydroxyamino acids, but the identity of the HexNAc (GlcNAc vs. GalNAc) and of the acceptor amino acid (HyPro vs. Ser/Thr) differ. The family GT27 enzymes attach GalNAc in α-linkage, so this may also be the linkage formed by GnT51 but this remains to be confirmed. Although clearly related to the GT27 sequences, GnT51 is sufficiently distinct that it defines a new family, GT60 [40].

Several sequences related to GnT51 can be detected in the genomes of prokaryotes and microbial eukaryotes using BLAST algorithms. The most similar predicted open reading frames are seen in the proteobacterium *Yersinia pestis* and the cyanobacterium *Synechococcus* spp. WH8012 [20, 35] (fig. 2A.i). These are predicted to en-

Figure 3. Alignment of sequences similar to GnT51. Six sequences related to GnT51 were initially identified by BLAST analyses of publicly accessible databases, and alignments were completed manually giving preference to the registration of hydrophobic (L, I, V, M, F, Y, W, sometimes A) residues. Gaps (-) introduced to optimize alignment of the six microbial sequences were usually adjacent to P or G residues, which are frequently found at surface turns. Deleted residues are denoted with a '.'. The position of the first amino acid of the alignment within the complete predicted coding sequence is given at the beginning (if known), and the length of the predicted catalytic domain followed by the length of the downstream C-terminal domain is given at the end. Regions I–III (denoted above the alignments) are



Dictyostelium discoideum GnT51, GB:AF509501
Synechococcus sp. WH8102 put. GT-A, DOE 84588, Contig48.gene455
Yersenia pestis put. GT-A, GB:AJ414154.1, GI:15980810
D.discoideum cis4c, GB:AF375997
Leishmania major put. GT-A, GB:AC084317.4, GI:18640648
Trypanosoma brucei put. GT-A, (assembled from GI:11836577, GB:AQ951144, GB:AQ942881, GB:AZ219465, GB:AQ639524)
Mus musculus GalNAc-T1, GB:AAB58477.1, GI:2149049

similar to the NRD2 domains found also in the GT2 and GT23 families, and region V corresponds to the previously described Gal/GalNAc domain of families GT27 and GT7 (see also fig. 2). Amino acids are color-coded in groups according to the properties of their side chains: green, hydrophobic (see above); blue, acidic; dark red, basic; bright red, structure breaking (G, P). To facilitate comparison of the sequences, identical residues are bolded and similar residues are highlighted when they have majority representation at a given position among the microbe-derived sequences (yellow, hydrophobic; teal, small or structure breaking; light gray, acidic or amide; dark gray, basic). The top three sequences are predicted to be soluble cytoplasmic proteins, and the middle three sequences are predicted to be type II Golgi membrane proteins based on the presence of candidate signal anchor sequences near their N termini. In the consensus row at the top of each panel, ':' denotes identical residues, '.' denotes similarity as defined above, ',' denotes identity within either the predicted cytoplasmic or Golgi groups and ';' denotes separate identities in both groups. Origins of the sequences are given at the bottom.

code proteins of similar size and domain architecture to GnT51 in the prokaryotic cytoplasm, and conserve amino acids predicted to be important for catalytic activity (fig. 3). More distantly related sequences are found in the genomes of microbial eukaryotes, including *Leishmania major*, *Trypanosoma brucei*, *Trypanosoma cruzi* (partial sequence not shown) and *D. discoideum* (fig. 3). These sequences also share a related overall architecture, including key residues thought to be important for catalysis based on comparison with the family GT27 polypeptide GalNAc-Tases of animals (figs. 2 A. ii, iii). These microbial eukaryotic sequences are, like the animal family GT27 enzymes, predicted to be type II membrane proteins that localize to the Golgi apparatus. Four Cys residues form a pattern not conserved in the predicted cytoplasmic proteins (fig. 3), consistent with possible roles in disulfide bonds in the oxidizing environment of the Golgi.

Dictyostelium and *T. cruzi* produce cell surface glycoproteins with mucin-like domains directly modified on Ser/Thr-residues by α-linked GlcNAc [41, 42], and recent results indicate that the *Dictyostelium* cis4c gene is necessary and sufficient for the addition of GlcNAc in this organism [43]. Interestingly, the *Dictyostelium* homolog was originally identified because its deletion confers resistance to the anticancer drug cisplatin [44], suggesting a role in drug cytotoxicity.

A phylogenetic analysis of the GnT51-related sequences suggests that they fall into two groups (fig. 4). The three predicted cytoplasmic proteins are deepest in the tree, suggesting that they originated in the prokaryotic cytoplasm and that GnT51 remained there during the evolution of eukaryotes (depicted as pathway 1 in fig. 5). The predicted Golgi-associated type II proteins occupy a distinct clade (in blue). The microbial members might represent a new family of polypeptide α-GlcNAc-Tases that modify Ser or Thr residues of cell surface and secretory proteins. The animal polypeptide α-GalNAc-Tases are predicted to have also descended from this clade.

The Skp1 GlcNAc-Tase and the family GT41 polypeptide-Ser/Thr O-β-GlcNAc-Tase that also modifies proteins in the cytoplasm and nucleus belong to different GTase superfamilies defined from structural studies, which correlates with the distinct structures formed and their differences in divalent cation dependence. GnT51 belongs to the metal ion-dependent, DxD motif-containing group that includes SpsA [42], whereas the *O*-β-GlcNAc-Tase belongs to the group that lacks DxD motifs involved in metal ion coordination typified by T4-DNA Glc-Tase [45]. Another cytoplasmic polypeptide GlcNAc-Tase, the family GT44 α-toxin of *Clostridium novyi* that catalyzes the formation of a GlcNAcα1-Ser/Thr linkage on Rho proteins [29], is also a member of the SpsA superfamily, but it too has a very dissimilar sequence.

