Fungal Genetics and Biology 47 (2010) 727–735

Contents lists available at ScienceDirect

Fungal Genetics and Biology

journal homepage: www.elsevier.com/locate/yfgbi

Review

Protein glycosylation in the phytopathogen *Ustilago maydis*: From core oligosaccharide synthesis to the ER glycoprotein quality control system, a genomic analysis

Alfonso Fernández-Álvarez, Alberto Elías-Villalobos, José I. Ibeas *

Centro Andaluz de Biología del Desarrollo, Universidad Pablo de Olavide – Consejo Superior de Investigaciones Científicas, Carretera de Utrera km.1, 41013 Seville, Spain

ARTICLE INFO

Article history:
Received 13 April 2010
Accepted 1 June 2010
Available online 8 June 2010

Keywords:
N-glycosylation
O-glycosylation
Ustilago maydis
Fungal pathogenesis
ER quality control
Glucosidases
Calnexin

ABSTRACT

The corn smut fungus *Ustilago maydis* has, over recent decades, become established as a robust pathogenic model for studying fungi–plant relationships. This use of *U. maydis* can be attributed to its biotrophic host interaction, easy culture and genetic manipulation in the laboratory, and the severe disease symptoms it induces in infected maize. Recent studies have shown that normal protein glycosylation is essential for pathogenic development, but dispensable for the saprophytic growth or mating. Given the relevance of protein glycosylation for *U. maydis* virulence, and consequently its role in the plant pathogenesis, here we review the main actors and events implicated in protein glycosylation. Furthermore, we describe the results of an *in silico* search, where we identify all the conserved members of the N- and O-glycosylation pathways in *U. maydis* at each stage: core oligosaccharide synthesis, addition of the core oligosaccharide to nascent target proteins, maturation and extension of the core oligosaccharide, and the quality control system used by the cell to avoid the presence of unfolded glycoproteins. Finally, we discuss how these genes could affect *U. maydis* virulence and their biotechnological implications.

1. Introduction

Most host–pathogen interactions involve a complex, and often poorly understood, crosstalk between the pathogen and its host, as the host tries to avoid pathogenic infection and the pathogen evolves in order to overcome the defense response (de Wit, 2007; van de Veerdonk et al., 2008; Stukenbrock and McDonald, 2009). There are a variety of questions specific to interactions between fungal pathogens and their plant hosts, such as how a fungus is able to penetrate into the plant, passing through several plant defense mechanisms such as the insulating plant cuticle (Tucker and Talbot, 2001; Gohre and Robatzek, 2008). This complex plant-fungi relationship becomes even more interesting in the case of biotrophic interactions where a pathosystem between fungal pathogens and colonized plant is established. Since the fungus needs to maintain plant cells alive in order to ensure its survival, this crosstalk is finely controlled by the fungus so that the plant response is attenuated. To better understand this type of crosstalk, we work with the corn smut fungus *Ustilago maydis*, an excellent model organism for the study of biotrophic relationships (Kamper et al., 2006; Brefort et al., 2009). *U. maydis* can grow as yeast in an asexual life cycle but when two sexually compatible cells meet on the plant surface, they mate to form a dikaryont filament and thus enter its sexual cycle. Sexually compatible Ustilago cells harbor different alleles at the mating type *loci a* and *b*. *Locus a* encodes a pheromone/receptor system, while the multi-allelic *locus b* encodes a pair of homeodomain proteins that are only active as a heterodimeric transcription factor when subunits are derived from different alleles, *bE* and *bW* (Banuett and Herskowitz, 1989; Kamper et al., 1995; Feldbrugge et al., 2004). Once formed as an active heterodimer, *bE*/*bW* acts as a master regulator in triggering the pathogenic program (Brachmann et al., 2001; Wahl et al., 2010). Under these conditions, cells undergo G2-phase cell cycle arrest and initiate polar growth (Garcia-Muse et al., 2003; Perez-Martin and Castillo-Lluva, 2008). When a combined physical–chemical signal is sensed by the fungus, it triggers development of an invasive morphogenic structure called the appressorium, which mediates fungal penetration into the plant (Mendoza-Mendoza et al., 2009). Once inside the plant, hyphae reactivate their cell cycle, permitting mycelium extension and colonization of plant tissues. An unknown signal is emitted by the fungus to induce the formation of plant tumors containing diploid fungal spores. *U. maydis* completes its pathogenic cycle following release from the plant and spore germination, where the fungus returns to a yeast-like morphology, (Kamper et al., 2006; Brefort et al., 2009). The infection process involves secreted fungal effectors which interact with a variety of plant responses. The *U. maydis* genome encodes a large number of effectors, many of which are organized into gene clusters that are upregulated during the pathogenic development. Effector genes generally encode secreted proteins

