



HHS Public Access

Author manuscript

Curr Opin Struct Biol. Author manuscript; available in PMC 2020 June 01.

Published in final edited form as:

Curr Opin Struct Biol. 2019 June ; 56: 204–212. doi:10.1016/j.sbi.2019.03.031.

Nucleocytoplasmic O-glycosylation in protists

Christopher M. West^{1,2,3,4}Hyun W. Kim¹

¹Department of Biochemistry & Molecular Biology, University of Georgia, Athens, GA 30602 USA

²Center for Tropical and Emerging Global Diseases, University of Georgia, Athens, GA 30602 USA

³Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602 USA

Abstract

O-glycosylation is an increasingly recognized modification of intracellular proteins in all kingdoms of life, and its occurrence in protists has been investigated to understand its evolution and its roles in the virulence of unicellular pathogens. We focus here on two kinds of glycoregulation found in unicellular eukaryotes: one is a simple O-fucose modification of dozens if not hundreds of Ser/Thr-rich proteins, and the other a complex pentasaccharide devoted to a single protein associated with oxygen sensing and the assembly of polyubiquitin chains. These modifications are not required for life but contingently modulate biological processes in the social amoeba *Dictyostelium* and the human pathogen *Toxoplasma gondii*, and likely occur in diverse unicellular protists. O-glycosylation that is co-localized in the cytoplasm allows for glycoregulation over the entire life of the protein, contrary to the secretory pathway where glycosylation usually occurs before its delivery to its site of function. Here we interpret cellular roles of nucleocytoplasmic glycans in terms of current evidence for their effects on the conformation and dynamics of protist proteins, to serve as a guide for future studies to examine their broader significance.

Introduction

Glycosylation is a prominent modification of proteins exposed to the extracellular environment – a theme that prevails throughout prokaryotes and eukaryotes. Though evidently a less common occurrence, O-glycosylated proteins are also found in the cytoplasm and nucleus [1]. A well-characterized example is the modification by O- β GlcNAc of thousands of nucleocytoplasmic proteins of animal and higher plant cells, at Ser and/or Thr residues, which renders a wide range of mostly protein specific effects [2, 3]. Less well known are the examples of glycoproteins that function in the nucleus or cytoplasm of prokaryotes and protists. The stepwise nature of O-glycosylation has important implications for these proteins as it allows for modulation during the life of the protein. This contrasts with conventional O- and N-glycosylation in the secretory pathway, where for the most part glycosylation occurs before delivery of the protein to its final site of function within an

⁴to whom correspondence should be addressed: Dept. of Biochemistry & Molecular Biology, 120 E. Green St., Davison Life Sciences – B129, University of Georgia, Athens GA 30602 USA, telephone +1-706-542-4259, westcm@uga.edu.

organelle, at the cell surface, or extracellularly. Unicellular organisms exhibit myriad specializations befitting their diverse lifestyles, providing fertile opportunities for the evolution of novel regulatory mechanisms such as glycosylation.

Evolutionarily, glycosylation appears to have originated in the cytoplasm of prokaryotes, which accordingly evolved elaborate mechanisms to translocate glycolipids, lipid-linked precursors for periplasmic glycosylation, and most product glycoproteins, across the plasma membrane. Some prokaryotic glycoproteins also remain in and function in the cytoplasm [4, 5]. Eukaryotes specialized the process by transferring much of the enzymology to the lumen of the secretory pathway, lessening the need for complex translocation apparatus but restricting the opportunity for dynamic glycosylation. Evidence for cytoplasmic glycosylation in extant protists has been recently explored both to gain insight regarding evolution of the compartmentalization of glycosylation across the kingdoms of life, and to search for novel targets for new control strategies toward pathogenic protists.

The protist kingdom harbors unicellular life forms that have evolved an incredible range of free-living, commensal and parasitic life styles. Studies in a model organism, the cellular slime mold *Dictyostelium discoideum*, and pathogenic agents including *Toxoplasma gondii*, trypanosomes, and oomycetes, are the primary sources of data for this report. *T. gondii* is a highly successful apicomplexan parasite that can infect all warm-blooded animal cells and a substantial fraction of the human population is considered to be latently infected [6]. Acute infections cause toxoplasmosis and progress to a latent phase, which poses a chronic risk to encephalitis and blindness upon immune suppression and for which there is no treatment. The oomycetes include species of *Phytophthora* and *Pythium* that are enormously expensive plant pathogens that affect human health through the food supply [7]. Finally, protists are important not only for the many other pathogens represented but as well for comprising a large fraction of the biomass on the planet through diatoms, green algae, and many other life forms that strongly influence carbon flow in the ocean.

The purpose of this review is to enumerate examples of nucleocytoplasmic glycoproteins in protists with a focus on how they are glycosylated and what is known about the structural and cellular consequences. Two examples that represent diverse themes are highlighted. We compare with instances of nucleocytoplasmic glycosylation in non-protists and, given limited research in this area, we draw on structural effects of related O-glycans from other compartments. Rather than directly serving as recognition determinants for other proteins, current evidence best supports the glycans functioning via effects on carrier protein conformation and dynamics that may then indirectly influence molecular interactions.

O-Fucosylation of nucleocytoplasmic proteins in protists

Two types of nucleocytoplasmic glycosylation are evidently conserved across diverse protist genera. The first is the simple modification of the hydroxyl groups of Ser and Thr residues with α -L-Fucose, which Bandini, Samuelson and Costello initially detected in *T. gondii* using *Aleuria aurantia* lectin (AAL) [8••]. Thirty-three different nuclear and cytoplasmic proteins, which include numerous putative nucleoporins, mRNA processing enzymes, transcription regulators, and signaling proteins, were confirmed to be directly fucosylated

using MS/MS methods. The fucosylated residues were found on isolated Ser or Thr, but were most abundant in and often clustered in tracts rich in Ser and Thr that are likely to lack secondary structure. Many sites were variably modified, suggesting that the carrier proteins are diversified by this modification. Immunolocalization studies using AAL show that many of the fucosylated proteins are found in assemblies that subtend the nuclear envelope possibly in register with nuclear pores. The linkage of α Fuc to these proteins is catalyzed by an O-fucosyltransferase (OFT) [9•] that was previously predicted, based on sequence similarity, to be an O- β GlcNAc-transferase (OGT), the enzyme responsible for the extensive O-GlcNAcylation of animal and higher plant nucleocytoplasmic proteins. Gene disruption studies show that OFT is important for optimal growth of *T. gondii* in a fibroblast monolayer growth model [9•]. A similar OFT was recently described in *Arabidopsis* where it modulates the function of a nuclear transcriptional regulator, possibly in opposition to the effect of nearby O-GlcNAcylation [10].

