Trends in Microbiology

CellPress

Review

Sweet New Roles for Protein Glycosylation in Prokaryotes

Jerry Eichler[1,*] and Michael Koomey[2]

Long-held to be a post-translational modification unique to Eukarya, it is now clear that both Bacteria and Archaea also perform protein glycosylation, namely the covalent attachment of mono- to polysaccharides to specific protein targets. At the same time, many of the roles assigned to this protein-processing event in eukaryotes, such as guiding protein folding/quality control, intracellular trafficking, dictating cellular recognition events and others, do not apply or are even irrelevant to prokaryotes. As such, protein glycosylation must serve novel functions in Bacteria and Archaea. Recent efforts have begun to elucidate some of these prokaryote-specific roles, which are addressed in this review.

Trends

Because many of the roles assumed by protein glycosylation in eukaryotes are not applicable to Bacteria or Archaea, this postmodification likely serves distinct roles in prokaryotes.

In Bacteria, protein glycosylation systems are found in nonpathogenic species, pointing to roles beyond virulence.

In Bacteria and Archaea, protein glycosylation contributes to the integrity and proper architecture of glycoprotein-containing assemblies.

Changes in protein glycosylation offers prokaryotes a rapid and reversible manner in which to respond to environmental changes.

Prokaryotic Protein Glycosylation – Understanding the How but Not the Why

As the list of completed genome sequences keeps growing, it is becoming increasingly clear that the number of protein-coding genes cannot alone account for the size of an organism's proteome. Sources of proteomic expansion include the various **post-translational modifications** (see Glossary) a given protein can undergo. Of the various protein-processing events that have been described, glycosylation, namely the covalent linkage of mono- to polysaccharides, is one of the most prevalent and probably the most complex (Box 1) [1,2]. Long thought to be restricted to Eukarya, it is now accepted that both Bacteria and Archaea also are capable of **N-glycosylation**, where glycans are amide-bonded to select Asn residues of a target protein, as well as **O-glycosylation**, where glycans are added to hydroxyl-presenting amino acids, particularly Ser and Thr [3–6] (Table 1). Despite the fact that numerous glycoproteins have been identified in Bacteria and Archaea [7,8], that the structures of many of the glycans decorating prokaryal glycoproteins have been solved [9–13], and that considerable progress has been made in delineating pathways of protein glycosylation in several bacterial and archaeal species [6,14,15], the roles served by the bacterial and archaeal versions of this universal post-translational modification remain poorly defined.

In Eukarya, numerous functions have been assigned to protein glycosylation. The glycosylation process begins in the endoplasmic reticulum, the first stop on the secretory pathway, where a lipid-bound polysaccharide core is transferred to target protein Asn residues. The N-linked glycan is then augmented by individual sugars, also transferred from lipid carriers, to yield a complex branched oligosaccharide [16–18]. The composition of the N-linked glycan dictates interactions of the modified protein with molecular chaperones, such as calnexin and calreticulin, and other enzymes that accommodate proper protein folding [19,20]. Indeed, thermodynamics-based studies have demonstrated the importance of protein glycosylation for protein folding [21,22]. At the same time, the same N-linked glycan structure is monitored by the quality-control system responsible for identifying aberrantly folded proteins and targeting them for degradation, if necessary [23–25]. Once an N-glycosylated protein has successfully navigated the coordinated protein folding and quality control steps, it may be delivered to the Golgi, the next station along the secretory pathway, via a sorting process that can also rely on N-linked

[1]Department of Life Sciences, Ben Gurion University of the Negev, Beersheva 84105, Israel
[2]Department of Biosciences, University of Oslo, 0316 Oslo, Norway

*Correspondence: jeichler@bgu.ac.il (J. Eichler).

Box 1. Glycosylation As a Source of Protein Diversity

Of the different post-translation modifications to which a given protein can be subjected, glycosylation introduces the most diversity. Several factors are responsible for the enormous variability associated with protein glycosylation. In addition to the variability derived from how a glycan is linked to a protein (e.g., N-linked or O-linked), considerable diversity is generated at the level of individual sugars comprising protein-linked glycans. For instance, the incorporation of sugars that differ in the number of backbone carbons (e.g., pentoses and hexoses), that can exist in different epimeric forms (e.g., glucose, mannose and galactose), and that can be distinguished via the addition of different chemical groups (e.g., amino or methyl groups) all contribute to glycan diversity. Further variablity arises when sugars start to oligomerize into a glycan due to the many possible linkages between any two sugars (in terms of both the position and stereochemistry of the connection), the possibility for branching, and the heterogeneity possible in a given oligosaccharide. Indeed, the variability of protein-linked glycans may be infinite because of the fact that no template limiting the size of an oligosaccharide seems to exist. Together, these considerations result in a plethora of protein-linked glycans unique in composition and/or architecture.

