Glycosyltransferases encoded by viruses

Nicolas Markine-Goriaynoff,1 Laurent Gillet,1 James L. Van Etten,2 Haralambos Korres,3 Naresh Verma3 and Alain Vanderplasschen1

1Immunology-Vaccinology (B43b), Department of Infectious and Parasitic Diseases, Faculty of Veterinary Medicine, University of Liège, B-4000 Liège, Belgium
2Department of Plant Pathology and Nebraska Center for Virology, University of Nebraska, Lincoln, NE 68583-0722, USA
3School of Biochemistry & Molecular Biology, Faculty of Science, Australian National University, Canberra, ACT 0200, Australia

Studies of cellular biology in recent decades have highlighted the crucial roles of glycans in numerous important biological processes, raising the concept of glycomics that is now considered as important as genomics, transcriptomics and proteomics. For millions of years, viruses have been co-evolving with their hosts. Consequently, during this co-evolution process, viruses have acquired mechanisms to mimic, hijack or sabotage host processes that favour their replication, including mechanisms to modify the glycome. The importance of the glycome in the regulation of host–virus interactions has recently led to a new concept called ‘glycovirology’.

One fascinating aspect of glycovirology is the study of how viruses affect the glycome. Viruses reach that goal either by regulating expression of host glycosyltransferases or by expressing their own glycosyltransferases. This review describes all virally encoded glycosyltransferases and discusses their established or putative functions. The description of these enzymes illustrates several intriguing aspects of virology and provides further support for the importance of glycomics in biological processes.

Introduction

After genomics, transcriptomics and proteomics, glycomics is an emerging field bringing a new insight into the biology of the cell. The glycome of a biological entity has been defined as all the sugars it makes, including glycans fixed on proteins, lipids or DNA (Hirabayashi et al., 2001). The glycome is characterized by non-templated and not strictly regulated biosynthesis, producing highly versatile molecules with intricate three dimensional structures. Until recently, saccharides bound to glycoproteins were considered little more than an irritation, increasing the difficulty of purifying and characterizing the ‘important part’, the protein moiety. We now realize that the saccharide is often as important as the protein itself, and that glycosylation can have many effects on the function, structure, physical properties and targeting of a protein. For excellent reviews on the synthesis of glycans and their biological importance, we recommend the special issue of Science devoted to the chemistry and biology of carbohydrates (http://www.sciencemag.org/content/vol291/issue5512/).

For millions of years, viruses have been co-evolving with their hosts. During this co-evolution process, viruses have had to deal with the most complex aspects of the host physiology, mimicking, hijacking and sabotaging host biological processes in their favour. Recently, a growing list of studies has highlighted the importance of the glycome in the regulation of host–virus interactions. These studies have led to the new concept of ‘glycovirology’, which was the main topic of the First International Meeting on Viral Glycobiology (June 15–18 2003, Göteborg, Sweden).

Viruses modify the glycome by two distinct mechanisms. Some viruses affect the expression of host glycosyltransferases (Cebulla et al., 2000; Hiraïwa et al., 2003). For example, human T-cell leukaemia virus (HTLV-1) transactivates the expression of cellular fucosyltransferase VII thereby increasing the expression of sialyl-Lewis x (sLex) on the surface of infected cells. Interestingly, the degree of expression of sLex on leukaemic cells in adult T leukaemia correlates significantly with the degree of extracellular infiltration. While some viruses affect the expression of host glycosyltransferases, other viruses modify the glycome by expressing their own glycosyltransferases. In this review, we describe all the glycosyltransferases encoded by viruses, and discuss their established or putative functions. The description of these enzymes illustrates several fascinating aspects of virology. (i) Viruses have the ability to avoid host anti-viral mechanisms. For example, some bacteriophages express α- and β-glucosyltransferases which
Glucosylate their DNA to make it resistant to host restriction endonucleases. (ii) Temperate bacteriophages infecting pathogenic bacteria confer additional virulence factors to their host. For example, some bacteriophages modify the host-cell serotype through expression of virally encoded glucosyltransferases during lysogenic infection. Presumably, this phenomenon has played a crucial role in the emergence of new serotypes and development of some pandemic or maintenance of some endemic bacterial diseases. (iii) Some viruses infecting multicellular organisms have the capability to alter host metabolism. For example, most baculoviruses encode an ecdyseroid glucosyltransferase that transfers glucose to ec dysyroid insect molting hormones. Expression of this enzyme allows the virus to block molting and pupation of infected insect larvae. (iv) Some viral genes have been acquired from the host at different times in the past during the co-evolution process. For example, the gene encoding the core 2-1,6-N-acetylgalactosaminyltransferase-M of bovine herpesvirus 4 (BoHV-4) was acquired approximately 1.5 million years ago from an ancestor of the African buffalo. These examples of virally encoded glycosyltransferases and others are detailed in this review.

Glycosyltransferase encoding genes have been reported in bacteriophages, phycodnaviruses, baculoviruses, poxviruses and herpesviruses (Table 1). They will be described in the order according to the biological complexity of their host. It is important to note that this review represents an exhaustive list of the viral glycosyltransferases described to date rather than a description of the best characterized representative.

Glycosyltransferases encoded by bacteriophages

Genes encoding two types of glycosyltransferases exist in bacteriophages. Some lytic bacteriophages glucosylate their DNA to protect it from host restriction endonuclease systems, while some temperate bacteriophages express glycosyltransferases inducing serotype conversion of the host bacteria during lysogeny. Descriptions of these two types of glycosyltransferases follow.

