Published in final edited form as:

Biochim Biophys Acta. 2010 February ; 1800(2): 160–171. doi:10.1016/j.bbagen.2009.11.006.

A cytoplasmic prolyl hydroxylation and glycosylation pathway modifies Skp1 and regulates O₂-dependent development in *Dictyostelium*

Christopher M. West^{*}, Zhuo A. Wang, and Hanke van der Wel

Department of Biochemistry and Molecular Biology, Oklahoma Center for Medical Glycobiology, University of Oklahoma Health Sciences Center, 975 Northeast Tenth Street, Oklahoma City, Oklahoma 73104 USA

Abstract

The soil amoeba *Dictyostelium* is an obligate aerobe that monitors O₂ for informational purposes in addition to utilizing it for oxidative metabolism. Whereas low O₂ suffices for proliferation, a higher level is required for slugs to culminate into fruiting bodies, and O₂ influences slug polarity, slug migration, and cell-type proportioning. *Dictyostelium* expresses a cytoplasmic prolyl 4-hydroxylase (P4H1) known to mediate O₂-sensing in animals, but lacks HIF α , a major hydroxylation target whose accumulation directly induces animal hypoxia-dependent transcriptional changes. The O₂-requirement for culmination is increased by P4H1-gene disruption and reduced by P4H1 overexpression. A target of *Dictyostelium* P4H1 is Skp1, a subunit of the SCF-class of E3-ubiquitin ligases related to the VBC-class that mediates hydroxylation-dependent degradation of animal HIF α . Skp1 is a target of a novel cytoplasmic O-glycosylation pathway that modifies HyPro143 with a pentasaccharide, and glycosyltransferase mutants reveal that glycosylation intermediates have antagonistic effects toward P4H1 in O₂-signaling. Current evidence indicates that Skp1 is the only glycosylation target in cells, based on metabolic labeling, biochemical complementation, and enzyme specificity studies. Bioinformatics studies suggest that the HyPro-modification pathway existed in the ancestral eukaryotic lineage and was retained in selected modern day unicellular organisms whose life cycles experience varying degrees of hypoxia. It is proposed that, in *Dictyostelium* and other protists including the agent for human toxoplasmosis *Toxoplasma gondii*, prolyl hydroxylation and glycosylation mediate O₂-signaling in hierarchical fashion via Skp1 to control the proteome, directly via degradation rather than indirectly via transcription as found in animals.

Keywords

cytoplasmic glycosylation; hydroxyproline; Skp1; hypoxia; *Dictyostelium*; *Toxoplasma*

INTRODUCTION

Dictyostelium is a model organism [1] that proliferates in the wild as a unicellular amoeba on a diet of bacteria and yeast in soils throughout the world. Starvation induces the solitary

^{*}Author to whom correspondence should be addressed: Chris West, Dept. of Biochemistry & Molecular Biology, OUHSC, BRC 417, 975 NE 10th St., Oklahoma City, OK 73104 USA; telephone: 405-271-4147; Fax: 405-271-3910; Cwest2@ouhsc.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

amoebae to aggregate into a multicellular slug, which in response to external cues migrates to the soil surface and culminates into a fruiting body (Fig. 1) [2,3]. Spores, perched at the top of the fruiting body stalk, have increased likelihood to be dispersed to distant locations where they can germinate to renew the cycle. Since culmination is irreversible [4], the migrating slug must continuously weigh the potential of finding a better location against the cost and risk of further searching. The slug-to-fruit decision is promoted by low NH_3 (a metabolic by-product), high O_2 , light, low humidity, and warmth, which are each characteristic of the above-ground environment where culmination would be optimal [4,5]. Environmental sensing may occur at the anterior slug tip (colored red in Fig. 1), which has properties of the primary organizer of animal embryos.

Prolyl 4-hydroxylases, first discovered in the rough endoplasmic reticulum as modifiers of collagen in animals and cell wall proteins in plants, have recently been identified in the cytoplasm as regulators of the stability of animal hypoxia inducible factor α (HIF α). HIF α is a subunit of the HIF α -HIF β transcriptional factor heterodimer which induces hypoxia response genes that support glycolysis, angiogenesis, and erythropoiesis [6]. Hydroxylation of HIF α , promoted by O_2 and other factors, allows recognition by the VHL subunit of the E3(VBC)-ubiquitin (Ub)-ligase followed by degradation within the 26S-proteasome (discussed below in Fig. 6A). PHD1-3 (also known as HPHs and EGLNs) are non-heme Fe(II)-dependent dioxygenases, and are thought to be able to function as direct O_2 -sensors because of their relatively high K_m s for O_2 [6]. They may also be activated by the co-substrate α -ketoglutarate (α -KG), a Krebs cycle intermediate, ascorbate, and Fe(II), and inhibited by the product succinate, other Krebs cycle metabolites, and reactive oxygen species (ROS). In addition, PHD levels are modulated transcriptionally and by turnover, and possibly also by nuclear compartmentalization. Much remains to be learned about the quantitative regulation of PHDs in cells.

Dictyostelium Skp1 was initially discovered as a novel fucosylated cytoplasmic protein (FP21) [7]. Skp1 is modified by a pentasaccharide, assembled by a novel O-glycosylation pathway (see Fig. 2 below) that is also compartmentalized in the cytoplasm, in contrast to the more usual location of protein O-glycosylation pathways in the rough endoplasmic reticulum/Golgi apparatus. Skp1 was subsequently discovered in yeast and mammalian cells as a critical subunit of the SCF-class of E3-Ub-ligases which target cell cycle regulatory proteins and transcriptional factors for degradation in the 26S-proteasome [8]. Skp1 probably has additional functions distinct from E3-Ub-ligases [9,10], though likely still in association with F-box proteins.

The *Dictyostelium* homolog (P4H1) of animal PHDs was discovered as a modifier of Skp1 and is required for its glycosylation [11]. P4H1 shares many biochemical properties with animal PHDs. As discussed below, genetic and biochemical evidence supports a model in which P4H1-dependent modification of Skp1 comprises an O_2 -sensing pathway for regulation of development. *Dictyostelium* lacks HIF α -type transcriptional factors, and Skp1 may be the only substrate of P4H1. The SCF complex, also present in *Dictyostelium*, is evolutionarily related to the VBC-type E3 Ub-ligase that targets HIF α in animals. We propose that by modifying the Ub-ligase itself rather than one of its substrates, an evolutionarily ancient mechanism of O_2 -signaling may have involved regulation of the proteome directly by protein turnover.

This article reviews underlying tenets of the model, including characteristics of *Dictyostelium* Skp1, the mechanism of its posttranslational modification, evidence for the role of the hydroxylation/glycosylation pathway in O_2 -signaling, and possible mechanisms of how the pathway modulates development. Finally, the significance of the Skp1-modification pathway for O_2 -regulation for other members of protist kingdom, including important vectors of human and agricultural disease, is evaluated.

Skp1

Studies in other organisms indicate Skp1 specifically binds the F-box motif found in the family of F-box proteins [8,9], which number from 15–700 members in different eukaryotes. A major subset includes SCF complexes, in which Skp1 also binds cullin-1 to form a class of E3-Ub-ligases. E3(SCF)Ub-ligases have been implicated in *Dictyostelium* development [12]. Cell fractionation shows that about half of *Dictyostelium* Skp1 is soluble and the remainder sedimentable with the particulate fraction in a salt-sensitive, detergent-insensitive fashion, consistent with a possible cytoskeletal or scaffold association [13]. Although Skp1 does not fractionate with nuclei, immunofluorescence studies reveal a clear nuclear enrichment that is dependent on at least partial glycosylation [14].

Dictyostelium Skp1 is glycosylated according to the scheme shown in Fig. 2. The linkage to HyPro143 and the glycan chain structure are supported by Edman degradation and MS fragmentation analyses; the glycan composition is supported by GC-MS and HPAEC [15]; the Gal α 1,3Fuc α 1,2Gal α 1,3 trisaccharide linkages are supported by exoglycosidase sensitivity [15]; the GlcNAc α 1,4Hy(trans)Pro linkage is supported by homologies of the GlcNAcT and P4H1 to known enzymes, and the ability of mAbs prepared against a synthetic GlcNAc 1,4Hy(trans)Pro-peptide to recognize Skp1 (see below); and the Gal α 1,3Fuc linkage is supported by analysis of the enzyme product by HPLC [16]. α Gal-2 is α -linked based on α -galactosidase sensitivity, and timed mild acid hydrolysis suggests that it is linked to either the other α -Gal or Fuc [15]. The core trisaccharide has the same structure as the human type 1 blood group H antigen, but is formed as a result of convergent evolution [17]. NMR studies and molecular dynamics simulations of model compounds show that the core trisaccharide adopts an overlapping ensemble of conformations best represented by a hairpin in which Fuc folds back toward the GlcNAc [18]. Crystal structures of SCF complexes from yeast, animals and plants suggest that the Pro-attachment site occurs at an exposed turn and is not covered by essential contacts with the F-box protein or cullin-1 [19]. However, structures of complexes that include full-length F-box proteins, or other proteins that may associate with Skp1, have not been investigated. Structure predictions suggest that the glycan projects away from the protein (<http://glycam.ccruc.uga.edu/>), where it may more likely be recognized by a separate binding protein than affect the folding of its own carrier polypeptide.

