Complicated N-linked glycans in simple organisms

Birgit Schiller, Alba Hykkolari, Shi Yan,
Katharina Paschinger and Iain B.H. Wilson*

Department für Chemie, Universität für Bodenkultur,
A-1190 Wien, Austria
*Corresponding author
e-mail: iain.wilson@boku.ac.at

Abstract

Although countless genomes have now been sequenced, the glycomes of the vast majority of eukaryotes still present a series of unmapped frontiers. However, strides are being made in a few groups of invertebrate and unicellular organisms as regards their N-glycans and N-glycosylation pathways. Thereby, the traditional classification of glycan structures inevitably approaches its boundaries. Indeed, the glycomes of these organisms are rich in surprises, including a multitude of modifications of the core regions of N-glycans and unusual antennae. From the actually rather limited glycomic information we have, it is nevertheless obvious that the biotechnological, developmental and immunological relevance of these modifications, especially in insect cell lines, model organisms and parasites means that deciphering unusual glycomes is of more than just academic interest.

Keywords: insects; molluscs; nematodes; N-linked oligosaccharides; protozoa; trematodes.

Introduction

The traditional classification of N-linked oligosaccharides into oligomannosidic, complex and hybrid, as, for instance, summarised by Kornfeld and Kornfeld (1985) in their classic review, is based on the glycan structures found in mammals and other vertebrates. The extensively studied complex N-glycans of vertebrates are exemplified by structures with N-acetylgalactosamine (GlcNAc) residues on both the α1,3- and α1,6-linked mannose (Man) residues of the common trimannosylchitobiosyl core region; there is a large range of subsequent antennal modifications, particularly with galactose (Gal) and sialic acid residues, on such oligosaccharides. The term ‘hybrid’ was defined for those N-glycans displaying features of both the complex and oligomannosidic types; in this case, there are only GlcNAc residues linked to the α1,3-linked mannose (sometimes also, if bisected, to the β1,4-linked mannose), but not to the α1,6-linked mannose. The most basic hybrid glycan, with the composition Man5GlcNAc1 (Man5Gn), is also a key intermediate on the route to complex N-glycans. It is clear, from studies on mutant mice with defects in N-acetylgalactosaminyltransferase I (GlcNAc-TI or GnTI), that the ability to produce hybrid and complex glycans is essential for mammalian development (Ioffe and Stanley, 1994; Metzler et al., 1994).

During the initial discovery process of what we now call ‘glycobiology’, there was an awareness that, in non-vertebrates, there exist glycans that cannot be assigned to the aforementioned classification. For instance, yeast (or perhaps more exactly Saccharomyces cerevisiae) produce extended structures consisting only of two core GlcNAc and polylan- nosidic extensions with, not just 9, but perhaps 100 mannose residues (Herscovics and Orlean, 1993); in addition, plants were known, as exemplified by the glycoprotein phytohaemagglutinin, to synthesise not just the ‘usual’ oligomannosidic oligosaccharides, but also ‘short’ structures containing β1,2-xylene (Xyl) and α1,3-fucose (Fuc) associated with themannosylchitobiosyl core region (Sturm et al., 1992). Around 1990, there was the first indication that insects also produced not only oligomannosidic N-glycans but also ‘paucimannosidic’ structures with up to two core fucose residues on the reducing-terminal (innermost) GlcNAc – something that was initially greeted with scepticism (Staudacher et al., 1992). The term ‘paucimannosidic glycan’, perhaps not yet familiar to the widest glycobiological audience, was introduced to describe those glycans, particularly found in plants and invertebrates with or without core modifications, but lacking antennal GlcNAc and possessing less than four mannose residues.

However, in recent years, it has become apparent that the N-glycans of lower organisms (especially invertebrates and protists) cannot be easily classified. The designations ‘complex’, ‘hybrid’, ‘oligomannosidic’ and even ‘paucimannosidic’ are no longer adequate to summarise these structures and so terminologies such as ‘complex core modifications’, ‘truncated complex’ or ‘pseudohybrid’ have been coined to supplement the traditional terms. Indeed, rather complicated glycans are found in nematode species, unusual glycans are present in amoebae, and partial mimics of mammalian glycans are expressed by parasites such as Schistosoma mansoni, Trichomonas vaginalis and Trypanosoma brucei. Here, therefore, we discuss neither the N-glycans of plants, yeasts and fungi nor the O-glycans and glycolipids of ‘lower’ animals, but summarise the knowledge about N-linked oligosaccharides of a range of protozoal and invertebrate species (see Figures 1 and 2) with a particular focus on parasitic and model organisms.
Figure 1 Structural elements in some N-glycans.