Glycosyltransferases encoded by lytic *E. coli* bacteriophages T2, T4 and T6. *E. coli* bacteriophages T2, T4 and T6 code for enzymes that transfer glucosyl residues to the hydroxymethylcytosines (HMC) of their DNA and are thus protected against C- and HMC-DNA-specific restriction endonucleases encoded by the chromosome of the host or its plasmids. These enzymes are the α-glycosyltransferases found in bacteriophages T2, T4 and T6 and the β-glycosyltransferases found in bacteriophage T4. They are responsible for the formation of α- and β-glycosidic linkages, respectively (Gram & Ruger, 1986; Freemont & Ruger, 1988; Winkler & Ruger, 1993). Another enzyme, β-glucosyl-HMC-α-glucosyl-transferase, produces the gentiobiosyl groups on the DNA of phages T2 and T6 (Gram & Ruger, 1986; Freemont & Ruger, 1988; Winkler & Ruger, 1993). The T4 bacteriophage β-glucosyltransferase (BGT) has been studied extensively, and its structure, catalytic and displacement mechanisms have been characterized (Vrielink et al., 1994; Morera et al., 2001; Lariviere & Morera, 2002; Lariviere et al., 2003).

The phage-encoded glucosyltransferases described above represent an efficacious protection system of the phage DNA against the host endonuclease restriction system. However, some bacteria have acquired glucosylated DNA-specific endonucleases enabling them to destroy glucosylated phage DNA (Ishaq & Kaji, 1980; Janosi et al., 1994). In addition to its protective role against host restriction endonucleases, DNA glucosylation could also influence the structure (Carlson et al., 1994) and expression (Sauerbier & Rautigam, 1970; Christiansen et al., 1973; Ruger, 1978; Dharmalingam & Goldberg, 1979) of the phage genome, as well as regulate some specific phage recombination systems (Levy & Goldberg, 1980). Glucosylation of DNA is a relatively rare process. Indeed, only one additional example of DNA glucosylation has been reported for *Trypanosoma brucei* where this process is believed to be involved in regulating the expression of variant surface glycoprotein (Gommers-Ampt et al., 1993). However, both for bacteriophages and trypanosomases the mechanisms by which DNA glucosylation affects the properties of the genome still need to be addressed.

Serotype-converting glycosyltransferases encoded by temperate bacteriophages. In permissive cells, temperate bacteriophages can lead to two different types of infection. The virus can either replicate, which results in cell lysis and release of progeny virions (lytic infection), or integrate into the host genome and express a limited set of genes establishing a quiescent infection (lysogenic infection). Eventually, lysogenic infection can be reactivated into a lytic infection. In many cases, temperate phages alter the phenotype of lysogenized cells resulting in the production of toxins or expression of modified cell surface antigens. The latter phenomenon is called lysogenic or antigenic conversion. Because temperate bacteriophages can confer a particular phenotype to their host, these viruses play a crucial role in the evolution of pathogenic bacteria (Barondess & Beckwith, 1990; Karaolis et al., 1999; Miao & Miller, 1999; Boyd & Brussow, 2002).

*Shigella flexneri* is the major cause of shigellosis or bacillary dysentery. The protective host immune response to *S. flexneri* is directed against the O-antigen component of the outer membrane lipopolysaccharide. The immune response to the O-antigen is serotype-specific and provides protection against further infection by an organism of the same serotype. The basic O-antigen of *S. flexneri* is referred to as serotype Y and the addition of glucosyl and/or O-acetyl groups to different sugars in the tetrasaccharide unit forms the basis of serotype conversion. The factors responsible for serotype conversion in *S. flexneri* are encoded by temperate bacteriophages. This process was first observed by Takita and Boyd, and Ewing in the late 1930s and 1950s, respectively (Takita, 1937; Boyd, 1938; Ewing, 1954). Later, the temperate phages responsible for this antigenic conversion were isolated from strains of...
<table>
<thead>
<tr>
<th>Virus</th>
<th>Glycosyltransferases</th>
<th>Biochemical effect:</th>
<th>Key reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bacteriophages</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>E. coli</em> bacteriophages T2, T4 and T6</td>
<td>α- and β-glucosyltransferases</td>
<td><strong>Glucosylation of viral DNA:</strong>&lt;br&gt;- protection of viral DNA from host endonuclease restriction system</td>
<td>Gram &amp; Ruger (1986); Freemont &amp; Ruger (1988); Winkler &amp; Ruger (1993); Vrielink et al. (1994); Lariviere et al. (2003)</td>
</tr>
<tr>
<td>Serotype converting bacteriophages</td>
<td>Gtr(type) (specific for each bacteriophage)</td>
<td><strong>Conversion of bacterial O-antigen:</strong>&lt;br&gt;- inhibition of cell superinfection by related phages&lt;br&gt;- inhibition of virion retention on cell debris after lytic infection&lt;br&gt;- host-cell serotype conversion</td>
<td>Allison &amp; Verma (2000); Vander Byl &amp; Kropinski (2000); Mmolawa et al. (2003)</td>
</tr>
<tr>
<td><strong>Phycodnaviruses</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chloroviruses:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBCV-1</td>
<td>ORFs a64r, a111r, a114r, a222-226r, a328l, a473l and a546l</td>
<td><strong>Post-translational modifications of the major capsid protein Vp54</strong></td>
<td>Wang et al. (1993); Graves et al. (2001); Van Etten et al. (2002)</td>
</tr>
<tr>
<td>PBCV-1</td>
<td>Hyaluronan synthase</td>
<td><strong>Synthesis of a dense hyaluronan network on infected cell surface</strong></td>
<td>DeAngelis et al. (1997); Graves et al. (1999)</td>
</tr>
<tr>
<td>CVK-2</td>
<td>Chitin synthase</td>
<td><strong>Synthesis of a dense chitin network on infected cell surface</strong></td>
<td>Kawasaki et al. (2002)</td>
</tr>
<tr>
<td>Phaeovirus:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EsV-1</td>
<td>ORF 84</td>
<td>- putative glycosyltransferase (putative hyaluronan, chitin or alginate synthase)</td>
<td>Delaroque et al. (2001)</td>
</tr>
<tr>
<td><strong>Baculoviruses</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All characterized baculoviruses, with exception of XcGV and PhopGV</td>
<td>Ecdysteroid transferase</td>
<td><strong>Glycosylation of insect ecdysteroid hormone:</strong>&lt;br&gt;- inhibition of insect moulting and pupation</td>
<td>O’Reilly &amp; Miller (1989); O’Reilly (1995)</td>
</tr>
<tr>
<td><strong>Poxviruses</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entomopoxviruses:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MsEPV</td>
<td>ORF MSV206</td>
<td>- putative glycosyltransferase</td>
<td>Afonso et al. (1999)</td>
</tr>
<tr>
<td>AmEPV</td>
<td>ORF AMV248</td>
<td>- putative glycosyltransferase</td>
<td>Bawden et al. (2000)</td>
</tr>
<tr>
<td>Chordopoxviruses:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MYXV and other leporipoxviruses</td>
<td>α-2,3-sialytransferase</td>
<td><strong>Post-translational modifications of SERP-1</strong></td>
<td>Jackson et al. (1999); Willer et al. (1999); Nash et al. (2000); Sujino et al. (2000)</td>
</tr>
<tr>
<td><strong>Herpesviruses</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhadinovirus:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BoHV-4</td>
<td>Core 2 β-1,6-N-acetylglucosaminyltransferase-mucin type</td>
<td><strong>Post-translational modifications of structural proteins</strong></td>
<td>Vanderplasschen et al. (2000); Markine-Goriaynoff et al. (2003, 2004)</td>
</tr>
</tbody>
</table>
various serotypes and characterized (Matsui, 1958; Okada et al., 1958; Iseki & Hamano, 1959; Gemski et al., 1975). The bacteriophage-encoded factors involved in O-antigen modification in *S. flexneri* are O-acetyltransferases or glycosyltransferases. During lysogeny, the phage Sf6 expresses an O-acetyltransferase responsible for the conversion of the unmodified serotype Y to serotype 3b (Verma et al., 1991). Other bacteriophages like Sfi, SfIII, SFV, SFX, and the assumed cryptic SfIV express glycosyltransferases responsible for the conversion of the unmodified serotype Y to serotypes 1a, 2a, 5a, X and 4a, respectively (Verma et al., 1993; Morona et al., 1995; Bastin et al., 1997; Huan et al., 1997a, b; Mavris et al., 1997; Adhikari et al., 1999; Guan et al., 1999; Allison & Verma, 2000; Adams et al., 2001; Allison et al., 2002). These phages are morphologically diverse and belong to several virus families, but common features occur especially in the organization of their glycosyltransferase genes. These genes are organized in a cluster located immediately downstream of the phage attachment site *attP*, which follows the integrase (*int*) and excisionase (*xis*) genes (Allison & Verma, 2000). This glycosylation cassette comprises three genes termed *gtrA*, *gtrB* and a serotype-specific glycosyltransferase [gtr(type)]; the expression products of those genes will be called hereafter GtrA, GtrB and Gtr(type), respectively. The two first proteins are highly conserved and interchangeable among serotypes, whereas the third protein appears to be unique to each bacteriophage. The mean GC content of the *gtrA* and *gtrB* genes is 42% while it is <35% in the gtr(type) genes. This is lower than the host GC content of 50% indicating that these genes may have been acquired from another organism or is directly related to the large number of hydrophobic amino acids in these proteins (Schnaitman & Klena, 1993).

