Glycosylation and the Immune System

Pauline M. Rudd, Tim Elliott, Peter Cresswell, Ian A. Wilson, Raymond A. Dwek

Almost all of the key molecules involved in the innate and adaptive immune response are glycoproteins. In the cellular immune system, specific glycoforms are involved in the folding, quality control, and assembly of peptide-loaded major histocompatibility complex (MHC) antigens and the T cell receptor complex. Although some glycopeptide antigens are presented by the MHC, the generation of peptide antigens from glycoproteins may require enzymatic removal of sugars before the protein can be cleaved. Oligosaccharides attached to glycoproteins in the junction between T cells and antigen-presenting cells help to orient binding faces, provide protease protection, and restrict nonspecific lateral protein-protein interactions. In the humoral immune system, all of the immunoglobulins and most of the complement components are glycosylated. Although a major function for sugars is to contribute to the stability of the proteins to which they are attached, specific glycoforms are involved in recognition events. For example, in rheumatoid arthritis, an autoimmune disease, agalactosylated glycoforms of aggregated immunoglobulin G may induce association with the mannos-binding lectin and contribute to the pathology.

A current model of the immune system classifies the processes by which foreign antigens are eliminated from their hosts into three interrelated systems, depending on whether the initial event takes place in the cytosol, the endocytic pathway or the extracellular space. In the cytosol, antigenic peptides generated by the proteasome are transported by membrane-bound transporters correlated with antigen processing (TAP1 and 2) into the ER where they bind to MHC class I molecules (1). In the endocytic pathway, protein antigens are degraded in specialized acidic compartments [MHC class II–containing compartments (MHC)] (2) into peptides prior to loading onto MHC class II molecules (3, 4). Also in the endocytic pathway, glycosylphosphatidylinositol (GPI) anchors and glycolipid antigens associate with CD1 (2, 5–7). MHC class I and class II molecules, as well as CD1, are recognized by T lymphocytes in the extracellular space, intact antigens are recognized either by antibody molecules or by the mannos-binding lectin (MBL), both of which subsequently activate the complement system. The diversity of protein glycosylation plays an important role in the biosynthesis and biological activity of the glycoproteins involved in antigen recognition. During transport of the glycoproteins through the secretory pathway, the sugar chains undergo successive modifications which are regulated by the glycosylation processing enzymes. This process ensures that each glycoprotein displays the relevant glycan structures for its range of essential functions. The sugars play a role in protein folding and assembly, quality control, ER-associated retrograde transport of misfolded proteins, the generation and loading of antigenic peptides into MHC class I, and also influence the range of antigenic peptides generated in the endosomal pathway for presentation by MHC class II.

By virtue of their size, glycans can shield large regions of the protein surfaces, providing protease protection for immune molecules in serum and secretions and in the T cell synapse, where they also limit nonspecific lateral protein-protein interactions. Glycans located close to the cell membrane, or in heavily glycosylated O-linked domains, can also determine the orientation and location of the binding faces of the proteins to which they are attached (8). In addition, the local three-dimensional structure of the individual protein directs its final glycoform pattern (9) such that a diverse repertoire of oligosaccharides will be displayed on the surface of a single cell. This makes it unlikely that the cell will bind lectins involved in the innate immune system, since these generally require repetitive arrays of specific sugars for functional recognition.

Some of these advantages conferred by glycosylation are exploited by viruses, which use the host glycosylation machinery to assemble their own envelope glycoproteins such that they help avoid immune detection (10). In autoimmune diseases, such as rheumatoid arthritis (RA) and systemic lupus erythematosus (11, 12), significant changes in the populations of immunoglobulin G (IgG) glycoforms have been noted. In RA, aggregated agalactosyl glycoforms of IgG are specifically recognized by the MBL such that inappropriate activation of the innate immune system may occur (13). Similarly on tumor cells, aberrant glycosylation can expose new epitopes for recognition by the immune system, such as for MUC1 (14), whereas mice deficient in β1,6 N-acetylgalcosaminyltransferase V (Mgat5) showed kidney autoimmunity and enhanced delayed-type hypersensitivity and increased susceptibility to experimental autoimmune encephalomyelitis (15).