A selection of some epitopes of N-glycans are shown: LacNAc, LacdiNAc, sialyl LacNAc, chitobiose (strictly N,N'-diacetyltchitobiose, as in the core region of N-glycans), Lewis A (present in plants and humans), Lewis X (Le^a; the fucosylated form of 'LacNAc' present in, e.g., schistosomes and vertebrates), LDNF (fucosylated LacdiNAc; i.e., fucosylated GalNAcβ1,4GlcNAc), difucosylation of the reducing-terminal (i.e., proximal or innermost) GlcNAc of N-glycans in many invertebrates, the 'GalFuc' epitope, the position of the reducing-terminal GlcNAc, the modification of the distal (second) core GlcNAc rather than on the core mannose (Levy-Ontman et al., 2011).

Perhaps the most studies on N-glycans of a non-parasitic, non-yeast unicellular organism have been performed on Dictyostelium discoideum – which is indeed a part-time multicellular organism (also known as either a cellular slime mould or social amoeba) due to its ability to form aggregates upon starvation and produce fruiting bodies. Although the overall carbohydrate composition in D. discoideum is similar to that of animals except for the absence of sialic acid (West et al., 2005), the N-glycans of this species are a good example of ‘complicated’ and unusual elaborations of typical oligomannosidic structures. The major neutral N-glycan in the amoebae has both ‘intersecting’ and ‘bisecting’ GlcNAc residues (see Figure 1) and core α1,3-fucose (Schiller et al., 2009); furthermore, charged glycans carrying sulphate and methylphosphate residues were reported first in the early 1980s (Freeze et al., 1980, 1983a) and their presence has been verified by mass spectrometry (Gabel et al., 1984; Feasley et al., 2010). While the presence of core xylose on slime mould N-glycans is not substantiated by the latest data, core α1,6-fucosylation has been recently detected by mass spectrometry on a single glycoprotein (Nakagawa et al., 2011).

It has become clear that the genetic basis for the glycosylation pathways of Dictyostelium shows many parallels to animal and plant pathways; indeed, Dictyostelium N-glycans are assembled through the common eukaryotic pathway using the standard eukaryotic precursor molecule GlcMan3GlcNAc2 whose biosynthesis is catalysed by the action of the 14 various alg gene products (Ivatt et al., 1984; Samuelson et al., 2005). However, the processing of the N-glycans is not dependent on GlcNAc-Ti, which in multicellular organisms is prerequisite for modifications such as addition of core fucose or bisecting GlcNAc. In the genome of D. discoideum, putative glycosyltransferase and glycohydrolase genes could be identified and their number compared with homologous genes in its relative D. purpureum (West et al., 2005; Sucgang et al., 2010). The prediction includes α1,3/4-fucosyltransferases from the CAZy family GT10 (not <10 homologues), β-GlcNAc transferases and one gene encoding a putative GlcNAc-P transferase. No glycosyltransferase involved in N-glycan biosynthesis has been characterised to date in recombinant form; however, a number of relevant transferase activities in crude extracts has been detected, e.g., intersecting and bisecting GlcNAc transferases (Sharkey
Figure 2  Examples of N-glycan structures from a selection of non-vertebrate eukaryotes.

In comparison to plants and vertebrates, examples of N-glycans from Dictyostelium discoideum (slime mould), Trichomonas vaginalis (protozoal parasite; the ‘biosynthetic’ Man5 structure being found also in Entamoeba histolytica, with the trimannosylchitobiosyl region being boxed with a dashed line), Schistosoma mansoni (trematode parasite), Dugesia japonica (planaria), Echinococcus granulosus (cestode parasite), Caenorhabditis elegans (nematode; the ‘GalFuc’ epitope being also found in some molluscs), Drosophila melanogaster (fruit fly; difucosylation also being found on bee venom glycoproteins) and Helix pomatia (mollusc) are shown. Incomplete lines indicate further structural possibilities. CCM, core chitobiose modification; EtNP, ethanolamine phosphate; Me, methyl; PC, phosphorylcholine; PMe, methylphosphate; S, sulphate. Monosaccharides are depicted according to the nomenclature of the Consortium for Functional Glycomics (see Figure 1).

and Kornfeld, 1991a), core α1,3-fucosyltransferase (Schiller et al., 2009), the GlcNAc phosphotransferase (Couso et al., 1986), and the S-adenosylmethionine-dependent methyltransferase, which modifies the Man-6-phosphate residues (Freeze and Wolgast, 1986; Freeze et al., 1992). Mutants defective in putative GlcNAc transferases and in phosphorylation of the N-glycans (specifically in the GlcNAc-P transferase) were recently identified (Pang et al., 2007; Qian et al., 2010), whereas defects in two enzymes of early N-glycan processing (a mannosyltransferase and a glucosidase) have been defined in earlier work (Freeze et al., 1983b, 1989).