The *gtrA* gene is highly conserved among the serotypes and encodes a 120 aa product. According to hydropathy data, GtrA is a four transmembrane protein with both N- and C-terminal ends located in the cytoplasm. Guan et al. (1999) proposed that GtrA functions as a flipase of the UndP-glucose precursor, allowing its transfer from the cytoplasm to the periplasm in order for it to be available to the Gtr(type) for the attachment of the glucose residue to the correct rhamnose unit conferring serotype specificity (Guan et al., 1999).

On the other hand, GtrB proteins are much larger in size than GtrA, ranging in length from 305 to 309 aa. As with GtrA, GtrB is highly conserved between serotypes with the greatest variation occurring in the last 20 aa. Hydrophathy data suggest that GtrB has two transmembrane domains with a large hydrophilic N terminus and a short C-terminal tail. *gtrB* encodes a bactoprenol glycosyltransferase that catalyses the transfer of glucose from UDP-glucose to bactoprenol phosphate in the cytoplasm to form UndP-β-glucose. This structure is then flipped by GtrA into the periplasm before the glucosyl residue is attached by the Gtr(type) to the growing O-antigen unit (Mavris et al., 1997; Allison & Verma, 2000). The gtr(type) is the third gene in the three-gene cluster involved in serotype conversion. Multiple alignments of these glycosyltransferases have failed to indicate any motifs that may be important in conserved function. The sizes of these proteins range from 416 to 506 aa. The secondary structure of the different Gtrs(type) is similar to flipases, O-antigen polymerases and O-antigen ligases (Schnaitman & Klena, 1993; Whitfield, 1995). Each of these proteins is functionally diverse and cannot be linked directly to the function of Gtrs. However, it has been postulated that GtrS might also play a role in recycling the lipid carrier back in the cytoplasm after transfer of the glucosyl group to the O-antigen unit as with Rfc (Wzy), an O-antigen polymerase of *S. flexneri* (Daniels et al., 1998; Guan et al., 1999; Allison & Verma, 2000). The diversity of the different Gtrs(type) may have evolved so that each Gtr protein can recognize a different target to which a glucosyl residue is attached via a specific linkage. Although they all seem to be diverse in function and structure, they all use a common substrate, UndP-β-glucose, indicating that at least one domain which interacts with the substrate is common.

The mechanism of glycosyltransferases-mediated antigenic conversion has been well studied in *S. flexneri*. However, similar processes have been observed in other bacteria. In particular, similar gtrA-gtrB-gtr(type) glycosylation cassettes have been observed in numerous strains of *Salmonella* spp. These genes may have a phage origin (Vander Byl & Kropinski, 2000). This was demonstrated for *S. enterica* serovar Typhimurium P22 (Iseki & Hamano, 1959; Rundell & Shuster, 1975; Weintraub et al., 1992; Steinbacher et al., 1996; Vander Byl & Kropinski, 2000) and ST64T (Mmolawa et al., 2003) phages. A similar mechanism of glycosyltransferases-mediated antigenic conversion has also been reported in *Salmonella* phages φ15 and φ34 (Bray & Robbins, 1967; Losick, 1969; Wright, 1971; Sasaki et al., 1974). These phages probably encode a similar three-gene glycosylation cassette. A glycosylation cassette that generates the serotype 4a in *S. flexneri* was described in *E. coli K-12* (Morona et al., 1995; Blattner et al., 1997; Mavris et al., 1997; Adams et al., 2001).