Glycosylation and Protein Folding

Almost all of the key molecules involved in antigen recognition and the orchestration of the subsequent events are glycoproteins consisting of two or more subunits. Some, such as the T cell receptor (TCR) and CD3, are assembled in the ER into multimolecular complexes. The proper folding and controlled assembly of many newly synthesized glycoproteins requires them to engage in a series of coordinated interactions with chaperones and enzymes, through the attachment of a common oligosaccharide precursor, GlcNAc3Man9Glc1 (Fig. 1A), to N-linked glycosylation sites. This sugar precursor is rapidly processed to GlcNAc2Man3Glc1 (Fig. 1B), which can bind two chaperones, the membrane-bound calnexin (Cnx) and soluble calreticulin (Crt) (Fig. 1B) (16, 17).

The lectin-like interactions of Cnx and/or Crt with nascent glycoproteins provide access to a folding pathway (18), allow the recruitment of the thiol oxidoreductase, ERP57, and assist the assembly of subunits. Cnx and Crt are also involved in the loading of antigenic peptides onto MHC class I from complexes of TAP and tapasin. In their role as quality control factors, Cnx and Crt retain unfolded glycoproteins in the ER until they are correctly folded and assembled, an event which is signaled by the permanent removal of the terminal glucose residue by glucosidase II (Fig. 2). The folded glycoprotein, or the assembled multimolecular complex, is then transferred to the Golgi apparatus where the oligomannose sugars may be further processed. Misfolded or unassembled subunits are reglucosylated by uridine 5′-diphosphate–glucose:glycoprotein glucosyltransferase. Reglucosylation of unfolded proteins allows them to rebind Cnx and enter a cyclical pathway until they either achieve their correctly folded structure and are released or they are targeted...
for retrograde transport and degradation (Fig. 2) (19, 20).

Glycosylation and ER-Associated Degradation

The TCR complex is composed of at least six protein subunit components (α, β, γ, δ, ε, ζ); neither partial complexes nor individual subunits can be expressed by themselves on the protein surface. Misfolded or unassembled glycoproteins can be eliminated directly from the ER in a process known as ER-associated degradation (ERAD) that is initiated by the action of mannosidase I. This enzyme removes a single mannose residue from the oligosaccharide precursor so that it carries the glycan motif, GlcNAc2Man (Fig. 2) (19, 20). If the glycoprotein is still misfolded or unassembled, the GlcNAc2Man, glycans are reglucosylated. Compared with GlcNAc2Man,Glc, GlcNAc2Man,Glc is a poor substrate for glucosidase II and the glycoprotein readily rebinds Ctx (21–23). The misfolded or unassembled glycoprotein subunit, attached to Ctx, may then be released from the ER through the Sec61 translocation channel into the cytosol where it is degraded by the proteasome. The release of the polypeptide into the cytoplasm may require ubiquitination of the cytosolic tail (24) or that of Ctx (25). Denatured polypeptides in the cytosol are degraded by the 20S proteasome while protein substrates, such as unassembled subunits, are ubiquitinated, unfolded, and processed by the 26S proteasome. Antigenic peptides from internally synthesized viral proteins are generated by the protease activity of the proteasome (26).

Glycosylation and Antiviral Strategies

Enveloped viruses, such as human immunodeficiency virus (HIV), can evade immune recognition by exploiting the host glycosylation machinery to protect potential protein antigenic epitopes (10). Enveloped viruses also use the host secretory pathway to fold and assemble their often heavily glycosylated coat proteins. A possible antiviral strategy involves the use of glycosylation inhibitors to interfere with the folding of viral envelope proteins, such as for hepatitis B virus (HBV) and the HIV.

HBV has three envelope glycoproteins (L, M, and S) which are derived from alternative translations of the same open reading frame. All three surface antigens contain one common N-linked glycosylation site, Asn146 (about 50% occupied), but the M protein contains a second at Asn185 (100% occupied) that is responsible for the interaction of M protein with Ctx. When glucosidases I and II are inhibited by n-butyl deoxyo-N-methylpyrrolidin (n-BuDNJ), M protein not only remains hyper-glycosylated, but is also misfolded (27). The inability of the virus to incorporate misfolded M protein into the viral coat has a destabilizing effect, and secretion of the virus is prevented. However, subviral particles (which do not contain DNA) are secreted and, in the presence of the inhibitor, consist only of S glycoproteins, rather than M and S (27).