β-Gal-Tase and α-Fuc-Tase

These two enzyme activities are responsible for the addition of the second and third sugars. They are found in the same protein FT85 after millionfold purification from the cytosolic fraction of cells [32]. FT85 is necessary and sufficient for each activity based on analysis of an FT85-null strain and recombinant expression in *E. coli* [46, 47]. The β-Gal-Tase activity transfers the first sugar, Gal, from UDP-α-Gal to GlcNAc in a β1,3-linkage on Skp1 isolated from the FT85-null mutant. Synthetic GlcNAc conjugates are not active as acceptors in vitro. The α-Fuc-Tase activity transfers Fuc from GDP-β-Fuc to the Gal in an α1,2-linkage on Skp1 isolated from a GDP-Fuc synthesis mutant [32]. This activity can also fucosylate the Skp1 disaccharide Galβ1,3GlcNAc conjugated to a hydrophobic aglycon. Both activities invert the configuration of the sugar linkage as it is transferred from the donor nucleotide to the sugar acceptor substrate.

The sequence of FT85, a soluble protein of 768 amino acids, contains two GTase-like domains separated by a short stretch of Asn residues near its middle (fig. 2 B. i) [46]. The first 250 amino acids of the N-terminal half are

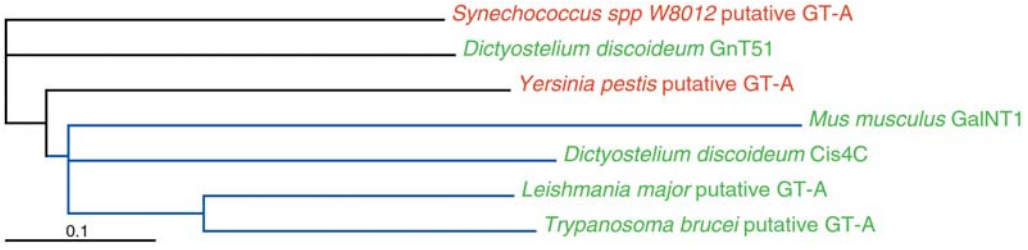


Figure 4. Phylogenetic analysis of GnT51-like sequences. The phylogenetic relationships of the sequences aligned in figure 3 were examined using a distance algorithm. Regions corresponding to amino acids 5–42, 66–123, 137–259 and 273–281 of GnT51 were included. The gaps represent insertions and deletions resulting from the divergent evolution of these homologous regions. Trees were generated using PAUP* [77] under the minimum evolution criterion, with tree-bisection reconnection branch swapping. All possible trees were searched and compared, and the tree with the best ME score is shown. The tree is rooted with the *Synechococcus* sequence, as phylogenetic studies suggest that this cyanobacterium branched earliest during prokaryotic diversification [78]. The blue clade consists of predicted Golgi type II membrane proteins. Prokaryotic and eukaryotic sequence names are in red and green, respectively.

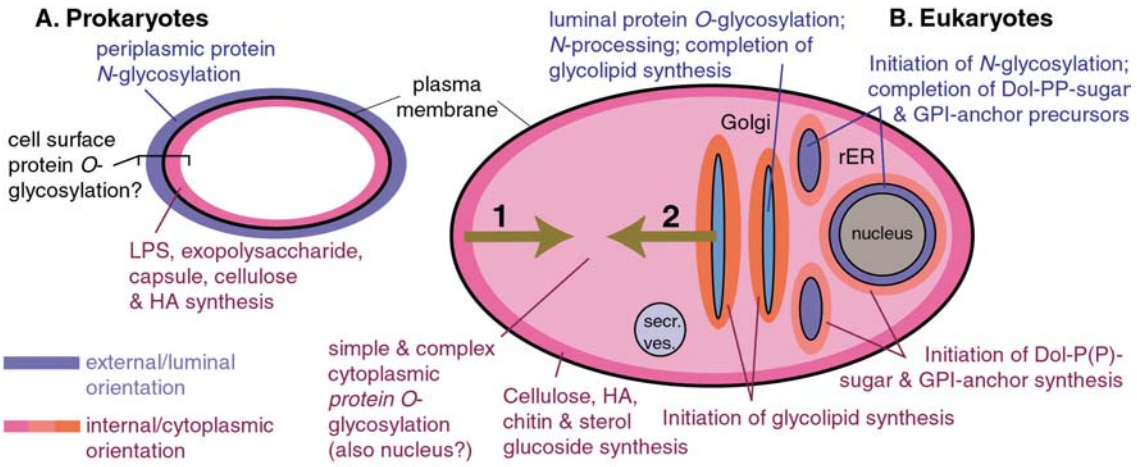


Figure 5. Possible compartmental origins of cytoplasmic O-GTases. (*A*) In prokaryotes, many lipid-linked and polysaccharide precursors of the cell wall and other cell surface structures are synthesized on the cytoplasmic surface of the plasma membrane, followed by translocation to the cell surface [73]. In Archeae, protein N-glycosylation appears to occur on the periplasmic surface of the plasma membrane, after membrane translocation [68]. The sidedness of protein O-glycosylation is under investigation, but the existence of potential prokaryotic GnT51 homologs lacking signal anchors (figs 2, 4) suggests that it can occur on the cytoplasmic side of the membrane. (*B*) This basic asymmetry is partially conserved in eukaryotes with the internalization of plasma membrane to form the rER and Golgi apparatus (external/luminal in blue shades; internal/cytoplasmic in magenta shades). Specific lipid O-glycosylation and polysaccharide synthesis remained internal but became mostly distributed among various membranes, and protein N-glycosylation localized to the luminal surface of the rER. The origin(s) of GTases oriented toward the Golgi lumen remains conjectural but may have accompanied eukaryotic evolution. Phylogenetic analyses (fig. 4 and [20]) suggest that FT85 and GnT51 of the Skp1 modification pathway possibly derived from the pool of cytoplasmically oriented, plasma membrane-associated prokaryotic GTases (scheme 1), rather than retrolocating from the Golgi (scheme 2).