Biochemical studies conducted *in vivo* and *in vitro* suggest that Skp1 is the only protein modified by the HyPro pathway (see sections below). The majority ($\geq 90\%$) of steady-state Skp1 appears to be constitutively glycosylated based on M_r analysis using SDS-PAGE [14, 5], except for possible variation of an outer α Gal [13]. 5–10% is not hydroxylated according to assays in which extracts are incubated with recombinant enzymes and radioactive sugar nucleotides [5]. However, this is a minimal estimate because of possible post-lysis modification. The high degree of modification of Skp1 in the steady state suggests a role in a global function of Skp1 such as the Ub(SCF)-ligase.

When modestly over-expressed in prespore cells of the slug (see Fig. 1), with a C-terminal c-myc epitope tag, little Skp1-myc is hydroxylated, but becomes hydroxylated/glycosylated when cells are induced to dedifferentiate [14]. In contrast, pre-existing Skp1 in the same cells is predominantly glycosylated, indicative of multiple pools. Therefore, Skp1-myc is stable in prespore cells and hydroxylation appears to be rate-limiting for modification of nascent Skp1 in slugs.

If Skp1 is expressed in proliferating cells with an N-terminal FLAG-tag, or with point mutations in its N-terminal domain (the contact region with cullin-1), unmodified and hydroxylated (but not glycosylated) Skp1, and glycoforms lacking one or both α Gal residues, accumulate [15; H. van der Wel, C.M. West, unpublished data]. In contrast, similar expression of Skp1 with its wild-type sequence except for a C-terminal tag is glycosylated normally. These results suggest

that modification of Skp1 is sensitive to its physical state, but whether this reflects a quality control or regulatory mechanism for the HyPro modification pathway remains to be determined.

Dictyostelium expresses two Skp1s, Skp1A and Skp1B, which differ only by the presence of Ser or Ala at codon 39. However, other amoebazoans possessing genes related to the Skp1 modification pathway (see Table 1) appear to have only a single Skp1 gene, with either Thr or Ala at this position. In addition, multiple Skp1s do not always occur in other protists with HyPro pathway-like genes. In *Dictyostelium*, no differences in Skp1A and Skp1B expression were detected by RT-PCR [14] or analysis of promoter::RFP constructs (H. van der Wel, C.M. West, unpublished data), and both isoforms are similarly glycosylated [20]. The lack of evolutionary conservation and the similar expression profiles suggest that the occurrence of 2 isoforms is not directly pertinent to O₂-regulation.

The equivalent of P143 (the attachment site) is conserved in Skp1s from yeast/fungi, protists, algae, plants, and invertebrates, but not vertebrates. Some organisms have additional Skp1 genes in which P143 is replaced by another amino acid [20]. The equivalent of P143 is replaced by Glu in vertebrate Skp1s which, interestingly, possess a nearby Pro in an identical TPEE sequence motif in a similar secondary structure context. The M_r of Skp1 purified from bovine liver, based on MALDI-ToF-MS, is consistent with the absence of a steady state modification (C. Newsome, C.M. West, unpublished data). A sequence alignment of Skp1s from a broad range of phylogenetic groups did not reveal amino acid identities correlating with the presence of Pro at codon 143, or the occurrence of HyPro modification-like genes in their genomes (unpublished data). The apparent absence of purifying selection suggests that recognition by HyPro pathway enzymes depends on motifs constrained by other conserved functions.

Skp1 prolyl hydroxylation

The Skp1 prolyl hydroxylase gene (P4H1) was predicted by sequence similarity to animal HIF α PHDs [17], and shown to be required for Skp1 glycosylation based on the effect of gene disruption [11]. Bacterially expressed P4H1 is specific for the modification of Pro143 of Skp1 (unpublished data). P4H1 is dispensable for proliferation but required for development in hypoxia (see below). P4H1 exhibits a K_m for O₂ of >40%, and is dependent on α KG, and ascorbate, as expected. P4H1 exhibits an unusually high affinity for Fe(II), and is inhibited by compounds that affect mammalian PHDs. Sequence comparisons show P4H1 to be more similar to HIF α -type than collagen-type PHDs, and more similar to 4-PHDs than 3-PHDs [11]. Although only 33% identical to the 120-amino acid catalytic domain of mammalian PHD2, P4H1 is considered to be the *Dictyostelium* ortholog of HIF α -type PHDs, suggesting that P4H1 catalyzes formation of 4(trans)-hydroxyproline.

Gnt1-mediated addition of GlcNAc

Gnt1 was purified as a UDP-GlcNAc-dependent polypeptide GlcNAcT capable of modifying HyPro143 of Skp1 *in vitro* [22]. Protein sequence data was used to clone *gnt1* and the recombinant protein exhibited similar activity [23]. However, *gnt1* has resisted disruption even in a P4H1-null background (C.M. West et al., unpublished data). Gnt1 was the defining member of its CAZy [24] sequence family (GT60), and mutagenesis studies indicate it to be homologous to the mucin-type polypeptide α GlcNAcT in the Golgi of *Dictyostelium*, and to CAZy GT27 α GalNAcTs that initiate mucin-type O-glycosylation in animals [25]. Based on these similarities, Gnt1 is predicted to catalyze addition of α GlcNAc to HyPro. In support of these assignments, a synthetic peptide containing 4(trans)HyPro is a specific albeit weak substrate for Gnt1 [22], and independent mAbs raised against a GlcNAc α 1,4(trans)HyPro containing Skp1 peptide recognize full-length Skp1 in crude extracts [26]. In addition, other than the Golgi homolog, there are no Gnt1-like sequences in the genome, and another known cytoplasmic

α GlcNAcT, an α -toxin from *Clostridium* [27], is also not observed. Gnt1 is evolutionarily unrelated to the OGT enzyme that catalyzes formation of O- β -GlcNAc in many organisms [28].

Gnt1 has predicted cytoplasmic homologs in a variety of other protists (Table 1). Each of these protist genomes also has a P4H1-like coding sequence and a Skp1-like sequence with the equivalent of Pro143, and therefore the P4H1- and Gnt1-like sequences may modify their Skp1s as in *Dictyostelium*. Interestingly, P4H1- and Gnt1-like sequences sometimes appear to encode separate domains of the same protein [17], as occurs for the other enzyme pairs in this pathway (see below). Biochemical complementation studies, in which purified recombinant P4H1 and Gnt1 are incubated with UDP-[3 H]GlcNAc in extracts of P4H1-null mutants, suggest that the P4H1/Gnt1 couple is dedicated solely to the modification of Skp1 in *Dictyostelium* (H. van der Wel, C.M. West, unpublished data).

PgtA (β 3GalT/ α 2FucT) is a two-domain diGT

PgtA was purified based on its fucosyltransferase activity and subsequently discovered to also possess a Skp1 galactosyltransferase activity. *pgtA* was cloned from amino acid sequence data obtained from the multi-million fold purified protein, and found to be required in cells for extension of the glycan chain beyond the first sugar GlcNAc. Cells are viable in its absence with only mild effects on proliferation [29]. PgtA is a soluble protein with separate domains mediating the different GT activities. Action of the β 3GalT activity, which is associated with the N-terminal domain, is followed by the α 2FucT activity of the C-terminal domain [29,30]. Formation of the Gal β 1,3GlcNAc-linkage was established based on its sensitivity to β 3-galactosidase. The N-terminal β 3GalT domain modifies Gn-O-Skp1 *in vitro* [30] and *in vivo* [26], indicating it possesses all determinants required for its enzymatic function. Both native and synthetic Skp1 GlcNAc-glycopeptide are both exceedingly poor acceptor substrates and poor inhibitors of the β 3GalT activity, and the acceptor activity of full-length Skp1 is strongly inhibited by alkylation of its Cys-residues with iodoacetamide [26]. These data indicate that Skp1 recognition depends on determinants distinct from the sugar modification site, implying sites of secondary interaction which might help differentiate Skp1 from other potential GlcNAc-modified cytoplasmic proteins. The α 2FucT activity was established based on product characterization of model substrates and α 2-fucosidase sensitivity [31]. The isolated C-terminal domain exhibits α 2FucT activity toward disaccharide acceptors *in vitro*, but only marginal activity toward full length GGn-O-Skp1 [29]. However, the full-length enzyme whose β 3GalT activity is mutationally inactivated has full activity toward GGn-O-Skp1. In the presence of both UDP-Gal and GDP-Fuc, the addition of Fuc is nearly as efficient as that of Gal [30], indicating that the enzyme can function processively. Unlike the β 3GalT activity, the α 2FucT activity of PgtA toward GGn-O-Skp1 is not affected by prior alkylation, indicating that determinants associated with the alkylated Cys residues are important for galactosylation *per se*. In addition to catalysis, the N-terminal domain is proposed to recognize Skp1 features distinct from its GlcNAc and present Skp1 to the C-terminal domain for fucosylation. Interestingly, concentrations of UDP-Gal above 10 μ M are inhibitory in both the single and coupled reactions [26], but the mechanism is not known.