The major component of the HIV-1 viral envelope is gp160, a noncovalently linked complex of gp120 and gp41. When gp120 is expressed in the presence of n-BuDNJ, the V1/V2 loops (residues 128 through 199) are misfolded. In contrast to HBV, where misfolding of the M-protein affects the assembly of the viral coat, misfolding of the V1 and V2 loops of gp120 does not affect the assembly of the viral envelope of HIV. This is because these loops do not form part of the surfaces which contact each other in the gp120 trimERIC complex (28). Binding of gp160 to CD4 is also unaffected by the presence of n-BuDNJ in the culture medium; however, the conformational change which follows the binding event involves the rearrangement of the V1/V2 loops to release gp120 from gp160. This rearrangement cannot take place (29), and, as a result, gp41 cannot mediate the fusion event, precluding entry of the virus into the cell.

Glycosylation and the Assembly of MHC I: Peptide Complexes

Mature MHC class I (MHC I) consists of three subunits: the heavy chain, which is a transmembrane glycoprotein; the small soluble nonglycosylated protein, β2-microglobulin (β2-M); and an antigenic peptide that is required for the subsequent transport of MHC I to the cell membrane. Class I assembly [Fig. 2, reviewed in (1)] requires multiple coordinated intra- and intermolecular events to ensure the continuous reporting of cellular contents to cytotoxic T lymphocytes. In the ER, unassembled heavy chains interact initially with membrane bound Ctx through the Man,GlcNAc2Man,Glc sugar which is attached (in human HLA) to Asn86 in the early stages of glycan processing (Fig. 1). Concomitant with β2-M association with heavy chain, the sugar is released from Ctx and binds to Ctx. One or both of these chaperones recruits the third oxidoreductase (ERp57) into the complex, facilitating the formation of the intrachain disulfide bonds of the class I heavy chain (30–32). The assembled class I–β2-M dimer, together with associated Ctx and ERp57, are associated with a larger complex that also contains the TAP transporter and the transmembrane glycoprotein tapasin (33). Tapasin provides the bridge that links MHC I, Ctx, and ERp57 to the TAP heterodimer (34).

Peptides that are destined for MHC I binding are generated by proteolysis of cytoplasmic proteins by the proteasome. The peptides are then translocated in an ATP-dependent fashion from the cytosol into the ER by the TAP heterodimer. Those peptides with the appropriate sequence may undergo further NH2-terminal trimming prior to binding to TAP-associated MHC I. Peptide binding induces dissociation of the class I–β2-M dimer from the tapasin-TAP complex, as well as from Ctx and ERp57, allowing glycan maturation and transport of the assembled MHC I from the ER (Fig. 2). The precise roles of tapasin, Ctx, and ERp57 in the late stages of MHC I assembly are poorly understood.

Generation of Peptides from Glycosylated T Cell Antigens

Class I MHC-restricted peptide antigens are generated primarily in the cytosol. Proteins are targeted for degradation by ubiquitination and degraded by the 26S proteasome and/or further broken down into smaller peptides by the 20S proteasome modified by the 11S regulator or PA28 complex (35) and by other downstream enzymes in the cytosol, such as leucine aminopeptidase (36), puromycin-sensitive aminopeptidase and bleomycin hydrolase as well as unidentified sugar-trimming enzymes within the ER (37).

The 20S proteasome is a cylindrical assembly approximately 150 Å in height and 110 Å in diameter. The four rings of the α and β subunits of the proteasome encase a channel that widens
into three large cavities in which two constrictions that are only 13 Å in diameter form a passage from the cytosol to the inside of the cylinder. A typical sugar, such as a complex biantennary sugar containing 12 monosaccharide residues, is about 30 Å by 10 Å by 10 Å. This prior to entry into the proteasome, the sugars are removed from glycosylated polypeptides by a proximal glycanase, cysteolic peptide N-glycosidase (38). A role for a cytosolic N-glycanase is supported by the identification of a peptide derived from tyrosinase which is transported into the ER bound to TAP (39). Tyrosinase contains six N-glycosylation sites. A tyrosinase peptide, YMNGTMSQV that in -glycanase, which removes N-linked glycans by cleaving the glycan from Asn at the N-glycosidic linkage, also converts the Asn residue to Asp (41).