similar in sequence to family GT2 inverting GTases, whose catalytic domains contain an NRD2-like motif at its N-terminus in association with a 125-amino acid motif referred to as cat2. This relationship is supported by the inactivating effects of site-directed mutagenesis of key predicted catalytic residues [47]. Expression of amino acids 1–430 in FT85-null *Dictyostelium* cells restores Skp1 β1,3Gal-Tase but not the Fuc-Tase activity, indicating that this domain can function independently as a βGal-Tase. In complementary fashion, expression of amino acids 400–768 restores the disaccharide α1,2Fuc-Tase but not Gal-Tase activity, suggesting that the C-terminal half of FT85 functions as the Fuc-Tase [46]. The C-terminal domain shows distant similarity to family GT2 sequences. Both activities are metal dependent, consistent with having DxD-like motifs near the expected positions.

Although the 2 GTase activities are capable of operating independently even in the intact protein, they also appear to be able to act processively based on the similar rates of both Gal and Fuc transfer when Skp1-GlcNAc is added in the presence of UDP-Gal and GDP-Fuc [47]. The processive, two-activity design of FT85 may ensure that the Skp1 glycan is extended rapidly to avoid accumulation of the β-Gal-terminated glycan, which might interact with β-galactoside binding lectins such as discoidins that also reside in the cytoplasm [48]. The two-domain architecture is also utilized for the synthesis of bacterial glycosaminoglycans (fig. 2B. ii) [49], which contain tandem arrays of disaccharide pairs. Although this design would seem to provide a mechanism for efficient, processive extension of the polymer, this has not been shown biochemically. Other proteins with multiple GT domains are predicted in plants and animals based on genomic analyses, but their activities are not defined [40]. The two-domain GTase architecture of FT85 and class 2 HA synthase is distinct from that of certain bifunctional GTases whose dual activities seem to reside in a single, modified family GT2 domain (fig. 2B. iii, iv) [49].

The first catalytic domain begins at the very N-terminus of FT85. There is no signal peptide or signal anchor indicating that FT85 is, like GnT51, a soluble cytoplasmic protein. The Fuc-Tase activity requires the presence of a reducing agent in vitro, concordant with the reducing potential of this compartment. These findings are consistent with the prediction that the catalytic domains of all known family GT2 GTases reside in or are oriented toward the cytoplasm [20]. FT85 is, however, unusual for GT2-domain-bearing proteins in that it is soluble rather than membrane associated, its acceptor substrate is a protein and its glycosylation product remains in the cytoplasm (or nucleus) rather than translocating across a membrane.

Function of the Skp1 glycan

Although the Skp1 glycan is novel, most of the sugar linkages can be formed by conventional GTases in the

Golgi apparatus. A potential precedent for Skp1 glycan function is therefore provided by Golgi-derived glycans, whose roles encompass myriad functions, sometimes specific to the protein modified, including stabilization of conformation and stability, and receptor-mediated targeting and signaling [50, 51]. Skp1's glycan may serve a protein-specific function, as Skp1 is the major recipient of this glycan type in vivo, based on the incorporation of [^{3}H]Fuc by metabolic labeling and into crude extracts of FT85-null cells reconstituted with purified FT85 [20]. However, related glycans may not have been detected if expressed at low levels or in other cell types. The phylogenetic generality of Skp1 glycosylation remains to be investigated, but the sheer evolutionary investment represented by the size of the six-enzyme modification pathway suggests that it will be found in other organisms as well. Information about the function of the Skp1 glycan is therefore available only in *Dictyostelium*, where it has been studied by mutating either Skp1 or specific steps in the glycosylation pathway.

Two mutant Skp1s with amino acid substitutions in their N-terminal regions are poorly glycosylated [27]. Matrix-assisted laser desorption ionization time-of-flight MS (MALDI-TOF-MS) analysis showed that all glycosylation intermediates accumulate except for the steps mediated by the FT85 β-Gal-Tase and the α-Fuc-Tase [26]. These isoforms have proved invaluable as substrates for purifying and characterizing the pathway enzymes, and indicate that the PH, GlcNAc-Tase and α-Gal-Tases are sensitive to the overall structure of Skp1 in vivo.

Normal Skp1 concentrates in the nucleus relative to the cytoplasm in both *Dictyostelium* and animal cells based on immunofluorescence studies [27], consistent both with the presence of proteasomes in this compartment and evidence that an animal E3^{SCF}-Ub ligase escorts its targets from the nucleus to the cytoplasm [52]. Association of both genetic isoforms of *Dictyostelium* Skp1 with the cell periphery and the contractile vacuole is also seen under certain conditions. The poorly glycosylated Skp1 mutants do not accumulate in the nucleus but show other normal localizations. Failure to concentrate in the nucleus was due to the glycosylation deficit, as mislocalization was also seen when glycosylation was inhibited by (i) introducing a P143A point mutation at the glycosylation site, (ii) inhibiting prolyl hydroxylation using inhibitors and (iii) expressing Skp1 in prespore cells where it is not glycosylated [27]. The mechanism by which glycosylation promotes nuclear concentration is not known.