A fascinating feature of D. discoideum is the shift in the N-glycome observed during development (Ivatt et al., 1981; Ivatt et al., 1984; Sharkey and Kornfeld 1991b): whereas N-glycans released from vegetative cells were partly resistant to the endoglycosidase Endo H, during aggregation and culmination they were sensitive to this treatment and smaller in size. Furthermore, the degree of modification with sulphate and/or phosphate decreased dramatically during late tip formation. Recent mass spectrometric studies confirm these trends and show a shift from Man,GlcNAc,Fuc, as the major neutral N-glycan to Man,GlcNAc,Fuc, (Schiller et al., 2009). Glycomic differences between the pre-spore and pre-stalk cells (i.e., those cells destined to later form the spore and stalk of the fruiting bodies) have also been observed (Riley et al., 1993).
**N-glycans of parasitic unicellular organisms**

A primary finding regarding N-glycosylation in obligate protist parasites is the loss of genes involved in the formation of the N-glycan precursor. The dolichol-linked oligosaccharides of these organisms range in size from Man$_6$GlcNAc$_2$ in *Trypanosoma cruzi* to Man$_5$GlcNAc$_2$ in *Entamoeba* and *Trichomonas*. Most extreme are the examples of *Plasmodium* and *Giardia* with just GlcNAc$_3$ as precursor, or *Theileria*, which, apparently, even though eukaryotic, lacks an N-glycosylation capacity entirely. The 'defects' in precursor formation are due to a lack of a variable number of alg genes (Samuelson et al., 2005). Those parasites synthesising at least Man$_5$GlcNAc$_2$ tend to also possess an endoplasmic reticulum (ER) glycan-dependent quality control mechanism involving glucosylation of nascent glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase (Banerjee et al., 2007); thereby, the presence of glucose on the final N-glycan structures is due to post-transfer glucosylation and not to Dol-P-Glc-dependent modifications of the precursor.

Compatible with the secondary loss of the alg-3, alg-5, alg-6, alg-8, alg-9, alg-10 and alg-12 genes during evolution, the causative agent of amoebic dysentery, *Entamoeba histolytica*, has the unprocessed ‘biosynthetic’ form of Man$_5$GlcNAc$_2$, containing two α1,2-mannose residues, as the most abundant glycan detectable (see Figure 2). In addition, some processing occurs to yield, e.g., glycans such as Glc$_3$Gal$_3$Man$_6$GlcNAc$_2$ or Gal$_3$Man$_5$GlcNAc$_2$. Neither incorporation of deoxysugoses nor additional GlcNAc residues could be detected (Magnelli et al., 2008). Similarly in several strains of *T. vaginalis*, a widespread sexually transmitted parasite, the major glycan detectable is also Man$_5$GlcNAc$_2$ (see Figure 2); however, much additional variation was also observed. The C1 strain is capable of the attachment of one or two pentose residues to its N-glycans, one of which is most likely xylose attached to the core mannose as in plants; another pentosylation site is on the N-glycans of the second GlcNAc of the core, as also described (see above) in a microalga. Additionally, some strains also exhibited modifications by N-acetyllactosamine and/or phosphorylthanolamine moieties, whereas the hybrid-like structures in this organism are a ‘mirror-image’ of those in multicellular organisms and thus we have proposed the term ‘pseudoxyborbide’ (see Figure 2) for such GlcNAc-modified N-glycans (Paschinger et al., 2012a). Recently, we have studied the N-glycans of the opportunistic amoebaal parasite *Acanthamoeba*; although pentosylation as in *T. vaginalis* has been detected, hexosylation of core fucose is present and the biosynthesis is based on a typical Glc$_3$Man$_5$GlcNAc$_2$ precursor (Schiller et al., unpublished data).

*Toxoplasma gondii* is primarily a feline parasite, but it can be passed to humans; it is incapable of synthesising glycans with more than five mannose and three glucose residues. However, the complication is that it can scavenge glucosylated dolichol-linked N-glycans from the cells in which it resides. Indeed, the N-glycans detected are dependent on the cells in which the parasite is cultivated. *Toxoplasma gondii* grown in normal mammalian cells possess oligomannose glycans ranging in size from Man$_6$GlcNAc$_2$ to Man$_5$GlcNAc$_2$; however, when grown in a cell line deficient in Dol-P-Man synthase and thus lacking glycans with the final four mannose residues, the parasite exhibited N-glycans, in part glucosylated, no different from its own endogenously produced forms (Garenaux et al., 2008).

The largest precursor synthesised in the trypanosomatids is Dol-P-P-Man$_5$GlcNAc$_2$; however, generally, the range of glycans reported in the literature is limited. For instance, in *T. cruzi* (causing Chagas disease in South America), after transient glucosylation, Man$_6$GlcNAc$_2$ structures are present on proteins (Parodi et al., 1983); some also galactose and sialic acid residues have been found in some studies (Couto et al., 1990). *Crichidia fasciculata*, an insect parasite, synthesises unglucosylated Dol-P-P-Man$_5$GlcNAc$_2$ glycan precursors and Man$_6$GlcNAc$_2$ is also the most abundant detected glycan on proteins; a second glycan of the composition Hex$_2$GlcNAc$_2$ was shown to contain galactofuranose (Parodi et al., 1981; Mendelzon and Parodi, 1986). *Leishmania mexicana* transfers unglucosylated Man$_5$GlcNAc$_2$ to proteins (Parodi et al., 1984) and Man$_5$GlcNAc$_2$, as well as Glc$_3$Man$_5$GlcNAc$_2$, are present on the Gp63 protease of both *L. mexicana* and *L. major* (Olafson et al., 1990; Funk et al., 1997). In the non-human parasite *T. tarentolae*, which infects a gecko but is also a potential expression system for recombinant proteins, an unsialylated biantennary, β,1-4-galactosylated, core α,1-6-fucosylated glycan has been detected (Breitling et al., 2002).