* Corresponding author. Fax: +34 954349376.
E-mail address: joibecor@upo.es (J.I. Ibeas).

1087-1845/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.fgb.2010.06.004

(Kamper et al., 2006). The majority of secreted proteins are subject to post-translational modification in the endoplasmic reticulum and Golgi apparatus, with protein glycosylation being the most frequent modification (Peberdy, 1999). Therefore, protein glycosylation is likely to have an important role in modulating the *U. maydis*-maize pathosystem.

Glycosylation consists of the addition of sugar cores to target proteins and is the most common post-translational protein modification in eukaryotes (Lehle et al., 2006). There are two main types of protein glycosylation, N-glycosylation and O-glycosylation, which differ in the type of core oligosaccharide and amino acid residue to which it is attached (Lehle et al., 2006). The synthesis of the polypeptide and the core oligosaccharide occur in the ER. The covalent link between these two structures and the maturation of the glycoprotein varies depending on the type of glycosylation. Once the glycoproteins have been correctly folded, and have got past the ER quality control mechanism in the case of N-glycoproteins, they are transported to the Golgi where they are modified further (Parodi, 2000; Helenius and Aebi, 2004; Lommel and Strahl, 2009). Glycosylation has diverse functions as it can be necessary for correct protein conformation, the stabilization of proteins against denaturation, and the mediation of host–pathogen protein interactions (Helenius and Aebi, 2001). In *Candida albicans* and *Cryptococcus neoformans* glycosylation is essential for virulence (Rouabhia et al., 2005; Olson et al., 2007; Lengeler et al., 2008). Although the role of glycosylation is poorly understood in the case of plant pathogens, it has recently been shown to be crucial for virulence in *U. maydis* (Fernandez-Alvarez et al., 2009; Fernandez-Alvarez et al., 2010). Surprisingly, the *U. maydis* genes *pmt4* and *gas1*, involved in protein O- and N-glycosylation respectively, are essential for pathogenic development but dispensable for normal completion of the saprophytic cycle. These mutants can grow in a yeast-like form and mate, but appressorium formation and plant penetration, in the case of the *pmt4* mutant, and the first steps of plant tissue invasion in *gas1* mutants, are defective (Schirawski et al., 2005; Fernandez-Alvarez et al., 2009). Although the molecular causes of these phenotypes are still being investigated, the fact that these *U. maydis* glycosylation mutants only affect plant-related processes, suggests a critical role for glycosylation in plant-fungus crosstalk.

In this paper, we have identified all of the *U. maydis* orthologs implicated in the synthesis of glycoproteins at different stages: glycosylation substrate production, core oligosaccharide synthesis, the addition, trimming and modification of the sugar structure to target proteins, and proteins implicated in the glycoprotein quality control mechanism.

We have carried out protein blast searches using *Saccharomyces cerevisiae* proteins from the Saccharomyces genome database as query sequences against the 6786 proteins defined in the MIPS *U. maydis* protein database. We used whole protein sequences for each query and required a degree of homology defined by the expectation value (*E* value) of candidate homologs. A low *E* value supposes a high degree of alignment/similarity between the query sequence and the candidate homolog. We have been very conservative in our selection, requiring *E* values of $\leqslant 10^{-20}$ for us to consider candidate sequences as putative *U. maydis* protein homologs. We will describe the level of conservation of the major components of the glycosylation pathways in *U. maydis* and discuss the possible implications of this conservation in terms of its virulence and biotechnological applications.