The N-terminal β 3GalT domain is related to the large CAZy GT2 family of catalytic domains based on sequence similarities and effects of mutagenesis [30]. GT2 domains are nearly always associated with cytoplasmic proteins in both prokaryotes and eukaryotes, though an exception that appears to reside in the Golgi lumen was recently described [32]. The C-terminal α 2FucT domain was the defining member of CAZy family GT74, and homologs have since been detected in other bacterial and eukaryotic genomes as predicted soluble proteins (Table 1). The protein sequences show similarity over the entire 260 amino acids thought to comprise the catalytic region. Three almost perfectly conserved sequence motifs are suggested by the

alignment shown in Fig. 3. Further studies are required to determine whether these motifs are specifically predictive of α 2FucT activity. The GT74 sequence cannot be aligned with CAZy GT11 or GT37 α 2FucT sequences, showing that formation of this linkage evolved independently. In eukaryotes, the GT74 sequence occurs in tandem with a β 3GalT-like GT2 sequence [17], but the order is reversed outside of the mycetezoan group.

Skp1 is the only acceptor substrate detected in biochemical reconstitution experiments in which PgtA and UDP-[3 H]Gal are added back to cytosolic, microsomal, or nuclear extracts from *pgtA*⁻ cells [26]. Similarly, Skp1 was the only protein labeled when GDP-[3 H]Fuc was added back to an extract of *gmd*⁻ mutant cells, which are unable to form GDP-Fuc from GDP-Man [7,21]. In addition, metabolic labeling of cells with [3 H]Fuc identifies predominantly Skp1 in the cytosolic fraction [7], and Skp1 is the only protein recognized in *pgtA*⁻ cells by any of the mAbs that recognize the GlcNAc-O-Skp1 [26]. These findings suggest that Skp1 is the major if not exclusive acceptor substrate for PgtA in cells, which is consistent with the recognition requirements of PgtA for Skp1 discussed above.

AgtA-mediated addition of α Gal-1

AgtA was originally purified as the enzyme that adds the fourth sugar α Gal-1 [16]. The AgtA gene was cloned based on amino acid sequence data from the purified enzyme protein, and shown to be required, *in vivo*, for addition of both α Gal-1 and the final sugar, α Gal-2 [33]. Highly purified AgtA catalyzes the addition of two Gal residues to Skp1 *in vitro* (A. Ercan, C. Feasley, H. van der Wel, C.M. West, unpublished data), suggesting that AgtA may also mediate the addition of α Gal-2 *in vivo*. The protein contains an N-terminal domain related to a sequence family that at the time had been predicted to have glycosyltransferase activity based on sequence/structure homology studies [34]. Purified AgtA modified synthetic Fuc α -Bn with a single Gal residue, forming Gal α 1,3Fuc α -Bn based on product characterization [16], making AgtA the founding member of the CAZy GT77 family. A large family of related sequences are expressed in the Golgi of plants, green algae and brown algae [35], and an example from *Arabidopsis* (AtRGXT1) has been shown to catalyze formation of a Xyl α 1,3Fuc linkage in the pectin rhamnogalacturonan-II.

BLAST searches reveal sequences similar to the GT77 catalytic domain in predicted cytoplasmic proteins of the distantly related *D. purpureum* [36], the true slime mold *Physarum polycephalum*, and viridiplantae, including *Arabidopsis*, bryophytes, and chlorophytes (Fig. 4). An alignment of the inferred catalytic domains of these sequences and of AtRGXT1 shows similarities, especially in hydrophobic residues that probably contribute to the folded core, throughout the length of the catalytic domain (Fig. 4). Four highly conserved sequence motifs are potentially indicative of an α 3-linkage for the sugar transferred from a UDP-precursor, though the type of sugar transferred cannot be inferred.

Dictyostelium AgtA possesses a second, C-terminal domain, consisting of 7 WD-40 repeats. These repeats probably fold into a β -propeller structure, which often mediates simultaneous, reversible, protein-protein interactions [37]. This domain is not required to galactosylate Fuc-Bn *in vitro* [16], or when truncated AgtA is overexpressed *in vivo* [26]. The -propeller-like domain is present in the AgtA homologs of other amoebozoans, suggesting that they are true orthologs (Table 1). Since the cytoplasmic *agtA*-like sequences of viridiplantae lack the β -propeller-like domain, and their genomes lack *pgtA*-like sequences, the predicted enzymes are unlikely to modify Skp1 in the same way.

Skp1 was the only acceptor substrate for AgtA detected in a biochemical reconstitution assay in which a cytosolic extract of stationary phase cells was incubated with purified AgtA and UDP-[3 H]Gal and incorporation analyzed by SDS-PAGE [33]. This is consistent with similar

findings regarding the specificity of PgtA, which modifies Skp1 to form the substrate trisaccharide for AgtA (see above).

Current evidence suggests that α -galactosylation of Skp1 may be conditionally regulated in the cell. Paradoxically, the native trisaccharide acceptor linked to a small aglycon is a much poorer substrate than the non-reducing terminal disaccharide, due mainly to a $V_{\max}$ disadvantage [16]. The hairpin conformation adopted by the free trisaccharide (see above) may sterically interfere with enzyme processing. Acceptor activity of natural trisaccharide-Skp1 (FGGn-O-Skp1) is dramatically reduced upon denaturation with urea or heat [16], suggesting conformational control of the reaction by the Skp1 polypeptide that may involve trisaccharide presentation. Folding may also be important *in vivo*, because an expressed mutant Skp1-myc mutant accumulates glycoforms lacking one or both α Gals [14]. Interestingly and in contrast to the other Skp1 GTs, AgtA is an abundant protein similar in amount to Skp1 itself, and copurifies with Skp1 during initial purification. The β -propeller domain of *Dictyostelium* myosin heavy chain kinase (MhkA) has been implicated in targeting of myosin for phosphorylation [38]. The fusion of a β -propeller domain to the AgtA catalytic domain may serve a similar function, by facilitating proper presentation of the acceptor trisaccharide to the AgtA active site, or recognizing the distal Skp1 determinants inferred from the effects of denaturation treatments.

Overall pathway properties and evolutionary setting

Skp1 modification must be post-translational since Pro143 is within 20 amino acids of the C-terminus. All four known HyPro modification pathway enzyme proteins are cytosolic after gentle cell lysis, consistent with the absence of detectable targeting sequences for organelles or the nucleus [17]. Although cell fractionation studies indicate the absence of the enzymes in vesicles where most GTs are located, an association with nuclei that is unstable to cell lysis has not been excluded. The enzymes are separable by anion exchange fractionation of the cytosolic fraction, and each of the single reactions is reasonably robust *in vitro*, before and after purification. Therefore, modification of Skp1 appears to involve sequential enzyme-substrate recognition events rather than processing in a multi-protein complex. The pathway thus resembles conventional O-glycosylation pathways in the rough endoplasmic reticulum and Golgi, except that the enzymes are not membrane associated and spatially distributed along the pathway coordinate.

The Skp1 modification enzymes may act in pairwise couples. The β 3GalT and α 2FucT activities are physically associated in PgtA, and the first enzyme (β 3GalT) acts like a gatekeeper to the follow-up action of the α 2FucT as described above. Since P4H1 and Gnt1 are apparently linked in a similar way in other organisms (Table 1), the first two steps may have retained a similar relationship as separate proteins. In view of the possibility that AgtA mediates attachment of both α -linked Gal residues, addition of the final sugars may follow the same paradigm following the conditional addition of α Gal-1. While the significance of the proposed enzyme couples is unclear, it is noted that the follow-up step 2 is likely subject to different kinds of constraints as described above for PgtA, and would have the effect of rendering the initial step 1 irreversible.

Current evidence suggests that Skp1 may be the only modification target of the HyPro modification pathway. As summarized above, reconstitution of extracts of mutant cells with a recombinant version of the deleted enzyme yields only Skp1 as the radiolabeled product. In general, the enzymes exhibit high affinity, μM - K_m values toward the Skp1 substrates and to most of the donor substrates. A caveat is in the case of P4H1 whose reaction product was detected in a coupled reaction with Gnt1, so potential substrates not recognized by Gnt1 would have been missed. In addition, each enzyme shows dependence on polypeptide determinants of Skp1 separate from the the region surrounding Pro143 in *in vitro* reactions, suggesting the

enzymes are restricted to substrates that share these features. When Skp1 is expressed *in vivo*, tags at the N-terminus or point mutations in its N-terminal domain interfere with efficient modification by P4H1, Gnt1, and AgtA. It is not known if incomplete modification is a direct effect on enzyme recognition, or consequences of effects on entry into SCF- or other complexes.

