Biochimica et Biophysica Acta 1790 (2009) 81–94

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

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

Review

Glycobiology in the cytosol: The bitter side of a sweet world

Yoko Funakoshi [a], Tadashi Suzuki [a,b,*]

[a] Glycometabolome Team, Systems Glycobiology Research Group, RIKEN Advanced Science Institute, 2-1 Hirosawa, Wako Saitama, 351-0198, Japan
[b] CREST (Core Research for Evolutionary Science and Technology), JST (Japan Science and Technology Agency), Honcho 4-1-8, Kawaguchi, Saitama 332-0012, Japan

ARTICLE INFO

Article history:
Received 5 May 2008
Received in revised form 3 August 2008
Accepted 11 September 2008
Available online 8 October 2008

Keywords:
Cytosol
Glycosylation
Lectin
Glycosidase
Glycosyltransferase
Glycobiology

ABSTRACT

Progress in glycobiology has undergone explosive growth over the past decade with more of the researchers now realizing the importance of glycan chains in various inter- and intracellular processes. However, there is still an area of glycobiology awaiting exploration. This is especially the case for the field of "glycobiology in the cytosol" which remains rather poorly understood. Yet evidence is accumulating to demonstrate that the glycoconjugates and their recognition molecules (*i.e.* lectins) are often present in this subcellular compartment.

1. Introduction

The involvement of carbohydrate moieties of the glycoconjugates in biological phenomena is essential in innumerable life forms ranging from bacteria to humans. For instance, *N*-glycosylation plays pivotal roles in various contexts including folding or inter- and intracellular trafficking of the proteins [1,2]. The biosynthetic pathways of *N*-glycosylation are well understood and genetically-modified animal models enable us to assess the importance of specific sugar modifications in even greater detail (plants, invertebrates [3–6], mice [7–9]). The cytosol is known to be involved in numerous glyco-related events, such as glycolysis, glycogen metabolism, biosynthesis of sugar nucleotides and hyaluronic acids. However, for example, insufficient attention has been paid to the occurrence of *N*-glycosyl proteins in the cytosol. This is especially the case after the elucidation of topological issues on major glycosylation pathways [9,10]. Accordingly, such studies have not been included in the main stream of glycobiological discourse and some considers them as experimental artifacts. The subcellular location of glycoproteins by purely biochemical fractionation is difficult to discuss. There have been, nevertheless, numerous reports on the occurrence of cytosolic glycoconjugates, as well as lectins and glycan-modifying enzymes, which now fit within a biological context. In this review we present an overview of our current knowledge on the occurrence of glycoproteins/lipids, with particular emphasis on *N*-glycosyl proteins. We also report on glycosidases/glycoamidases and lectins which are involved in metabolism/recognition of *N*-glycosyl proteins and/or free *N*-glycans in the cytosol.

2. Occurrence of glycoconjugates in the cytosol (1): cytosol-origin

2.1. 'Primordial' glycoconjugates

Because most of well-characterized glycosylation events are known to occur in the lumen of vesicles such as the ER or Golgi, it may be seen as unusual to find glycoconjugates in the cytosol. However, it is widely accepted that some of the early biosynthetic glycoconjugate intermediates ('primordial' glycoconjugates) are indeed formed on the cytosolic face of vesicles. For example, the biosynthesis of glucosyl ceramide (Glc-Cer), the precursor for the vast majority of glycolipids, is known to take place on the cytosolic face of the ER membrane. Subsequent steps of Glc-Cer biosynthesis occur on the lumenal side of the membrane after the flipping of the target Glc-Cer into the lumen [11,12].

Similarly, dolichol-linked oligosaccharide (OS-PP-Dol) precursors with structures up to $Man_5GlcNAc_2$, are also known to be formed on the cytosolic face of the ER membrane [2]. The $Man_5GlcNAc_2$-PP-Dol is flipped into the lumenal side of the ER by the action of a putative "flippase" [13,14] for further modifications to form completely-assembled donor substrates for oligosaccharyltransferase, *i.e.*, Glc_3-$Man_9GlcNAc_2$-PP-Dol for most eukaryotes. These early Dol-PP-OS