The biological functions of serotype-converting glycosyltransferases encoded by temperate bacteriophages are diverse. Because they modify the O-antigen polysaccharide chains during lysogeny, and because these structures constitute the receptor for the phage adsorption on the cell surface; this mechanism provides to the infected bacteria an immunity against superinfection by other related phages (Weintraub et al., 1992; Huan et al., 1997a, b; Mavris et al., 1997; Guan et al., 1999; Vander Byl & Kropinski, 2000; Allison et al., 2002). Similarly, during lytic infection, the receptor conversion could prevent retention of progeny phages on the host-cell debris. As protective host immune response to *S. flexneri* infection is directed against the O-antigen, serotype conversion represents an important virulence factor for the bacteria. Indeed, creating antigenic variation enhances bacteria survival because the host has
to mount a specific immune response to each serotype. Presumably, horizontal transfer of glycosylation cassettes within bacterial populations has led to new serotypes and development of some pandemic or maintenance of some endemic bacterial diseases.

**Glycosyltransferases encoded by phycodnaviruses**

Phycodnaviruses and putative phycodnaviruses are large (150 to 190 nm in diameter) polyhedral, dsDNA-containing viruses (genomes up to 560 kb) that infect eukaryotic algae (Van Etten, 2000, 2003; Van Etten & Meints, 1999; Van Etten et al., 2002). They exist throughout the world and play a dynamic, albeit largely unknown, role in regulating phytoplankton communities in aqueous environments, such as the termination of massive algal blooms commonly referred to as red and brown tides (Fuhrman, 1999; Wilhelm & Suttle, 1999; Suttle, 2000; Wommack & Colwell, 2000). Some phycodnaviruses encode more than 600 proteins, meaning they have more genes than the smallest free-living organism.

The phycodnaviruses probably have a common ancestor with the poxviruses, iridoviruses, and African swine fever virus (Iyer et al., 2001), and accumulating evidence indicates that the phycodnaviruses and their genes have a long evolutionary history, possibly from the time eukaryotes arose from prokaryotes (more than one billion years ago) (Villarreal & DeFilippis, 2000). Some algal virus genes encode commercially important enzymes such as DNA restriction endonucleases, whereas other viral genes encode enzymes that are the smallest in their class and may represent the minimal catalytic unit. Consequently, these ‘small’ proteins often serve as models for mechanistic and structural studies.

Members of the genus Chlorovirus are plaque-forming viruses that infect certain isolates of unicellular, chlorellalike green algae. The prototype chlorella virus, *Paramecium bursaria* chlorella virus (PBCV-1), infects *Chlorella NC64A* that is normally a symbiont of the protozoan *P. bursaria*. The 330744 bp PBCV-1 genome encodes 697 potential open reading frames (ORFs) of 65 codons or larger (Van Etten, 2003). About 375 of these ORFs probably code for proteins and ~50 % of these 375 gene products have been tentatively identified. Pertinent to this review is that some of the PBCV-1 genes encode enzymes that manipulate sugars. At least seven of the enzymes are putative glycosyltransferases involved in glycosylation of the PBCV-1 major capsid protein Vp54. Three enzymes are involved in the synthesis of the extracellular matrix polysaccharide hyaluronan and two more are involved in nucleotide sugar metabolism. At least five PBCV-1-encoded enzymes are involved in polysaccharide degradation and one enzyme is a glycosylase that initiates pyrimidine photodimer excision (Van Etten, 2003). These last six enzymes will not be discussed in this review.

**Protein glycosylation.** Typically, viral proteins are glycosylated by host-encoded glycosyltransferases located in the endoplasmic reticulum (ER) and Golgi (Doms et al., 1993; Knipe, 1996; Olofsson & Hansen, 1998). Consequently, the glycan portion of virus glycoproteins is host-specific. However, glycosylation of PBCV-1 major capsid protein Vp54 differs from this paradigm. This conclusion arose originally from antibody studies. Polyclonal antiserum prepared against intact PBCV-1 virions inhibits virus plaque formation by agglutination of particles. Spontaneously derived, antiserum-resistant variants of PBCV-1 arise at a frequency of ~10^{-6}. These antiserum-resistant variants fall into four serologically distinct classes (Wang et al., 1993). Polyclonal antiseras prepared against members of each of these antigenic classes react exclusively with the Vp54 equivalent from the viruses in the class used for the immunization. Western blot analyses of Vp54 proteins isolated from variants, before and after removing the glycans, established that the antigenic variants reflect differences in the Vp54 glycans. In addition, the ratio of the seven neutral sugars [glucose, fucose, galactose, mannose, xylose, rhamnose and arabinose (Wang et al., 1993)] associated with PBCV-1 Vp54 and the variants change in a manner that correlates with antigenicity and Vp54 migration on SDS-PAGE. Variants from different classes can complement and recombine in dual infection experiments to produce wild-type progeny, indicating that the enzymes involved in glycosylation reside in different virus-encoded complementation groups (Graves et al., 2001).

Additional observations indicate that Vp54 glycosylation is unusual. Unlike most glycoproteins that exhibit size microheterogeneity, Vp54 appears homogeneous both by SDS-PAGE and mass spectrometry analysis. The ability to easily crystallize Vp54 as a homotrimer provides additional evidence that the protein is homogeneous (Nandhagopal et al., 2002). Other observations indicating that Vp54 glycosylation is unusual are described in a recent review (Van Etten, 2003).

The identification of four glycan-linked Asn residues in Vp54 provided additional evidence that Vp54 glycosylation probably does not involve host glycosyltransferases. None of these glycan-linked Asn residues reside in a Nx(T/S) sequence commonly recognized by the ER- and Golgi-located glycosyltransferases (Reuter & Gabius, 1999). Taken together, the results suggest that PBCV-1 encodes the enzymes involved in constructing the glycans attached to Vp54.