The HyPro modification pathway of *Dictyostelium* offers a precedent for potential, yet-to-be-discovered complex glycosylation pathways in the cytoplasm. As reviewed recently [39,40], there are numerous clues for their more widespread existence. The isolated occurrences in bacteria of potential homologs of individual HyPro-pathway GTs indicate these examples do not form the same pathway [41,17]. Though most prokaryotic glycosyltransferases modify lipids and possibly proteins for export to the cell surface, several examples of potential cytoplasmic glycoproteins [42-44] hint at functions in intracellular regulation. In addition to the ubiquitous eukaryotic polypeptide β GlcNAcT (OGT) that modifies many cytoplasmic, nuclear, and possibly mitochondrial proteins with a single sugar [28], bacterial toxins and effectors modify eukaryotic host G-proteins and a GTPase elongation factor with α GlcNAc or α Glc [27]. Perhaps the best well-characterized complex O-glycosylation pathway other than the HyPro-pathway is encoded by PBCV-1, which modifies its own viral precursor capsid proteins in the cytoplasm of the *Chlorella* alga [45]. Evidence also exists for enzymes that can form a peptidoglycan-like structure in association with plastids in the cytoplasm of certain algae [46]. The enzymes that have been postulated or shown to be involved in these processes are so dissimilar to the Skp1 GTs that they almost certainly represent independent lines of evolution. Other less well documented predictions of cytoplasmic glycosylation [47,41,39] originally inspired initial interest on the Skp1 O-glycosylation pathway in *Dictyostelium*. Numerous plant and animal genomes that lack Skp1 modification GT genes harbor other GT-like sequences predicted to be expressed as cytoplasmic proteins, though their compartmentalization remains to be confirmed [41] (Fig. 4). Since each of the Skp1 GTs were discovered using traditional biochemical purification approaches, and their domains (except for the β 3GalT) were each the first in its CAZy sequence family [24] to be described, it is likely that many cytoplasmic GTs remain to be discovered. In support of this idea, accumulating evidence for proteins with carbohydrate-binding activity in the nucleus and cytoplasm [48] suggests that novel functions for glycosylation in these compartments are likely to be discovered.

Functions of the HyPro modification pathway in O₂-dependent signaling

Dictyostelium is an obligate aerobe, depending on at least 2.5% O₂ to proliferate and aggregate in response to starvation. However, subsequent development is delayed at this level of O₂, and execution of the slug-to-fruit switch, the decision point for culmination, does not occur if the O₂ level is <12% [49]. The critical time for higher O₂ is at ~12 h of development, during formation of the tipped aggregate and elongation into a slug [5], and the critical duration of higher O₂ is 1-3 h (Z.A. Wang, C.M. West, unpublished data). Consistent with a need for higher O₂, cells must be exposed at an air-water interface (on an agar or filter surface) to progress to later phases of development. If developing cells are submerged as for mammalian cell culture, a hyperoxic atmosphere, which may overcome the diffusion barrier to O₂-transport posed by water, is required to complete development [50].

Though cell respiration rates decline during development [51], higher external O₂ may still be required to drive diffusion into the interior of the cell aggregate. However, high O₂ is only required for a short critical period and, in hypoxia, extensive slime trails were formed indicative of persistent slug migration. These and other (see below) observations suggest that O₂ serves an instructional role in addition to driving oxidative metabolism. For example, as listed in Fig.

1, O₂ influences cell polarity and motility, slug migration, and cell type proportioning in the slug.

Slugs composed of cells lacking P4H1 require even higher levels of O₂, 18–21% or greater, to culminate [5]. In contrast, overexpression of P4H1 under the control of developmental promoters in a wild-type background reduces the O₂-requirement to 2.5–5%. These effects suggest that P4H1 itself mediates O₂-regulation of culmination, and that an override pathway detectable in the absence of P4H1 requires yet higher O₂. P4H1[−] cells just below their O₂-threshold differentiate the normal pattern of prestalk and prespore cell differentiation markers, but not markers of later differentiation [5], just like wild-type slugs formed just below their O₂-threshold.

The exact O₂ threshold for culmination of normal and mutant slugs can differ slightly between trials, which might be explained by variations in other factors that also influence culmination, including light and NH₃ [5]. Near the O₂ threshold, breakthrough often occurs in patches especially near the periphery of the filter, suggesting that an unknown positive feedback process also influences culmination.

The failure of P4H1[−] slugs to culminate below the O₂-threshold appears to be due to an inability to generate rather than respond to a culmination signal, based on the observation that the presence of a small percentage of wild-type cells is sufficient to allow culmination of P4H1[−] cells at the wild-type O₂-level [5]. Interestingly, many of the normal cells in these chimeras migrate to the anterior end of the slug, based on analysis of RFP-labeled strains, and differentiate into prestalk cells. Similarly, cells overexpressing P4H1 migrate to the tip in mixtures with wild-type cells. Tip cells are considered to be an important signaling center for initiating culmination [56,57] (Fig. 1). Complementation of P4H1-null cells by expression of P4H1 under a prestalk cell (tip) promoter also rescues culmination. Expression under a prespore promoter rescues too, though the complemented strain also sorts to the anterior end in chimeras with P4H1-null cells. Finally, P4H1 more effectively reduces the O₂-threshold when expressed in prestalk compared to prespore cells, though notably this effect requires expression in the wild-type background. These findings are consistent with a model in which P4H1 signals culmination from the anterior tip, where the full response may require activity in prespore cells as well.

Dictyostelium P4H1 has a high K_m for O₂ that is well above the ambient level [11]. This suggests that P4H1 activity may be rate-limited by O₂ as described in animals, but this has not been demonstrated directly. P4H1 appears to be expressed evenly in prestalk and prespore cells based on P4H1 promoter::RFP expression studies and Western blotting [5], indicating that different cell types or the tip are not predisposed to respond to O₂. Based on the sorting behavior of cells expressing higher levels of P4H1 (see above), cells exposed to higher O₂ may migrate to the tip and generate a culmination signal. Owing to the narrower diameter of the anterior region of the slug (see Fig. 1), resulting in a higher surface:volume ratio conducive to higher O₂ penetration, prestalk cells might naturally exhibit greater O₂-signaling. Interestingly, slug surface cells also have a propensity to become prestalk cells [50,58]. Although the biological significance of these processes is currently not known, they might underlie a hypersensitivity that allows the entire slug to respond to changes in ambient O₂ sensed by any cell as a sentinel. In the native soil environment, this system may help the slug to determine that it has translocated from a subterranean or submerged location where O₂ is limiting to an above-ground, normoxic position where fruiting body formation would be optimal for spore dispersal. It is interesting that an O₂ gradient may guide slugs to the soil surface [3], but it is not known if P4H1 is required for this response as well.

The significance of HyPro-dependent glycosylation for O₂-signaling has been investigated in the Skp1 glycosylation mutant strains. *pgtA*⁻ cells, which accumulate Gn-O-Skp1, exhibit the wild-type (12%) requirement for O₂, initially suggesting that glycosylation beyond the first sugar is dispensable [5]. The separate contributions P4H1 and Gnt1 have not yet been analyzed because *gnt1* has resisted disruption (C.M. West et al., unpublished data). Surprisingly, subsequent studies showed that *agtA*⁻ cells, which accumulate FGn-O-Skp1, require ~16–18% O₂ to culminate, almost as high as required by P4H1-null cells [26]. Thus PgtA has a function not revealed by the *pgtA*-knockout experiment. Mutants that assemble only 2 sugars, due either to expression of the N-terminal domain of PgtA in a *pgtA*⁻ background or a mutation in *gmd* required to form GDP-Fuc, have an intermediate O₂-dependence between that of *pgtA*- and *agtA*-knockout cells. Therefore, whereas PgtA negates the action of P4H1/Gnt1, AgtA negates the action of PgtA, restoring the initial effect of P4H1/Gnt1. This suggests the existence of multiple signals that control development, contingent upon sufficient O₂ and other factors to initiate the pathway.

Since Skp1 is the only detectable substrate of PgtA-β3GalT or -α2FucT in extracts, the reversal of the P4H1⁻ phenotype by the action of PgtA provides genetic evidence that Skp1 is the functional substrate of P4H1-dependent signaling in cells. This is consistent with unpublished data (Z.A. Wang, C.M. West) that the O₂-threshold for culmination is proportional to Skp1 levels in genetically manipulated cells. Though the majority of steady-state Skp1 appears to be modified throughout the life cycle, nascent Skp1-myc, induced under developmental control of a prespore cell promoter, is very inefficiently hydroxylated [14], suggesting that modification of newly synthesized Skp1 may be rate limiting at the time of O₂-sensing. In addition, proteins that potentially interact with Skp1 in SCF complexes, including CulA, CulB, and FbxA, are required for normal development [12,60–62], though effects on the O₂-threshold remain to be investigated. Future work is needed to determine whether these proteins interfere with or promote Skp1 modifications, and if O₂ signals via the SCF complex(es) to regulate culmination.