Homologs of the OFT gene are present in many protists and can be evolutionarily traced back to the prokaryotic kingdom, where an ancient gene duplication may have allowed for the divergence of the OGT (Secret Agent, or SEC) and OFT (Spindly, or SPY) lineages from a common ancestor [11]. The high degree of conservation in both the N-terminal TPR repeat and C-terminal catalytic domains suggests a conserved mechanism of regulation and action. In accordance with the phylogenetic analysis and some experimental evidence [11, 12], it is likely that O-GlcNAc and O-Fuc will be found in numerous different protists. Recent studies document the role of the OFT/Spy-dependent O-fucosylation of nucleocytoplasmic protein homologs in another branch of protist evolution represented by the social amoeba *D. discoideum* (van der Wel *et al.*, unpublished data), confirming the widespread distribution of nucleocytoplasmic O-Fuc. The possibility that O-Fuc is dynamic like O-GlcNAcylation remains to be investigated. Extensive monosaccharide modifications of nucleocytoplasmic proteins in regions rich in Ser, Thr and Pro are an emerging theme across all kingdoms of life. A third monosaccharide, O- α Man, was recently reported on numerous nucleocytoplasmic proteins of yeast, though a responsible nucleocytoplasmic glycosyltransferase has not been identified [13], and hints for nuclear O- α GalNAc in mammalian cells have recently appeared [14, 15]. Taken together, these findings suggest the importance of monosaccharide modifications of nucleocytoplasmic proteins throughout eukaryotic evolution. While the reason that different organisms adopted distinct neutral sugars remains to be explained, the finding of both types (O- β GlcNAc and O- α Fuc) in the same cells of higher plants suggests distinct functions for these modifications [10].

A structural effect of a single O- α -L-Fuc was previously characterized based on an NMR solution structure of the extracellular PMP-C protease inhibitor [16]. An α Fuc appended to Thr9 was stably oriented via hydrophobic interactions and hydrogen bonds with Thr16 and Arg18 (Fig. 1a). Less understood is how, based on hydrogen-deuterium exchange and thermal stability studies, this local conformation change resulted in decreased flexibility and increased stabilization of the protein fold. In an analysis of an EGF-domain, a Thr-linked α Fuc residue resulted in a modest increase in Ca²⁺-binding, though a correlative effect on structure was not detected [17]. In an X-ray crystallographic analysis of the interface between the extracellular EGF-domains of Notch and Delta-like 4 [18], O-Fuc directly contributed to the protein-protein interaction. These examples are from disulfide stabilized

domains of extracellular proteins so might not, however, be representative of effects on intrinsically disordered regions of nucleocytoplasmic proteins.

Effects of monosaccharides on intracellular proteins have relied extensively on model peptides. Attachment of O- β GlcNAc to Thr residues in the carboxy-terminal repeat domain of RNA polymerase II, which dynamically oscillates with the transcription cycle, induces a transition from a random coil to a β -turn-like structure (Fig. 1b), based on CD, MD and NMR analysis of peptides [19]. Related conformational changes (Fig. 1c) were reported for a peptide from the estrogen receptor- β [20]. In peptides from the tau protein, whose aggregation is associated with Alzheimer's disease, O-GlcNAcylation was found to oppose polyproline helix formation [21]. These effects were largely distinctive from effects of phosphorylation (Fig. 1c). Installation of O-GlcNAc near the N-terminus of another model peptide stabilized α -helix formation but destabilized helix formation if positioned internally or at the C-terminus, mimicking effects of phosphorylation [22]. Addition of a single O- β GlcNAc within the turn of a helix-turn-helix of another model peptide decreased the rate of fibrillization without evidently altering its average structure or stability [23]. In an NMR study of a capped Thr-hydroxyproline dipeptide from Skp1 (see below), substitution of the hydroxyproline with O- α GlcNAc promoted the *trans*-isomer of the internal peptide bond [24]. Although studies are few in number, these *in vitro* examples document the many different *cis*-effects by which simple neutral monosaccharides can influence the intrinsic behavior of peptide structures.

Examples of other peptides and domains modified by single sugars, taken from extracellular proteins, reinforce this concept. Substitution of mucin-type peptides with O- α GalNAc can result in restricted rotation and flexibility around adjacent peptide bonds and orientation of the sugar owing to hydrogen-bonding of the N-acetyl moiety to the peptide backbone [25–27], as highlighted in an example of a multiply substituted glycopeptide (Fig. 1f). In another context, O- α GalNAc on Thr promoted the *trans*-isomer of the peptide bond with Pro [30]. A series of papers reported the effect of O- α Man on peptides and domains from extracellular yeast proteins, though the findings may be pertinent based on the recent evidence for cytoplasmic O-mannosylation in yeast [13]. Modification of Ser/Thr rich peptides with multiple α Man residues invoked a constraint on peptide bond rotations that was specific to this sugar owing to the axial configuration of its C2-hydroxyl group, and correlated with resistance to protease digestion [31•]. Examination of a disulfide-stabilized cellulose-binding domain showed that a single α Man mediates intramolecular contacts (Fig. 1d) whose effects propagated to distant regions of the polypeptide resulting in increased stability and resistance to proteolysis [32]. Thus, an observed resistance to protease digestion was attributed to a conformational effect rather than steric interference. Conversely, a crystal structure of E-cadherin [33] revealed multiple solvent exposed residues of O- α Man with no evidence for an effect on the underlying peptide β -strand organization (Fig. 1e). Nevertheless, the Man occupies space which may sterically hinder motion of the backbone even in the absence of specific interactions.

While more studies are needed to understand how effects on peptide organization manifest at the level of entire intracellular proteins, these findings show remarkable consequences of single neutral hexoses that are sugar and amino acid-context specific, and likely relevant to

physiological protein function. These *cis* effects do not rely on the involvement of carbohydrate recognition mechanisms as represented by lectins and other carbohydrate binding proteins acting in *trans*. The exocyclic methyl moiety, which is unique among common sugars, suggests specialized roles for O- α Fuc. Thus, it can be anticipated, for the O-fucosylated proteins characterized in *T. gondii*, that many interesting protein specific functions will emerge from future studies. Furthermore, the frequent occurrence in polySer/Thr tracts invites speculation that polymer avidity effects might concentrate the α -fucosylated proteins in the nuclear assemblies via sugar-sugar interactions [e.g., 34] or by promoting liquid-liquid phase separation [35, 36], a process that can be regulated by known posttranslational modifications including phosphorylation, methylation and Sumoylation.

Hydroxyproline-dependent O-glycosylation of protist Skp1

The second example is a complex linear O-glycan, discovered by West to be present only in protists [37], that consists of a core trisaccharide equivalent to the type 1 blood group H trisaccharide and which is further modified at its non-reducing end by a species-specific disaccharide (Figure 2a). The glycan is found on Skp1, which is best known as a binding partner of F-box proteins (FBPs) that mediates docking to cullin-1 in Skp1/cullin-1/F-box protein (SCF)-type E3 polyubiquitin ligases. Skp1's flexible C-terminal region of ~70 amino acids can associate with 20–300 different FBPs (Figure 2b), many of which are substrate receptors for polyubiquitination, in yeast, higher plants, and animals [38]. The glycosylation of Skp1 was discovered in *Dictyostelium* as an unusual nucleocytoplasmic protein labeled with [3 H]Fuc. Subsequent studies, based on mass spectrometry, exoglycosidase sensitivity, characterization of glycosyltransferase specificities, and finally NMR, established the structure of the glycan as Gal α 1,3Gal α 1,3Fuca1,2Gal β 1,3GlcNAc α 1-, linked to 4-hydroxyproline at residue 143 of the 162-residue Skp1 polypeptide [41●●]. Glycosylation of hydroxyproline is common in the plant secretory pathway but the reducing terminal sugar is either Gal or Ara [42]. MD simulations together with solution NMR studies support a model in which the Skp1 pentasaccharide forms a relatively stable conformation with <15% rotational freedom around each glycosidic linkage (Figure 2a) [41●●].