**Glycosylation Protects Cleavage Sites in Class II Antigen Processing**

Class II MHC molecules draw on a different pool of peptides that are generated in the endocytic pathway from proteins that have either been diverted to endosomes from the secretory pathway or internalized by receptor-dependent or -independent endocytosis. Several endosomal and lysosomal proteases have been implicated in generating class II restricted peptides, including the endopeptidases cathepsin D, L, and S and the exopeptidases cathepsin A, B, and H (3). A carbohydrate-dependent, HLA class II–restricted human T cell response to the bee venom allergen phospholipase A2 has been described in allergic patients (42). For microbial tetanus toxin, the dominant processing enzyme is an asparagine-specific cysteine endopeptidase (AEP) (43). N-glycosylation blocks the action of AEP, indicating that glycans can protect self proteins and some viral glycopeptides from degradation. In contrast, microbial organisms which do not contain N-linked sugars can be fully processed to antigenic peptides.

**Glycosylation and T Cell Recognition of Antigen Presenting Cells**

Multiple events and stages are involved in the T cell recognition of antigen presenting cells (APCs). These include the formation of a junction between the two cells known as the immunological synapse (43, 44), recognition of antigenic peptide-loaded MHC molecules by the TCRs (Fig. 3) and signal transduction. Sugars are likely to be important in many of these processes. For example, receptor molecules of a similar length are recruited into a specialized junction between the T cell and APC (45). Oligosaccharides located on the membrane proximal domains tend to be conserved across species and seem likely to restrict the orientations of the cell adhesion molecules CD2 and CD48 contributing to the alignment of the opposing cell surfaces (46) [(Fig. 3)].

In addition, oligosaccharides may play a role in the transport of peptide-loaded MHCs (pMHCs) into the center of the junction where they are recruited by TCRs (Fig. 3). In vivo, this movement takes place over a 5-min time scale and may involve serial engagements of pMHC with the TCR, mediated by actin-based transport mechanisms (44). The size and location of the glycans prevent nonspecific protein-protein interactions, such as the aggregation of TCRs on the membrane, as well as possibly limiting the geometry of the interactions of the proteins in the central clusters (47). Extensive glycosylation makes it unlikely that specific interactions of the protein components of αβTCR-pMHC

---

**Fig. 2.** A proposed model for the folding, and peptide loading of MHC I and the degradation of unfolded glycoproteins from the ER. Nascent glycoproteins are translated across the lumen of the ER and within the cytosol may become immediately associated with membrane-bound calnexin (Cnx). Association of glycopeptides with Cnx is dependent upon the presence of a single terminal glucose residue bound to the GlcNAc-Man9 glycan precursor. In this example, interaction of Cnx is shown with newly synthesized MHC class I glycoproteins. The thiol oxidoreductase Erp57 also binds and facilitates disulfide bond formation, which facilitates the interaction with β2M. This induces dissociation of class I from Cnx, which is replaced by the soluble chaperone calreticulin (Clr). MHC class I-β2M dimers become loaded with peptide while associated with the TAP transporter, an interaction mediated by tapasin. The peptide-MHC complex is released and transported to the cell surface. Misfolded glycoproteins are recognized as such by enzymes, which convert the GlcNAc-Man9 structure to GlcNAc-Man3Glc, in a stepwise manner. The GlcNAc-Man3Glc structure is recognized by Cnx, and the glycoprotein is subsequently targeted for proteolytic degradation through the Sec 61 pathway. It is not yet proven whether the terminal Glc remains attached during this process.
can form laterally on the cell surface. In support of this notion, no dimers or specific higher order oligomers of soluble TCR-pMHC molecules have been found in the crystal lattices of mouse (48, 49) and human (50, 51) class I TCR-pMHC complexes. Within the synapse, specific protein-protein contacts must arise either from interaction of the CD4 or CD8 co-accessory molecules with the TCR-pMHC or from lateral interactions of the CD3 components.

During signal transduction, the sugars can also play a role in stabilizing the individual molecules (52) in the complexes in the synapse by protecting them from the action of proteases during T cell engagement in a process which, for class II APCs, may take several hours.