At the organismic level, Skp1 overexpression interferes with terminal culmination, resulting in the appearance of fruiting bodies whose spores fail to rise to the apex of the stalk and exhibit a round shape defect [C. West et al., unpublished data]. A similar 'half-mast' phenotype has been described for cullin-1-null cells [21], suggesting that the effect of Skp1^{oe} is mediated via the SCF complex. This phenotype has also been observed in a mutant that does not express SP85, a spore coat structural protein that is thought to regulate a checkpoint for terminal sporulation to ensure that spore differentiation is properly coordinated with morphogenesis [53]. Thus SP85 checkpoint signaling appears to involve the SCF complex, possibly controlling the turnover of an unknown intracellular target protein. This may be analogous to SCF regulation of outside-in integrin-mediated signaling in animal cells [54]. Interestingly, overexpressed Skp1 mutants that are poorly glycosylated do not have this phenotype, indicating that glycosylation may be important for the dominant-negative effect of Skp1 on terminal culmination.

An independent approach to study glycan function has been to target the glycosylation pathway directly. FT85 (β-Gal-Tase/α-Fuc-Tase) mutants, which accumulate Skp1-GlcNAc [47], have a smaller average cell size and grow to higher maximal cell density [46]. Clearly cytoplasmic glycosylation, possibly of Skp1, is important for general cellular regulation, but its mechanism of action remains to be identified.

With its five sugars, the Skp1 glycan contains enough information to be recognized by a hypothetical receptor. For example, the glycan might directly interact with a nuclear importin, or bind to a nuclear receptor or to known carbohydrate binding proteins present in the cytoplasm. Numerous such proteins have been identified, and with some notable exceptions [55, 56], their natural ligands have not been defined [48, 57–59]. A search for potential Skp1-glycan receptors must await isolation of sufficient amounts of native or synthetic material for biochemical studies. Mapping of the glycan attachment site onto the crystal structure of Skp1 in the human SCF complex suggests that it would be oriented away from the contact surfaces with cullin-1 and the F-box protein (fig. 1). An argument against a role for the glycan in receptor recognition comes from the finding that Skp1 in a GDP-Fuc synthesis mutant, which appears to express only Skp1-GlcNAc-Gal, also concentrates in the nucleus [27]. If a receptor is involved, the core disaccharide is the primary determinant for recognition.

Alternatively, glycosylation may enable nuclear concentration by an indirect mechanism. The PH, GlcNAc-Tase and both α-Gal-Tase activities are sensitive to Skp1 structure in vivo and possibly in vitro, and thus can be inferred to sample the structure of Skp1 prior to modifying it. Therefore, full glycosylation signifies that the features of Skp1 structure recognized by these enzymes are present in the finally modified protein, features which might be required for Skp1 to engage in complex formation and/or concentrate in the nucleus. There is evidence that prolyl hydroxylases of the rER have this kind of quality-control function [60], and the UDP-Glc:N-glycan α-Glc-Tase of the rER operates in the reverse mode by attaching Glc to a terminal Man-residue of an N-glycan if it detects that

the attached protein is not folded [61]. Formation of the Ctf13p/Skp1 heterodimer in yeast may be dependent on the Hsp90 chaperone [13], suggesting that Skp1 complex formation is in fact monitored or assisted. Another possibility is that modification of Pro143 stabilizes its cis or trans isomer, analogous to the role of phosphorylation in certain cell cycle proteins [62]. Evidence for two pools of Skp1 that appear to be differentially modified by an outer α-Gal is consistent with terminal glycosylation being rate limiting [20]. This viewpoint suggests a model in which each glycosylation intermediate directs the Skp1 molecule to the next enzyme in the pathway for further quality control. The pathway is directional, because once modified, Skp1 is no longer a substrate for the previous enzyme. This is a variation of a quality control model to explain how incompletely or improperly folded proteins are not further processed in the rER [63], based on recognition by the UDP-Glc:N-glycan α-Glc-Tase referred to above.

General significance of the *Dictyostelium* Skp1 modification

4-Prolyl hydroxylation has long been thought to occur only in the rER on secreted proteins, including animal collagens where it stabilizes the triple helix, and plant and algal extensins where it serves as an attachment site for glycans [64, 65]. Prolyl hydroxylation is now known to occur in the cytoplasm in association with ubiquitination. O_2-dependent prolyl hydroxylation mediates recognition of HIF-1α by the $E3^{VHL}$-Ub ligase, resulting in its polyubiquitination and degradation [6, 33], and such a mechanism might explain why Skp1 is not hydroxylated during multicellular development when cells are crowded. Unlike for HIF-1α, the HyPro residue in *Dictyostelium* Skp1 is modified by GlcNAc, which potentially represents a mechanism to block hydroxylation-dependent recognition by an E3-Ub ligase.

Like prolyl hydroxylation, complex glycosylation is also traditionally associated with proteins passing through the secretory pathway. How general is complex glycosylation, like that seen on Skp1, in the cytoplasm and nucleus? Certainly the cytoplasm is no stranger to complex glycosylation per se, as evidenced by the synthesis [66] and accumulation of glycogen, a Glc polymer, and the initiation of glycolipid and polysaccharide synthesis on the cytoplasmic side of many membranes [20] (see fig. 5). However, nearly all of these glycoconjugates are destined to translocate to the cell exterior. For example, in prokaryotes, precursors for the peptidoglycan, lipooligosaccharide and capsule are initially glycosylated at the cytoplasmic surface of the plasma membrane and subsequently flipped to the cell surface [67, 68]. The polysaccharides cellulose, hyaluronan and chondroitin are translocated across the plasma membrane as they are polymerized [49, 69]. In eukaryotes, precursors for N-glycan biosynthesis and early intermediates in GPI-anchor synthesis and glycolipid synthesis are similarly flipped into the rER or Golgi interiors, and cellulose, chitin and hyaluronan are synthesized as in prokaryotes (fig. 5) [20]. Even *Chlorella* algal, viral capsid proteins, which appear to be glycosylated in the cytoplasm [70], ultimately have extracellular fates after cell lysis.