Identification and characterization of each of the responsible glycosyltransferases revealed that glycosylation is rendered by a series of cytoplasmically-localized sugar nucleotide-dependent enzymes that are evolutionarily related to bacterial cytoplasmic or eukaryotic Golgi glycosyltransferases [37]. There is no evidence for glycosyltransferases forming a single enzymatic complex, however, gene fusions of successive glycosyltransferases occur in different protists allowing for processive addition of sugars. All evidence points to Skp1 being the only glycoprotein target, according to radiotracer and biochemical complementation studies, and genetic interactions with Skp1 mutants that affect oxygen-sensing (see below) [43].

Skp1 from the unrelated *T. gondii* was found also to be modified by a linear pentasaccharide with conserved glycosidic linkages [44●] though, interestingly, the fourth sugar is α Glc rather than α Gal [45]. Furthermore, the glycosyltransferases that mediate the addition of the final two sugars in *T. gondii* are distinct from the single enzyme that completes assembly of the *Dictyostelium* glycan. This remarkable replacement in amoebozoa attests to an

advantage for the complete pentasaccharide. Recently, recombinant glycosyltransferases from the plant pathogen *Pythium ultimum*, which causes root and stem rot of important crops, were found to have activities comparable to those of *T. gondii* homologs (van der Wel *et al.*, submitted; Mandalasi *et al.*, unpublished data). These examples confirm the widespread occurrence of Skp1 glycosylation across protist phylogeny. There is no evidence for Skp1 glycosylation outside of protists, save for possible examples in the fungal kingdom, suggesting that this modification is selectively advantageous for unicellular life.

The attachment of the first sugar, α GlcNAc, requires the hydroxylation of Skp1 Pro143 (*Dictyostelium*), which is contingent on the availability of sufficient levels of environmental O₂ based on *in vitro* and *in vivo* studies [37]. The enzyme that hydroxylates Skp1, PhyA, is the likely protist ortholog of PHD2, the O₂ sensor that regulates the stability of Hypoxia Inducible Factor- α (HIF α) in animal cells [46]. The 4(*trans*)-hydroxylation of two Pro residues in HIF α induces an endo-to-exo transition in the pyrrolidine ring that, together with the hydroxyl moiety *per se*, comprises a structure that is not glycosylated but instead forms a degron that is recognized by the Von Hippel-Lindau (VHL) protein, a substrate receptor for the VHL class of E3 polyubiquitin ligases. PhyA is also an O₂ sensor in *Dictyostelium* and *T. gondii*, based on effects of disrupting Skp1 modification on the O₂-dependence of development [37] and fibroblast monolayer infection, respectively [47]. The same hydroxylation event occurs on Skp1 Pro as on HIF α Pro [48], but the genetic evidence indicates that glycosylation is required to render the full benefit of hydroxylation for O₂-sensing at the cellular level [49].

Glycosylation has substantial effects on Skp1 conformation. Addition of α GlcNAc to HO-Skp1 induced increased α -helical content based on circular dichroism [50]. The increased α -helical content was mapped by NMR [51] to the region that is C-terminal to the Pro143 glycosylation site which forms helix-8 when complexed with some FBPs (Figures 2b-d) [38]. Full glycosylation resulted in additional changes in the organization of the predominantly disordered region surrounding the Pro143 glycosylation site based on chemical shift perturbation analysis [51], which was associated with an overall shape change of the Skp1 homodimer based on small angle X-ray scattering studies [50]. Thus, peripheral as well as core sugars of the glycan appear to be important for its function, consistent with the genetic analyses.

NMR studies on the full-length pentasaccharide form of Skp1 (denoted GGFGGn-Skp1), in which individual sugars were ¹³C-labeled, showed similar motions (τ_c values inferred from R_1 and R_2 measurements) for the first and last sugars, consistent with evidence that the pentasaccharide is organized as a single unit with motions relative to well folded N-terminal domain deriving from disorder in the local polypeptide chain itself [41●●]. All-atom MD simulations of monomeric GGFGGn-Skp1 yielded several 250 ns trajectories whose calculated R_1 , R_2 , and τ_c values correlated with NMR observables for the individual sugars. In these simulations, the pentasaccharide exhibited consistent interactions with helix-7 and the loop connecting to helix-8 of the polypeptide that was mediated by 3–4 frequent hydrogen bonds and optimal van der Waals contacts (Figures 2d, 2e). Analysis of *T. gondii* Skp1, in which α Glc replaced the penultimate α Gal [45], yielded a similar result except that an additional hydrogen bond involving the terminal sugar was detected (Mandalasi *et al.*,

unpublished data). Correlating with this intramolecular association was a tendency of helix-8 to orient away from the protein core with an open and extended conformation (compare Figs. 2d, 2c), in a manner that potentially provides access to an incoming F-box domain likely attended by a chaperone. Preliminary modeling indicates that the glycan exerts this effect without interfering with F-box binding, though its final disposition remains to be examined.

The Skp1 structural changes correlate with effects of glycosylation on the interaction between Skp1 and several FBPs. Solution studies with the soluble, heterologous mammalian FBP Fbs1 showed that addition of α GlcNAc promoted association with Fbs1 to a degree not achieved by hydroxylation alone [50]. Full extension of the glycan to the pentasaccharide was required for maximal interaction with two FBPs as determined by co-immunoprecipitation studies on mutant extracts of *Dictyostelium* [52]. It remains to be determined whether differential glycosylation affects association with certain or all FBPs. Since at steady state >90% of Skp1 is fully modified and there is no evidence for sugar removal, the slower rate of glycosylation of nascent Skp1 at lower O₂ levels [52] may affect the timing of Skp1/FBP assembly more than the final equilibrium. While it is currently unclear whether intermediate glycoforms persist long enough to be functionally significant, the incremental effects revealed by the glycosyltransferase mutants suggest a plausible basis for a stepwise evolution of this complex pathway.

The proposed influence of a relatively stable glycan conformation on a predominantly disordered polypeptide region before intermolecular protein docking is a novel mechanism, but is in need of further validation. Skp1 has lent itself well to this analysis owing to its relatively small size (324 amino acids as the dimer), minimal glycan microheterogeneity, and a single glycan attachment site. *Cis*-acting mechanisms may become more frequently observed as studies progress from peptides to whole proteins, taking advantage of advances in NMR and MD techniques. It is intriguing to speculate that glycans are well suited, on account of their limited intrinsic flexibility, to impart graded influence on the mobility of disordered domains, where most posttranslational modifications tend to be applied in proteins [53]. Although it is possible that the Skp1 glycan functions by additional mechanisms such as recognition by a carbohydrate binding domain, the interactome studies provided no evidence for a protein with this activity [52], and the highly abundant cytoplasmic β -galactoside binding lectins (discoidins) of *Dictyostelium* may have a role in bacterial interactions [54].

