Glycoconjugate Journal (1992) **9**:63

GLYCOPINION

Editor: RAYMOND A. DWEK

The glycoproteins of enveloped viruses provide interesting and important challenges to glycobiologists intent on determining structure–function relationships. The authors of this article discuss the role of glycosylation in the further diversification of an already heterogenous population of molecules. The implications of the presence of such a spectrum of structures are discussed with regard to host range, infectivity, efficacy of vaccines and the use made by the virus of the host cell glycosylation and transport systems.

Some of the issues which this article raises include:

- Does glycosylation enhance the functional heterogeneity of the glycoproteins on enveloped viruses? Could this promote the survival of such viruses in different ecological niches or when the immune response or antiviral agents change the conditions under which these viruses replicate?
- What are the consequences for glycosylation of the mutations which lead to the quasi species distribution of the RNA virus populations?
- What are the major factors which determine the oligosaccharide profiles of the envelope glycoproteins?
- What are the effects of viral infection on the normal glycosylation pathways in the cell if viruses can alter host cell protein synthesis?
- Can specific glycosylation changes alter the degree of virulence of a virus?

Clearly, understanding the glycosylation of viruses is one of the emerging challenges for glycobiology.

Viral glycoprotein heterogeneity-enhancement of functional diversity

IRENE T. SCHULZE* and IAN D. MANGER

Department of Microbiology, St Louis University School of Medicine, 1402 South Grand Blvd, St Louis, MO 63104, USA

Variations in the amino acid sequence of RNA virus envelope glycoproteins can cause changes in their antigenicity and can alter the host-cell tropism of the virus and the degree of virulence which it exhibits. Such changes may alter the course and outcome of viral diseases, either directly because of changes in the biological properties of the glycoproteins or indirectly through effects on immune surveillance and vaccine efficacy. The nature and extent of glycosylation of the surface glycoproteins of RNA viruses have also been implicated in such phenotypic alterations. It follows therefore that the 'plasticity' of the viral genome and the host-encoded glycosylation machinery combine to create populations of highly diverse viruses. This diversity is considered to be responsible for survival of these viruses in a variety of biological niches and for their ability to overcome the inhibitory effects of neutralizing antibodies and antiviral agents. In this article we discuss the implications of the inter-relationship between these two mechanisms for the generation of diversity.

RNA virus genome heterogeneity and the quasi-species concept

The accumulation of nucleotide substitutions within the genomes of RNA viruses appears to occur at rates up to a million times faster than in cellular (DNA) genes. This enormous difference in the rate of evolution can be attributed to the short generation time of viral genomes and to the low fidelity of the viral RNA-dependent polymerases. These enzymes lack the exonucleolytic proof-reading capabilities associated with cellular DNA polymerases. Current estimates of the fidelity of the RNA-dependent RNA polymerases suggest nucleotide mis-incorporation rates of 10^{-3} to 10^{-4} nucleotides per replication cycle. Similar estimates have been obtained for the frequency of nucleotide mis-incorporation by the RNA-dependent DNA polymerase (reverse transcriptase) required for replication of retroviruses, including the human immunodeficiency

* To whom correspondence should be addressed.

0282-0080

virus (HIV). These estimates suggest a mis-incorporation rate of about one nucleotide per genome per replication cycle. Thus, after a number of cycles, there is significant sequence diversity within the virus population both at the nucleotide and protein coding levels. Uncloned populations of most RNA viruses therefore consist of heterogeneous mixtures of related genomes (quasi-species) [1, 2] rather than a single genome species. These genomes share a consensus sequence which identifies the nucleotide found most frequently at each position in the sequence. Since the individual genomes within the population usually differ from each other and from the consensus sequence at one or more positions, the proteins which they encode are also heterogeneous in sequence.

HIV virus populations provide a prime example of this heterogeneity and of the quasi-species concept. Isolates of HIV obtained from an infected individual at a single time point show significant variation in all viral genes but particularly in the sequence of the viral envelope glycoprotein, gp120. Sequence diversity increases in sequential isolates from a single individual [3]. The progression from HIV infection to symptomatic disease (AIDS) correlates with the emergence of viral isolates of increasing virulence [4]. These changes in phenotype can be more easily observed with HIV than with most other RNA viruses because HIV enjoys an exceptionally long period of time during which it can grow and accumulate mutations within a single patient. However, with other enveloped RNA viruses, similar patterns of heterogeneity can be detected; virus populations, both those obtained directly from clinical material and those grown in cell culture, contain particles which differ from each other in antigenic and/or host-cell binding properties and minority forms within the populations can be selected as monoclonal antibody escape mutants or by growth in a different host [5–7].