Dictyostelium Skp1 is a lead example of complex glycosylation that occurs in the cytoplasm on a protein that remains in this or the nuclear compartment. There are numerous other candidate cytoplasmic glycoproteins based on the results of indirect studies that deserve confirmation, including the phosphoglucomutase family protein parafusin, cytokeratins, and MARCKS protein [20]. The possibility that in animal cells these proteins, including Skp1, are subject to complex glycosylation is supported by the existence in animal genomes of a family GT2-like sequence that is distantly related to that of N-terminal domain of FT85. In addition, the phylogenetic analysis of eukarytotic GnT51-like sequences in figure 4 suggests that since the predicted cytoplasmic and Golgi sequences lie in separate clades, related GlcNAc-Tases are likely to exist in other lower eukaryotes. However, the structures that might occur elsewhere may differ, as O-glycosylation carried out in the Golgi is remarkably varied both evolutionarily and developmentally.

A special class of glycoconjugates includes those that are secreted and secondarily translocated to the cytoplasm or nucleus. An example is hyaluronan, which can be localized immunocytochemically in the cytoplasm and nucleus together with specific receptors [71]. A second example includes glycoproteins subject to ER-associated degradation. It is thought that improperly folded glycoproteins are transported back out of the rER into the cytoplasm [72], where they are either deglycosylated [73] and polyubiquitinated, or polyubiqutinated based on recognition of their high-mannose N-glycans [5], and subsequently degraded in the 26S proteasome. This rER exit pathway may also allow entry of cholera toxin into the cell, implying that glycoproteins derived from the secretory pathway might avoid degradation and function in the cytoplasm [72]. This might be the avenue by which a putative glycosylated prion protein [55] and the N-glycosylated nuclear lectin CB70 [74] enter the nucleus, and N-glycosylated glycoproteins accumulate in the mitochondrion [75]. A variation on this concept has been suggested to mediate signaling by certain growth factors and cytokines. For example, there is evidence that the connective tissue growth factor glycoprotein directly signals new gene expression in the nucleus after endocytosis and exit from perinuclear endosome-related vesicles [76]. Instead of the cytoplasm being a transient station for glycoconjugates, it is the end compartment in this model. If

correct, these newly arrived glycoproteins are susceptible to remodeling by GTases like those that construct the outer residues of the Skp1 glycan, because the Skp1 core glycan (blood group H, type 1) is similar to structures produced in the Golgi of animal cells.

As outlined in the first section, Skp1 participates in many pathways. While diverse functions may be met by multiple Skp1-like gene products in lower animals and plants, the glyco- and phosphomodifications described here may generate additional structural isoforms in specific organisms. However, posttranslational modifications may play a more global role in regulating complex formation, compartmentalization or catalytic cycles. Although the simplest model is that the pentasaccharide contributes functionality via a specific receptor, other models to be considered include (i) steric interference of other interactions, (ii) effects on Skp1 conformation and (iii) perhaps most interestingly, a role for the modifying enzymes in monitoring cellular metabolism and/or Skp1 folding. For example, the prolyl hydroxylase and GTases might withhold Skp1 when (i) concentrations of precursors of hydroxylation and glycosylation are suboptimal or when inhibitory nucleotide concentrations [32, 34] are too high, or (ii) Skp1 is not properly folded or assembled into complexes. Once these constraints are overcome, the glycan structure may be completed. In this model, the glycan may serve no further function other than to render Skp1 no longer subject to quality control by its modification enzymes.

Acknowledgements. We are grateful to Hanke van der Wel for aligning the sequences and to Eric Gaucher for performing the phylogenetic analysis. Studies in the author's lab have been supported by NIH grant GM-37539 and the Florida Division of the American Cancer Society.