The findings suggest the signaling model shown in Fig. 5. Skp1 is predicted to exert a negative effect, via E3(SCF)Ub-ligase-mediated polyubiquitylation and degradation of a hypothetical activator of culmination, based on increased signaling (reduced O₂-requirement) when Skp1 is decreased. Protein kinase A (PKA) acts downstream of Skp1 modification based on suppressive effects of its overexpression on the P4H1-null phenotype [5]. P4H1 is interpreted to exert a negative effect on Skp1 based on the positive effect of O₂ on the overall pathway. P4H1 and Gnt1 are grouped together as their separate roles are unknown. PgtA reverses this effect, and AgtA corrects the reversal, thereby enabling the wild-type level of inhibition of Skp1 that normally activates the pathway. How the glycosylation reactions are regulated is not known, but it is noted that the synthesis of UDP-Gal is developmentally up-regulated near culmination [59], and the PgtA-β3GalT and AgtA-α3GalT activities show sharply contrasting dependence on UDP-Gal concentration in *in vitro* assays (see above). Hyperoxia overrides the P4H1-null effect indicative of another O₂-dependent pathway, as depicted on the right.

At the cellular level, we propose that even 2.5% O₂ is sufficient to support modification of the majority of Skp1 in the steady state. This degree of modification dampens Skp1 and SCF activity to a level appropriate for SCF-regulated functions. During development, access to O₂ is limited as a physical consequence of multicellularity, so that a higher level of O₂ is required to modify nascent Skp1. Considering the high stability of the Skp1:F-box protein interaction [e.g., 63], we suggest that developmentally regulated F-box proteins which are translated during the critical period for high O₂, and potentially regulate the activity of a hypothetical culmination activator, will be most affected by the posttranslational status of coordinately synthesized Skp1. The biochemical mechanisms of Skp1 inhibition depicted by the model remain to be determined, but may involve changes in SCF subunit assembly, SCF

dimerization, SCF activity, SCF cellular compartmentalization, and/or F-box protein selection. Glycosylation was previously shown to enhance nuclear accumulation of Skp1 in *Dictyostelium* [14].

Whereas down-regulation of PHDs leads to accumulation of HIF α in animals (see Fig. 6 below), knockout of P4H1 does not materially alter Skp1 levels in *Dictyostelium* [5]. Possibly, hydroxylated Skp1 is protected by GlcNAc capping, which might explain why it has not been possible to knock out *gnt1*. If homologs of subunits of the VBC-type Ub-ligase that target hydroxylated HIF α are present in *Dictyostelium*, they are highly diverged (C.M. West, unpublished studies). Modification may be regulated for selected subpopulations of Skp1, such as a nuclear pool, a prestalk/tip-associated pool, or nascently synthesized Skp1. Interestingly, the majority of Skp1-myc expressed in prespore cells is non-hydroxylated [14], suggesting that hydroxylation does not keep pace with Skp1 synthesis during development.

Evidence for Skp1 modification and O₂-regulation in other organisms

As summarized above and in Table 1, sequences related to HyPro modification genes are selectively yet broadly distributed across diverse protist phyla including other amoebazoa, excavates, stramenopiles (heterokonts), green algae, and alveolates (Table 1). Qualifications for inclusion in Table 1 include presence of key motifs identified by comparison with sequences from other amoebazoa or proteins with known activity, and absence of an N-terminal signal peptide motif for targeting to the secretory pathway. Biochemical studies support the bioinformatics evidence for Gnt1 in *T. gondii*, as soluble extracts from tachyzoites have a Skp1 GlcNAcT activity [64].

All eukaryotic genomes with a homolog of Gnt1 also have a homolog of P4H1 and a Skp1 with the equivalent of Pro143 (Table 1). A subset, which thus far includes only examples from other amoebazoe, alveolates and stramenopiles, also possesses a potential homolog of PgtA. Currently, AgtA-like sequences are restricted to the amoebazoe. Though this suggests that the pathway is most highly specialized in *Dictyostelium*, it is equally likely that the core glycan is modified in different ways in the other organisms, potentially by GTs from new CAZy families yet to be discovered. Interestingly, the *Phytophthora* genomes each harbor a second Skp1 gene that lacks the Pro143 equivalent, suggesting the ability to bypass the regulatory pathway.

Despite their apparent absence from animals and plants, the broad phylogenetic distribution of pathway genes in protists suggests existence of the glycosylation pathway in the ancestral eukaryotic lineage and deletion from organisms for which O₂-sensing is not important or for which other O₂-sensing mechanisms are more suitable. The pathway genes seem to be more likely found in unicellular organisms with life cycle stages exposed to varying degrees of hypoxia. *Dictyostelium* and *Toxoplasma* experience a range of normoxic and hypoxic environments, and sporulate or encyst in response to stress. In contrast, the pathway does not occur in another amoebazoan, *Entamoeba*, which lives anaerobically and may have no need for an O₂-response pathway. In the apicomplexan lineage, pathway genes are found in *Toxoplasma gondii* and *Neospora caninum*, related parasites that encounter normoxic, hypoxic and anoxic environments during their life cycles. In contrast, the genes are absent from related *Plasmodium* or *Cryptosporidium* species, which exist in well-oxygenated tissues where O₂ levels vary less. The Skp1 O₂-sensing pathway may not be suited for unicellular organisms that can function in anoxic conditions (e.g., fungi) or do not exist outside of an animal host (e.g., trypanosomatids). However, trypanosomatids do hydroxylate selected thymidine bases of DNA (base J), and cap these with β -D-Glc [65], a paradigm similar to Skp1 modification except for the targeting of DNA instead of protein. The responsible glucosyltransferase remains to be identified, and the potential relation of the modification to hypoxic regulation is unknown.

O₂-sensing pathway components known in other organisms do not occur in *Dictyostelium* and *Toxoplasma*. For example, in bacteria and selected unicellular eukaryotes, PAS domains associated with two-hybrid kinases bind O₂ which activates kinase activity that transcriptionally induce various response genes [66,67]. This class of PAS domains is not found in the *Dictyostelium* and *Toxoplasma* genomes. An O₂-sensing pathway based on O₂-dependent sterol biosynthesis, Sre1, and a PHD-like enzyme Ofd1 that may modify Sre1, is important in fission yeast [68], but the genes for this pathway are absent in *Dictyostelium* and *Toxoplasma*. Similarly, an O₂-dependent heme biosynthetic pathway modulates Hap1 transcriptional factor/repressor activity in yeast [69], but Hap1 is not found in the *Dictyostelium* genome. The class of basic helix-loop-helix transcriptional factors that includes HIF α , the target of the animal ortholog of P4H1, is also absent. Rapid 'bioenergetic' responses to hypoxia, based on O₂-dependent heme oxygenases, cytochrome P450s and NADPH-oxidases that modulate K⁺ and Ca⁺⁺ channels [70], have evolved in vertebrates to control ventilation, but the potential significance of these pathways in unicellular organisms is unexplored.

Skp1 is potentially the sole target of *Dictyostelium* P4H1 whereas the spectrum of targets of the three animal HIF α -type PHDs is probably broader. Since many plant and invertebrate Skp1s have the equivalent of Pro143 in a highly conserved sequence context, Skp1 might be a target of HIF α -type PHDs [71,72] in these organisms even in the absence of a glycan cap as occurs in *Dictyostelium*. The absence of Pro143 in vertebrate Skp1s argues against this possibility unless an alternative Pro residue is targeted. Nevertheless, covalent regulation of Skp1 in unicellular organisms may have evolutionarily fixed Skp1 as a discrete adaptor protein in SCF Ub-ligases.

The comparison of animal O₂-signaling via the PHD- HIF α -VBC module with *Dictyostelium* O₂-signaling via the P4H1-Skp1-SCF module suggests an evolutionarily ancient role for protein turnover in O₂-regulation. The PHDs are apparent orthologs, and the SCF-type and VBC-type Ub-ligases are evolutionarily related with Skp1 of the SCF-complex being conserved with elongin C of the VBC-complex [8] though, notably, elongin C lacks the C-terminal domain of Skp1 that contains Pro143. By modifying the Ub-ligase itself, instead of a target of the Ub-ligase as in animals (depicted in Fig. 6A), *Dictyostelium* P4H1 may directly control the targets of all the E3(SCF)Ub-ligases, which may number up to 50 based on the number of F-box proteins predicted in the *Dictyostelium* genome (C.M. West, H. van der Wel, unpublished data). Therefore prolyl hydroxylation may directly regulate a broader subset of the proteome via polyubiquitination and degradation (Fig. 6B). The proposed *Dictyostelium*-type mechanism may represent a more primitive paradigm, which became focused on a specific transcriptional factor in animals. Hierarchical regulation by subsequent glycosylation suggests new druggable targets involved in developmental transitions of *Toxoplasma* and other parasites but not their host cells.

