Review

CellPress

Special Issue: The Magic of the Sugar Code

Congenital disorders of glycosylation: a concise chart of glycocalyx dysfunction

Thierry Hennet and Jürg Cabalzar

Institute of Physiology, University of Zurich, CH-8057 Zurich, Switzerland

Glycosylation is a ubiquitous modification of lipids and proteins. Despite the essential contribution of glycoconjugates to the viability of all living organisms, diseases of glycosylation in humans have only been identified over the past few decades. The recent development of next-generation DNA sequencing techniques has accelerated the pace of discovery of novel glycosylation defects. The description of multiple mutations across glycosylation pathways not only revealed tremendous diversity in functional impairments, but also pointed to phenotypic similarities, emphasizing the interconnected flow of substrates underlying glycan assembly. The current list of 100 known glycosylation disorders provides an overview of the significance of glycosylation in human development and physiology.

Glycosylation disorders

Glycosylation is by far the most complex form of protein [1] and lipid modification [2,3] in all domains of life. The tremendous diversity of glycoconjugate structures resulting from intricate biosynthetic pathways is a major factor hampering the assignment of functions to glycans chains. Much has been learnt from the study of disrupted glycosylation genes in model organisms, thereby establishing numerous essential contributions of glycans in regulating cell and organ functions [4]. The study of human diseases of glycosylation brings additional insights by providing a more differentiated view on glycan functions. Indeed, most human mutations are hypomorphic, thus causing partial loss of glycosylation reactions that lead to variable clinical manifestations.

Diseases of glycosylation are also referred to as congenital disorders of glycosylation (CDG). Given the heterogeneity of glycans, the clinical scope of CDG is considerable, ranging from nearly normal phenotypes to severe multi-organ dysfunctions causing infantile lethality. CDG are rare diseases. The prevalence among CDG types differs from one type to another, but is largely unknown. The difficulty in identifying patients is another reason behind the rarity of CDG. Unspecific symptoms and the lack of simple laboratory tests make the recognition of CDG cases challenging. The identification of CDG has long relied on the detection of underglycosylated serum transferrin by isoelectric focusing [5]. While easy to perform and requiring only a few microliters of blood, this test exclusively reveals alterations of *N*-glycosylation. Unfortunately, similar blood tests have not been established to reliably diagnose defects in other classes of glycosylation. The simplicity of the serum transferrin test also explains why disorders of *N*-glycosylation account for the majority of known CDG.

Recent developments in genome-wide DNA sequencing technology enable the identification of mutations without a priori knowledge of candidate genes. As in other fields of biology, next-generation sequencing approaches have increased the pace of discovery for new types of CDG [6]. The barrier of 100 genes defects impairing glycosylation has just been passed (see http://www.physiol.uzh.ch/Glycosylation for a graphical overview). These defects encompass nearly all glycosylation pathways and affect different molecular processes, from substrate biosynthesis up to protein trafficking [7,8]. The recent application of unbiased strategies, such as exome and whole-genome sequencing, have further revealed CDG-causing mutations in genes previously not associated with glycosylation, thereby expanding our view on these complex pathways.

CDG were originally classified in two groups. So-called 'CDG type I' included defects of lipid-linked oligosaccharide assembly from the formation of dolichol-PP-*N*-acetyl-*D*-glucosamine (GlcNAc) up to the transfer of oligosaccharides to asparagine residues on nascent proteins. CDG type II, by contrast, included defects of *N*-glycan trimming and elongation as well as defects in any other class of glycosylation [9]. Given that several defects affect multiple glycosylation pathways, the artificial distinction between CDG type I and II has been replaced by a flat nomenclature simply associating implied genes with the suffix 'CDG' [10]. Functionally, defects can also be grouped based on their contribution to glycosylation reactions (Figure 1). Accordingly, here we discuss glycosylation disorders through five functional categories, featuring: (i) genes encoding glycosyltransferase enzymes; (ii) genes involved in donor substrate biosynthesis; (iii) genes mediating the translocation of donor substrates; (iv) genes regulating glycosyltransferase localization; and (v) genes affecting the homeostasis of secretory organelles.

