Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Metabolic flux control in glycosylation

Andrew G McDonald, Jerrard M Hayes and Gavin P Davey

Glycosylation is a common post-translational protein modification, in which glycans are built onto proteins through the sequential addition of monosaccharide units, in reactions catalysed by glycosyltransferases. Glycosylation influences the physicochemical and biological properties of proteins, with subsequent effects on subcellular and extracellular protein trafficking, cell–cell recognition, and ligand–receptor interactions. Glycan structures can be complex, as is the regulation of their biosynthesis, and it is only recently that the systems biology of metabolic flux control and glycosyltransferase networks has become a study in its own right. We review various models of glycosylation that have been proposed to date, based on current knowledge of Golgi structure and function, and consider how metabolic flux through glycosyltransferase networks regulates glycosylation events in the cell.

Address
School of Biochemistry and Immunology, Trinity College Dublin, Dublin 2, Ireland

Corresponding author: Davey, Gavin P (gdavey@tcd.ie)

Current Opinion in Structural Biology 2016, **40**:97–103

This review comes from a themed issue on **Carbohydrate–protein interactions and glycosylation**

Edited by **Nagasuma Chandra** and **Harry Gilbert**

http://dx.doi.org/10.1016/j.sbi.2016.08.007

Introduction

Glycosylation is a major cellular activity in which oligosaccharides are covalently attached to proteins and lipids to form glycoconjugates. The type of glycosylation can vary, from a single monosaccharide to the formation of complex, highly branched structures known as glycans. The biosynthesis of glycans involves a large family of enzymes known as glycosyltransferases, which transfer the monosaccharide portion of a nucleotide sugar to its glycoprotein or glycolipid acceptor. In the case of protein, glycans typically fall into two major classes: N-linked, in which the carbohydrate is linked via the terminal (amide) nitrogen of the side chain of an asparagine residue; and O-linked, in which the attachment point is the side-chain oxygen of serine or threonine. The carbohydrate content of a glycoprotein can contribute significantly to its overall mass, and alter its biophysical properties. Glycosylation is a key determinant of many biological processes, such as intracellular and extracellular trafficking, cellular differentiation and development. In disease states, such as cancer, where alterations to the cellular glycosylation profile are present, a deeper knowledge of metabolic flux and how it regulates glycosyltransferase activity is needed. Improved control of glycosylation networks during bioprocessing is also fundamental to the development of biopharmaceuticals. Of increasing relevance is the use of mathematical modelling for understanding and prediction of glycoform heterogeneity, leading to new insights and to new directions for experimental work. We review some of the systems-biological models that have been developed, focusing in particular on the concept of flux-based competition amongst the glycosyltransferases, as well as among the substrates common to a given enzyme.

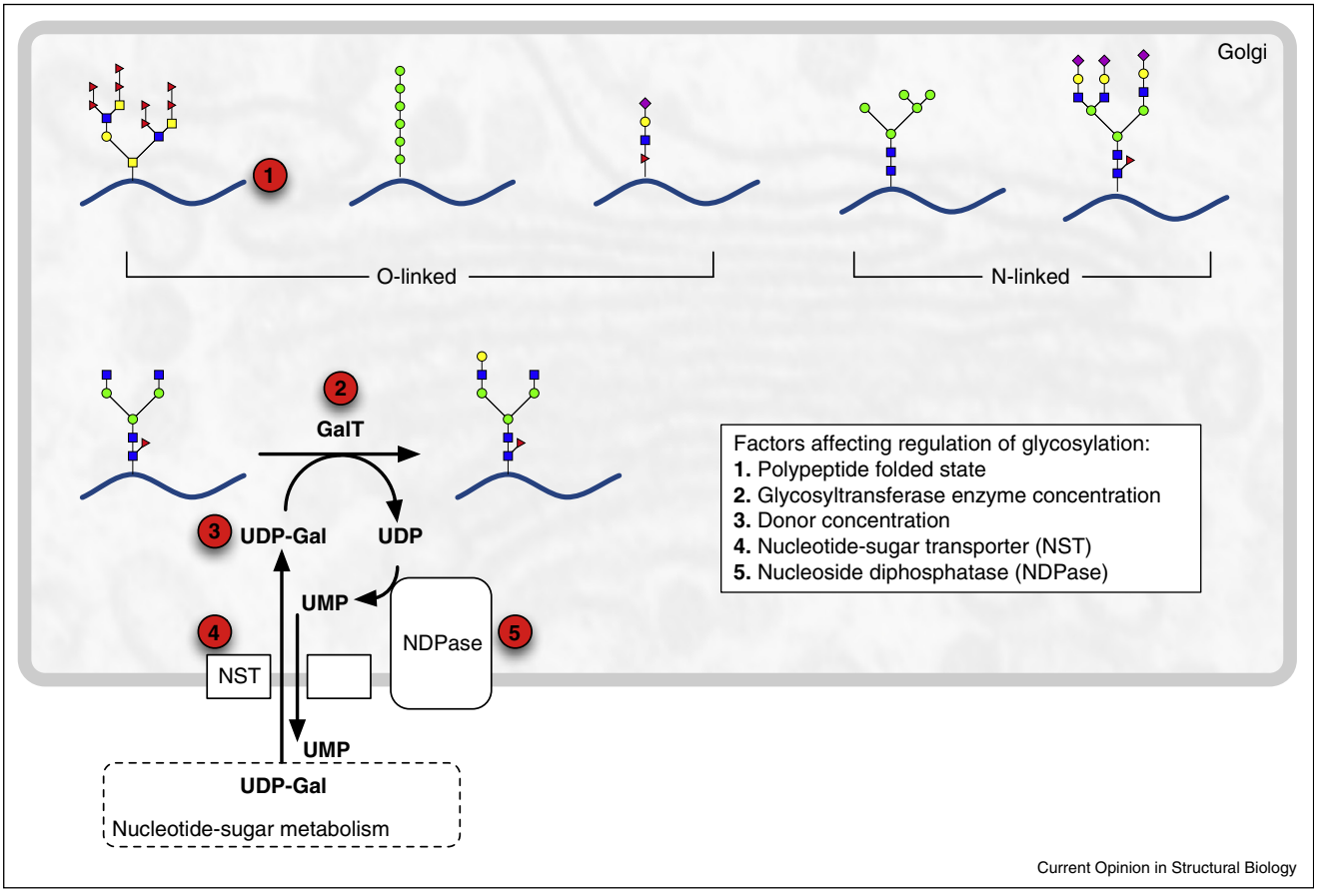
Enzymes of glycosylation and their substrates