glycan composition [26,27]. Once in the Golgi, the N-linked glycan is subjected to further processing through the addition and/or removal of constituent sugars to yield a range of N-linked oligosaccharides [28–30]. The Golgi is also the site of O-glycosylation, a second major protein glycosylation event that can also introduce considerable diversity into the glycosylation profile of a glycoprotein [31,32]. In functional terms, the heterogeneity in glycan content generated in the Golgi can be exploited for targeting different glycoproteins to distinct subcellular compartments [33,34] or, in the case of cell-surface-exposed glycoproteins, can contribute to various cell–cell or other recognition events important for the development, differentiation, or physiology of a particular cell, tissue, or organism [35–38]. Such hetereogeneity can, moreover, reflect different diseased states [39–42].

At the same time, it would appear that many of the roles assumed by protein glycosylation in eukaryotes are not relevant for prokaryotes. For instance, whereas protein folding and N-glycosylation are linked in the eukaryal secretory pathway, these processes occur on either side of the plasma membrane in the case of bacterial and archaeal proteins secreted by the twin arginine translocation pathway. Here, such proteins fold in the cytoplasm [43,44], while **oligosaccharyltransferase**-based N-glycosylation transpires on the outer surface of the bacterial and archaeal cell [45,46] (as do some versions of bacterial O-glycosylation [47]). Likewise, the need to sort proteins to distinct subcellular compartments is extremely limited in prokaryotes. Finally, the number of recognition events required by a prokaryotic cell is likely to be far less than its eukaryal counterpart. Therefore, Bacteria and Archaea must rely on protein glycosylation for other purposes than how this post-translational modification is used in Eukarya (Figure 1, Key Figure). In this review, recent works addressing these roles are discussed.

Bacterial Protein Glycosylation – Not for Virulence Alone

The number of bacterial protein glycosylation systems recognized continues to grow. This process has been fueled in part by a few serendipitous discoveries followed by comparative genomics that allows one to immediately see how broadly distributed protein glycosylation systems (and genes) truly are. The take-home messages here are that (i) bacterial protein glycosylation is much more prevalent than one could have imagined, and (ii) it is not strictly associated with pathogenic species. There are thus a number of outstanding questions relating to the biological significance of bacterial protein glycosylation that are emphasized here.

In studying protein glycosylation in Bacteria, the predominant emphasis has been placed on pathogens. However, systems related to those found in pathogens abound in commensal and environmental isolates. Thus, bacterial protein glycosylation is not a canonical virulence factor as defined by the criteria established by Falkow [48]. Nonetheless, clear defects in colonization and virulence in mammalian, insect, and plant model systems are seen for glycosylation null mutants. This is particularly true of so-called dedicated systems in which the glycosylation

Glossary

Archaellum: the motility structure of Archaea, functionally equivalent to the bacterial flagellum.
Autotransporters: found in a broad range of Gram-negative bacteria, autotransporters comprise a family of outer membrane or secreted proteins that facilitate their own transport to the cell surface. In such proteins, the autotransporter domain, comprising the C-terminal portion of the protein, forms a beta-barrel structure in the outer membrane through which the N-terminal domain is presented on the cell surface. Autotransporters are associated with virulence, contributing to adhesion, aggregation, invasion, biofilm formation, and toxicity.
N-glycosylation: the covalent linkage of glycans to select Asn residues in a target protein through an amide bond.
O-glycosylation: the covalent linkage of glycans to hydroxyl-presenting amino acids, particularly Ser and Thr.
Oligosaccharyltransferase: oligosaccharyltransferases catalyze the transfer of glycans from the lipid carriers upon which they are assembled onto selected residues in glycoproteins. In Bacteria, Archaea, and lower eukaryotes, the oligosaccharyltransferase acts alone, whereas in higher Eukarya, the enzyme exists as a multimeric complex.
Post-translation modification: an event that follows translation, designed to create variants of a given protein through the covalent attachment of one or more of several classes of molecules (e.g., sugars, lipids, or small chemical groups, like acetyl or methyl groups), the formation of intra- or intermolecular linkages (e.g., disulfide bonds), proteolytic cleavage (e.g., signal peptide removal), and/or any combination thereof.
Sequon: sequence motifs in a polypeptide denoting sites where glycans are attached. In N-glycosylation, the Asn of a sequon, namely an Asn-Xaa-Ser/Thr sequence, where Xaa is any residue but Pro, is modified. Variations to the canonical sequon have been observed in prokaryotes. The sequons processed in O-glycosylation are less well defined.