Acknowledgments

Work described from the authors' laboratory was supported by NIH grants R01-GM-37539 and R01-GM-83483.

ABBREVIATIONS

CAZy	carbohydrate active enzyme database
FucT	fucosyltransferase
GalT	galactosyltransferase
GT	glycosyltransferase

HIF	hypoxia-inducible factor
HyPro	hydroxyproline
PHD	prolyl hydroxylase domain
ROS	reactive oxygen species
SCF	protein complex consisting of Skp1, cullin-1 (or A), and an F-box protein
VBC	protein complex consisting of the von Hippel-Lindau protein (VHL), elongin B, elongin C, and cullin-2
Ub	ubiquitin

References

- Williams RS, Boeckeler K, Gräf R, Müller-Taubenberger A, Li Z, Isberg RR, Wessels D, Soll DR, Alexander H, Alexander S. Towards a molecular understanding of human diseases using *Dictyostelium discoideum*. *Tr Mol Med* 2006;12:415–424.
- West CM. Comparative analysis of spore coat formation, structure, and function in *Dictyostelium*. *Int Rev Cytol* 2003;222:237–293. [PubMed: 12503851]
- Bonner JT, Lamont DS. Behavior of cellular slime molds in the soil. *Mycologia* 2005;97:178–184. [PubMed: 16389969]
- Singleton CK, Kirsten JH, Dinsmore CJ. Function of ammonium transporter A in the initiation of culmination of development in *Dictyostelium discoideum*. *Eukaryot Cell* 2006;5:991–996. [PubMed: 16835443]
- West CM, van der Wel H, Wang ZA. Prolyl 4-hydroxylase-1 mediates O₂-signaling during development of *Dictyostelium*. *Development* 2007;134:3349–3358. [PubMed: 17699611]
- Kaelin WG Jr, Ratcliffe PJ. Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway. *Mol Cell* 2008;30:393–402. [PubMed: 18498744]
- Gonzalez-Yanes B, Cicero JM, Brown RD Jr, West CM. Characterization of a cytosolic fucosylation pathway in *Dictyostelium*. *J Biol Chem* 1992;267:9595–9605. [PubMed: 1577798]
- Willems AR, Schwab M, Tyers M. A hitchhiker's guide to the cullin ubiquitin ligases: SCF and its kin. *Biochim Biophys Acta* 2004;1695:133–170. [PubMed: 15571813]
- Hermant D. F-box proteins: more than baits for the SCF? *Cell Div* 2006;1:30. [PubMed: 17166256]
- Seol JH, Shevchenko A, Shevchenko A, Deshaies RJ. Skp1 forms multiple protein complexes, including RAVE, a regulator of V-ATPase assembly. *Nat Cell Biol* 2001;3:384–391. [PubMed: 11283612]
- van der Wel H, Ercan A, West CM. The Skp1 prolyl hydroxylase from *Dictyostelium* is related to the hypoxia-inducible factor-alpha class of animal prolyl 4-hydroxylases. *J Biol Chem* 2005;280:14645–14655. [PubMed: 15705570]
- Mohanty S, Lee S, Yadava N, Dealy MJ, Johnson RS, Firtel RA. Regulated protein degradation controls PKA function and cell-type differentiation in *Dictyostelium*. *Genes Dev* 2001;15:1435–1148. [PubMed: 11390363]
- Kozarov E, van der Wel H, Field M, Gritzali M, Brown RD Jr, West CM. Characterization of FP21, a cytosolic glycoprotein from *Dictyostelium*. *J Biol Chem* 1995;270:3022–30230. [PubMed: 7852383]
- Sassi S, Sweetinburgh M, Erogul J, Zhang P, Teng-umnuay P, West CM. Analysis of Skp1 glycosylation and nuclear enrichment in *Dictyostelium*. *Glycobiology* 2001;11:283–295. [PubMed: 11358877]
- Teng-umnuay P, Morris HR, Dell A, Panico M, Paxton T, West CM. The cytoplasmic F-box binding protein Skp1 contains a novel pentasaccharide linked to hydroxyproline in *Dictyostelium*. *J Biol Chem* 1998;273:18242–18249. [PubMed: 9660787]

16. Ketcham C, Wang F, Fisher SZ, Ercan A, van der Wel H, Locke RD, Doulahk S, Matta KL, West CM. Specificity of a UDP-galactose:fucoside α 1,3galactosyltransferase that modifies Skp1 in the cytoplasm of *Dictyostelium*. *J Biol Chem* 2004;279:29050–29059. [PubMed: 15123660]
17. West CM, van der Wel H, Sassi S, Gaucher EA. Cytoplasmic glycosylation of protein-hydroxyproline and its relationship to other glycosylation pathways. *Biochim Biophys Acta* 2004;1673:29–44. [PubMed: 15238247]
18. Almond A, Petersen BO, Duus JØ. Oligosaccharides implicated in recognition are predicted to have relatively ordered structures. *Biochemistry* 2004;43:5853–5863. [PubMed: 15134459]
19. Zheng N, Schulman BA, Song L, Miller JJ, Jeffrey PD, Wang P, Chu C, Koepp DM, Elledge SJ, Pagano M, Conaway RC, Conaway JW, Harper JW, Pavletich NP. Structure of the Cul1-Rbx1-Skp1-F box^{Skp2} SCF ubiquitin ligase complex. *Nature* 2002;416:703–709. [PubMed: 11961546]
20. West CM, Kozarov E, Teng-umnay P. The cytosolic glycoprotein FP21 of *Dictyostelium discoideum* is encoded by two genes resulting in a polymorphism at a single amino acid position. *Gene* 1997;200:1–10. [PubMed: 9373134]
21. Schiller B, Hykollari A, Voglmeir J, Pörtl G, Hummel K, Razzazi-Fazeli E, Geyer R, Wilson IBH. Development of *Dictyostelium discoideum* is associated with alteration of fucosylated N-glycan structures. *Biochem J* 2009;423:41–52. [PubMed: 19614564]
22. Teng-Umnay P, van der Wel H, West CM. Identification of a UDP-GlcNAc:Skp1-hydroxyproline GlcNAc-transferase in the cytoplasm of *Dictyostelium*. *J Biol Chem* 1999;274:36392–36402. [PubMed: 10593934]
23. van der Wel H, Morris HR, Panico M, Paxton T, Dell A, Kaplan L, West CM. Molecular cloning and expression of a UDP-N-acetylglucosamine (GlcNAc):hydroxyproline polypeptide GlcNAc-transferase that modifies Skp1 in the cytoplasm of *Dictyostelium*. *J Biol Chem* 2002;277:46328–46337. [PubMed: 12244115]
24. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res* 2009;37(Database issue):D233–238. [PubMed: 18838391]
25. Ercan A, West CM. Kinetic analysis of a Golgi UDP-GlcNAc:polypeptide-Thr/Ser N-acetyl-alpha-glucosaminyltransferase from *Dictyostelium*. *Glycobiology* 2005;15:489–500. [PubMed: 15616122]
26. Wang ZA, van der Wel H, Vohra Y, Buskas T, Boons GJ, West CM. Role of a cytoplasmic dual-function glycosyltransferase in O₂-regulation of development in *Dictyostelium*. *J Biol Chem* 2009;284:28896–28904. [PubMed: 19687007]
27. Belyi Y, Aktories K. Bacterial toxin and effector glycosyltransferases. *Biochim Biophys Acta*. 2009 in press (this issue).
28. Olszewski NE, West CM, Sassi SO, Hartweck LM. O-GlcNAc protein modification in plants: evolution and function. *Biochim Biophys Acta*. 2009 in press.
29. van Der Wel H, Morris HR, Panico M, Paxton T, North SJ, Dell A, Thomson JM, West CM. A non-Golgi alpha 1,2-fucosyltransferase that modifies Skp1 in the cytoplasm of *Dictyostelium*. *J Biol Chem* 2001;276:33952–33963. [PubMed: 11423539]
30. van der Wel H, Fisher SZ, West CM. A bifunctional diglycosyltransferase forms the Fuc α 1,2Gal β 1,3-disaccharide on Skp1 in the cytoplasm of *Dictyostelium*. *J Biol Chem* 2002;277:46527–46534. [PubMed: 12244067]
31. West CM, Scott-Ward T, Teng-umnay P, van der Wel H, Kozarov E, Huynh A. Purification and characterization of an alpha1,2-L-fucosyltransferase, which modifies the cytosolic protein FP21, from the cytosol of *Dictyostelium*. *J Biol Chem* 1996;271:12024–12035. [PubMed: 8662604]
32. West CM, Nguyen P, van der Wel H, Metcalf T, Erdos GW. Dependence of spore stress resistance on a spore coat branched heteropolysaccharide in *Dictyostelium*. *Euk Cell* 2009;8:27–36.
33. Ercan A, Panico M, Sutton-Smith M, Dell A, Morris HR, Matta KL, Gay DF, West CM. Molecular characterization of a novel UDP-galactose:fucoside alpha3-galactosyltransferase that modifies Skp1 in the cytoplasm of *Dictyostelium*. *J Biol Chem* 2006;281:12713–12721. [PubMed: 16495217]
34. Egelund J, Skjöt M, Geshi N, Ulvskov P, Petersen BL. A complementary bioinformatics approach to identify potential plant cell wall glycosyltransferase-encoding genes. *Plant Physiol* 2004;136:2609–2620. [PubMed: 15333752]

35. Egelund J, Petersen BL, Motawia MS, Damager I, Faik A, Olsen CE, Ishii T, Clausen H, Ulvskov P, Geshi N. *Arabidopsis thaliana* RGXT1 and RGXT2 encode Golgi-Localized (1,3)- α -D-xylosyltransferases involved in the synthesis of pectic rhamnogalacturonan-II. *Plant Cell* 2006;18:2593–2607. [PubMed: 17056709]
36. Schaap P, Winckler T, Nelson M, Alvarez-Curto E, Elgie B, Hagiwara H, Cavender J, Milano-Curto A, Rozen DE, Dingermann T, Mutzel R, Baldauf SL. Molecular phylogeny and evolution of morphology in the social amoebas. *Science* 2006;314:661–663. [PubMed: 17068267]
37. Yu L, Gaitatzes C, Neer E, Smith TF. Thirty-plus functional families from a single motif. *Protein Sci* 2000;9:2470–2476. [PubMed: 11206068]
38. Steimle PA, Naismith T, Licate L, Egelhoff TT. WD repeat domains target *Dictyostelium* myosin heavy chain kinases by binding directly to myosin filaments. *J Biol Chem* 2001;276:6853–6860. [PubMed: 11106661]
39. Hart, GW.; West, CM. Unique forms of nucleocytoplasmic glycosylation. In: Varki, A., et al., editors. *Essentials of Glycobiology*. 2. Cold Spring Harbor Laboratory Press; 2009.
40. Funakoshi Y, Suzuki T. Glycobiology in the cytosol: the bitter side of a sweet world. *Biochim Biophys Acta* 2009;1790:81–94. [PubMed: 18952151]
41. West CM, van der Wel H, Gaucher EA. Complex glycosylation of Skp1 in *Dictyostelium*: implications for the modification of other eukaryotic cytoplasmic and nuclear proteins. *Glycobiology* 2002;12:17R–27R.
42. Hanna ES, Roque-Barreira MC, Bernardes ES, Panunto-Castelo A, Sousa MV, Almeida IC, Brocchi M. Evidence for glycosylation on a DNA-binding protein of *Salmonella enterica*. *Microb Cell Fact* 2007;6:11. [PubMed: 17407574]
43. Weart RB, Lee AH, Chien AC, Haeusser DP, Hill NS, Levin PA. A metabolic sensor governing cell size in bacteria. *Cell* 2007;130:335–347. [PubMed: 17662947]
44. Fischer W, Haas R. The RecA protein of *Helicobacter pylori* requires a posttranslational modification for full activity. *J Bacteriol* 2004;186:777–784. [PubMed: 14729704]
45. Van Etten JL, Gurnon JR, Yanai-Balser GM, Dunigan DD, Graves MV. Chlorella viruses encode most, if not all, of the machinery to glycosylate their glycoproteins independent of the endoplasmic reticulum and Golgi. *Biochim. Biophys Acta*. 2009 in press.
46. Takano H, Takechi K. Plastid peptidoglycan. *Biochim Biophys Acta*. 2009 in press.
47. Hart GW, Haltiwanger RS, Holt GD, Kelly WG. Glycosylation in the nucleus and cytoplasm. *Annu Rev Biochem* 1989;58:841–874. [PubMed: 2673024]
48. Lannoo N, Van Damme EJM. Nucleocytoplasmic plant lectins. *Biochim Biophys Acta*. 2009 in press.
49. Sandona D, Gastaldello S, Rizzuto R, Bisson R. Expression of cytochrome c oxidase during growth and development of *Dictyostelium*. *J Biol Chem* 1995;270:5587–5593. [PubMed: 7890678]
50. Sternfeld J, Bonner JT. Cell differentiation in *Dictyostelium* under submerged conditions. *Proc Natl Acad Sci, USA* 1977;74:268–271. [PubMed: 264680]
51. Liddel GU, Wright BE. The effect of glucose on respiration of the differentiating slime mold. *Dev Biol* 1961;3:265–276. [PubMed: 13761944]
52. Sternfeld J, David CN. Oxygen gradients cause pattern orientation in *Dictyostelium* cell clumps. *J Cell Sci* 1981;50:9–17. [PubMed: 7320079]
53. Bonner JT. Evolution of development in the cellular slime molds. *Evol Develop* 2003;5:305–313.
54. Sawada Y, Maeda Y, Takeuchi I, Williams J, Maeda Y. Rapid patterning of *Dictyostelium discoideum* cells under confined geometry and its relation to differentiation. *Develop Growth Differ* 1998;40:113–120.
55. Sternfeld J. Proportion regulation in *Dictyostelium* is altered by oxygen. *Differentiation* 1988;37:173–179.
56. Smith E, Williams KL. Evidence for tip control of the ‘slug/fruit’ switch in slugs of *Dictyostelium discoideum*. *J Embryol Exp Morphol* 1980;57:233–240. [PubMed: 7430932]
57. Kirsten JH, Xiong Y, Dunbar AJ, Rai M, Singleton CK. Ammonium transporter C of *Dictyostelium discoideum* is required for correct prestalk gene expression and for regulating the choice between slug migration and culmination. *Dev Biol* 2005;287:146–156. [PubMed: 16188250]

58. Wilkins MR, Williams KL. The extracellular matrix of the *Dictyostelium discoideum* slug. *Experientia* 1995;51:1189–1196. [PubMed: 8536806]
59. Telser A, Sussman M. Uridine diphosphate galactose-4-epimerase, a developmentally regulated enzyme in the cellular slime mold *Dictyostelium discoideum*. *J Biol Chem* 1971;246:2252–2257. [PubMed: 4252223]
60. Nelson MK, Clark A, Abe T, Nomura A, Yadava N, Funair CJ, Jermyn KA, Mohanty S, Firtel RA, Williams JG. An F-Box/WD40 repeat-containing protein important for *Dictyostelium* cell-type proportioning, slug behaviour, and culmination. *Dev Biol* 2000;224:42–59. [PubMed: 10898960]
61. Ennis HL, Dao DN, Pukatzki SU, Kessin RH. *Dictyostelium* amoebae lacking an F-box protein form spores rather than stalk in chimeras with wild type. *Proc Natl Acad Sci USA* 2000;97:3292–3297. [PubMed: 10725352]
62. Tekinay T, Ennis HL, Wu MY, Nelson M, Kessin RH, Ratner DI. Genetic interactions of the E3 ubiquitin ligase component FbxA with cyclic AMP metabolism and a histidine kinase signaling pathway during *Dictyostelium discoideum* development. *Eukaryot Cell* 2003;2:618–626. [PubMed: 12796307]
63. Xu K, Belunis C, Chu W, Weber D, Podlaski F, Huang KS, Reed SI, Vassile LT. Protein protein interactions involved in the recognition of p27 by E3 ubiquitin ligase. *Biochem J* 2003;371:957–964. [PubMed: 12529174]
64. West CM, van der Wel H, Blader JJ. Detection of cytoplasmic glycosylation associated with hydroxyproline. *Meth Enzymol* 2006;417:389–404. [PubMed: 17132515]
65. Borst P, Sabatini R. Base J: discovery, biosynthesis, and possible functions. *Annu Rev Microbiol* 2008;62:235–251. [PubMed: 18729733]
66. Gilles-Gonzalez MA, Gonzalez G, Sousa EHS, Tuckerman J. Oxygen-sensing histidine-protein kinases: assays of ligand binding and turnover of response-regulator substrates. *Meth Enzymol* 2008;437:173–189. [PubMed: 18433629]
67. Müllner M, Hammel O, Mienert B, Schlag S, Bill E, Uden G. A PAS domain with an oxygen labile [4Fe-4S](2+) cluster in the oxygen sensor kinase NreB of *Staphylococcus carnosus*. *Biochemistry* 2008;47:13921–13932. [PubMed: 19102705]
68. Lee CY, Stewart EV, Hughes BT, Espenshade PJ. Oxygen-dependent binding of Nro1 to the prolyl hydroxylase Ofd1 regulates SREBP degradation in yeast. *EMBO J* 2009;28:135–143. [PubMed: 19158663]
69. Hon T, Dodd A, Dirmeier R, Gorman N, Sinclair PR, Zhang L, Poyton RO. A mechanism of oxygen sensing in yeast. *J Biol Chem* 2003;278:50771–50780. [PubMed: 14512429]
70. Ward JP. Oxygen sensors in context. *Biochim Biophys Acta* 2008;1777:1–14. [PubMed: 18036551]
71. Asif MH, Trivedi PK, Misra P, Nath P. Prolyl-4-hydroxylase (AtP4H1) mediates and mimics low oxygen response in *Arabidopsis thaliana*. *Funct Integr Genomics*. 2009 Mar 10; [Epub ahead of print].
72. McDonough MA, Li V, Flashman E, Chowdhury R, Mohr C, Liénard BM, Zondlo J, Oldham NJ, Clifton IJ, Lewis J, McNeill LA, Kurzeja RJ, Hewitson KS, Yang E, Jordan S, Syed RS, Schofield CJ. Cellular oxygen sensing: Crystal structure of hypoxia-inducible factor prolyl hydroxylase (PHD2). *Proc Natl Acad Sci USA* 2006;103:9814–9819. [PubMed: 16782814]

Fig. 1.