Comparison of PBCV-1 ORFs to proteins in the databases identified seven possible glycosyltransferase encoding genes, a64r, a111r, a114r, a222-226r, a328l, a473l and a546l. None of these putative PBCV-1-encoded glycosyltransferases have an identifiable signal peptide that would target them to the ER. Furthermore, cellular protein localization programs predict that all of these proteins are cytoplasmic, with the exception of a473l expression product predicted to be in a membrane.

The a64r gene encodes a 638 aa protein that has four motifs conserved in ‘Fringe type’ glycosyltransferases. Drosophila
and mammalian Fringe homologues encode glycosyltransferases that modulate the function of Notch family receptors (Bruckner et al., 2000). Analysis of 13 PBCV-1 antigenic variants revealed mutations in a64r that correlated with a specific antigenic variation. Dual infection experiments with different antigenic variants indicated viruses that contained wild-type a64r could complement and recombine with viruses that contained variant a64r to form wild-type virus. Therefore, it was concluded that a64r encodes a glycosyltransferase contributing to Vp54 glycans (Graves et al., 2001).

Collectively, the results indicate that glycosylation of the PBCV-1 major capsid protein differs from that of other viruses and that glycosylation is independent of the ER and Golgi apparatus. Could Vp54 glycosylation reflect an ancestral pathway that existed prior to ER and Golgi formation? Among many questions to be addressed, it will be important to determine if the Vp54 glycan precursors are attached to a lipid carrier such as undecaprenyl-phosphate, which is the carrier for bacterial peptidoglycans and cell surface polysaccharides (Raetz & Whitfield, 2002) or dolichol diphosphate, which is the carrier in eukaryotic cells (Reuter & Gabius, 1999).

PBCV-1-encoded hyaluronan synthesis. Surprisingly, three PBCV-1-encoded enzymes, glutamine:fructose-6-phosphate amidotransferase (GFAT), UDP-glucose dehydrogenase (UDP-GlcDH) and hyaluronan synthase (HAS) (the gene encoding HAS will be called has hereafter), are involved in the synthesis of hyaluronan, a linear polysaccharide composed of ~20000 alternating β1,4-glucuronic acid and β1,3-N-acetylglucosamine residues (DeAngelis, 1999). All three genes are transcribed early in PBCV-1 infection and a dense fibrous hyaluronan network accumulates on the extracellular surface of the infected chlorella cells (DeAngelis et al., 1997; Landstein et al., 1998; Graves et al., 2001) (Fig. 1). Hyaluronan was unexpected because, hitherto, it had only been found in the extracellular matrix of vertebrates and the extracellular capsules of a few pathogenic bacteria (DeAngelis, 1999, 2002). The biosynthesis of a single hyaluronan chain requires considerable energy; five ATP equivalents, two NAD cofactors, one acetylCoA group and the two monosaccharide components are required to form each individual disaccharide unit (DeAngelis, 1999). Therefore, hyaluronan presumably serves some essential purpose or provides a selective advantage to the virus.

Three biological functions have been considered for the extracellular hyaluronan. First, the polysaccharide prevents uptake of virus-infected chlorella by the paramecium. Presumably, such infected algae would lyse inside the paramecium and the released virions would be digested by the protozoan. This scenario would be detrimental to virus survival. Second, the viruses have another host that acquires the virus by taking up the hyaluronan-covered algae. Third, virus infected cells aggregate, presumably because of the extracellular polysaccharide. This aggregation, which can trap uninfected cells, may aid the virus to spread to its next host.

![Mock- infected](https://example.com/image1.png)

![240 min p.i.](https://example.com/image2.png)

![240 min p.i. + HA lyase](https://example.com/image3.png)

**Fig. 1.** Ultrastructural changes in the algal cell wall (CW) induced by phycodnavirus PBCV-1 infection. Chlorella NC64A cells were observed by quick-freeze deep-etch electron microscopy. Upper, middle and lower panels show a mock-infected cell, an infected cell 240 min post-inoculation (p.i.) and an infected cell 240 min p.i. after treatment with hyaluronan-lyase, respectively. Cyto, Cytoplasm.

A complicating factor in understanding the biological importance of hyaluronan is the discovery that not all chlorella viruses encode a has gene and at least some virus-infected cells lack the surface polysaccharide (Graves et al., 1999). The story became more complicated with the recent discovery that some chlorella viruses encode a chitin synthase gene (chs) instead of a has gene (Kawasaki et al., 2002). Chitin is a linear polysaccharide composed of β1,4-linked N-acetylglucosamine residues. Cells infected with viruses with a chs gene accumulate chitin on their surface. A few viruses contain both has and chs genes and both polysaccharides appear on the surface of cells infected with these viruses (Kawasaki et al., 2002).

Fucose synthesis. PBCV-1 is the first virus known to encode enzymes involved in nucleotide sugar metabolism. The virus encodes two enzymes, GDP-D-mannose 4,6 dehydratase (GMD) and the bifunctional GDP-4-keto-6-deoxy-D-mannose epimerase/reductase (GMER), which comprise the highly conserved pathway that converts GDP-D-mannose to GDP-L-fucose. Both virus recombinant proteins have properties comparable to human and bacterial enzymes except that, in addition to the dehydratase activity, the PBCV-1 GMD also catalyses NADPH-dependent reduction of the intermediate GDP-4-keto-6-deoxy-D-mannose,
forming GDP-D-rhamnose (Tonetti et al., 2003). As a consequence, in vitro reconstruction of the biosynthetic pathway using PBCV-1 GMD and GMER produces both GDP-L-fucose and GDP-D-rhamnose in the presence of NADPH.

It is interesting that both fucose and rhamnose are present in the glycans attached to the PBCV-1 major capsid protein Vp54. However, the relevance of this finding is tempered by the fact that cell walls of the uninfected host contain rhamnose and trace amounts of fucose (Meints et al., 1988). Thus, the uninfected chlorella synthesizes these two sugars during normal growth. Presumably the virus-encoded pathway could circumvent a limited supply of these GDP-sugars by the host.