1 Pickart C. M. (2001) Mechanisms underlying ubiquitination. Annu. Rev. Biochem. **70:** 503–533
2 Deshaies R. J. (1999) SCF and cullin/RING H2-based ubiquitin ligases. Annu. Rev. Cell Dev. Biol. **15:** 435–467
3 Zheng N., Schulman B. A., Song L., Miller J. J., Jeffrey P. D., Wang P. et al. (2002) Structure of the Cul1-Rbx1-Skp1-F box^{Skp2} SCF ubiquitin ligase complex. Nature **416:** 703–709
4 Winston J. T., Koepp D. M., Zhu C., Elledge S. J. and Harper J. W. (1999) A family of mammalian F-box proteins. Curr. Biol. **9:** 1180–1182
5 Yoshida Y., Chiba T., Tokunaga F., Kawasaki H., Iwai K., Suzuki T. et al. (2002) E3 ubiquitin ligase that recognizes sugar chains. Nature **418:** 438–442
6 Min J. H., Yang H., Ivan M., Gertler F., Kaelin W. G. and Pavletich N. P. (2002) Structure of an HIF-1alpha -pVHL complex: hydroxyproline recognition in signaling. Science **296:** 1886–1889
7 Gagne J. M., Downes B. P., Shin S. H., Durski A. M. and Vierstra R. D. (2002) The F-box subunit of the SCF E3 complex is encoded by a diverse superfamily of genes in *Arabidopsis.* Proc. Natl. Acad. Sci. USA **99:** 11529–11524
8 Galan J. M. and Peter M. (1999) Ubiquitin-dependent degradation of multiple F-box proteins by an autocatalytic mechanism. Proc. Natl. Acad. Sci. USA **96:** 9124–9129
9 Verma R., Chen S., Feldman R., Schieltz D., Yates J., Dohmen J. et al. (2000) Proteasomal proteomics: identification of nucleotide-sensitive proteasome-interacting proteins by mass spectrometric anaysis of affinity-purified proteasomes. Mol. Biol. Cell **11:** 3425–3439
10 Verma R., McDonald H., Yates J. R. and Deshaies R. J. (2001) Selective degradation of ubiquitinated Sic1 by purified 26S proteasome yields active S-phase cyclin-cdk. Mol. Cell **8:** 439–448
11 Seol J. H., Shevchenko A., Shevchenko A. and Deshaies R. J. (2001) Skp1 forms multiple protein complexes, including RAVE, a regulator of V-ATPase assembly. Nat. Cell Biol. **3:** 384–391
12 Kaplan K. B., Hyman A. A. and Sorger P. K. (1997) Regulating the yeast kinetochore by ubiquitin-dependent degradation and skp1p-mediated phosphorylation. Cell **91:** 491–500
13 Stemman O., Neidig A., Kocher T., Wilm M. and Lechner J. (2002) Hsp90 enables Ctf13p/Skp1p to nucleate the budding yeast kinetochore. Proc. Natl. Acad. Sci. USA **99:** 8585–8590
14 Galan J. M., Wiederkehr A., Seol J. H., Haguenauer-Tsapis R., Deshaies R. J., Riezman H. et al. (2001) Skp1p and the F-box protein Rcy1p form a non-SCF complex involved in recycling of the SNARE Snc1p in yeast. Mol. Cell. Biol. **21:** 3105–3117
15 Farras R., Ferrando A., Jasik J., Kleinow T., Okresz L., Tiburcio A. et al. (2001) SKP1-SnRK protein kinase interactions mediate proteasomal binding of a plant SCF ubiquitin ligase. EMBO J. **20:** 2742–2756
16 Mathias N., Johnson S., Byers B. and Goebl M. (1999) The abundance of cell cycle regulatory protein Cdc4p is controlled by interactions between its F-box and Skp1p. Mol. Cell. Biol. **19:** 1759–1767
17 Yamanaka A., Yada M., Imaki H., Koga M., Ohshima Y. and Nakayama K.-I. (2002) Multiple Skp1-related proteins in *Caenorhabditis elegans*: diverse patterns of interaction with cullins and F-box proteins. Curr. Biol. **12:** 267–275
18 Nayak S., Santiago F. E., Jin H., Lin D., Schedl T. and Kipreos E. T. (2002) The *Caenorhabditis elegans* Skp1-related gene family: diverse functions in cell proliferation, morphogenesis and meiosis, Curr. Biol. **12:** 277–287
19 West C. M., Kozarov E. and Teng-umnuay P. (1997) The cytosolic glycoprotein FP21 of *Dictyostelium discoideum* is encoded by two genes resulting in a polymorphism at a single amino acid position. Gene **200:** 1–10
20 West C. M., van der Wel H. and Gaucher E. A. (2001) Complex glycosylation of Skp1 in *Dictyostelium*: implications for the modification of other eukaryotic cytoplasmic and nuclear proteins. Glycobiology **12:** 17R–27R
21 Mohanty S., Lee S., Yadava N., Dealy M. J., Johnson R. J. and Firtel R. A. (2001) Regulated protein degradation controls PKA function and cell-type differentiation in *Dictyostelium*. Genes Dev. **15:** 1435–1448
22 Kawakami T., Chiba T., Suzuki T., Iwai K., Yamanaka K., Minato N. et al. (2001) NEDD8 recruits E2-ubiquitin to SCF E3 ligase. EMBO J. **20:** 4003–4012
23 Li F. N. and Johnston M. (1997) Grr1 of *Saccharomyces cerevisiae* is connected to the ubiquitin proteolysis machinery through Skp1: coupling glucose sensing to gene expression and the cell cycle. EMBO J. **16:** 5629–5638
24 Henzl M. T., Thalmann I. and Thalmann R. (1998) OCP2 exists as a dimer in the organ of Corti. Hear. Res. **126:** 37–46
25 Ganoth D., Bornstein G., Ko T. K., Larsen B., Tyers M., Pagano M. et al. (2001) The cell-cycle regulatory protein Cks1 is required for SCF(Skp2)-mediated ubiquitinylation of p27. Nat. Cell Biol. **3:** 321–324
26 Teng-umnuay P., Morris H. R., Dell A., Panico M., Paxton T. and West C. M. (1998) The cytoplasmic F-box binding protein Skp1 contains a novel pentasaccharide linked to hydroxyproline in *Dictyostelium*. J. Biol. Chem. **273:** 18242–18249
27 Sassi S., Sweetinburgh M., Erogul J., Zhang P., Teng-umnuay P. and West C. M. (2001) Analysis of Skp1 glycosylation and