MHC I interacts with TCRs on CD8 T cells, whereas MHC II is recognized by CD4 T cells. Although CD4 and CD8 are structurally very distinct molecules, they interact with structurally homologous sites on their MHC ligands. The binding site of CD4 is held at the appropriate distance from the T cell membrane by protein domains. In CD8, the globular head containing the pMHC binding site is tethered to the T cell membrane by an extended polypeptide stalk that contains four O-linked sugars. These sugars are expected to confer some rigidity to this extended peptide chain, to protect it from digestion by proteases, and to inhibit nonspecific interactions of the stalk with the TCR.

Recognition of Glycopeptide Epitopes by T Cells

Around 0.5% of peptides that are bound to MHC class I molecules on the surface of normal cells carry O-linked N-acetylgalactosamine (GlcNAc) residues (53), which are presumably derived from O-glycosylated cytosolic and nuclear proteins (54). In contrast, peptides carrying N-linked GlcNAc residues are probably not presented by MHC class I (55), although N-GlcNAc–substituted peptides are naturally presented on the surface of cells bound to MHC class II molecules (56).

Both CD8+ and CD4+ T cells can recognize glycopeptides carrying mono- (55, 57–59) and disaccharide substitutions (58, 60) in a classical MHC-restricted way. In these cases, specificity of the T cell depends on contacts with both the glycan and the peptide moieties of the antigenic glycopeptide (61, 62). In the context of the MHC molecule, the crystal structures of three MIC class I–glycopeptide complexes (61, 63) clearly show that recognition of the additional carbohydrate moieties is within the physicochemical capability of an αβTCR, and certain motifs in the TCR CDR3 region recur in T cells recognizing O-GlcNAc–substituted glycopeptides (64, 65). In addition, titration of sugars of varying lengths has shown that the αβTCRs can recognize up to two sugars plus an extensive homocysteine linker that itself would correspond in length to approximately two sugars (63). Longer sugar chains can elicit γδ TCR responses which are not MHC restricted (63).

That such large antigens can be presented to T cells has been confirmed by the structure of a rat MHC class I molecule that is bound to a 13-residue peptide oligomer (13-mer) (66). The peptide bulges out extensively up to 13 Å above the peptide-binding groove on the MHC class I molecule. The immunodominant arthritogenic MHC class II–restricted T cell epitope in collagen-induced arthritis in mice is a glycopeptide with one or two hydroxylsine residues substituted with galactose (67), and T cells recognize carbohydrates on type II collagen (67). The presence of larger glycan chains has been shown to lead to the loss of MHC-restricted recognition of glycan by glycopeptide–specific T cells, and the tendency to induce γδ T cells which recognize the nominal glycan antigen, regardless of whether it is tethered to the APC surface via an MHC-binding peptide or via a lipid tail (58).

Other studies have also indicated that glycopeptides containing more complex glycans can be recognized by classical αβ T cells in an MHC class I (68) and class II–restricted manner. Moreover, in a detailed study of the CD1-restricted recognition by both CD8+ and CD4+ αβ T cells of the ganglioside GM1, the pentasaccharide moiety was shown to be the minimal structure for forming the epitope. Whether the entire pentasaccharide can be accommodated by the central TCR antigen combining region is unknown, but the crystal structure of the disaccharide-substituted glycopeptide:MHC complex (69) suggested that the longer oligosaccharides may be accommodated outside the normal “footprint” of the TCR.

Glycolipids and GPI Anchors Are Presented by CD1

CD1 [reviewed in (2, 5–7)] is a nonpolymorphic, β2,M-associated cell surface glycoprotein with antigen-presenting properties. In contrast to MHC class I and II molecules, the five isoforms of CD1 (CD1a, CD1b, CD1c, CD1d, CD1e), do not present peptides to TCR but lipids or glycolipids which are acquired within the acidic endosomal MIIIC compartments (2). While CD1a, CD1b, and CD1c present glycolipids derived from the cell walls of...
mycobacteria, such as *M. leprae* and *M. tuberculosis* (5), murine CD1d has been shown to present GPI anchors (69) and human CD1e, isoprenoids (70). The exact details of the molecular interactions which allow bulky glycolipids to be accommodated in the presentation groove of CD1 are not known. However, significant insights have been gained from the crystal structure of murine CD1d that revealed a deep, voluminous binding groove consisting of two very hydrophobic pockets (71). On the basis of this structure, it is predicted that the long fatty acid tails of the glycolipid bind in these pockets with the hydrophilic glycan moiety oriented so that it protrudes from the binding groove making contact with the TCR (5).