The situation in *T. brucei*, which causes African sleeping sickness, is more complicated as there are two pools of precursor (Bangs et al., 1988); it seems that the bloodstream form of *T. brucei* can transfer both Man$_6$GlcNAc$_2$ and Man$_5$GlcNAc$_2$ to the variant surface glycoprotein (VSG) in a site-specific manner (Jones et al., 2005). Indeed, both the blood-borne and the procyclic form of the parasite express two paralogous oligosaccharyltransferases (ThSTT3A and ThSTT3B) with different specificity (Izquierdo et al., 2009). The resulting glycosylation patterns differ for VSG types I, II and III in a site- and protein-specific manner; the structures include typical oligomannose types, such as Man$_5$GlcNAc$_2$, paucimannosidic glycans with the compositions Man$_4$GlcNAc$_2$, and 'hybrid' and biantennary complex types, some of which are modified with terminal α1,3-linked galactose residues. Particularly striking are glycans with sometimes highly extended and branched poly N-acetyllactosamine chains (Zamze et al., 1990, 1991; Mehlert et al., 2002, 2010; Atrih et al., 2005). In an alg$_2$ null mutant strain lacking the sixth ER mannosyltransferase, mannosidase inhibition results in the presence of some pseudohybrid and glucosylated glycans (Manthri et al., 2008) akin to those found in *T. vaginalis*.

In contrast, confirming the existence of N-glycosylation in the malaria parasite *Plasmodium falciparum* has proven a difficult task (Davidson and Gowda, 2001). Indeed, it is known that *P. falciparum* is missing all of the ALG glycosyltransferases except ALG7 (UDP-N-acetylglucosamine-1-phosphotransferase), ALG13 (the second GlcNAc transferase) and the STT3 oligosaccharyltransferase catalytic subunit (Samuelson et al., 2005). Both *P. falciparum* and *Giardia lamblia* are capable of synthesising Dol-P-P-GlcNAc$_2$ and transferring this to proteins (Ratner et al., 2008; Bushkin et al., 2008).
et al., 2010). Earlier reports of larger structures in these species are probably to be explained by contamination with glycans derived from the host or the medium.

N-glycans of platyhelminths

Our knowledge of platyhelminth (flatworm) N-glycosylation is focussed primarily on the parasitic trematodes Schistosoma mansoni and S. japonica. During the life cycle of the parasite, some shifts in the N-glycosylation pattern occur and the glycosylation of some specific glycoproteins have also been investigated. It would appear that a real ‘mix’ of plant- and animal-type core modifications are present, in that xylosylation\(^1\) of the core β-mannose as well as α1,3- and α1,6-fucosylation of the reducing GlcNAc can occur in various combinations on glycoproteins derived from eggs, cercariae, miracidia, adults, or their secretions with core α1,3-fucose being apparently absent from adults and cercariae (Kho et al., 1997a, 2001; Wuhrer et al., 2006a,b; Hokke et al., 2007; Jang-Lee et al., 2007; Meevissen et al., 2010, 2011). Furthermore, up to three antennae have been identified on S. mansoni N-glycans; these antennae can consist of LacNAc (Galβ1,4GlcNAc) and LacdiNAc (GalNAcβ1,4GlcNAc) units (sometimes repeats), which may be decorated with fucose residues to result in, e.g., Le\(^a\), LDNF (see Figure 1) and difucosyl epitopes (Srivatsan et al., 1992; Kho et al., 2001; Wuhrer et al., 2006a; Jang-Lee et al., 2007; Meevissen et al., 2010, 2011). Although fucosyl- and xylosyltransferase activities have been found in schistosome extracts (DeBose-Boyd et al., 1996; Faveeuw et al., 2003; Paschinger et al., 2005), these have not yet been proven in three studies on S. mansoni N-glycans; of the reducing GlcNAc can occur in various combinations on glycoproteins derived from eggs, cercariae, miracidia, adults, or their secretions with core α1,3-fucose being apparently absent from adults and cercariae (Kho et al., 1997a, 2001; Wuhrer et al., 2006a,b; Hokke et al., 2007; Jang-Lee et al., 2007; Meevissen et al., 2010, 2011). Furthermore, up to three antennae have been identified on S. mansoni N-glycans; these antennae can consist of LacNAc (Galβ1,4GlcNAc) and LacdiNAc (GalNAcβ1,4GlcNAc) units (sometimes repeats), which may be decorated with fucose residues to result in, e.g., Le\(^a\), LDNF (see Figure 1) and difucosyl epitopes (Srivatsan et al., 1992; Kho et al., 2001; Wuhrer et al., 2006a; Jang-Lee et al., 2007; Meevissen et al., 2010, 2011). Although fucosyl- and xylosyltransferase activities have been found in schistosome extracts (DeBose-Boyd et al., 1996; Faveeuw et al., 2003; Paschinger et al., 2005), these have not yet been correlated with the relevant homologues in the schistosome genome; however, variations in their transcript levels have been found (Fitzpatrick et al., 2009).