nuclear enrichment in *Dictyostelium*. Glycobiology **11:** 283–295
28 Comer F. I. and Hart G. W. (2000) O-glycosylation of nuclear and cytosolic proteins. J. Biol. Chem. **275:** 29179–29182
29 Busch C. and Aktories K. (2000) Microbial toxins and the glycosylation of Rho family GTPases. Curr. Opin. Struct. Biol. **10:** 528–535
30 Kozarov E., van der Wel H., Field M., Gritzali M., Brown R. D. and West C. M. (1995) Characterization of FP21, a cytosolic glycoprotein from *Dictyostelium*. J. Biol. Chem. **270:** 3022–3033
31 Kivirikko K. I. and Myllyla R. (1982) Posttranslational enzymes in the biosynthesis of collagen: intracellular enzymes. Methods Enzymol. **82:** 245–304
32 West C. M., Scott-Ward T., Teng-umnuay P., van der Wel H., Kozarov E. and Huynh A. (1996) Purification and characterization of an α1,2-L-fucosyltransferase, which modifies the cytosolic protein FP21, from the cytosol of *Dictyostelium*. J. Biol. Chem. **271:** 12024–12035
33 Bruick R. K. and McKnight S. L. (2001) A conserved family of prolyl-4-hydroxylases that modify HIF. Science **294:** 1337–1340
34 Teng-umnuay P., van der Wel H. and West C. M. (1999) Identification of a UDP-GlcNAc:Skp1-hydroxyproline GlcNAc-transferase in the cytoplasm of *Dictyostelium*. J. Biol. Chem. **274:** 36392–36402
35 van der Wel H., Morris H., Panico M., Paxton T., Dell A., Kaplan L. et al. (2002) Molecular cloning and expression of a UDP-GlcNAc:hydroxyproline polypeptide GlcNAc-transferase that modifies Skp1 in the cytoplasm of *Dictyostelium*. J. Biol. Chem. **277:** 46328–46337
36 Breton C. and Imberty A. (1999) Structure/function studies of glycosyltransferases. Curr. Opin. Struct. Biol. **9:** 563–571
37 Hagen F. K., Hazes B., Raffo R., deSa D. and Tabak L. A. (1999) Structure-function analysis of the UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase. J. Biol. Chem. **274:** 6797–6803
38 Tarbouriech N., Charnock S. J. and Davies G. J. (2001) Three-dimensional structures of the Mn and Mg dTDP complexes of the family GT-2 glycosyltransferase SpsA: a comparison with related NDP-sugar glycosyltransferases. J. Mol. Biol. **314:** 655–661
39 Hassan H., Reis C. A., Bennett E. P., Mirgorodskaya E., Roepstorff P., Hollingsworth M. A. et al. (2000) The lectin domain of UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase-T4 directs its glycopeptide specificities. J. Biol. Chem. **275:** 38197–38205
40 Henrissat B. and Davies G. J. (2000) Glycoside hydrolases and glycosyltransferases. Families, modules, and implications for genomics. Plant Physiol. **124:** 1515–1519 http://afmb.cnrs-mrs.fr/CAZY/index.html
41 Jung E., Gooley A. A., Packer N. H., Karuso P. and Williams K. L. (1998) Rules for the addition of O-linked N-acetylglucosamine to secreted proteins in *Dictyostelium discoideum* – in vivo studies on glycosylation of mucin MUC1 and MUC2 repeats. Eur. J. Biochem. **253:** 517–524
42 Previato J. O., Sola-Penna M., Agrellos O. A., Jones C., Oeltmann T., Travassos L. R. et al. (1998) Biosynthesis of O-N-acetylglucosamine-linked glycans in *Trypanosoma cruzi*. Characterization of the novel uridine diphospho-N-acetylglucosamine:polypeptide N-acetylglucosaminyltransferase-catalyzing formation of N-acetylglucosamine alpha1→O-threonine. J. Biol. Chem. **273:** 14982–14988
43 West C. M., van der Wel H., Metcalf T., Kaplan L. and Gaucher E. (2002) Identification of a Golgi-associated UDP-GlcNAc: polypeptide mucin-type alpha-N-acetylglucosaminyltransferase that modifies cell surface proteins in *Dictyostelium*. Glycobiology **12:** abstract 455
44 Li G., Alexander H., Schneider N. and Alexander S. (2000) Molecular basis for resistance to the anticancer drug cisplatin in *Dictyostelium*. Microbiology **146:** 2219–2227
45 Wrabl J. O. and Grishin N. V. (2001) Homology between O-linked GlcNAc transferases and proteins of the glycogen phosphorylase superfamily. J. Mol. Biol. **314:** 365–374
46 van der Wel H., Morris H. R., Panico M., Paxton T., Dell A., Thomson J. M. et al. (2001) A non-Golgi α1,2-fucosyl-transferase that modifies Skp1 in the cytoplasm of *Dictyostelium*. J. Biol. Chem. **276:** 33952–33963
47 van der Wel H., Fisher S. Z. and West C. M. (2002) A bifunctional diglycosyltransferase forms the Fucα1,2Galβ,3-disaccharide on Skp1 in the cytoplasm of *Dictyostelium*. J. Biol. Chem. **277:** 46527–46534
48 Erdos G. W. and Whitaker D. (1983) Failure to detect immunocytochemically reactive endogenous lectin on the cell surface of *Dictyostelium discoideum*. J. Cell Biol. **97:** 993–1000
49 deAngelis P. L. (2002) Microbial glycosaminoglycan glycosyltransferases. Glycobiology **12:** 9R–16R
50 Varki A. (1993) Biological roles of oligosaccharides: all of the theories are correct. Glycobiology **3:** 97–130
51 Van den Steen P., Rudd P. M., Dwek R. A. and Opdenakker G. (1998) Concepts and principles of O-linked glycosylation. Crit. Rev. Biochem. Mol. Biol. **33:** 151–208
52 Fukuchi M., Imamura T., Chiba T., Ebisawa T., Kawabata M., Tanaka K. et al. (2001) Ligand-dependent degradation of Smad3 by a ubiquitin ligase complex of ROC1 and associated proteins. Mol. Biol. Cell. **12:** 1431–1443
53 West C. M., Zhang P., McGlynn A. C. and Kaplan L. (2002) Outside-In signaling of cellulose synthesis by a spore coat protein in *Dictyostelium*. Euk. Cell **1:** 281–292
54 Carrano A. C. and Pagano M. (2001) Role of the F-box protein Skp2 in adhesion-dependent cell cycle progression. J. Cell Biol. **153:** 1381–1389
55 Rybner C., Finel-Szermanski S., Felin M., Sahraoui T., Rousseau C., Fournier J. G. et al. (2002) The cellular prion protein: a new partner of the lectin CBP70 in the nucleus of NB4 human promyelocytic leukemia cells. J. Cell. Biochem. **84:** 408–419
56 Wang J., Stuckey J. A., Wishart M. J. and Dixon J. E. (2002) A unique carbohydrate binding domain targets the Lafora disease phosphatase to glycogen. J. Biol. Chem. **277:** 2377–2380
57 Park J. W., Voss P. G., Grabski S., Wang J. L. and Patterson R. J. (2001) Association of galectin-1 and galectin-3 with Gemin4 in complexes containing the SMN protein. Nucleic. Acids Res. **27:** 3595–3602
58 Yu F., Finley R. L., Raz A. and Kim H.-R. C. (2002) Galectin-3 translocates to the perinuclear membranes and inhibits cytochrome c release from mitochondria. J. Biol. Chem. **277:** 15819–15827
59 Barre A., Van Damme E. J., Peumans W. J. and Rouge P. (1999) Homology modelling of the core domain of the endogenous lectin comitin: structural basis for its mannose-binding specificity. Plant Mol. Biol. **39:** 969–978
60 Lamande S. R. and Bateman J. F. (1999) Procollagen folding and assembly: the role of endoplasmic reticulum enzymes and molecular chaperones. Sem. Cell Dev. Biol. **10:** 455-464
61 Parodi A. J. (2000) Protein glucosylation and its role in protein folding. Annu. Rev. Biochem. **69:** 69–93
62 Yaffe M. B., Schutkowski M., Shen M., Zhou X. Z., Stukenberg P. T., Rahfeld J. U. et. al. (1997) Sequence-specific and phosphorylation-dependent proline isomerization: a potential mitotic regulatory mechanism. Science **278:** 1957–1960
63 Helenius A. and Aebi M. (2001) Intracellular functions of N-linked glycans. Science **291:** 2364–2369
64 Kieliszewski M. J. and Shpak E. (2001) Synthetic genes for the elucidation of glycosylation codes for arabinogalactan-proteins and other hydroxyproline-rich glycoproteins. Cell. Mol. Life Sci. **58:** 1386–1398
65 Kivirikko K. I. and Pihlajaniemi T. (1998) Collagen hydroxylases and the protein disulfide isomerase subunit of prolyl-4-hydroxylases. Adv. Enzymol. Relat. Areas Mol. Biol. **72:** 325–398