Other examples of nucleocytoplasmic glycosylation

Evidence exists for more extensive nucleocytoplasmic glycosylation beyond these two prototype examples. A type of fucosylation associated with mitochondria and the cytoplasm was described in *Leishmania*, a parasite group that inflicts a range of visceral and cutaneous infections in humans, though dependence on a candidate essential fucosyltransferase has yet to be established [55]. Bioinformatics studies suggest the occurrence of several cytoplasmically localized glycosyltransferases in *Dictyostelium* and *T. gondii* that are yet to be assigned [56, 9•]. In a divergence from protein targets, a nuclear β -glucosyltransferase modifies a hydroxylated base in trypanosome DNA and regulates transcription termination

[57]. In prokaryotes, which represent a rich evolutionary source of glycosylation mechanisms, numerous cytoplasmic glycoproteins have been described [4]. In particular an arginine rhamnosylated elongation factor is important for regulation of protein translation [5]. Several prokaryotes secrete toxins whose glycosyltransferase activities inactivate Ras in the cytoplasm of human cells by attaching a monosaccharide to a discrete Thr or Tyr residue in the GTP-binding site [58, 59], or inactivate a death domain by arginine-GlcNAcylation [60]. Yeast and animals anchor the reducing terminus of their glycogen polysaccharides to a Tyr-residue of a glycosyltransferase that is closely related to the final glycosyltransferase that modifies *T. gondii* Skp1 [9]. A novel type of complex N-glycosylation is assembled in the cytoplasm of algal cells by up to 6 virally encoded cytoplasmic glycosyltransferases, though the capsid glycoproteins probably have their major function following lytic release of the virions [61]. The glycans were characterized by crystallographic studies of the viral particles, indicating their assumption of a low energy conformation associated with the carrier proteins. Thus, there are many indications of new classes of nucleocytoplasmic glycosylation yet to be discovered and/or structurally analyzed.

Summary and implications for the future

Emerging evidence from numerous systems on the role of nucleocytoplasmic glycans is currently best interpreted as self-effects on the modified proteins. Progress in this area requires specialized approaches because the regions of proteins that are glycosylated, and the sugars themselves, tend to have a dynamic nature that requires time-resolved techniques such as NMR to analyze. Nevertheless, it is clear that much cellular regulation occurs in intrinsically disordered regions of proteins via changes in the ensemble of conformations that are sampled over time. The nature of these changes is protein specific and can be expected to affect any of the types of intra- and inter-molecular interactions in which proteins participate. This can be contrasted with examples where glycans themselves constitute determinants that are directly bound by carbohydrate recognition domains. While protists, like other eukaryotes, are known to express nucleocytoplasmic lectins that might function in this manner, an emerging concept is that cytoplasmic lectins are more likely to recognize extracellular self-glycans and foreign glycans [54, 62, 63]. *Cis*-effects are understood for extracellular glycoproteins as well, an excellent recent example being effects of N-glycans on IgG and Fc receptors [64, 65]. Indications from current research are that nucleocytoplasmic glycosylation diversifies the structures of its target proteins allowing for ongoing glycoregulation over the life of the protein. Further investigations of these mechanisms are warranted not only to explore the diversity of regulatory solutions that have evolved in protists, but also for the promise that new information holds for controlling the virulence of parasitic members of this group.

Acknowledgments

Research from the authors' labs summarized in this article was supported by grants from the US NIH (R01GM037539, R01GM084383, 5R21AI123161), and the Mizutani Foundation for Glycoscience (Japan). We thank Alex Eletsy for his comments on the manuscript, and David Live for discussions.

Abbreviations

CD	circular dichroism
MD	molecular dynamics
FBP	F-box protein

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
 - of outstanding interest
1. West CM, Hart GW: Nucleocytoplasmic glycosylation. In: Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, Seeberger PH, editors. *Essentials of Glycobiology* [Internet] 3rd edition. Cold Spring Harbor Laboratory Press; 2017, Chapter 18
 2. Hart GW: Nutrient regulation of signaling and transcription. *J Biol Chem* 2019, 294:2211–2231. [PubMed: 30626734]
 3. Eustice M, Bond MR, Hanover JA: O-GlcNAc cycling and the regulation of nucleocytoplasmic dynamics. *Biochem Soc Trans* 2017, 45:427–436. [PubMed: 28408483]
 4. Fredriksen L, Moen A, Adzhubei AA, Mathiesen G, Eijsink VG, Egge-Jacobsen W: *Lactobacillus plantarum* WCFS1 O-linked protein glycosylation: an extended spectrum of target proteins and modification sites detected by mass spectrometry. *Glycobiology* 2013, 23:1439–1451. [PubMed: 24000282]
 5. Lassak J, Keilhauer EC, Fürst M, Wuichet K, Gödeke J, Starosta AL, Chen JM, Søggaard-Andersen L, Rohr J, Wilson DN, Häussler S, Mann M, Jung K: Arginine-rhamnosylation as new strategy to activate translation elongation factor P. *Nat Chem Biol* 2015, 11:266–270. [PubMed: 25686373]
 6. Blader IJ, Coleman BI, Chen CT, Gubbels MJ: Lytic cycle of *Toxoplasma gondii*: 15 years later. *Annu Rev Microbiol* 2015, 69:463–485. [PubMed: 26332089]
 7. Kamoun S, Furzer O, Jones JD, Judelson HS, Ali GS, Dalio RJ, Roy SG, Schena L, Zambounis A, Panabières F, Cahill D, Ruocco M, Figueiredo A, Chen XR, Hulvey J, Stam R, Lamour K, Gijzen M, Tyler BM, Grünwald NJ, Mukhtar MS, Tomé DF, Tör M, Van Den Ackerveken G, McDowell J, Daayf F, Fry WE, Lindqvist-Kreuzer H, Meijer HJ, Petre B, Ristaino J, Yoshida K, Birch PR, Govers F: The Top 10 oomycete pathogens in molecular plant pathology. *Mol Plant Pathol* 2015, 16:413–434. [PubMed: 25178392]
 - 8••. Bandini G, Haserick JR, Motari E, Ouologuem DT, Lourido S, Roos DS, Costello CE, Robbins PW, Samuelson J: O-fucosylated glycoproteins form assemblies in close proximity to the nuclear pore complexes of *Toxoplasma gondii*. *Proc Natl Acad Sci USA* 2016, 113:11567–11572. [PubMed: 27663739] This article establishes a highly novel form of nucleocytoplasmic glycosylation in *Toxoplasma* and characterizes its localization in the parasite.
 - 9•. Gas-Pascual E, Ichikawa HT, Sheikh MO, Serji MI, Deng B, Mandalasi M, Bandini G, Samuelson J, Wells L, West CM: CRISPR/Cas9 and glycomics tools for *Toxoplasma* glycobiology. *J Biol Chem* 2019, 294:1104–1125 [PubMed: 30463938] This resource article defines the biosynthetic basis for assembly of nucleocytoplasmic O- α Fuc as defined in ref. 8 above, documents its importance for parasite growth in a fibroblast growth assay, and nominates genes that encode unknown cytoplasmic glycosyltransferases.
 10. Zentella R, Sui N, Barnhill B, Hsieh WP, Hu J, Shabanowitz J, Boyce M, Olszewski NE, Zhou P, Hunt DF, Sun TP: The *Arabidopsis* O-fucosyltransferase SPINDLY activates nuclear growth repressor DELLA. *Nat Chem Biol* 2017, 13:479–485. [PubMed: 28244988]
 11. Olszewski NE, West CM, Sassi SO, Hartweck LM: O-GlcNAc protein modification in plants: Evolution and function. *Biochim Biophys Acta* 2010, 1800:49–56. [PubMed: 19961900]