Dictyostelium development and O₂-dependence. Starvation induces *Dictyostelium* cells to aggregate by chemotaxis and cell-cell adhesion. As depicted here, the tipped aggregate elongates into a slug, and the inherited tip (colored red) coordinates subsequent migration and culmination into a fruiting body. Cell rearrangements are indicated by arrows. During culmination, prestalk cells (Pst, 3 subtypes are denoted as shades of red) migrate down and become stalk cells, and prespore cells (Psp, blue) ascend and become spores. Culmination requires a high level of O₂, which evidence suggests signals via the HyPro-modification pathway. ALC: anterior-like cells residing in the posterior region. *Dictyostelium* normally requires only 2.5% O₂ for growth, but needs elevated O₂ to form slugs in timely fashion, culminate and sporulate (as indicated by green arrows). O₂ is also involved in other developmental functions as listed at bottom.

Fig. 2.

Skp1 hydroxylation/glycosylation pathway in *Dictyostelium*. The enzymes for the first 5 modification steps have been identified and their genes now cloned. *gmd*, mutated in strain HL250 [21], is required in concert with an epimerase (not shown) to form GDP-Fuc from GDP-Man. α Gal-2, whose addition may be mediated by AgtA (see text), is most likely attached to α Gal-1 or Fuc. The dotted line indicates that P4H1 and Gnt1 appear to be expressed as a protein fusion in some species (Table 1).

Fig. 3.

$\alpha 2$ FucT sequence motifs, from an alignment of PgtA(C-terminal domain)-like sequences from CAZy family GT74. Each eukaryotic and prokaryotic example is a predicted cytoplasmic protein based on the absence of apparent N-terminal signal sequences. Amino acid names are colored according to their chemical class (hydrophobic, small, positive charge, negative charge or amide/hydrophilic/other). Based on the assumption that all three amoebazoan sequences have the same $\alpha 2$ FucT function, positions occupied by amino acids of the same chemical class are highlighted in color according a different color scheme (hydrophobic, small/hydrophilic/Pro, positive charge, negative charge/amide). Positions conserving this quality in other sequences are highlighted similarly. Three conserved sequence motifs predicted to be associated with $\alpha 2$ FucT activity are indicated (M1–M3). Positions of amino acid identity in all eukaryotic sequences are bolded. Sequence origins are listed at the bottom.

Fig. 4.
 α GalT sequence motifs. Alignment of *Dictyostelium* AgtA with related sequences from proteins of CAZy family GT77, predicted to be expressed in the cytoplasm based on absence of N-terminal targeting motifs. In addition, a known α 3XyIT (AtRGXT1) expressed in the Golgi of *Arabidopsis* is included on the bottom row for comparison. Four conserved sequence motifs from the predicted catalytic domains, including AtRGXT1, are indicated (M1–M4). Additional motifs not highlighted distinguish the amoebazoan sequences from the others. There are currently no motifs detected that differentiate predicted cytoplasmic viridiplantae sequences from predicted Golgi-associated sequences, and therefore the assignments as cytoplasmic should be considered provisional. Only the amoebazoan sequences (top 3 rows) are associated with WD-40 repeats predicted to fold into a 7-bladed -propeller. See Fig. 3 for color-coding information. Sequence origins are listed at the bottom.

Fig. 5.

Signaling pathway model. A proposed scheme for O_2 -signaling via the Skp1 modification enzymes, Skp1 (as a subunit of E3(SCF)Ub-ligases), a hypothetical activator, and protein kinase A (PKA). See text for explanation. The dependent sequence of posttranslational modifications are shown vertically at the left, and the dependent sequence of developmental transitions is shown vertically on the right. Speculative aspects of the model, such as regulation of glycosyltransferases by Fuc and Gal (at the left), are denoted in pink. High O_2 can override the pathway, as shown at the right.

A. Animal Model:

B. Protist Model:

Fig. 6.

Relation of Skp1 hydroxylation/glycosylation to HIF1 hydroxylation. (A) Model depicting the role of O_2 -dependent prolyl hydroxylation of animal HIF1 α in E3(VBC)Ub-ligase dependent polyubiquitination and proteasomal degradation. (B) Model describing how, by shifting the target of hydroxylation from a substrate of the Ub-ligase to a subunit of the Ub-ligase, hydroxylation (and subsequent glycosylation) might regulate a larger number of targets. For culmination, hydroxylation or α -galactosylation of Skp1 is inferred to weaken SCF activity (see Fig. 5). Note that the SCF- and VBC-type Ub-ligases are evolutionarily related. The SCF complex consists of cullin-1, Skp1, and an F-box protein; the VBC complex consists of cullin-2, elongin-B, elongin-C, and the von Hippel-Lindau protein. Skp1 is related to elongin-C, but possesses a unique C-terminal extension which carries the glycosylation site.

Table 1

Phylogenetic distribution of HyPro-modification pathway-like genes

Supergroup/organism	Subgroup	Skp1	phyA	gntI (Gnt51)	pgtA (FT85)	agtA (GT78) ⁶
		Pro143 ¹	P4HI ²	αGlcNAcT ³	β3GalT ⁴	α2FucT ⁵ α3GalT/αGalT
<u>Amoebozoa</u>						
<i>Dictyostelium discoideum</i>	mycetezoan	++	++ ⁷	++	++ ++ ⁸	++
<i>D. purpureum</i>	mycetezoan	+	+ ⁹	+	+	++
<i>Physarum polycephalum</i>	mycetezoan	+	+	+	+	++
<i>Entamoeba histolytica</i>	archamoebae	+				
<u>Alveolates</u>						
<i>Toxoplasma gondii</i>	apicomplexan	+	+	++	++	+ ¹⁰ + ¹¹
<i>Neospora caninum</i>	apicomplexan	+	+	+	+	+
<i>Plasmodium falciparum</i>	apicomplexan	+			+	+
<i>Tetrahymena thermophila</i>	ciliate	+		+		
<i>Karlodinium micrum</i>	dinoflagellate	+	+	+		
<u>Stramenopiles (Heterokonts)</u>						
<i>Thalassiosira pseudonana</i>	brown alga/diatom	+	+	+	+	+
<i>Phytophthora sojae</i>	oomycete	+	+	+	+	+
<i>P. infestans</i>	oomycete	+	+	+	+	+
<i>P. ramorum</i>	oomycete	+	+	+	+	+
<i>Aureococcus anophagefferens</i>	pelagophyte	+		+	+	+
<u>Plantae (viridiplantae)</u>						
<i>Chlamydomonas reinhardtii</i>	green alga	+	+	+		
<i>Volvox carteri</i>	green alga	+	+	+	+	+
<i>Arabidopsis thaliana</i>	vascular plant	+	+	+	+	+
<u>Excavates</u>						
<i>Malawimonas californiana</i>	malawimonad	+	+	+		
<i>Naegleria gruberii</i>	heterolobosea	+	+	+		
<u>Opisthokonts</u>						
Fungi/yeast		+	+			
Invertebrate animals		+	+			

Supergroup/organism	Subgroup	Skp1	phyA	gntI (Gnt51)	pgtA (FT85)	agtA (GT78) ⁶
Vertebrates		Pro143 ¹	P4HI ²	α GlcNAcT ³	β 3GalT ⁴	α 3GalT/ α GalT
Bacteria		-	+			
<i>Trichodesmium erythraeum</i> (strain IMS101)	marine cyanobacterium	+	+	+		
<i>Desulfovibrio desulfuricans</i>	gram positive				+	+ ¹

¹ updated from [20]

² updated from [17]

³ updated from [17]

⁴ updated from [41]

⁵ criteria based on Fig. 3

⁶ criteria based on Fig. 4

⁷ ++ genomic and biochemical evidence

⁸ ____ denotes activities are linked within same (predicted) protein

⁹ genomic evidence

¹⁰ r: physical order of domains is reversed relative to *Dictyostelium* *pgtA*

¹¹ biochemical (not genomic evidence)