2. N-glycosylation in *U. maydis*

N-glycosylation, the most common type of eukaryotic protein glycosylation, involves the linkage of an oligosaccharide core to Asparagine (Asn) residues (Spiro, 2002). The sugar core is composed of two N-acetylglucosamines, nine mannoses and three glucoses ($Glc_3Man_9GlcNAc_2$) assembled together by various types of linkage (Supplementary Fig. 1). This structure is widely conserved from yeast to humans and is the result of a complex pathway involving several proteins, many of which are essential for cell viability (Helenius and Aebi, 2004).

The *U. maydis* pathways involved in the synthesis of mannose and glucose residues required to assemble the sugar core have been described previously (Hewald et al., 2005; Saavedra et al., 2008) and will not be discussed further here.

2.1. Production of protein glycosylation substrates in U. maydis

The hexosamine pathway that produces UDP-N-acetylglucosamine (UDP-GlcNAc) from fructose 6-phosphate is highly conserved, being essential for many cellular process in yeast, *Caenorhabditis elegans* and mammalian cells (Love and Hanover, 2005; Johnston et al., 2006; Milewski et al., 2006). UDP-GlcNAc is not only used in protein N-glycosylation, since chitin and UDP-N-acetylgalactosamine (UDP-GalNAc) are also obtained from UDP-GlcNAc. In *S. cerevisiae,* this pathway is composed of four sequential steps catalyzed by glutamine-fructose-6-phosphate amidotransferase, glucosamine-6-phosphate acetyltransferase, N-acetylglucosamine-phosphate mutase and UDP-N-acetylglucosamine pyrophosphorylase (Milewski et al., 2006). Analysis of the *U. maydis* genome reveals that the corn fungus contains putative protein homologs for all of the enzymes which catalyze this pathway. UDP-glucose-4-epimerase (Um11600 in *U. maydis*) catalyzes the reversible UDP-GlcNAc to UDP-N-acetylgalactosamine step which is also a common substrate in some types of protein glycosylation (Table 1 and Supplementary Fig. 2). Thus, all the core metabolic pathways involved in the production of the different of substrates used in protein glycosylation seem to be conserved in *U. maydis.*

2.2. Sugar assembly: core oligosaccharide synthesis

In all eukaryotes, the synthesis of the core oligosaccharide, the first stage of N-glycosylation pathway, occurs on both sides of the ER membrane. The pathway starts with the synthesis of dolichol, which is further phosphorylated to generate dolichol phosphate (Dol-P), a lipid carrier. The formation of the initial core oligonucleotide is achieved by associating Dol-P with N-acetylglucosamines, derived from UDP-GlcNAc, followed by the addition of mannoses and glucoses from the nucleotide-activated sugar donors, GDP-mannose (GDP-man) and UDP-glucose (UDP-Glc) (Burda and Aebi, 1999). First, two N-acetylglucosamines and five mannoses are associated with Dol-P on the cytosolic surface of the ER. This structure is moved to the lumen side of the ER by an ATP-independent, bi-directional flippase (Burda and Aebi, 1999; Helenius and Aebi, 2002). Core oligosaccharide synthesis finishes with the addition of four more mannoses and three glucoses. Only after the addition of these sugars is the core transferred to selected Asn residues of nascent proteins by the oligosaccharyltransferase (OST) (Kaplan et al., 1987; Burda et al., 1996; Burda and Aebi, 1998; Knauer and Lehle, 1999; Helenius and Aebi, 2002; Jones et al., 2005).

A protein blast search of the *U. maydis* protein database using *S. cerevisiae* query sequences shows that the *U. maydis* genome contains a sequence homologous to the gene encoding dolychol phosphate mannose synthase (Um06329). Although the *U. maydis* protein Um04779 shows a low *E* value (8.2×10^{-29}) in BLAST analysis when the whole protein is compared to *S. cerevisiae* dolichol kinase, their kinase domains appear highly conserved. We also identified a UDP-N-acetylglucosamine-1-P transferase (Um10484) which adds the first GlcNAc to dolychol phosphate. In *S. cerevisiae,* Alg13/Alg14p attaches a second GlcNAc in a process that is thought