Glycosylation of viral proteins

Given this diversity in amino acid sequence, what role can glycosylation play in promoting phenotypic diversity? Do functional viral glycoproteins require highly defined oligosaccharide structures at specific sites? How is this achieved if the population of molecules to which the sugars are attached is heterogeneous in amino acid sequence?

Firstly, all viruses make use of the host cell machinery for the synthesis, folding, and transport of proteins to the site of virus assembly at the cell surface. This machinery includes the array of biosynthetic and trimming enzymes responsible for attachment and processing of the oligosaccharides on their glycoproteins. No virus has been found to encode enzymes which can affect the glycosylation of its proteins by controlling commitment to particular processing pathways. Although the influenza A and B viruses encode a neuraminidase, this enzyme is critical for release of progeny virus from the host cell rather than for the processing of its oligosaccharides. The absence of virally encoded enzymes indicates that functional glycoproteins can be made using the oligosaccharide structures specified by the cell.

Processes by which viruses control the structure of the oligosaccharides on their glycoproteins

Although the cell is the major determinant of the oligosaccharide structure of viral glycoproteins, the oligosaccharides found on viruses usually represent only a subset of those found on the host cell. How is this subset selected or determined?

The spectrum of structures found on viral glycoproteins reflects first of all the complement of processing and trimming enzymes present in the infected host cell. Since some enveloped viruses substantially reduce host cell protein synthesis, the copy-number and activity of these enzymes may be changed during infection. Changes in the availability of dolichol precursors and sugar nucleotide substrates may further affect glycosylation.

A second major determinant of glycosylation is the viral polypeptide itself. The utilization of glycosylation sequons and the spectrum of oligosaccharide structures found at specific sites will be determined in part by the primary amino acid sequence, polypeptide folding, and the accessibility of the oligosaccharides to the various processing enzymes. All three of these determinants are obviously highly specific for each polypeptide independent of whether it is encoded by the cell or the virus.

A third and highly important determinant of the spectrum of glycosylated proteins found on virions is the selection of appropriately glycosylated polypeptides during their folding, transport, and assembly into particles. The diversity of oligosaccharide structures can be limited at any of these steps by the requirement to produce a functional protein.

What will be the effect of polypeptide sequence heterogeneity on these processes? Since individual copies of the glycoprotein will differ in amino acid sequence, they can be expected to differ in the number and position of both *N*-linked and *O*-linked glycosylation sites. This can be clearly demonstrated with *N*-linked glycosylation sites. We have shown that influenza virus populations recovered directly from a patient by nasopharyngeal swab contain some hemagglutinin (HA) genes which lack sequons indicated by the consensus sequence for the population as a whole (unpublished observations). The effect of sequence heterogeneity on the positioning of *O*-linked glycosylation sites is difficult to ascertain because the amino acid sequence requirements for the attachment of these carbohydrates are not clearly defined. Secondly, amino acid sequence-mediated differences in the secondary structure of the protein or in interactions between carbohydrate chains adjacent to one another in the folded structure can be expected to influence the processing of oligosaccharides at specific locations on the glycoprotein.

Biological consequences of viral glycoprotein heterogeneity

What are the biological consequences of having a population of glycoproteins which are not identical in amino acid sequence? Analysis of the accumulation of amino acid substitutions in the influenza virus HA molecule over a period of years indicates that certain regions of the three-dimensional structure are highly conserved. Substitutions in the more variable regions of the molecule, including those which add or delete glycosylation sites, give rise to the antigenic drift observed in all of the human influenza subtypes. Analysis of the distribution of amino acid substitutions in gp120 of HIV isolates indicate the same general pattern of conserved and variable regions. Whilst the sequences which define the tertiary structure and functional regions such as the CD4 binding domain are highly conserved, they are interspersed with hypervariable antigenic loops in which a variety of amino acid substitutions are tolerated. Amino acid substitutions in these loops frequently delete or introduce sequons for *N*-linked glycosylation, whereas sequons in the constant regions are highly conserved [3].

The conservation of certain glycosylation sites points to their having important roles in determining or maintaining the structure and functional integrity of the molecule. The oligosaccharides at the more highly conserved *N*-linked glycosylation sites appear to be important in the folding and/or assembly of the glycoproteins into their biologically active forms. They also appear to play an important role in protection of the molecule by masking proteolytically sensitive sites.

On the other hand, there is clear evidence that some regions of viral glycoproteins need to be free of oligosaccharides. Certain regions of the influenza virus HA have been found not to contain sequons for *N*-linked glycosylation despite years of antigenic drift. The addition of oligosaccharides to these regions by site directed mutagenesis reduces transport of the HA to the cell surface and/or reduces the structural stability and function of the HA [8].