12. Banerjee S, Robbins PW, Samuelson J: Molecular characterization of nucleocytoplasmic O-GlcNAc transferases of *Giardia lamblia* and *Cryptosporidium parvum*. *Glycobiology* 2009, 19:331–336. [PubMed: 18948359]
13. Halim A, Larsen IS, Neubert P, Joshi HJ, Petersen BL, Vakhrushev SY, Strahl S, Clausen H: Discovery of a nucleocytoplasmic O-mannose glycoproteome in yeast. *Proc Natl Acad Sci USA* 2015, 112:15648–15653. [PubMed: 26644575]
14. Cejas RB, Lorenz V, Garay YC, Irazoqui FJ: Biosynthesis of O-N-acetylgalactosamine glycans in the human cell nucleus. *J Biol Chem* 2019, 294:2997–3011. [PubMed: 30591584]
15. Deng B, Tarhan YE, Ueda K, Ren L, Katagiri T, Park JH, Nakamura Y: Critical role of estrogen receptor alpha O-Glycosylation by N-Acetylgalactosaminyltransferase 6 (GALNT6) in its nuclear localization in breast cancer cells. *Neoplasia* 2018, 20:1038–1044. [PubMed: 30208353]
16. Mer G, Hietter H, Lefèvre JF: Stabilization of proteins by glycosylation examined by NMR analysis of a fucosylated proteinase inhibitor. *Nat Struct Biol* 1996, 3:45–53. [PubMed: 8548454]
17. Kao YH, Lee GF, Wang Y, Starovasnik MA, Kelley RF, Spellman MW, Lerner L: The effect of O-fucosylation on the first EGF-like domain from human blood coagulation factor VII. *Biochemistry* 1999, 38:7097–7110. [PubMed: 10353820]
18. Luca VC, Jude KM, Pierce NW, Nachury MV, Fischer S, Garcia KC: Structural basis for Notch1 engagement of Delta-like 4. *Science* 2015, 347:847–853. [PubMed: 25700513]
19. Simanek EE, Huang DH, Pasternack L, Machajewski TD, Seitz O, Millar DS, Dyson HJ, Wong CH: Glycosylation of threonine of the repeating unit of RNA polymerase II with β -linked N-acetylglucosamine leads to a turnlike structure. *J Am Chem Soc* 1998, 120:11567–11575.
20. Chen YX, Du JT, Zhou LX, Liu XH, Zhao YF, Nakanishi H, Li YM: Alternative O-GlcNAcylation/O-phosphorylation of Ser16 induce different conformational disturbances to the N terminus of murine estrogen receptor beta. *Chem Biol* 2006, 13:937–944. [PubMed: 16984883]
21. Brister MA, Pandey AK, Bielska AA, Zondlo NJ: O-GlcNAcylation and phosphorylation have opposing structural effects in tau: phosphothreonine induces particular conformational order. *J Am Chem Soc* 2014, 136:3803–3816. [PubMed: 24559475]
22. Elbaum MB, Zondlo NJ: O-GlcNAcylation and phosphorylation have similar structural effects in α -helices: post-translational modifications as inducible start and stop signals in α -helices, with greater structural effects on threonine modification. *Biochemistry-US* 2014, 53:2242–2260.
23. Liang FC, Chen RP, Lin CC, Huang KT, Chan SI: Tuning the conformation properties of a peptide by glycosylation and phosphorylation. *Biochem Biophys Res Commun* 2006, 342:482–488. [PubMed: 16487934]
24. Karunaratne CV, Weldegiorghis TK, West CM, Taylor CM: Conformational changes associated with post-translational modifications of Pro(143) in Skp1 of *Dictyostelium*--a dipeptide model system. *J Am Chem Soc* 2014, 136:15170–15175. [PubMed: 25250945]
25. Coltart DM, Royyuru AK, Williams LJ, Glunz PW, Sames D, Kuduk SD, Schwarz JB, Chen XT, Danishefsky SJ, Live DH: Principles of mucin architecture: structural studies on synthetic glycopeptides bearing clustered mono-, di-, tri-, and hexasaccharide glycodomains. *J Am Chem Soc* 2002, 124:9833–9844. [PubMed: 12175243]
26. Barb AW, Borgert AJ, Liu M, Barany G, Live D: Intramolecular glycan-protein interactions in glycoproteins. *Meth Enzymol* 2010, 478:365–388. [PubMed: 20816490]
27. Borgert A, Heimbürg-Molinario J, Song X, Lasanajak Y, Ju T, Liu M, Thompson P, Ragupathi G, Barany G, Smith DF, Cummings RD, Live D: Deciphering structural elements of mucin glycoprotein recognition. *ACS Chem Biol* 2012, 7:1031–1039. [PubMed: 22444368]
28. Meyer B, Möller H: Conformation of glycopeptides and glycoproteins. *Top Curr Chem* 2007, 267: 187–251.
29. Schuster O, Klich G, Sinnwell V, Kränz H, Paulsen H, Meyer B: 'Wave-type' structure of a synthetic hexaglycosylated decapeptide: a part of the extracellular domain of human glycoporphin A. *J Biomol NMR* 1999, 14:33–45. [PubMed: 10382304]
30. Narimatsu Y, Kubota T, Furukawa S, Morii H, Narimatsu H, Yamasaki K: Effect of glycosylation on cis/trans isomerization of prolines in IgA1-hinge peptide. *J Am Chem Soc* 2010, 132:5548–5549. [PubMed: 20355726]

