NOTES

Identification of the Archaeal alg7 Gene Homolog (Encoding N-Acetylg glucosamine-1-Phosphate Transferase) of the N-Linked Glycosylation System by Cross-Domain Complementation in *Saccharomyces cerevisiae*

Hosam Shams-Eldin,† Bonnie Chaban,‡ Sebastian Niehus, Ralph T. Schwarz, and Ken F. Jarrell

*Institute for Virology, AG Parasitologie BMFZ, Philipps-University Marburg, Hans-Meerwein-Str. 2, 35043 Marburg, Germany*; *Department of Microbiology and Immunology, Queen’s University, Kingston, Ontario K7L 3N6, Canada*; and *Unité Glycobiologie Structurale et Fonctionnelle UMR CNRS/USTL n 8576 - IFR 147, Université Sciences et Technologies de Lille, 59655 Villeneuve D’Ascq cedex, France*

Received 9 November 2007/Accepted 23 December 2007

The Mv1751 gene product is thought to catalyze the first step in the N-glycosylation pathway in *Methanococcus voltae*. Here, we show that a conditional lethal mutation in the alg7 gene (N-acetylglucosamine-1-phosphate transferase) in *Saccharomyces cerevisiae* was successfully complemented with Mv1751, highlighting a rare case of cross-domain complementation.

The process of covalently attaching a carbohydrate glycan to asparagine residues within a protein is termed N-linked glycosylation. The vast majority of eukaryotes (fungi, plants, animals, slime molds, and euglenas) synthesize asparagine-linked glycosylation (ALG) by initiation of the synthesis of lipid-linked oligosaccharides in a cyclic pathway, the dolichol cycle (11, 12, 14, 19, 22). Although the body of evidence recognizing that *Archaea* also N glycosylate proteins has existed for some time (9, 17), only recently have genes responsible for assembling glycans and attaching them to archaeal proteins been experimentally identified (1, 4). These genes, termed agl (archaeal glycosylation) genes, are recognized as possessing many similarities to their counterparts in the eukaryotic and/or bacterial domains. For example, the oligosaccharyl transferase which transfers the complete glycan to its target protein and creates the N-glycosidic bond is a STT3 homolog in all three domains, although in prokaryotes it acts as a single protein and in eukaryotes it acts as part of a multiprotein oligosaccharyl transferase complex (3, 20).

Research into the N-glycosylation pathway in *Archaea* in our laboratory has used the methanogen *Methanococcus voltae* as a model organism. The flagella and surface layer proteins in this species are known to possess a trisaccharide composed of N-acetylg glucosamine (GlcNAc) (linked to asparagine residues), followed by a diacetylated glucuronic acid and an acetylated mannuronic acid with the amino acid threonine covalently attached at position 6 (21). The genes encoding the glycosyl transferases responsible for the assembly of the second and terminal sugar, as well as for the oligosaccharyl transferase, have been experimentally confirmed (4; B. Chaban and K. F. Jarrell, unpublished results).

The glycol transferase gene Mv1751 was identified in *M. voltae* as the most likely gene to encode the enzyme to carry out the first step in this pathway, i.e., the attachment of a GlcNAc residue to a dolichol lipid carrier in the membrane. This is the only protein in the *M. voltae* genome that belongs to Pfam family PF00953 (glycosyl transferase family 4), which includes the eukaryotic N-acetylglucosamine-1-phosphate (GlcNAc-1-P) transferase (encoded by the alg7 gene). This enzyme catalyzes the conversion of UDP-N-acetyl-D-glucosamine and dolichyl phosphate to UMP and N-acetyl-D-glucosaminyl-diphosphodolichol. It is known that alg7 is an essential gene that is responsible for the first committed step in eukaryotic N glycosylation (3). Unlike bacteria that assemble N-linked glycans on undecaprenol as the lipid carrier, both archaea and eukaryotes use dolichol as a carrier (9).

Conceptual translation of Mv1751 (317 amino acids) revealed 25% amino acid sequence identity and 41% similarity with the *S. cerevisiae* ALG7/GPT protein, where blocks of homology were distributed over the entire length of the protein (Fig. 1). Furthermore, the hydrophobicity plots (13) revealed multiple hydrophobic regions that could function as transmembrane segments. Analysis of the sequence of Mv1751 using the Hsaur software program package (Unix Sequence Analysis Resources, DKFZ, Heidelberg, Germany) revealed the presence of seven predicted transmembrane segment domains interspersed by hydrophilic/less-hydrophobic areas (data not shown). Two consensus motifs for N-linked glycosylation (N-X-T/S) were also detected. Similar motifs were shown in other
GPT proteins, but there is no direct evidence as to whether or not these sites are glycosylated in the native protein. Also, the Mv1751 protein contains a conserved pair of aspartate residues, D-101 and D-102, in a position topologically similar to that of other GPT proteins, which are part of the highly conserved DDXXD motif (Fig. 1) that binds Mg$^{2+}$ (5, 16). These motifs are predicted to be located on the cytoplasmic loop 2 in yeast GPT as well as Mv1751 (data not shown). As found in most cases, a basic residue (Arg or Lys) (18) is located a few residues downstream of the DDXXD motif (Fig. 1), presumably implicated also in binding of metal ions (2). Furthermore, the archaeal VFPGDT motif is highly conserved, with similarity to the equivalent region around Asp-287 in the yeast GPT (283-VFVGDT-288), and it is believed to be the GPT active-site nucleophile region (18) (Fig. 1).

In view of the sequence similarity to the yeast ALG7 and the Pfam motif, topology, and hydrophobicity comparisons, we reasoned that Mv1751 may be a functional homolog of the yeast gene and might be able to complement a yeast alg7 mutant. Such an approach has already proven successful for identification of the human alg7 gene (8). Therefore, we tested the ability of Mv1751 to suppress the functional loss of a conditional lethal yeast alg7 mutant.