Abbreviations: Asn, asparagine; CBP, carbohydrate binding protein; Cer, ceramide; CRT, calreticulin; Dol, dolichol; ENGase, endo-β-*N*-acetylglucosaminidase; ER, endoplasmic reticulum; ERAD, ER-associated degradation; ERQC, ER-quality control; FITC, fluorescein isothiocianate; Glc, glucose; GlcNAc, *N*-acetylglucosamine; Hsc, heat shock cognate; HSP, heat shock protein; HyPro, hydrokylated proline, hydroxyl proline; IGOT, isotope-coded glycosylation site-specific tagging; OGT, *O*-GlcNAc transferase; OS, oligosaccharide; P4H1, prolyl 4-hydroxylase; PNGase, peptide:*N*-glycanase; PP'ase, pyrophosphatase; SCF, Skp1/Cul1/F-box protein

* Corresponding author.
E-mail address: tsuzuki_gm@riken.jp (T. Suzuki).

0304-4165/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbagen.2008.09.009

intermediates are also the glycoconjugates where glycans have access to the cytosol.

2.2. O-GlcNAc modification

Aside from the biosynthetic intermediates, *O*-GlcNAc modification, where β-linked GlcNAc is attached to Ser or Thr residues of cytosolic/nuclear proteins, is the best-known example of cytosolic/nuclear glycosylation. This modification is found to occur exclusively in the cytosol or nucleus and unlike other types of glycosylation, remains as a single GlcNAc residue through a β-glycosidic linkage to Ser/Thr residues and is not elongated further by the addition of other sugars.

This structure was first discovered as a terminal *O*-linked GlcNAc structure residing inside the cells, in which ^{3}H-galactose labeling was achieved using a soluble exogenously-added galactosyltransferase [15]. In addition, it was found that proteins with this modification were enriched in the nuclei and the soluble fraction [16]. *O*-GlcNAc is evolutionally conserved in eukaryotes, and even in the prokaryote *Listeria* [17], but not in yeast, although the possibility remains that this organism may substitute *O*-GlcNAc with other sugars for cellular functions [18].

Thus far, the biological functions of *O*-GlcNAcylation in cytosolic events can be summarized as follows: (1) it inhibits protein phosphorylation on Ser/Thr residues, by reciprocal site occupation [19–21](for review [22,23]); (2) it affects protein degradation, by influencing PEST sequences known as the rapid degradation signal motif [24,25] or by directly modifying the proteasome complex to decrease the proteasomal function [26–28]; (3) it regulates of intracellular localization of the carrier proteins [29–34] (for review [16,35]): in many cases, *O*-GlcNAc is known to function as an 'alternative' nuclear localization signal; (4) it is involved in protein interaction, some *O*-GlcNAcylated transcription factors are known to interact for transactivation [31,36]; and (5) it affects the activity of transcription factors or repressors [37–40].

The regulation of *O*-GlcNAcylation is mediated by the *O*-GlcNAc transferase (OGT) [41,42] and *O*-GlcNAcase [43–45]. This modification is truly dynamic, and drastic changes in *O*-GlcNAc levels were found to occur as a result of various stimulations/stresses on cells [46,47]. Nutrient levels, especially glucose content, could directly influence the UDP-GlcNAc levels, which subsequently affected the *O*-GlcNAc levels through the hexosamine biosynthetic pathway in the cells [48]. More than 500 of the cytoplasmic/nuclear substrates for this modification have been characterized [49], and almost all identified *O*-GlcNAcylated proteins are also phosphorylated [18,48]. Because some *O*-GlcNAc sites have been mapped to the same amino acid or residues that are only a few away from the phosphorylation site, it has been postulated that these two modifications may exist in a reciprocal relationship, or to at least have some mutual effects [49] (Fig. 1).

There are some reports on the function of *O*-GlcNAc modifications *in vivo*. In *Drosophila melanogaster*, the polytene chromosome was examined for fluorescein isothiocianate (FITC)-WGA binding and it was found that staining mainly occurred at the transcriptionally active "puff" regions. This observation implied the existence of many *O*-GlcNAcylated nuclear factors involved in transcription [50]. In *Caenorhabditis elegans*, a null mutation of the *O*-GlcNAc transferase gene (*ogt-1*) did not influence viability or fertility, but affected the insulin signaling pathway by changing the metabolic rate and suppressed the dauer-forming phenotype of the constitutive *daf-2* mutation [51]. In contrast, enhancement of the dauer formation was observed with the *O*-GlcNAcase (*oga-1*) mutation [52]. Although the detailed molecular mechanism has not been revealed, *O*-GlcNAcylation was therefore predicted to "fine-tune" the insulin signaling pathway. Interestingly, the phosphorylation levels were upregulated in a few proteins by the *oga-1* mutation. Those proteins did not appear to be *O*-GlcNAcylated [52].

The mutation of the *O*-GlcNAc transferase seems to have a stronger effect in mammalian systems. Because the gene was found to be essential to stem cell viability [53], the *O*-GlcNAc transferase mutation in animals was examined using a Cre-loxP conditional knockout strategy [54]. As indicated by the varied *O*-GlcNAcylated proteins with pivotal roles in various cellular physiologies, loss of *O*-GlcNAc caused by the clonal *OGT* mutation induced multiple phenotypes in each tissue type. These phenotypes included T-cell apoptosis, neuronal dysfunction, and fibroblast growth arrest even following stimulus by serum [54]. Interestingly, hyper-phosphorylated tau was found to accumulate in neuron-specific *OGT* mutant cells. Therefore, it was proposed that the accumulation of the tau protein might be responsible for neuronal apoptosis, which in turn could lead to early lethality.

For further reading on this modification to proteins, the readers are directed to recent reviews [18,49,55].

2.3. Dictyostelium hydroxyproline O-glycosylation

Skp1 in *Dictyostelium* is the only protein which is characterized as a substrate for the peculiar *O*-glycosylation described below in this section. This protein is best known as an adaptor in the SCF-family of E3-ubiquitin ligases, whose subunits include Skp1, cullin-1 and an F-box protein that selects target proteins for polyubiquitination [56,57]. The glycosylation of *Dictyostelium* Skp1 was first discovered through the incorporation of metabolically labeled Fuc [58]. Subsequent studies revealed the irregular glycosylation on a hydroxylated proline (HyPro, residue 143) with the structure, Galα1-6Galα1-3Fucα1-2Galβ1-3GlcNAcα1-HyPro (Fig. 2) [59].

The mechanisms and the cytosolic enzymes involved in the formation of such glycan structures have been extensively studied by West and colleagues. The glycosylation event begins with hydroxylation of Pro143 by an enzyme which is only distantly related to a known prolyl 4-hydroxylase(P4H1) [60,61]. This enzyme lacks the N-terminal targeting sequences for the secretory pathway [62].

The enzyme which transfers GlcNAc onto HyPro was purified from the *Dictyostelium* cytosol [63] and termed Dd-ppGnT1. The gene encoding this enzyme was identified by molecular approaches [64]. The enzyme has a neutral pH optimum, is most efficient in low salt conditions, and requires divalent metal ions for activity. This enzyme is distantly related to the enzymes that initiate mucin-type *O*-linked glycosylation (GalNAcT) rather than the cytosolic *O*-GlcNAc transferase.

The next galactosylation and fucosylation steps are carried out by FT85, a 768-residue protein with the N-terminal domain containing galactosyltransferase activity [65]. The C-terminal region has fucosyltransferase activity [66,67]. Consequently FT85 is a bifunctional glycosyltransferase similar to glycosaminoglycan synthases [68]. Galactosylation of the Fuc residue occurs by the action of UDP-galactose:fucoside α3-galactosyltransferase, which has no similarity to known α3GTs. The soluble enzyme was biochemically purified and named GT78 [69]. The authors also identified the gene encoding this enzyme and found that the enzyme was identical to one of the

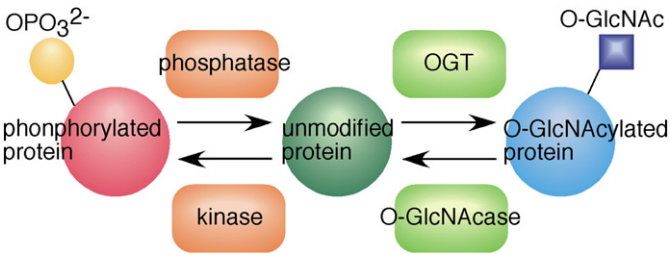

Fig. 1. The reciprocal regulation of protein function by phosphorylation and *O*-GlcNAc. Figure according to Zachara and Hart [55]. In this scheme, *O*-GlcNAcylation competes for the Ser/Thr residue in the target protein with phosphorylation.