- 31●. Prates ET, Guan X, Li Y, Wang X, Chaffey PK, Skaf MS, Crowley MF, Tan Z, Beckham GT. The impact of O-glycan chemistry on the stability of intrinsically disordered proteins. *Chem Sci* 2018, 9:3710–3715. [PubMed: 29780502] This article illustrates how modification of mucin-like domains with O- α -Man, as might occur on yeast nucleocytoplasmic proteins, can influence their organization.
32. Chaffey PK, Guan X, Chen C, Ruan Y, Wang X, Tran AH, Koelsch TN, Cui Q, Feng Y, Tan Z: Structural Insight into the Stabilizing Effect of O-Glycosylation. *Biochemistry* 2017, 56:2897–2906. [PubMed: 28494147]
33. Harrison OJ, Jin X, Hong S, Bahna F, Ahlsen G, Brasch J, Wu Y, Vendome J, Felsovalyi K, Hampton CM, Troyanovsky RB, Ben-Shaul A, Frank J, Troyanovsky SM, Shapiro L, Honig B: The extracellular architecture of adherens junctions revealed by crystal structures of type I cadherins. *Structure* 2011 19:244–256. [PubMed: 21300292]
34. Sletmoen M, Gerken TA, Stokke BT, Burchell J, Brewer CF: Tn and STn are members of a family of carbohydrate tumor antigens that possess carbohydrate-carbohydrate interactions. *Glycobiology* 2018, 28:437–442. [PubMed: 29618060]
35. Banani SF, Lee HO, Hyman AA, Rosen MK: Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* 2017, 18:285–298. [PubMed: 28225081]
36. Rai AK, Chen JX, Selbach M, Pelkmans L: Kinase-controlled phase transition of membraneless organelles in mitosis. *Nature* 2018, 559:211–216. [PubMed: 29973724]
37. West CM, Blader IJ: Oxygen sensing by protozoans: how they catch their breath. *Curr Opin Microbiol* 2015, 26:41–47. [PubMed: 25988702]
38. Zimmerman ES, Schulman BA, Zheng N: Structural assembly of cullin-RING ubiquitin ligase complexes. *Curr Opin Struct Biol* 2010, 20:714–721
39. Leber V, Nans A, Singleton MR: Structural basis for assembly of the CBF3 kinetochore complex. *EMBO J* 2018, 37:269–281. [PubMed: 29212814]
40. Mizushima T, Yoshida Y, Kumanomidou T, Hasegawa Y, Suzuki A, Yamane T, Tanaka K: Structural basis for the selection of glycosylated substrates by SCF(Fbs1) ubiquitin ligase. *Proc Natl Acad Sci USA* 2007, 104:5777–5781. [PubMed: 17389369]
- 41●●. Sheikh MO, Thieker D, Chalmers G, Schafer CM, Ishihara M, Azadi P, Woods RJ, Glushka JN, Bendiak B, Prestegard JH, West CM: O₂ sensing-associated glycosylation exposes the F-box-combining site of the Dictyostelium Skp1 subunit in E3 ubiquitin ligases. *J. Biol. Chem* 2017, 292:18897–18915. [PubMed: 28928219] This article definitively establishes the structure of the Skp1 glycan in Dictyostelium, and generates a model for how it controls the conformation of Skp1 in a manner that correlates with Skp1's interactions with F-box proteins.
42. Taylor CM, Karunaratne CV, Xie N: Glycosides of hydroxyproline: some recent, unusual discoveries. *Glycobiology* 2012, 22:757–767. [PubMed: 22190471]
43. Wang ZA, Singh D, van der Wel H, West CM: Prolyl hydroxylation- and glycosylation-dependent functions of Skp1 in O₂-regulated development of Dictyostelium. *Dev Biol* 2011, 349:283–295. [PubMed: 20969846]
- 44●. Rahman K, Zhao P, Mandalasi M, van der Wel H, Wells L, Blader IJ, West CM: The E3 ubiquitin ligase adaptor protein Skp1 is glycosylated by an evolutionarily conserved pathway that regulates protist growth and development. *J Biol Chem* 2016, 291:4268–4280. [PubMed: 26719340] This article defines basic precepts regarding the structure, biosynthesis, and cellular function of the Skp1 glycan in *Toxoplasma gondii*, a pathogen that latently infects up to one half of the human population.
45. Rahman K, Mandalasi M, Zhao P, Sheikh MO, Taujale R, Kim HW, van der Wel H, Matta K, Kannan N, Glushka JN, Wells L, West CM: Characterization of a cytoplasmic glucosyltransferase that extends the core trisaccharide of the *Toxoplasma* Skp1 E3 ubiquitin ligase subunit. *J Biol Chem* 2017, 292:18644–18659. [PubMed: 28928220]
46. Islam MS, Leissing TM, Chowdhury R, Hopkinson RJ, Schofield CJ: 2-oxoglutarate-dependent oxygenases. *Annu Rev Biochem* 2018, 87:585–620. [PubMed: 29494239]
47. Xu Y, Brown KM, Wang ZA, van der Wel H, Teygong C, Zhang D, Blader IJ, West CM: The Skp1 protein from *Toxoplasma* is modified by a cytoplasmic prolyl 4-hydroxylase associated with

- oxygen sensing in the social amoeba *Dictyostelium*. *J Biol Chem* 2012, 287:25098–25110. [PubMed: 22648409]
48. van der Wel H, Johnson JM, Xu Y, Karunaratne CV, Wilson KD, Vohra Y, Boons GJ, Taylor CM, Bendiak B, West CM: Requirements for Skp1 processing by cytosolic prolyl 4(trans)-hydroxylase and α -N-acetylglucosaminyltransferase enzymes involved in O₂ signaling in *Dictyostelium*. *Biochemistry* 2011, 50:1700–1713. [PubMed: 21247092]
49. Zhang D, van der Wel H, Johnson JM, West CM: Skp1 prolyl 4-hydroxylase of *Dictyostelium* mediates glycosylation-independent and -dependent responses to O₂ without affecting Skp1 stability. *J Biol Chem* 2012, 287:2006–2016. [PubMed: 22128189]
50. Sheikh MO, Schafer CM, Powell JT, Rodgers KK, Mooers BHM, West CM: Glycosylation of Skp1 affects its conformation and promotes binding to a model F-box protein. *Biochemistry* 2014, 53:1657–1669. [PubMed: 24506136]
51. Xu X, Eletsky A, Sheikh MO, Prestegard JH, West CM: Glycosylation promotes the random coil to helix transition in a region of a protist Skp1 associated with F-box binding. *Biochemistry* 2018, 57:511–515. [PubMed: 29251491]
52. Sheikh MO, Xu Y, van der Wel H, Walden P, Hartson SD, West CM: Glycosylation of Skp1 promotes formation of Skp1/Cullin-1/F-box protein complexes in *Dictyostelium*. *Mol Cell Proteomics* 2015, 14:66–80. [PubMed: 25341530]
53. Beltrao P, Albanese V, Kenner LR, Swaney DL, Burlingame A, Villen J, Lim WA, Fraser JS, Frydman J, Krogan NJ: Systematic functional prioritization of protein posttranslational modifications. *Cell* 2012, 150:413–425. [PubMed: 22817900]
54. Dinh C, Farinholt T, Hirose S, Zhuchenko O, Kuspa A: Lectins modulate the microbiota of social amoebae. *Science* 2018, 361:402–406. [PubMed: 30049880]
55. Guo H, Novozhilova NM, Bandini G, Turco SJ, Ferguson MAJ, Beverley SM: Genetic metabolic complementation establishes a requirement for GDP-fucose in *Leishmania*. *J Biol Chem* 2017, 292:10696–10708. [PubMed: 28465349]
56. Sugang R I, Kuo A, Tian X, Salerno W, Parikh A, Feasley CL, Dalin E, Tu H, Huang E, Barry K, Lindquist E, Shapiro H, Bruce D, Schmutz J, Salamov A, Fey P, Gaudet P, Anjard C, Babu MM, Basu S, Bushmanova Y, van der Wel H, Katoh-Kurasawa M, Dinh C, Coutinho PM, Saito T, Elias M, Schaap P, Kay RR, Henrissat B, Eichinger L, Rivero F, Putnam NH, West CM, Loomis WF, Chisholm RL, Shaulsky G, Strassmann JE, Queller DC, Kuspa A, Grigoriev IV: Comparative genomics of the social amoebae *Dictyostelium discoideum* and *Dictyostelium purpureum*. *Genome Biol* 2011, 12:R20. [PubMed: 21356102]
57. Bullard W, Lopes da Rosa-Spiegler J, Liu S, Wang Y, Sabatini R: Identification of the glucosyltransferase that converts hydroxymethyluracil to base J in the trypanosomatid genome. *J Biol Chem* 2014, 289:20273–20282. [PubMed: 24891501]
58. Belyi Y, Jank T, Aktories K: Cytotoxic glucosyltransferases of *Legionella pneumophila*. *Curr Top Microbiol Immunol* 2013, 376: 211–226. [PubMed: 23900830]
59. Jank T, Eckerle S, Steinemann M, Trillhaase C, Schimpl M, Wiese S, van Aalten DM, Driever W, Aktories K: Tyrosine glycosylation of Rho by *Yersinia* toxin impairs blastomere cell behaviour in zebrafish embryos. *Nat Commun* 2015, 6:7807. [PubMed: 26190758]
60. Li S, Zhang L, Yao Q, Li L, Dong N, Rong J, Gao W, Ding X, Sun L, Chen X, Chen S, Shao F: Pathogen blocks host death receptor signalling by arginine GlcNAcylation of death domains. *Nature* 2013, 501: 242–246. [PubMed: 23955153]
61. De Castro C, Molinaro A, Piacente F, Gurnon JR, Sturiale L, Palmigiano A, Lanzetta R, Parrilli M, Garozzo D, Tonetti MG, Van Etten JL: Structure of N-linked oligosaccharides attached to chlorovirus PBCV-1 major capsid protein reveals unusual class of complex N-glycans. *Proc Natl Acad Sci USA* 2013, 110:13956–13960. [PubMed: 23918378]
62. Bleuler-Martínez S, Butschi A, Garbani M, Wälti MA, Wohlschlager T, Potthoff E, Sabotić J, Pohleven J, Lüthy P, Hengartner MO, Aebi M, Künzler M: A lectin-mediated resistance of higher fungi against predators and parasites. *Mol Ecol* 2011, 20:3056–3070. [PubMed: 21486374]
63. Stowell SR, Arthur CM, McBride R, Berger O, Razi N, Heimbürg-Molinario J, Rodrigues LC, Gouridine JP, Noll AJ, von Gunten S, Smith DF, Knirel YA, Paulson JC, Cummings RD: Microbial