66 Gibbons B. J., Roach P. J. and Hurley T. D. (2002) Crystal structure of the autocatalytic initiator of glycogen biosynthesis, glycogenin. J. Mol. Biol. **319:** 463–477
67 Whitfield C. and Roberts I. S. (1999) Structure, assembly and regulation of expression of capsules in *Escherichia coli*. Mol. Microbiol. **31:** 1307–1319
68 van Heijenoort J. (2001) Formation of the glycan chains in the synthesis of bacterial peptidoglycan. Glycobiology **11:** 25R–36R
69 Williamson R. E., Burn J. E. and Hocart C. H. (2001) Cellulose synthesis: mutational analysis and genomic perspectives using *Arabidopsis thaliana*. Cell. Mol. Life Sci. **58:** 1475–1490
70 Graves M. V., Bernadt C. T., Cerny R. and van Etten J. L. (2001) Molecular and genetic evidence for a virus-encoded glyosyltransferase involved in protein glycosylation. Virology **285:** 332–345
71 Lee J. Y. and Spicer A. P. (2000) Hyaluronan: a multifunctional, megadalton, stealth molecule. Curr. Opin. Cell Biol. **12:** 581–586
72 Tsai B., Ye Y. and Rapaport T. A. (2002) Retro-translocation of proteins from the endoplasmic reticulum into the cytosol. Nat. Rev. Mol. Cell Biol. **3:** 246–255
73 Suzuki T., Park H. and Lennarz W. J. (2002) Cytoplasmic peptide:N-glycanase (PNGase) in eukaryotic cells: occurrence, primary structure and potential functions. FASEB J. **16:** 635–641
74 Lefebvre T., Cieniewski C., Lemoine J., Guerardel Y., Leroy Y., Zanetta J. P. et al. (2001) Identification of N-acetyl-D-glucosamine-specific lectins from rat liver cytosolic and nuclear compartments as heat-shock proteins. Biochem. J. **360:** 179– 188
75 Chandra N. C., Spiro M. J. and Spiro R. G. (1998) Identification of a glycoprotein from rat liver mitochondrial inner membrane and demonstration of its origin in the endoplasmic reticulum. J. Biol. Chem. **273:** 19715–19721
76 Wahab A. A., Brinkman H. and Mason R. M. (2001) Uptake and intracellular transport of the connective tissue growth factor: a potential mode of action. Biochem. J. **359:** 89–97
77 Swofford D. L. (2000) PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods). Version 4. Sinauer Associates, Sunderland, MA
78 Gupta R. S. (2001) The branching order and phylogenetic placement of species from completed bacterial genomes, based on conserved indels found in various proteins. Int. Microbiol. **4:** 187–202