This article focuses on the most commonly encountered N-linked and O-linked glycans and the enzymes responsible for their synthesis. N-linked glycosylation starts on the cytoplasmic side of the endoplasmic reticulum through the transfer of a dolichol phosphate-linked glycan consisting of two *N*-acetylglucosamine, nine mannose and three glucose residues ($Glc_3Man_9GlcNAc_2$) to a vacant glycosylation site on a protein catalysed by oligosaccharide transferase (OST). As the protein translocates through the ER/Golgi network, a combination of membrane-bound glycosyltransferases and glycosidases act thereafter to remodel the *N*-glycan structure to form a wide range of possible glycoforms. The different types of glycosylation are shown in Figure 1.

The consensus sequence (sequon) for the initiation of asparagine-linked glycosylation is the Asn-X-Ser/Thr motif, where X is not proline, whereas for O-linked glycosylation the sequon is an unoccupied side chain of a Ser/Thr residue. O-linked glycans of mucin type are formed initially by the addition of a *N*-acetylgalactosamine (GalNAc), catalysed by a family of 20 GalNAc-transferases distributed throughout the Golgi [1]. Whereas all *N*-glycans have in common a single, trimannosyl pentasaccharide core structure, *O*-glycans are known to form as many as eight different core structures [2]. With the exception of glucosidases and mannosidases, which hydrolyse single glucose and mannose residues, respectively, from the non-reducing end of a glycan, the other enzymes of glycosylation are transferases, catalysing the general reaction:

$$Ax + B = A + Bx$$

Figure 1



Types of glycosylation and some key points of regulation. Examples are shown of N-linked and O-linked glycan, the latter including *O*-GalNAc, *O*-mannose and *O*-fucose variants. The major regulation points include availability of substrate, enzyme activity, and transport. Additional regulation points include localisation of enzymes within the Golgi, and their co-association.

where A*x* is a nucleotide sugar donor (NSD), B is a proteoglycan acceptor, and *x* is the monosaccharyl unit transferred. Donor substrates are transported into the Golgi by specific transport proteins, as illustrated in Figure 1 for UDP-galactose (UDP-Gal). In addition to the glycosidases involved in *N*-glycan processing, GlcNAc-transferases (GnTs), GalNAc-transferases (GalNAcTs), galactosyltransferases (GalTs), sialytransferases (SiaTs) and fucosyltransferases (FucTs) are localised to separate regions of the Golgi. Glycans can be modified further by phosphorylation and sulfation through the action of kinases, sulfotransferases and *O*-acetyltransferases. There is also evidence of homomeric and heteromeric associations of the glycan processing enzymes [3], which may influence their kinetic properties and their localisation.

Models of glycosylation

A comparative study of the major systems glycobiology models in relation to glycoengineeering has recently appeared [4]. Here we give a brief account of the models of glycosylation that have been proposed. The earliest theoretical study, by Shelikoff and co-workers [5] studied the initiation of N-linked glycosylation and modelled the co-translational attachment of oligosaccharides to glycosylation sites. An important aspect of this initiation process, that of sequon occupancy, has been receiving increased attention of late due to the difficulty in predicting the initiation of protein glycosylation from a knowledge of the primary sequence alone, although some patterns are discernable in the case of *N*-glycans [6]. The location and decision to glycosylate a protein at a particular site is likely to be influenced by many factors, including the presence of neighbouring glycosylation sites. Murray *et al.* have shown that any aromatic amino acid at position *n*-2 relative to the Asn-X-Thr/Ser sequon results in glycoproteins possessing *N*-glycans of lower complexity [7•]. A kinetic model of mucin Core 1 formation was developed by Gerken in 2004 to explain the influence of neighbouring glycosylation sites on *O*-glycan formation [8]. Currently, machine-learning methods are most commonly used to predict the location of glycan