glycan microarrays define key features of host-microbial interactions. *Nat Chem Biol* 2014, 10:470–476. [PubMed: 24814672]

64. Subedi GP, Barb AW: CD16a with oligomannose-type N-glycans is the only “low-affinity” Fc γ receptor that binds the IgG crystallizable fragment with high affinity in vitro. *J Biol Chem* 2018, 293:16842–16850. [PubMed: 30213862]
65. Falconer DJ, Subedi GP, Marcella AM, Barb AW: Antibody fucosylation lowers the Fc γ RIIIa/CD16a affinity by limiting the conformations sampled by the N162-glycan. *ACS Chem Biol* 2018, 13:2179–2189. [PubMed: 30016589]

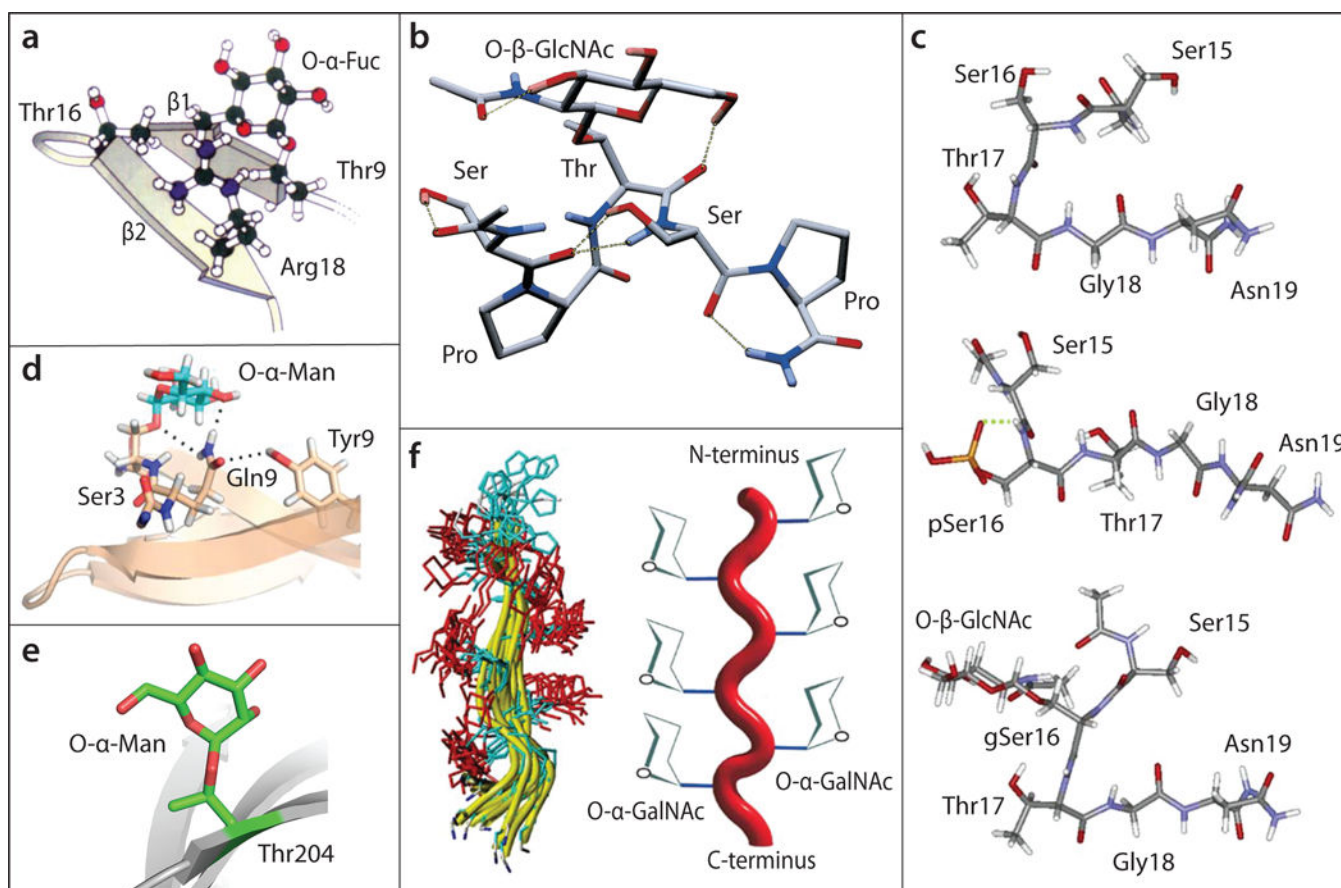
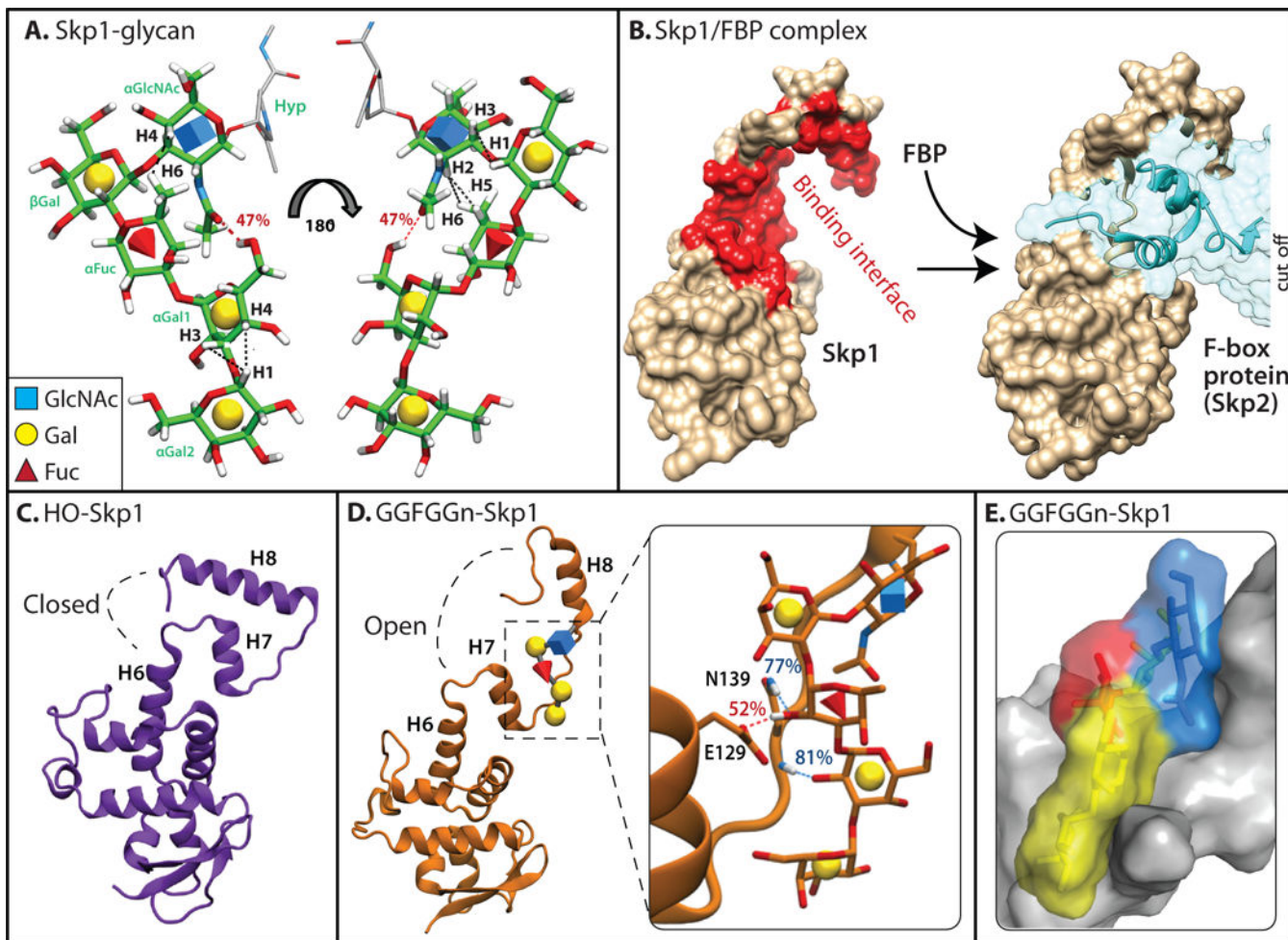


Figure 1.

Examples of monosaccharide modifications on peptides. **(a)** α -L-Fuc on Thr9 of the β 1 strand of PMP-C protease inhibitor, illustrating stabilizing contacts with Thr16 and Arg18 of the β 2 strand [modified from 16, with permission of the publisher]. **(b)** Ac-Ser-Pro-Thr(O- β -GlcNAc)-Ser-Pro-NH₂ from the repeating sequence motif of RNA polymerase-II, showing a turn induced by the sugar [modified from 19, with permission of the publisher]. **(c)** Distinct effects of phosphorylation or O- β -GlcNAcylation of Ser16 of a peptide from the estrogen β -receptor [modified from 20, with permission of the publisher]. **(d)** α -D-Man linked to the cellulose binding domain of a cellobiohydrolase, depicting stabilizing contacts with Gln2 that indirectly affect distal peptide organization [modified from 32, with permission of the publisher]. **(e)** α -D-Man on Thr204 of a β -strand exposed at the surface of EC2 of E-cadherin, illustrating solvent exposure and negligible contact with the polypeptide.[from PDB 3Q2V; 33]. **(f)** α -D-GalNAc substitution of consecutive hydroxyamino acid residues (His-Thr(GalNAc)-Ser(GalNAc)-Thr(GalNAc)-Ser(GalNAc)-Ser(GalNAc)-Ser(GalNAc)-Val-Thr-Lys) from glycophorin A; MD snapshots and cartoon illustrate impact on polypeptide folding and orientation of the GalNAc residues [taken from 28, 29, with permission of the publisher].