**Antibody-Antigen Recognition by Immunoglobulins in the Extracellular Space**

In the extracellular space, two forms of antigen-binding are currently recognized. The first is the recognition of repetitive arrays of sugars, for example on bacteria or yeast, by the MBL. The second is the classical recognition of antigen by immunoglobulins.

Structural roles for glycosylation in IgG are well documented. In human IgG Fc (Fig. 4A), crystallographic studies (72) have shown that the two C_H2 domains do not interact by protein-protein contacts, but instead through an interstitial region that is formed by oligosaccharides, attached at Asn^297 on each heavy chain. Protein-oligosaccharide and oligosaccharide-oligosaccharide interactions play a role in maintaining the relative geometry of the C_H2 domains (73), consistent with less efficient binding of both aglycosylated and degalactosylated IgG to the Fcγ receptor (72-74).

By contrast, IgA1 (Fig. 4A) does not contain an interstitial space large enough to accommodate the Fc sugars and the Asn^258 side chains are fully exposed and point out into solution (74). Their ER and Golgi processing is unhindered by interaction with the protein surface, and most serum IgA1 N-linked sugars, which include tri- and biantennary structures, are fully galactosylated and sialylated. The hinge region of IgA1, which contains five O-glycosylation sites, is resistant to many com-
mon proteases (75). The sugars, which on se-
ruum IgA1 mainly consist of sialylated GalGal-
NAc (74, 76, 77), shield large areas of the hinge
region (Fig. 4A) and produce an extended rigid
structure (78).

In IgG, the constraint imposed by the inter-
chain location of the sugars within the C12-2
domains restricts them to complex biantennary
structures. A further constraint arises when a
galactose residue is added to the terminal
GlcNAc residue on the α1,6 arm during the
biosynthetic process of elongating the sugar
chain. This galactose residue forms a tight in-
teraction with a lectin-like binding pocket on
the antibody (Lys246, Glu258, and Thr260),
which results in further immobilization of the
glycan chain. However, sugars that are first
agalactosylated on the α1,3 arm (G1α1,3) can
more readily become digalactosylated, as re-
lected in the final ratio of G1α1,3 to Glcα1,6
glycans, which is about 1:4 (79). Another fac-
tor that significantly influences the distribution
of glycoforms in IgG is the level of galactosyl
transferase that transfers galactose to terminal
N-acetylgalcosamine residues; its levels deter-
mine the relative proportions of zero-, mono-, and
digalactosylated IgG (IgG0, IgG1, and
IgG2 respectively). The arm-specific distribu-
tion of monogalactosylated IgG between the
G1α1,6 and G1α1,3 glycoforms continues to be
determined by the protein-sugar interactions.

In rheumatoid arthritis, galactosyl trans-
ferase levels are decreased such that IgG0 is
increased. The lack of terminal galactose resi-
dues on the α1,6 arm result in the exposure of
the terminal GlcNAc residues to the MBL.

Recognition of Repetitive Arrays of Sugars by the MBL
Sugar-protein interactions that involve single
monosaccharide residues have low binding affinities. In general, to trigger biological
events, multiple presented sugar ligands are
required to interact with multivalent recep-
tors. Thus, in carbohydrate–lectin–type inter-
actions, or in carbohydrate antigen–antibody
interactions, several monosaccharide moi-
eties must normally be presented in the cor-
rect conformation in order to bind to the
receptor with high affinity (80). Alterna-
tively, in C-type lectins, clustering of multiple
copies of the same sugar epitope on a surface
in a particular geometry, can provide a mul-
tivalency surface for recognition by the multi-
ple carbohydrate recognition domains (81).

In the extracellular space, many bacteria and yeasts present arrays of mannose residues on
their surfaces to the MBL (82) that leads to
activation of the complement pathway and

**Carbohydrates and Glycobiology**

Glycoproteins are key components of the
immune system effectors. Their oligosaccha-
rides are important in synthesis, stability, rec-
ognition and regulation of the proteins them-
selves and in many of their diverse interac-
tions. Establishing the extent and variability of
glycosylation in the context of the three-
dimensional structure of the proteins to which
they are attached will lead to further insights
into immune processes in health and disease.
It is already becoming evident that the man-
ner in which oligosaccharides are presented
provides a mechanism for distinguishing self
from non-self. Pathogens which present spe-
cific oligosaccharides as repetitive arrays
can, if these arrays have a suitable geometry,
activate multivalent receptors. One example
is the MBL, which is involved in the recog-
nition and subsequent elimination of patho-
gens. Others may yet be discovered. Sugars
attached to mammalian glycoproteins (self)
will not normally be presented in an appro-
priate homogeneous geometrical array to ac-
tivate the immune response. The ubiquity and
diversity of protein glycosylation is not a
paradox, but consistent with functional rules.