Corresponding author: Hennet, T. (thierry.hennet@uzh.ch).
Keywords: CDG; disease; glycoprotein; endoplasmic reticulum; Golgi.

0968-0004/
 http://dx.doi.org/10.1016/j.tibs.2015.03.002

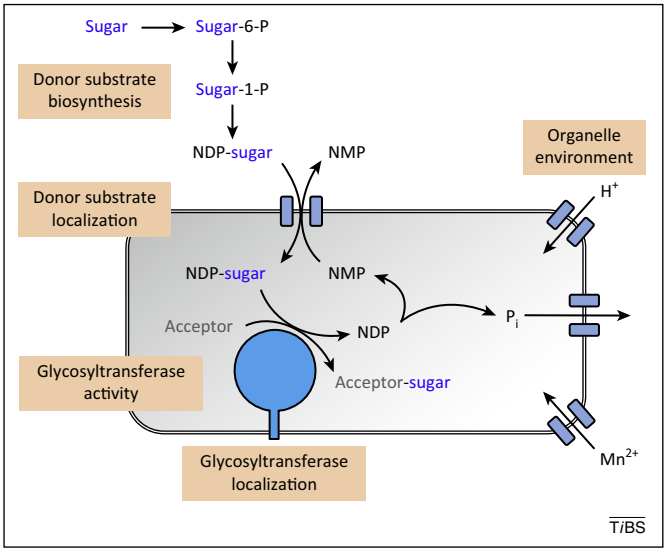

Figure 1. Glycosylation reaction. Schematic representation of the key players required for glycosylation reactions occurring in the Golgi apparatus. The biosynthesis of nucleotide-activated sugars (NDP-sugar) occurs in the cytosol, whereas glycosyltransferase enzymes are localized on the luminal side of the endomembranes of the secretory pathway. Transporter systems are required for the import of nucleotide-activated sugars into the Golgi apparatus and for maintaining optimal ionic conditions in the organelle, thereby regulating pH, Mn^{2+} import, and P export.