What are the effects of the presence or absence of oligosaccharides in the less well-conserved regions of the viral glycoproteins? It is in these regions that oligosaccharide diversity can have a major effect since, depending on the environment in which the virus finds itself, the presence or absence of an oligosaccharide may determine viral fitness and survival. Glycosylation can affect the functional integrity of the glycoprotein in a number of ways. It can modulate the proteolytic cleavage-dependent maturation events which lead to the acquisition of fusion activity required for infection of some viruses [9]. Secondly, it affects the receptor binding properties and cellular host range of some viruses [7]. These changes can be lethal if they interfere with the infectious process within one type of host cell. However, they also promote the sampling of new host cell environments.

Glycosylation also results in modification of the antigenic properties of the virus. The presence of *N*-linked carbohydrates in or near epitopes appears to interfere directly with the binding of antibodies [6, 10] and with the stimulation of the cellular immune response by peptide epitopes [11]. The importance of this effect of glycosylation on viral survival is suggested by the observation that all H1 influenza viruses isolated since 1977 have HAs with eight conserved glycosylation sites on the HA-1 subunit whereas viruses of the same subtype which have been grown for many years in embryonated eggs and in cell cultures contain as few as four. During virus replication in nature, host cell mimicry, resulting from the use of the glycosylating machinery of the cell, apparently fosters evasion of immune surveillance.

The effect of glycosylation on population dynamics

Because of the large size of virus populations, individual virions with glycoproteins altered at a specific site are likely to represent only a minority component within a quasi-species population which is in dynamic equilibrium. They are therefore unlikely to be detected in the population at large, even though they are phenotypically distinct from the remainder of the population. How then, does glycosylation diversity affect the virus population as a whole?

The amplitude of diversity within a quasi-species population (the mutant spectrum) and its consensus sequence will remain constant only until the selective pressures which determine the composition of the population are changed. When this occurs, selection of one or more of the minority forms from within the population can result in the establishment of a new mutant spectrum and the emergence of a new consensus sequence.

Environmental factors which are likely to change the population equilibrium include such processes as:

1) Transfer of a few infectious particles to a susceptible host or isolation of the progeny from a single particle *in vitro*,

2) Antibody neutralization of the dominant antigenic subsets of the virions within the population,

3) Substantial reduction in the size of the virus population by chemotherapeutic agents,

4) Changes in the receptor specificity of the viral glycoprotein responsible for attachment to host cells,

5) Changes in the receptors available to the virus due to interspecies transfer,

6) Changes in the distribution of a mixed population of host cells by the cytocidal activity of the virus.

Based on the above discussion, it should be clear that the number, position and type of oligosaccharides on a viral glycoprotein will play a role in determining the fitness of the virus for its environment and therefore in the dynamics of the population. Thus, while the importance of the marked heterogeneity of *N*- and *O*-linked carbohydrates in modifying

the bioactivities of cellular proteins is a subject for debate, viruses appear to make use of this heterogeneity to enhance their own survival. The combination of a rapidly mutating genome, the evolution of envelope proteins which tolerate amino acid sequence variations whilst retaining essential structural features, and the utilization of host-encoded post-translational modifications enable RNA viruses to present rapidly changing 'faces' to their hosts.

Lastly, we submit that the information presented here emphasizes the importance of considering the properties of the viral population rather than focussing on its isolated elements when evaluating therapeutic options and vaccine design.

References

1. Eigen M and Biebricher CK (1988) in *RNA Genetics* (Domingo E, Holland JJ, and Ahlquist P, eds) Vol. 3, pp. 211–45. Boca Raton, Fla.: CRC Press Inc.
2. Steinhauer DA and Holland JJ (1987) *Ann Rev Microbiol* **41**:409–33.
3. Klotman ME and Wong-Staal F (1991) in *The Human Retroviruses* (Gallo RC and Jay G, eds) pp 54–9. New York: Academic Press Inc.
4. Cheng-Mayer C, Seto D, Tateno M, and Levy JA (1988) *Science* **240**:80–2.
5. Katz JM and Webster RG (1988) *Virology* **165**:446–56.
6. Aytay S and Schulze IT (1991) *J Virol* **65**:3022–8.
7. Crecelius DM, Deom CM and Schulze IT *Virology* **139**: 164–77.
8. Gallagher P, Henneberry J, Wilson I, Sambrook J and Gething MJ (1988) *J Cell Biol* **107**:2059–73.
9. Kawaoka Y and Webster RG (1989) *J Virol* **63**:3296–300.
10. Skehel JJ, Stevens DJ, Daniels RS, Douglas AR, Knossow M, Wilson IA and Wiley DC (1984) *Proc Nat Acad Sci USA* **81**:1779–83.
11. Botarelli P, Houlden BA, Haigwood NL, Servis C, Montagna D, and Abrignani S (1991) *J Immunol* **147**:3128–32.

Letters or comments relating to this article would be received with interest by Pauline Rudd, Assistant to the Special Advisory Editor, R. A. Dwek.