Additional phycodnavirus-encoded putative glycosyltransferases. Glycosyltransferase-encoding genes are apparently common in the phycodnavirus family. The sequencing of three other phycodnaviruses genomes are either finished or nearly so. The Ectocarpus siliculosus virus (EsV) encodes a putative chitin synthetase (Delaroque et al., 2001), Heterosigma akashiwo virus (HaV) encodes at least three putative glycosyltransferases (K. Nagasaki, personal communication) and Emiliania huxleyi virus (EhV) encodes at least one putative glycosyltransferase (W. Wilson, personal communication).

Ecdysteroid glucosyltransferase encoded by baculoviruses

Baculoviruses are members of the Baculoviridae (Blissard et al., 2000). This family comprises large circular dsDNA viruses restricted to arthropods (Granados & Federici, 1986; O’Reilly et al., 1992a) and that package their virions in large, proteinaceous capsules called occlusion bodies (OBs). The baculovirus infection cycle (Federici, 1997) begins when a larva of a susceptible host species consumes food contaminated with OBs. Virus particles are released into the gut and replicate in the cells of the midgut epithelium. Generally, a wide range of tissues become infected so that, by the end of the infection process, the larva is completely lysed by the viral infection. After death, the larval cadaver disintegrates and the progeny OBs are released into the environment to spread the virus to other hosts.

Most baculoviruses have been isolated from Lepidoptera. These insects undergo a very specific biological cycle characterized by transformations of the caterpillar occurring during their growth by means of successive moults followed by a final pupation (Fig. 2a) (Chapman, 1998). Ecdysteroids are important caterpillar hormones that initiate moulting both from one larval instar to the next and from the last larval instar to the pupa. These hormones are secreted into the haemolymph by the prothoracic gland.

One of the remarkable characteristics of many baculoviruses is their ability to suppress the development of the infected host, preventing it from moulting or pupating (Subrahmanyam & Ramakrishnan, 1981; Dougherty et al., 1987; O’Reilly & Miller, 1989, 1990; Burand & Park, 1992; O’Reilly, 1995). A study of Autographa californica multiply embedded nucleopolyhedrovirus (AcMNPV) revealed that

![Fig. 2.](http://vir.sgmjournals.org)
this fascinating aspect of the baculovirus–host interaction is attributable to expression of a viral ecdysteroid UDP-glycosyltransferase (egt) gene (O’Reilly & Miller, 1989). During baculovirus infection, the viral egt expression product (EGT) is secreted from infected cells into the haemolymph where it catalyses the conjugation of glucose or galactose to circulating ecdysteroid hormone (Fig. 2b) (O’Reilly & Miller, 1990; O’Reilly et al., 1991, 1992b; Kelly et al., 1995; Toister-Achituv & Faktor, 1997; Evans & O’Reilly, 1998, 1999). This conjugation inactivates the biological functions of the ecdysteroids (O’Reilly & Miller, 1990; O’Reilly et al., 1991, 1992b; Kelly et al., 1995; Toister-Achituv & Faktor, 1997; Evans & O’Reilly, 1998, 1999). The developmental outcome of the infection depends on the balance between host synthesis of active ecdysteroids and their EGT-mediated inactivation (O’Reilly et al., 1998).

There are several benefits to baculoviruses that encode an EGT. Ecdysteroids induce physiological modifications that decrease insect permissivity to baculovirus infection (O’Reilly et al., 1995, 1998); by altering these hormones, baculoviruses prevent the developmental resistance to virus replication observed in caterpillars (Keeley & Vinson, 1975; Murray et al., 1991; Mikhaylov et al., 1992; Engelhard & Volkman, 1995; Washburn et al., 2001). Moults and pupation monopolize important insect biological resources and are associated with an arrest of food ingestion (O’Reilly & Miller, 1991; Flipsen et al., 1995; Shikata et al., 1998; Slavicek et al., 1999; Bianchi et al., 2000; Chen et al., 2000; Dai et al., 2000; Wilson et al., 2000; Sun et al., 2002). Ecdysteroid inactivation by EGT prevents these effects and preserves all the resources of the caterpillar for virus replication. In addition to its detrimental effect on the virus yield described above, pupation is also an obstacle to virus spread. By preventing this process, baculoviruses avoid progeny virions being encased in a rigid and hermetic pupa. Ecdysteroids also affect the caterpillar behaviour. For example, many lepidopteran larvae leave the host plant and burrow into the soil prior to pupation. This behaviour is not observed when caterpillars are infected by EGT expressing baculoviruses. This modification of behaviour thus increases the chance that infected caterpillars die on plant leaves where OBs will later be ingested by naive caterpillars (O’Reilly, 1995).

In conclusion, baculoviruses provide one of the most striking examples of how a single virus gene can induce systemic alterations in the host physiology.

**Glycosyltransferases encoded by poxviruses**

The *Poxviridae* consists of large DNA viruses sharing several unique properties (Moss, 2001). The family contains two subfamilies, the *Entomopoxvirinae* and the *Chordopoxvirinae*, that infect invertebrate insects and vertebrates, respectively.

Sequencing of two entomopoxvirus genomes revealed putative glycosyltransferase gene homologues, named ORF MSV206 and ORF AMV248 in the genome of *Mellanoputus sanguinipes* Entomopoxvirus (MsEPV) (a grasshopper-derived virus) (Afonso et al., 1999) and *Anasacta moorei* Entomopoxivirus (AmEPV) (a red hairy caterpillar-derived virus) (Bawden et al., 2000), respectively. ORF MSV206 and ORF AMV248 resemble bacterial glycosyltransferases involved in lipopolysaccharide capsule biosynthesis and virulence. It has been proposed that the MSV206 protein is responsible for the cell surface carbohydrate changes induced on grasshopper haemocytes by MsEPV infection. However, the biological functions and relevance of these ORFs remain to be determined.

Some chordopoxviruses encode another type of glycosyltransferase. The South American myxoma virus (MYXV) is the prototype of the *Leporipoxvirus* genus of the *Chordopoxvirinae*. MYXV naturally infects the Brazilian tapeti [forest rabbit (*Sylvilagus brasiliensis*)], causing the development of a small localized tumour that can persist for months. In contrast to the trivial symptoms induced in its natural host, infection of the European rabbit (*Oryctolagus cuniculus*) causes the often fatal disease myxomatosis.

In 1999, Jackson and colleagues demonstrated that MYXV encodes and expresses a functional α,2,3-sialyltransferase (Jackson et al., 1999). Inactivation of the MST3N gene encoding this viral glycosyltransferase revealed that it is not essential for replication *in vitro* or for induction of clinical myxomatosis in susceptible rabbits (Jackson et al., 1999). However, in the absence of sialyltransferase expression, disease symptoms are delayed, suggesting that MYXV α,2,3-sialyltransferase may act synergistically with other virulence factors.

Two hypotheses that are not mutually exclusive have been proposed for the role(s) of MST3N in MYXV infection. First, the glycosyltransferase could act by post-translationally modifying glycoproteins secreted by infected cells and/or glycoproteins expressed at the surface of the infected cells or the virions. Second, the glycosyltransferase could also act as a lectin. A discussion of these hypotheses follows.

**MYXV α,2,3-sialyltransferase contributes to post-translational modifications of the myxoma virus-encoded serpin (serine-protease inhibitior) SERP-1 (Nash et al., 2000), a secreted viral anti-inflammatory protein.** SERP-1 is the only virus-encoded anti-inflammatory protein. SERP-1 is the only virus-encoded serpin that is secreted as a soluble glycoprotein. SERP-1 within the infected tissues. Sialylation of SERP-1 could decrease the antigenicity of SERP-1 or it could increase the half-life of SERP-1 within the infected tissues. Sialylation of SERP-1 could also affect its ability to bind to the surface of specific
target cells. This latter property could increase its local concentration or sequester SERP-1 away from the plasma environment. These hypotheses could apply to any other secreted or cell-associated viral immunomodulating molecule.

Sialylation of cell surface glycoproteins is involved in diverse immunological processes and immune evasion strategies developed by different pathogens (Jackson et al., 1999). Terminal Sia of glycoconjugates exposed on the surface of MYXV-infected cells or virions could mask them from components of the host innate immune response or contribute to interactions with the surface of cells expressing some receptors for sialylated molecules. These receptors could be members of the sialoadhesin family (Crocker et al., 1996, 1997). Alternatively, in association with a GlcNAcX1,3/4-fucosyltransferase, the MYXV α-2,3-sialyltransferase could form the structures sialyl-Lewis a [Lea; Siaα2,3Galβ1,3(Fucα1,4)GlcNAc] and sialyl-Lewis x [Leα; Siaα2,3Galβ1,4(Fucα1,3)GlcNAc], which are ligands for E-, P- and L-selectins involved in regulating cell traffic in inflammation and immune responses (Bevilaqua & Nelson, 1993; Tedder et al., 1995; Varki, 1997). Interestingly, pMST3N transfers sialic acid to the fucosylated acceptors from Lewis a and Lewis x (Sujino et al., 2000), a unique activity that distinguishes pMST3N from all known mammalian sialyltransferases. This mechanism could, for example, assist in virus spread into the host.

In addition to its putative biological functions mediated by its enzymatic products, pMST3N could also act by itself, as a lectin at the surface of the virus and/or the infected cells. Some observations suggest that vaccinia virus (another chordopoxvirus) particles contain ligands for asialo-glycoconjugates (Vanderplasschen & Smith, 1997). As the trans Golgi network or endosomal membranes contribute to the outermost membrane of the extracellular enveloped virus (EEV) (Schmelz et al., 1994; Tooze et al., 1993), these ligands could be cellurally derived sialyltransferases that are usually resident in the TGN (Jackson et al., 1999). Expression of the virus-encoded α-2,3-sialyltransferase on the outer membrane of the MYXV EEV could contribute to its binding onto target cells.

In addition to MYXV, other leporipoxviruses encode an α-2,3-sialyltransferase, suggesting that this gene is common in this genus (Jackson et al., 1999; Willer et al., 1999). The presence of the gene has been demonstrated in Shope fibroma virus (SFV), hare and squirrel fibroma virus. Moreover, infection by SFV was shown to induce an α-2,3-sialyltransferase activity in infected cells (Jackson et al., 1999).

Core 2 β1,6-GalT activity generates core 2 (Galβ1→3GlcNAc β1→6GalNAc) structures, and the synthesis of the 1 structure [GlcNAcβ1→3(GlcNAc β1→6)Gal] among the β1,6GalTs described to date, C2GnT-M is the only enzyme able to perform all three enzymic activities described in this family, i.e. synthesizes core 2, core 4 and I structures. The core 2 structure is produced by core 2 β1,6GalT (C2GnT) activity using core 1 (Galβ1→3GlcNAc) as an acceptor substrate (Fukuda et al., 1986; Piller et al., 1988). In many cells, the latter is the major constituent of O-glycans (Fukuda, 1994). The biological importance of the core 2 structure is explained by the various ligand carbohydrates that are formed in core 2 branched oligosaccharides. For example, sialyl Leα and sulphated sialyl Leα present in core 2 branched oligosaccharides are the preferred ligands for P- and L-selectins (Lowe, 1994; Hemmerich et al., 1995; Wilkins et al., 1996; Bstrup et al., 1999; Hiraoka et al., 1999; Fukuda, 2002). The core 4 β1,6GalT (C4GnT) activity generates core 4 from core 3 (GlcNAcβ1→3GlcNAc). Core 4 is mainly expressed in mucin-producing tissues (Schacht & Brockhausen, 1992). In contrast to core 2 and core 4, which are O-glycan structures, the I structure is found in N-glycans, O-glycans and glycolipids. I β1,6GnT (IGNt) activity converts linear poly N-acetyllactosamine (Galβ1→4GlcNAcβ1→3)α to branched poly N-acetyllactosamine, Galβ1→4GlcNAcβ1→3(Galβ1→4GlcNAcβ1→6)Gal→R (Fukuda et al., 1979; Piller et al., 1984).

Because BoHV-4 is the only known virus to date that encodes a homologue of the cellular C2GnT-M, it provides a unique model to study the biological functions of a β1,6GnT in the context of viral infection. For this reason, considerable efforts were made to study the origin and the roles of this unique viral gene. The study of the origin of the Bo17 gene revealed that it was acquired from an ancestor of the African buffalo around 1.5 million years ago, implying that cattle subsequently acquired BoHV-4 by cross-species transmission (Markine-Goriaynoff et al.,

Viral glycosyltransferases

http://vir.sgmjournals.org

African buffalo (Syncerus caffer) is also a natural reservoir of BoHV-4 in Africa (Rossiter et al., 1989). Sequencing of the BoHV-4 genome revealed that, in comparison to the other members of the Rhadinovirus genus, BoHV-4 has a reduced set of ORFs homologous to cellular genes (Zimmermann et al., 2001). However, the Bo17 gene of BoHV-4 is the only known viral gene that encodes a functional homologue of the cellular core 2 β1,6-N-acetylgalactosaminyltransferase-mucin type (C2GnT-M) (Vanderplasschen et al., 2000; Choi et al., 2004), a member of the β1,6-N-acetylgalactosaminyltransferase (β1,6GnT) gene family.

β1,6GnTs are involved in the synthesis of (GlcNAcβ1→6) GalNac linkages and play crucial roles in glycan synthesis (Fukuda, 1994; Schacht, 1994). Expression of their products changes during development, immunodeficiency and oncogenesis. Three activities have been described in the β1,6GnT gene family based on the acceptor substrate: the synthesis of the O-glycan core 2 [Galβ1→3(GlcNAc β1→6)GalNAc] and core 4 [GlcNAcβ1→3(GlcNAc β1→6)GalNAc] structures, and the synthesis of the I structure [GlcNAcβ1→3(GlcNAc β1→6)Gal]. Among the β1,6GnTs described to date, C2GnT-M is the only enzyme able to perform all three enzymic activities described in this family, i.e. synthesizes core 2, core 4 and I structures.
2003). This observation makes Bo17 the most recent known gene to be acquired by herpesvirus. This phylogenetic study also revealed that, despite its recent acquisition, Bo17 has spread to fixation in the viral population and that since the transfer of the cellular gene to the viral genome the sequence has been subjected to a strong constraint against nonsynonymous substitutions. Taken together, these data suggest that Bo17 encodes a function that is important for BoHV-4.

Bo17 expression product (pBo17) is a functional homologue of cellular C2GnT-M. pBo17 expressed transiently in CHO cells has all three enzymic activities exhibited by cellular C2GnT-M, i.e. core 2, core 4 and I branching activities (Vanderplasschen et al., 2000). Moreover, BoHV-4 infection of a permissive C2GnT-M-negative cell line led to the production of core 2 branched oligosaccharides (Vanderplasschen et al., 2000). Both transcription and expression product activity of Bo17 are conserved in nine representative strains of BoHV-4 (Markine-Goriaynoff et al., 2004). Deletion of Bo17 revealed that this gene is not essential for virus replication in vitro despite contributing to post-translational modifications of structural proteins (Markine-Goriaynoff et al., 2004). Consistent with this function, Bo17 is expressed as an early gene (Markine-Goriaynoff et al., 2004).

Several hypotheses that are not mutually exclusive could be made concerning the role(s) of Bo17 in the biology of BoHV-4 infection. In vivo, the post-translational modifications of structural proteins could affect the tropism of the virion and/or its sensitivity to antibody and/or complement neutralization. Supporting the latter hypothesis, the resistance of Sindbis virus virions to complement is affected by the composition of the glycosaminoglycans, and notably the level of sialylation expressed on the cells in which the virus was grown (Hirsch et al., 1983; Meri & Pangburn, 1990).

In addition to its effect on structural proteins, pBo17 could also affect glycosylation of viral and/or cellular glycoproteins exposed on the plasma membrane of the infected cells, thus affecting the biological properties of the infected cell. These modifications could affect the sensitivity of the infected cells to antibody and/or complement lysis and/or to cell-mediated cytotoxicity. Consistent with this hypothesis, several studies indicate that increasing the level of C2GnT activity, and of the resulting core 2 branched oligosaccharides, significantly decreases interactions between the expressing cells and cells of the immune system (Fukuda & Tsuboi, 1999). Alternatively, it is possible that by expression of Bo17, BoHV-4 could direct the exposure of cell surface ligands for selectins, thus affecting the tropism of the infected cells to secondary sites of replication and/or re-excretion. Further experiments are required to identify BoHV-4 and cellular proteins subjected to pBo17 enzymic activities and to determine the effect of these post-translational modifications on the biological properties of the virion and/or the infected cell.

Conclusions

As outlined in this review, virus-encoded glycosyltransferases are either involved or predicted to be involved in a variety of virus–host interactions and are more common than generally believed. It is unlikely that the list of known viral glycosyltransferases described above is complete. Indeed, we predict that many other virus-encoded glycosyltransferases will be discovered as additional viruses are sequenced. Supporting this prediction, sequencing of the 1100 kb amoeba virus (also called mimivirus for mimicking microbe virus) genome indicates that the virus encodes approximately 1100 genes among which at least six are putative glycosyltransferases (I.-M. Claverie, personal communication). Finally, natural selection through evolution of viruses able to affect the glycome is additional evidence of the importance of glycomics in biological processes.

Acknowledgements

A.V., N. M.-G. and L. G. are Senior Research Associate, Postdoctoral Researcher and Research Fellow of the ‘Fonds National Belge de la Recherche Scientifique’ (FNRS), respectively. Authors are grateful to Dr. R. O’Reilly for critical reading of the section devoted to ecdysteroid transferase encoded by baculoviruses, and to J. Heuser for providing electron micrographs of PBCV-1 infected chlorella cells (Fig. 1).

References


and three AcMNPV deletion mutants against Spodoptera exigua larvae (Lepidoptera: noctuidae). J Invertebr Pathol 75, 28–35.