**References and Notes**
30. J. A. Lindquist, O. N. Jensen, M. Mann, G. J. Hammer-
34. B. Sadasivan et al., Immunol. 5, 103 (1996).
50. Y. H. Ding et al., Immunity 8, 403 (1998).
**Carbohydrates and Glycobiology**

**Glycosylation of Nucleocytoplasmic Proteins: Signal Transduction and O-GlcNac**

Lance Wells,* Keith Vosseller,* Gerald W. Hart†

The dynamic glycosylation of serine or threonine residues on nuclear and cytosolic proteins by O-linked β-N-acetylglucosamine (O-GlcNac) is abundant in all multicellular eukaryotes. On several proteins, O-GlcNac and O-phosphate alternatively occupy the same or adjacent sites, leading to the hypothesis that one function of this sarachine is to transiently block phosphorylation. The diversity of proteins modified by O-GlcNac implies its importance in many basic cellular and disease processes. Here we systematically examine the current data implicating O-GlcNac as a regulatory modification important to signal transduction cascades.

Cells respond to their environment through dynamic posttranslational modification of their existing proteins. There are more than 20 posttranslational modifications that occur on eu-

karyotic proteins (1). Several of these modifications, with phosphorylation being the hallmark, participate in signal transduction. Generally, glycosylation is not thought to participate directly in signaling. Complex N- and O-linked glycosylation occurs on membrane-bound or secreted proteins that are synthesized in the endoplasmic reticulum and Golgi apparatus. The luminal or extracellular localization of these glycans restricts their potential for dynamic responsiveness to signals. In contrast, O-GlcNac is a simple monosaccharide modification that is abundant on serine or threonine residues of nucleocytoplasmic proteins (2, 3). An O-GlcNac site consensus motif has not yet been identified. However, many attachment sites are identical to those used by serine/threo-

nomic kinases, and a neural network program has been developed to predict O-GlcNac sites (4). Unlike phosphorylation, O-GlcNac modification of tyrosine residues has yet to be observed. Many proteins have been identified that carry this modification, including transcription factors, cytoskeletal proteins, nuclear pore proteins, oncogene products, and tumor suppressors (5–7). O-GlcNac appears to modify a large number of nucleocyto-

plasmic proteins (Fig. 1). The attributes of the modified polypeptide chain (5) of the modification have been purified, characterized, and cloned (9–12). The OGT enzyme is modified by both O-GlcNac and tyrosine phosphorylation and has 11 protein-

protein interaction domains known as tetratricopeptide (TPR) repeats (13). OGT specifically interacts with a variety of other proteins. Because the transducer has a variety of binding partners and is itself Viewpointposttranslationally modified, OGT’s substrate specificity, locali-

zation, and/or activity are likely to be regulated by signal transduction cascades. Furthermore, the activity of this enzyme is exquisitely responsive to intracellular uridine 5’-diphos-

phate (UDP)-GlcNac and UDP concentrations, which are in turn highly sensitive to glucose concentrations and are known to fluctuate in response to a number of stimuli (14). Disruption of OGT activity is lethal in mouse embryonic stem cells, underscoring the importance of this modification (15). O-GlcNacase, which specifically catalyzes the removal of O-

GlcNac from proteins, is a cytosolic neutral β-N-acetylgalactosaminidase, unlike the general acidic lysosomal hexosaminidas.

Both OGT and O-GlcNacase are highly conserved from Caenorhabditis elegans to humans (9, 10, 12). Regulation of O-GlcNac levels by these two enzymes may be analogous to the regulation of phosphorylation by kinases and phosphatases.

---

*Deceased. ViewpointDepartment of Biological Chemistry, Johns Hopkins School of Medicine, 725 North Wolfe Street, Baltimore, MD 21205 USA.

†To whom correspondence should be addressed. E-mail: gwhart@jhmi.edu