Among the cestodes, the N-glycans may be less complex than those of the schistosomes; core α1,6-fucosylation has been proven in three studies on Echinococcus glycoproteins (Kho et al., 1997b; Hulsmeier et al., 2010; Paschinger et al., 2012b); possible antennal modifications include galactose or phosphorylcholine – the latter accounting for the immunogenicity of the protein antigen Ag5. In another tapeworm, Taenia crassiceps, core fucosylation and terminal galactose is also a feature, but antennal fucose was also found (Lee et al., 2005).

Not all flatworms are parasitic, and the planaria have gained a status as a model for pluripotency; due to the stem cells present throughout the animal, regeneration of any amputated tissue is possible. The N-glycans of one species, Dugesia japonica, have been studied by two groups – the major N-glycan is a ‘processed’ Man3GlcNAc, structure with all three non-reducing terminal mannoses being methylated (Natsuka et al., 2011; Paschinger et al., 2011); however, a small proportion of glycans is also core α1,6-fucosylated and the fucose residue is further modified by galactose and even a further methylhexose residue (Paschinger et al., 2011).

N-glycans of nematodes

As recently summarised (Paschinger et al., 2008), the simple roundworm Caenorhabditis elegans synthesises a wide range of N-glycans: with the oligomannosidic, paucimannosidic, fucosylated, ‘truncated complex’ and phosphorylcholine-modified types being supplemented by those with the so-called core chitobiose modifications. However, C. elegans is a non-parasitic model organism with a large number of parasitic relatives; unfortunately, information about the actual covalent structures of the N-glycans is available for only a limited number of species: Haemonchus contortus, Ostertagia ostertagi, Dityococcus viviparus, Parelaphostrongylus tenuis, Ascaris suum, Onchocerca volvulus, Acanthocheilonema vitaeae and Trichinella spiralis (Reason et al., 1994; Haslam et al., 1996, 1997, 1999, 2000; Duffy et al., 2006; Pöld et al., 2007; Meyvis et al., 2008).

There are two interesting features that C. elegans seems to share with its parasitic relatives, the presence of multiple fucoses (Haslam et al., 2002; Paschinger et al., 2004; Zhu et al., 2004) with at least three of them bound to the chitobiose core of the N-glycans (Hanneman et al., 2006; Struve and Reinhold, 2012) and the modifications of the N-glycan antennae with phosphorylcholine bound to the GlcNAc residues (Haslam et al., 2002; Paschinger et al., 2006) (see Figure 2). C. elegans shares the former feature with the sheep parasite H. contortus (Haslam et al., 1996); this includes the presence of core α1,3-fucose, which is an epitope for anti-horseradish peroxidase (anti-HRP) as well as for IgE from H. contortus-infected sheep (Van Die et al., 1999). On the other hand, N-glycans carrying the phosphorylcholine epitope have been found not only in filarial nematodes such as A. viteae and O. volvulus (Haslam et al., 1997, 1999) but also in parasites with larvae migrating through different tissues in animal hosts, such as in T. spiralis and A. suum (Morelle et al., 2000b; Pöld et al., 2007). This modification is of special interest as phosphorylcholine is associated with immunomodulation by nematode parasites (Harnett and Harnett, 2001).

A particularly unusual feature of a portion of C. elegans N-glycans is the capping of core fucose with galactose (Hanneman et al., 2006; Gutternigg et al., 2007a; Takeuchi et al., 2008); to date, there is no report in the literature that these ‘GalFuc’ epitopes are also present in nematode parasites (although our on-going work indeed indicates their occurrence in at least two parasitic species). The presence of galactose residues, on up to all three fucoses associated with the chitobiosyl core, confers a definite complexity to these glycans, which are recognised by both worm and fungal galectins (Takeuchi et al., 2008; Butschi et al., 2010); in terms of the evolutionary context, it is as if the galactose was perhaps first associated with the core and only later ‘migrated’ to the antennal positions familiar in vertebrate glycans.