**Figure 2.**

The *Dictyostelium* pentasaccharide and its effect on Skp1 conformation. **(a)** The sequence and linkages of the glycan were determined by analysis of the ^{13}C -sugar labeled glycopeptide and confirmed on ^{13}C -sugar labeled versions of Gal-Gal-Fuc-Gal-GlcNAc-Skp1 (GGFGGn-Skp1). Its conformation was inferred from calculation of its lowest energy state (Glycan), molecular dynamics simulations, and NMR studies of ^{13}C -GGFGGn-Skp1. Black dashed lines represent NOEs observed from NMR analysis of GGFGGn-Skp1; red dashed line represents a hydrogen bond observed during MD simulations. **(b)** The conformation of unmodified human Skp1 in a crystallographic complex (PDB 2ASS) with the FBP Skp2 (at right). The largely hydrophobic binding interface is highlighted in red (left) using PISA analysis in the excerpted surface rendition at the left. The highly conserved C-terminal region (upper half) can adopt at least two different conformations with other FBPs [38–40]. Images were generated using Chimera. **(c)** Back-side view (relative to panel c) of the average conformation of *Dictyostelium* HO-Skp1 during MD simulations initiated from the structure in panel c. **(d)** Average conformation of GGFGGn-Skp1, with the inset showing hydrogen bonds which are present during over 50% of the simulation time. **(e)**

Space-filling model of glycosylated region of Skp1, depicting van der Waals contacts of the glycan with the polypeptide. Panels a, c, d, and e are modified from ref. 41●●.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript