Biochimica et Biophysica Acta xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

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

Short O-GalNAc glycans: regulation and role in tumor development and clinical perspectives

Joanne Chia [a,1], Germaine Goh [a,1], Frederic Bard [a,b,*]

[a] Institute of Molecular and Cell Biology, 61 Biopolis Drive, Proteos, 138673, Singapore
[b] Department of Biochemistry, National University of Singapore, 21 Lower Kent Ridge, Road, 119077, Singapore

ARTICLE INFO

Article history:
Received 4 January 2016
Received in revised form 3 March 2016
Accepted 3 March 2016
Available online xxxx

Keywords:
O-GalNAc glycosylation
Cancer
Membrane trafficking
GALNTs

ABSTRACT

Background: While the underlying causes of cancer are genetic modifications, changes in cellular states mediate cancer development. Tumor cells display markedly changed glycosylation states, of which the O-GalNAc glycans called the Tn and TF antigens are particularly common. How these antigens get over-expressed is not clear. The expression levels of glycosylation enzymes fail to explain it.
Scope of Review: We describe the regulation of O-GalNAc glycosylation initiation and extension with emphasis on the initiating enzymes ppGalNAcTs (GALNTs), and introduce the GALA pathway — a change in GALNTs compartmentation within the secretory pathway that regulates Tn levels. We discuss the roles of O-GalNAc glycans and GALNTs in tumorigenic processes and finally consider diagnostic and therapeutic perspectives.
Major conclusions: Contrary to a common hypothesis, short O-glycans in tumors are not the result of an incomplete glycosylation process but rather reveal the activation of regulatory pathways. Surprisingly, high Tn levels reveal a major shift in the O-glycoproteome rather than a shortening of O-glycans. These changes are driven by membrane trafficking events.
General Significance: Many attempts to use O-glycans for biomarker, antibody and therapeutic vaccine development have been made, but suffer limitations including poor sensitivity and/or specificity that may in part derive from lack of a mechanistic understanding. Deciphering how short O-GalNAc glycans are regulated would open new perspectives to exploit this biology for therapeutic usage. This article is part of a Special Issue entitled "Glycans in personalised medicine" Guest Editor: Professor Gordan Lauc.

1. Introduction

1.1. Glycosylation in health and cancer

Cancer remains one of the leading causes of mortality and a major loss of years of potential life, with an estimated 8.2 million deaths worldwide in 2012 [1]. Cancer continues to pose tremendous challenges for treatment and diagnosis. This is due to the complicated pathology of the disease that involves a whole panel of dysregulated cellular processes, which are interconnected and often vital for proper cell functioning. Identifying the key alterations specific to the neoplastic tissues and understanding the underlying mechanisms would allow precise targeting and detection of the disease. Current detection and treatment methods still suffer insufficient selectivity between tumor and normal tissues. This could be due in part to a primary focus on genetic aspects of the disease. In tumor formation, genetic changes in tumor cells and interaction with normal cells lead to emerging cellular states that often cannot easily be traced to specific changes in DNA. These cellular states ultimately define the behavior and evolution of a tumor and their understanding could be exploited for better specificities in diagnosis and treatment.

Among the processes defining cellular states, the most frequently occurring and also the most complicated post-translational modification (PTM) is glycosylation. Glycosylation is the enzymatic process that adds carbohydrate chains or glycans on protein and lipids. Glycosylation occurs in a complex and concerted series of steps taking place in the endoplasmic reticulum (ER) and more predominantly, in the Golgi apparatus [2,3]. Unlike other biopolymers, glycan synthesis is not template-driven, not directly encoded in the genome and as a consequence it is not well understood how this synthesis is controlled. Yet, the result is a vast diversity of glycan structures with many important functions.

In mammals, glycan structures are assembled from ten monosaccharides: fucose (Fuc), galactose (Gal), glucose (Glc), N-acetylglucosamine (GlcNAc), N-acetyl-galactosamine (GalNAc), glucuronic acid (GlcA), iduronic acid (IdoA), sialic acid (Sia), mannose (Man) and xylose (Xyl)

Abbreviations: GalNAc, N-acetylgalactosamine; Tn antigen, T antigen nouvelle; S-Tn, sialylated Tn; TF antigen, Thomsen-Friedenreich antigen; S-TF, sialylated TF; ppGalNAcT (GALNT), polypeptide N-acetylgalactosaminyltransferase; GALA pathway, GALNT Activation pathway; C1GALT1, core 1 β3-galactosyltransferase; EMT, epithelial-mesenchymal transition; sLe^x, sialyl Lewis x.

* Corresponding author at: Department of Biochemistry, National University of Singapore, 21 Lower Kent Ridge, Road, 119077, Singapore.
E-mail address: fbard@imcb.a-star.edu.sg (F. Bard).

[1] These authors contributed equally.

http://dx.doi.org/10.1016/j.bbagen.2016.03.008
0304-4165/

[4–6]. Combinations of these monosaccharides coupled with differences in linkages (1–3 versus 1–4, etc.), anomeric states (α versus β), branching, length and substituted components (phosphate, sulfate, etc.) creates further diversity.

Glycan addition to substrate proteins and lipids also further generates different subsets of glycoconjugates. In proteins, glycans can be attached to at least nine out of 20 types of amino acids and the two prevailing processes involve amide linkages to asparagine residues (N-glycosylation) and glycosidic linkages to serine and threonine (Ser/Thr) side chains (O-glycosylation) [7,8]. Even within O-glycosylation, there are subtypes. Several sugars can be added to the Ser/Thr to yield different classes of O-linked glycans, such as α-linked O-GalNAc, α- or β-linked O-galactose, α- or β-linked O-glucose, α-linked O-fucose, α-linked O-mannose, β-linked O-GlcNAc and β-linked O-xylose. O-GalNAc glycans represent the most abundant O-glycans type. In this review, we will sometimes refer to O-GalNAc glycosylation as the process of sugar addition and O-GalNAc glycans as the various products of this process. We will sometimes use simply O-glycosylation or O-glycans when there is no risk of confusion.

Hence, glycosylation exponentially expands and diversifies the encoded genomic information [9,10], increasing the biochemical complexity of eukaryotes [11]. However, it should not be perceived that glycosylation occurs randomly with huge variations even within a specific protein. Normally, only specific sites of a given protein are glycosylated, and with a limited number of glycan structures at each site (microheterogeneity) [12]. This suggests precise regulatory mechanisms in place to control glycan biosynthesis, particularly within the Golgi, where most of the glycosylation machinery resides [13].

Glycans contribute a wide range of biological functions in the organism, particularly as part of glycoproteins. This is emphasized by the fact that some glycan functions are well conserved in the evolution of multicellular vertebrates [14]. Glycosylation by GlcNAc can occur in the cytosol. By contrast, most glycosylation pathways occur within the secretory pathway and generate the complex glycans.

In the early secretory pathway, glycans are essential for protein stability and secretion by regulating the folding of newly synthesized proteins, for quality control in the ER and for protein targeting in the secretory pathway [15,16]. At the cell surface, glycans contribute to multiple processes such as cell-cell communication, cell adhesion and migration, signal transduction, immune surveillance and host-pathogen interactions [14,17]. In fact, nearly all cell surface and secreted proteins are glycosylated [5,18–20].

Recent studies have found that O-GalNAc addition occurs on at least one Ser/Thr residue in more than 85% of secreted proteins [19]. Cell surface glycans can modulate their carrier protein conformation as well as provide ligands for glycan-binding proteins such as selectins, galectins and siglecs. Glycans are often required for physiological processes such as cell-matrix adhesion and cell–cell interactions [21–24]. Altogether, this stresses the importance of glycosylation for coordinating multi-cellular life. Interestingly, many glycoproteins have been implicated in tumor pathology [2,25,26]. It therefore makes sense that glycosylation would be significantly perturbed in cancer.

Altered glycosylation was first described more than 60 years ago and has since been recognized as a hallmark in oncogenic transformation [2]. Various glycan changes including under- or over-expression of specific glycan structures, expression of unprecedented or incomplete/truncated glycan structures or increased levels of precursor structures have been observed on tumor cells compared to their normal counterparts [27]. It is apparent that different glycan structures affect the cellular processes, as well as the tumor microenvironment, that play a pivotal role in cancer progression, angiogenesis, metastasis, cell-cell contact and epithelial-mesenchymal transition (EMT) in cancer cells. A few glycans are markedly associated with malignant transformation and progression. Given that a key aspect in tumor progression is clonal selection of the fittest cells from a genetically heterogeneous population, it suggests that these cancer-specific glycans are selected for and are likely to promote tumor cell survival [27].

1.2. Cancer-specific O-GalNAc glycans: the Thomsen-Friedenreich related antigens

Among the cancer-specific glycans are the Thomsen-Friedenreich (TF)-related antigens, which comprise several short O-GalNAc glycans: T antigen nouvelle (Tn antigen), TF (aka T antigen or core 1) and their downstream sialylated counterparts (S-Tn and S-TF). The Tn antigen consists of a monosaccharide GalNAc alpha-O-linked to Ser/Thr in the polypeptide chain (GalNAc-α1-Ser/Thr) and the TF antigen is formed from the subsequent addition of Gal to GalNAc (Gal-β1–3GalNAc-α1-Ser/Thr). Sialylation of Tn and TF antigens involves the addition of terminal sialic acid to carbon 6 on GalNAc (S-Tn) and to carbon 3 on galactose (S-TF) respectively which prevents further elongation of the structure (Fig. 1). Because they tend to be observed also during development, they were initially named oncofetal antigens [28].

The initial discovery of these antigens came from the observation of occasional agglutination of stored blood cells in 1930. The agglutination was due to contaminating bacterial neuraminidases that exposed the TF antigen on blood cells and the TF antigen was recognized by anti-TF antibodies in the sera, contributing to hemagglutination [29]. Tn was subsequently discovered in 1957 to be expressed in the subpopulations of blood cells of lineages in patients with Tn syndrome, a rare hematological disorder [30]. Tn on the cell surface leads to increased polyagglutinability of erythrocytes and consequent hemolytic anemia, possibly due to the anti-Tn IgM [31].

The link between Tn and cancer was first observed in 1969 based on the binding of tumor cells to the snail *Helix pomatia* lectin (HPL) [32], and subsequent work by Springer and colleagues showing that Tn was highly expressed in around 90% of breast tumors [33]. Subsequent studies in the 1970–1980s established Tn to be a pan-carcinoma antigen as it is frequently expressed in cancers. Tn is expressed in 70–90% of most human solid tumor tissues, such as breast, colon, lung, bladder, cervix, ovary, stomach, and prostate while there is very low expression in the corresponding normal tissues [34–36]. Only the embryonic brain has been reported to express high levels of Tn [37]. Tn appears to be expressed mainly in epithelial carcinomas and less in blood cancers [38,39]. Expression of Tn appears in early tumor stages [40–42] and correlates with cancer progression, tumor metastasis and poor patient prognosis [43–46].

The mechanistic understanding of Tn expression has long been unclear. As high levels of Tn suppose large amounts of unmodified GalNAc residues, a long-standing hypothesis is loss of activity of the downstream core 1 β3-galactosyltransferase (C1GALT1) and/or the Core 3 synthase Core 3 β1–3 N-acetylglucosaminyltransferase (B3GNT6) [47–49]. As most tissues display Core1 glycans, loss of C1GALT1 activity was the driving hypothesis and proposed to be due to defects in Cosmc, a dedicated molecular chaperone required for proper folding of C1GALT1 [47,50]. As Cosmc is X-linked, its loss in cancers could arise more easily through gene mutations, chromosomal deletions and epigenetic silencing. These changes were indeed observed in a few specimens of cervical, colon, pancreatic tumor samples and cell lines derived from leukemia and melanoma [47,51,52]. However, the proposed mechanism is unlikely in most cancer types for various reasons described below. The main argument is that the TF antigen is also highly prevalent in many of the same cancers [53–57]. Yet, loss of Cosmc/C1GALT1 would abolish its expression.

In this review, we examine the regulation of O-GalNAc glycosylation, focusing on the initiating enzymes GALNTs. We describe current knowledge on the roles of O-GalNAc glycans and GALNTs in cancer progression. Finally, we illustrate the clinical perspectives of the use of cancer-specific O-GalNAc glycans for cancer diagnosis and therapeutics.

Please cite this article as: J. Chia, et al., Short O-GalNAc glycans: regulation and role in tumor development and clinical perspectives, Biochim. Biophys. Acta (2016), http://dx.doi.org/10.1016/j.bbagen.2016.03.008