As a first step, we amplified the M. voltae Mv1751 coding region (GenBank accession no. DQ372942.1) with the primers MvALG7_forward (5'-GGGGAATTCATGGCTAAAAGCGGTGAATTTATG-3') and MvALG7_reverse_FLAG (GGGCTCGAGCTTGTCATCGTCGCTTGTAGTCAATTATTTAACCGTCGAATTTAC). The primers not only introduced EcoRI or XhoI sites at either end but also incorporated a FLAG tag on the C-terminal end of the amplification product, which was then digested with EcoRI/XhoI and ligated into a similarly restricted yeast shuttle vector containing URA3 as a selective marker (carried on pRS426Met) (6), yielding the pRS426Met-Mv1751 plasmid. This plasmid was then transformed into a conditional lethal yeast mutant of the ALG7.
gene, YPH499-HIS-GAL-ALG7 (15), as previously described (8), and transformants were streaked on minimal medium lacking histidine and uracil. The conditional lethal mutation was made by replacing the native \textit{alg7} promoter with the selection marker/promoter HIS3/GAL1 cassette, thereby eliminating the strain’s requirement for histidine supplementation and placing the \textit{alg7} gene under the regulation of GAL1. GAL1 has been shown to be induced in the presence of galactose and to be tightly repressed in the presence of glucose. The expression of the endogenous \textit{ALG7} can be turned off by shifting the yeast cells from galactose-containing medium (YPGR medium [1% yeast extract, 2% peptone, 4% galactose, 2% raffinose]), in which the promoter is not repressed, to glucose-containing medium (synthetic dextrose minimal medium containing 5% glucose). Mutant cells transformed with plasmids carrying the human \textit{ALG7} (Fig. 2, quadrant C) (positive control [8]) or Mv1751 (Fig. 2, quadrant D) genes displayed slow but sustained growth and were able to suppress the functional loss of the \textit{ALG7} mutant and grow on glucose. Conversely, cells transformed with the control plasmid pRS426Met failed to grow (Fig. 2, quadrant B). This result indicates that the transformed \textit{M. voltae} Mv1751 gene is indeed functional in yeast.

To visualize the Mv1751 protein and its localization within the yeast cells, indirect immunofluorescence microscopy of methanol/acetone-treated yeast spheroplasts using mouse anti-FLAG and goat anti-mouse immunoglobulin G conjugated with rhodamine was employed. \textit{ALG7} is an endoplasmic reticulum (ER) membrane-bound enzyme in eukaryotes (10), and Mv1751 also has multiple potential transmembrane domains, suggesting a membrane location. Consistent with the growth data shown in Fig. 2, the immunofluorescence data demonstrate that the yeast cells were synthesizing detectable levels of the Mv1751 protein and the expressed protein displayed predominantly a reticular pattern, including staining of the nuclear rim, a distribution typically observed for proteins localized to the ER. When costained with the ER marker protein BiP (binding protein), a nearly complete overlap was observed (Fig. 3, top).

Interestingly, the percent identity between the entire human and yeast \textit{alg7} proteins is 38% (53% similarity), while the percent identity between the archaeal and yeast proteins is only 25% (41% similarity). However, much of the conservation is found within the glycosyl transferase 4 domain, implying that these positions might represent important catalytic or structural residues (Fig. 1, between position markers 127 to 327). The important fact highlighted is that such little direct sequence conservation can result in the complementation of an essential gene.

The result of this complementation demonstrates that Mv1751 is indeed a GlcNAc-1-P transferase that is capable of...
replacing the essential \textit{alg7} gene in \textit{S. cerevisiae}. For the first time, this allows for an experimentally verified function to be assigned to this gene and offers strong support to the proposal that Mv1751 catalyzes the first step in the \textit{N}-glycosylation pathway in \textit{M. voltae}. The linking sugar of both eukaryotic and \textit{M. voltae} glycans is GlcNAc, and both use dolichol as the anchoring lipid to assemble \textit{N}-glycans.

It therefore follows that a homologous protein would carry out the enzymatic attachment of nucleotide-activated GlcNAc to dolichol in both domains. As such, in recognition of its place in the archaeal \textit{N}-glycosylation pathway, we propose to rename the \textit{M. voltae} Mv1751 gene as \textit{aglH} and assign it the function of GlcNAc-1-P transferase.

It is interesting that attempts to knock out this gene have proven unsuccessful in \textit{M. voltae} (3) as well as in the related methanogen, \textit{Methanococcus maripaludis} (D. VanDyke and K. F. Jarrell, unpublished data), while later steps in the \textit{N}-linked pathway, including the oligosaccharyl transferase final step, have been disrupted. This suggests that AglH may play an essential role in additional pathways. One possibility is in glycolipid biosynthesis. It is known that \textit{M. voltae} contains a GlcNAc-1-P diether glycolipid, where the attachment of the sugar is directly to archaeol (10). The pathways for glycolipid biosynthesis are poorly understood and it may be the case that AglH is involved in this process.

Of final note is the fact that the Mv1751 gene was able to complement an essential gene in another domain of life. It is rare to find two genes from different domains of life, especially essential genes, that are interchangeable. Because of the conservation of many aspects of the \textit{N}-linked glycosylation systems in bacteria, archaea, and eukaryotes, the deciphering of the roles and interchangeability of various components may be advanced by consideration of the use of cross-domain complementation.

B.C. was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Post-Graduate Scholarship. This work was supported by a Discovery Grant from NSERC (to K.F.J.) and by the Deutsche Forschungsgemeinschaft (DFG), Bonn, Germany (to R.T.S.).

REFERENCES