---

\(^1\)Xylose (derived from the Greek ἔξυλος, wood) is a monosaccharide normally considered to be a major component of plant N-glycans and polysaccharides; however, in animals, this monosaccharide is present in proteoglycans (as the ‘core’ linkage to protein of chondroitin and heparan sulphates) and in some O-glucose-based glycans present on vertebrate EGF domains as well as in the N-glycans of some trematodes and molluscs.
Despite the various α1,3-fucosyltransferase homologues in *C. elegans* and their Lewis-type activity *in vitro* (Nguyen et al., 2007), there is no sign that these are generating Lewis-type antennal modifications *in vivo*; the only nematode thus far shown to actually possess Le^a^ is *D. viviparus* (Haslam et al., 2000). Indeed, the underlying (unfucosylated) LacNac motif is not a general feature of nematode N-glycans, and reports on the occurrence of related LacdiNac epitope in nematode N-glycans are scarce; however, this feature is found in *T. spiralis* (Morelle et al., 2000a,b) and *Dirofilaria immitis* (Kang et al., 1993). Chito-oligomers (GlcNAC_3_), which were only detected after hydrofluoric acid treatment, are a feature of filarial worms (Haslam et al., 1999). In all these species, multiantennary N-glycans are present; based on homologies and actual activity assays, *C. elegans* possesses GlcNAC-Tl, -TI and -TV genes required for the synthesis of triantennary N-glycans (Chen et al., 2002; Warren et al., 2002). Otherwise, the core fucosyltransferases FUT-1 and FUT-8 and the ‘capping’ galactosyltransferase GALT-1 have demonstrated N-glycan-modifying activity correlating with N-glycan structures (Paschinger et al., 2004, 2005; Titz et al., 2009). Other enzymes required for galactose, fucose or phosphorylcholine modifications of N-glycans in *C. elegans* are yet to be identified; certainly, as surprises as to the specificities of such enzymes may well occur – the GlcNAC-Tl independence of the core α1,3-fucosyltransferase FUT-1 (see Figure 3 for the biosynthetic scheme) was unexpected and a possible indication of convergent evolution regarding the formation of anti-HRP epitopes in different species (as mentioned above, core α1,3-fucose is an epitope for anti-HRP).

**N-glycans of molluscs**

The N-glycans of a few species of molluscs, including slugs, snails, limpet, octopus and squid, have been studied over the years. In some cases, mollusc extracts were examined, in others, specific proteins such as rhodopsins or haemocyanins. In some species, such as gastropods (Gutternigg et al., 2004, 2007b), core β1,2-xyllose and α1,6-fucose are present as well as methylation of terminal mannose residues and a low degree of core α1,3-fucosylation (see also the review by Staudacher in this issue) are a feature, whereas in others, e.g., squid and octopus rhodopsins (Zhang et al., 1997; Takahashi et al., 2004), the same ‘GalFuc’ motif (Galβ1,4Fuc on the reducing terminal GlcNac; see Figure 1) as in planaria and nematodes has been found; in the squid, as is sometimes the case in *C. elegans*, the GalFuc motif is in the context of difucosylation of the core GlcNac. In the snail *Biomphalaria glabrata*, the intermediate host of *S. mansoni*, >100 N-glycan structures have been isolated from haemolymph proteins, including biantennary glycans with core xylose, core α1,6-fucose, methylmannose and fucosylated LacdiNac (e.g., Fucβ1,3GalNacα1,4GlcNac) motifs; other than methylation, these features are shared with the tetraolde and cross-react with anti-schistosome antibodies (Lehr et al., 2007).

Among the various haemocyanins examined, the one from keyhole limpet (Megathura crenulata; KLH) also features glycan cross-reactive with anti-schistosome antibodies. Fucosylated LacdiNac, core xylose, galactosylated ‘GalFuc’ (i.e., Galβ1,4Galβ1,4Fucα1,6) and Galβ1,6Man motifs.
have been detected on KLH glycans by mass spectrometry (Kurokawa et al., 2002; Wuher et al., 2004; Geyer et al., 2005). In other species, the glycans present on the haemocyanins vary from the ‘less exciting’ hybrid and Man$._{5}$GlcNAc$_{2}$ structures in Panulirus interruptus (Van Kuik et al., 1986a) through to oligosaccharides with methylated mannosse in Hippopus hippopus (Puanglarp et al., 1995); methylated Lewis-like motifs in Haliotis tuberculata (Velkova et al., 2011); disubstituted antennal fucose in Rapana thomasiensis (Gielen et al., 2005); and sulphated mannosse, methylated GlcNAc and methylated galactose in Rapana venosa (Dolashka-Angelova et al., 2003). Methylated galactose and core xylose have been found on the haemocyanins from invertebrates (Pan Kuik et al., 1983, 1986b, 1987; Lommerse et al., 1997). Thereby, it is more than likely that many more specific peculiarities; it is more than likely that many more types of modification remain to be discovered.

**N-glycans of insects**

The N-glycosylation capacity of insects (Rendić et al., 2008) is of interest for both academic and biotechnological reasons with, on the one hand, the fruit fly Drosophila melanogaster as an important model organism and, on the other, the various insect cell lines used to produce recombinant proteins; also, the immunogenicity of insect venom glycoproteins is another factor. The first studies on the N-glycans of insects indicated the presence of oligomannosidic glycans and also of a core α1,6-fucosylated paucimannosidic structure (MMF; see Figure 3 for related glycans) (Butters and Hughes, 1981; Williams et al., 1991). However, it was also obvious that an until then unknown modification was also present in insects: difucosylation of the core reducing-terminal GlcNAc – i.e., its modification by both α1,3- and α1,6-fucose. Core difucosylation (see Figure 1) was first observed on bee venom phospholipase A$_{1}$ (Kubelka et al., 1993), but also, e.g., on bee and wasp venom hyaluronidase (Kubelka et al., 1995; Kolarich et al., 2005) and glycoproteins from D. melanogaster adults and neuronal cells as well as on the pheromone DUP99B (Fabini et al., 2001; Saudan et al., 2002; Rendić et al., 2006). Also, recombinant glycoproteins produced in Trichoplusia ni (High Five) cells (Ailor et al., 2000; Palmberger et al., 2011) can be core difucosylated. Thereby, core α1,3-fucose, and not xylose, is responsible for the cross-reactivity of insect glycoproteins towards antiresponse plant glycans, including anti-HRP.

In general, insect N-glycans are not normally possessing extended antennae. However, there are exceptions, such as the fucosylated LacdiNAc (LDNF; see Figure 1) found on a proportion of bee venom phospholipase (Kubelka et al., 1993), Galβ1,3GlcNAc modifications of royal jelly glycoproteins (Kimura et al., 2003) and sialyl-LacNAc (see Figure 1) on some Drosophila embryonal glycans (Aoki et al., 2007); among these examples are even triantennary forms. In a locust apolipophorin, the rather unusual modification by aminoethylphosphonate was also proposed (Hård et al., 1993). Thus, insects do possess the ability to initiate the processing of glycans in a ‘complex’ manner (Geisler and Jarvis, 2012), even though pauci- and oligomannosidic forms dominate the spectra of those insect samples analysed to date. It may seem a paradox that even the biosynthesis of paucimannosidic glycans requires the prior action of GlcNAc-TI. However, due to the hexosamidase activity encoded by the fused lobes gene in their secretory pathways (Léonard et al., 2006), removal of the GlcNAc first transferred by GlcNAc-TI is an integral part of N-glycan processing not just in insects but in many invertebrates (Figure 3); thereby, the action of the fused lobes enzyme (named on the basis of the morphology of the brain in the corresponding Drosophila mutant) results in a lack of antennal elongation. Nevertheless, GlcNAc-TI activity generates the necessary ‘GO’ signal for core fucosylation and Golgi mannosidase II (Schachter, 2009). However, a major interest in the exploitation of insect cells as expression systems is indeed to circumvent the removal of this GO signal by the fused lobes hexosamidase (Fdi) by, e.g., overexpressing mammalian glycosyltransferases, which cap GlcNAc residues (Aumiller et al., 2012).

**N-glycans of ascidians**

Ascidians or sea squirts are chordates and thus are considered to be evolutionarily close to vertebrates. Despite the potential phylogenetic interest, only one study regarding their N-glycans has been published. Specifically, a glycan in the neural tissue of Ciona intestinalis has been described as co-eluting with plant glycans containing xylose and core α1,3-fucose; otherwise, oligomannosidic and fucosylated triantennary glycans were detected in other tissues of this organism (Yagi et al., 2008).

**Conclusion**

The large diversity in glycan structures and the incredible glycogenic potential of the so-called lower organisms, whether unicellular or multicellular, are obvious. However, although the N-glycans of a wide range of ‘simple’ organisms have been studied over the years, this work has not been tackled systematically. This means there are many holes in our knowledge. Nevertheless, there are some trends to consider: one is the frequent lack of charged modifications of their N-glycans (in particular, sialic acid); however, there are exceptions (e.g., sulphation in slime moulds or phosphorylcholine in nematodes) and methodological constraints may lead to an underestimation of their occurrence. Another is the presence of unusual modifications of the core region; however, a more general point to consider is the modification of N-linked oligosaccharides by the classic GlcNAc-TI. This is quite probably a hallmark of multicellular organisms – it is not quite clear how this enzyme evolved, but it is probable that even in unicellular organisms (such as trypanosomatids),
which also synthesise biantennary glycans, the transfer of GlcNAc to the N-glycan in the Golgi does not take place through the same mechanism as in vertebrates, and it appears that homologues of the ‘multicellular’ GlcNAc-TI are absent from these species. Indeed, it may well be that in unicellular parasites first the α1,6-mannose is modified before the α1,3-arm; at least, in T. vaginalis, there is probably only transfer of GlcNAc to the α1,6-arm to form pseudohybrid glycans.