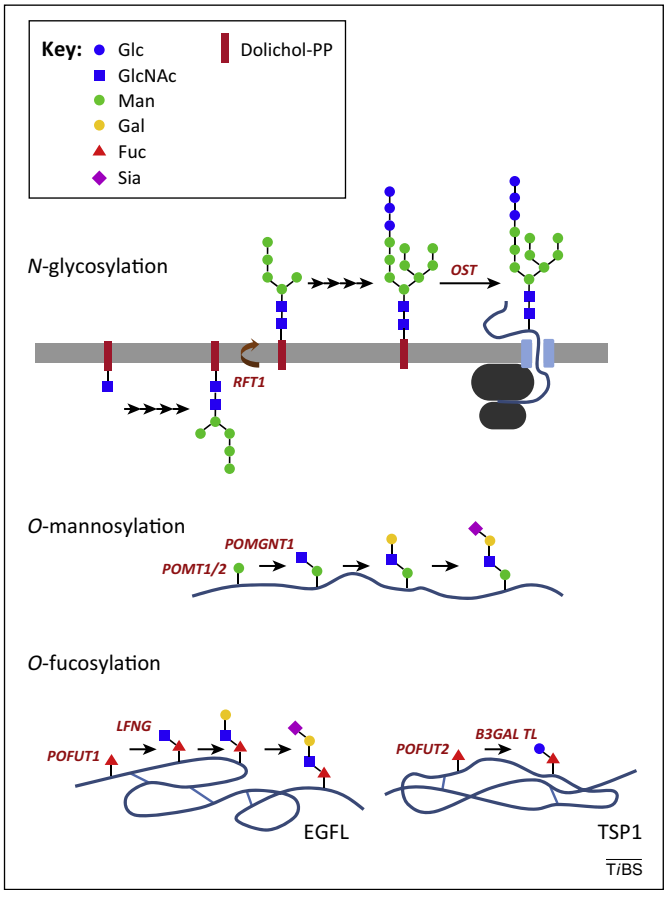

Figure 2. Biosynthesis of core structures for *N*-glycosylation, *O*-mannosylation, and *O*-fucosylation. *N*-glycosylation begins at the endoplasmic reticulum (ER) membrane by the stepwise assembly of dolichol-PP-$GlcNAc_2Man_9Glc_3$, which is transferred to the selected Asn residues of nascent glycoproteins by the oligosaccharyltransferase complex (OST). *O*-Mannosylated and *O*-fucosylated glycans are shaped by the sequential addition of different monosaccharides based on the acceptor specificity of glycosyltransferases. See main text for additional definitions.

Glycosyltransferases

The human genome includes close to 200 glycosyltransferase genes [11]. Glycosyltransferases are the enzymes shaping glycans through the formation of glycosidic linkages. The majority of these glycosyltransferases are transmembrane proteins anchored in the endoplasmic reticulum (ER) and Golgi membranes [12]. Defects of ER glycosyltransferases involved in the assembly of the lipid-linked oligosaccharide $GlcNAc_2Man_9Glc_3$ limit the availability of this substrate for transfer to *N*-glycosylation sites on acceptor proteins during translation (Figure 2). Such defects of *N*-glycosylation lead to glycoproteins lacking whole *N*-glycan chains. Depending on the glycoproteins affected, nonoccupancy of *N*-glycosylation sites can impair protein folding, secretion, and stability. At the level of the organism, such defects lead to multiple organ dysfunctions. Neurological symptoms are frequent, including psychomotor retardation, ataxia, and hypotonia. Liver and cardiac dysfunctions are also frequently observed, as are endocrine disorders, which mainly affect the sexual maturation of female patients [13].

The functional impairments associated with some glycosyltransferase deficiencies reflect the functional relevance of the involved glycoproteins. For example, *O*-mannosylation [14] is an essential modification of α-dystroglycan that ensures proper interactions between the dystroglycan complex and proteins of the extracellular matrix [15]. Such interactions are essential for the integrity of muscular fibers, for the migration of neurons in the cortex, and for the retinal architecture [16]. Given that α-dystroglycan is the main carrier of *O*-mannose chains, the manifestations of *O*-mannosylation disorders relate to α-dystroglycan functions and, therefore, encompass muscular degeneration, brain abnormality, and blindness. Clinically, these disorders belong to the congenital muscular dystrophies and are known as Walker–Warburg syndrome, Muscle–Eye–Brain disease, Fukuyama-type congenital muscular dystrophy, and limb-girdle muscular dystrophy. The most severe cases are usually associated with mutations in the core mannosyltransferase genes *POMT1* [17] and *POMT2* [18] and in the β1-2 GlcNAc-transferase gene *POMGNT1* [19] (Figure 2), but other gene defects also account for severe cases of Walker–Warburg syndrome and Muscle–Eye–Brain disease. To date, defects in 12 genes are known to cause congenital muscular dystrophies, although the functions of some of these genes are still unclear. For example, the fukutin (*FKTN*) and fukutin-related protein gene (*FKRP*) genes encode putative glycosyltransferases involved in *O*-mannosylation, but their exact substrate specificity and activity remain unknown [20].

Another form of *O*-linked glycosylation is characterized by the addition of fucose (Fuc) to serine and threonine in the context of the epidermal growth factor (EGF)-like domains and thrombospondin-1 (TSP1) domains. Typical acceptor proteins are members of the Notch family, including the ligands Jagged and Delta-like, which are signaling proteins involved in morphogenetic processes [21]. Complete deficiency of core *O*-fucosyltransferases POFUT1 and POFUT2 has not been described yet, but heterozygous mutations in the *POFUT1* gene have been identified in