This is just one example where it becomes obvious that the classic division of N-glycan types does not hold up when considering non-vertebrate species. It is even difficult to consider what the term ‘complex’ means when presented with some of the glycan structures – such as those carrying the ‘core chitobiose modifications’ in nematodes. Traditionally, ‘complex’ glycosylation refers to N-glycans with GlcNAc residues modifying both the α1,3- and α1,6-linked mannose residues of the conserved pentasaccharide core. However, a glycan lacking such residues but possessing three capped fucose substitutions of the chitobiosyl region is also structurally complex (never mind its biosynthesis) – or should we just say it is ‘complicated’? It is also obvious that previous attempts to ‘name’ glycans on the basis of their terminal sugars (e.g., the Schacht nomenclature featuring names such as ‘GnGnF’ or ‘GalGal’; Schacht, 1986) also cannot deal with the structures observed in lower organisms. We end up with referring to glycans by their mass or their composition or by referring to diagrams featuring squares, triangles and circles whose meanings are not even accepted by all, never mind understood by non-specialists; we also do not have so many simple and/or abbreviated names for the epitopes in lower organisms unless they are, such as Lewis-type glycans, shared with those in mammals. Currently, it seems that the flood of glycomic information from non-vertebrates has exhausted the normal human desire to name objects; it should certainly not mean that mammalian-centric researchers should ignore the nature of these unusual glycans or that we oversimplify or overgeneralise because we lack ‘nuttshell’ summaries.

Both Rudolf and Hildegard Geyer have made substantial contributions to our knowledge about the glycosylation of non-vertebrate species: not just about their N-glycans (especially, as cited above, a number of studies on glycans cross-reacting with anti-schistosome antibodies) but also in the realm of glycolipids, a topic that is not addressed here. Their unique knowledge in glycan analysis has aided many glyco-biologists, including ourselves; there would be indeed be a continued need for this internationally respected expertise (apparently and unfortunately not a future focus in Giessen) as even two scientific lifetimes are insufficient for an exploration of the glyco-universe. What is certain is that their work has partially paved the way for others to explore further galaxies of glycymes.

Acknowledgements

Work in our laboratory related to the topic of this review has been and is primarily funded by the Austrian Fonds zur Förderung der wissenschaftlichen Forschung [grants to I.B.H.W. (P13810, P15475, P17681, P18447, P19615, P20565 and P23922) and K.P. (P21946)].

References


Crithidia harmosa contain galactofuranose residues. J. Biol. Chem. 261, 2129–2133.


Received March 12, 2012; accepted April 7, 2012.
Birgit Schiller was born in Vienna (Austria) where she studied food-science and biotechnology at the Universität für Bodenkultur Wien (BOKU). She soon focussed on the field of biochemistry, acquiring expertise in the area of glycoanalysis during her Master thesis and PhD project both in the laboratory of Professor Iain Wilson. Birgit Schiller recently completed her PhD during which she ventured into the unknown world of Acanthamoeba glycosylation. Thus she was able to describe for the first time the composition of the glycosidic precursor molecule and the processed N-glycans from this opportunistic pathogen.

Alba Hykollari studied Biotechnology at the Universität für Bodenkultur (BOKU) in Vienna. She did her master thesis on structural characterization of native and recombinant S-layer proteins and secondary cell wall polymers from selected Bacillus strains at the Department of Nanobiotechnology and graduated in 2007. For her PhD thesis, she joined the ‘Molecular Glycobiology’ group at the University (Department of Chemistry) working on the N-glycan analysis and glycogenomics of the social amoeba Dictyostelium discoideum.

Shi Yan studied biochemistry and molecular biology at Soochow University, China and obtained his Master of Science degree in 2008, working on the biological function of a glycosyltransferase in human gastric carcinoma. In 2009, he joined the Department of Chemistry at the BOKU, Vienna, Austria and started his PhD on glycosylation of Caenorhabditis elegans, mainly working on comparative N-glycome analyses of worm mutants and parasitic nematodes. He has also contributed to the EuroGlycoArrays network on producing variety of glycosyltransferases for chemo-enzymatic synthesis of glycan compounds.

Katharina Paschinger gained her degree (Diplom-Ingenieur) in 1998 from the Universität für Bodenkultur Wien after performing work on the peptide:N-glycosidase from almonds as her ‘Diplomarbeit’. After a period in industry, she returned to academic science and gained her doctorate on the Glycosylation of Caenorhabditis elegans in 2008 under the supervision of Iain Wilson. She has been postdoc in the same laboratory and is currently leader of her own project on comparative nematode glycomics. She is not only co-author of over twenty-five papers, but is the mother of two daughters.

Iain B.H. Wilson studied biochemistry at the University of Edinburgh and obtained his honours degree in 1988 before going ‘up’ to Oxford for his DPhil at the Glycobiology Institute under the supervision of Prof. Raymond Dwek and Dr. Thomas Rademacher. After postdoc stays in Oxford, Vienna and Dundee, he returned to Vienna (1999) first as a university assistant prior to ‘habilitating’ in 2001. He leads the ‘Molecular Glycobiology’ research group within the Department of Chemistry at the BOKU and has been awarded the Neose Technologies Glycoscience Research Award in 1999, the Novartis (Austria) Prize for Chemistry in 2007 and the GlycoThera Award in 2011.