Mass Spectrometry in the Elucidation of the Glycoproteome of Bacterial Pathogens

Robert L.J. Graham and Sonja Hess*

Proteome Exploration Laboratory, Beckman Institute, California Institute of Technology, Pasadena, California, USA

Abstract: Presently some three hundred post-translational modifications are known to occur in bacteria *in vivo*. Many of these modifications play critical roles in the regulation of proteins and control key biological processes. One of the most predominant modifications, N- and O-glycosylations are now known to be present in bacteria (and archaea) although they were long believed to be limited to eukaryotes. In a number of human pathogens these glycans have been found attached to the surfaces of pilin, flagellin and other surface and secreted proteins where it has been demonstrated that they play a role in the virulence of these bacteria. Mass spectrometry characterization of these glycosylation events has been the enabling key technology for these findings. This review will look at the use of mass spectrometry as a key technology for the detection and mapping of these modifications within microorganisms, with particular reference to the human pathogens, *Campylobacter jejuni* and *Mycobacterium tuberculosis*. The overall aim of this review will be to give a basic understanding of the current 'state-of-the-art' of the key techniques, principles and technologies, including bioinformatics tools, involved in the analysis of the glycosylation modifications.

Keywords: Proteomics, mass spectrometry, post-translational modifications, glycosylation, eukaryota, prokaryota *Campylobacter jejuni, Mycobacterium tuberculosis*.

INTRODUCTION

The oligomerization of different classes of biomolecules lays the foundation for the construction of a coding system for information storage [1]. This is best exemplified by the genome, that with the exception of epigenetic changes may be considered static, and the proteome, which is considered much more complex and dynamic [2].

It is only in the last few decades that the functional importance of the carbohydrates as a third class of biomolecules involved in information storage has been recognized [3]. The oligomerization of carbohydrates can give rise to an almost boundless array of structures. A hexasaccharide comprising only four different hexose sugars is capable of giving rise to 2.7 billion different and distinct structures [4], whilst the number of possible isomers of a hexasaccharide comprising a set of six hexose's is greater than 1.05×10^{12} [5]. In contrast a sequence of six amino acids only generates ~ 50,000 different structures. Thus carbohydrates have the potential to carry information content several orders of magnitude higher in a short sequence than any other biomolecule. This vast array of structures in carbohydrates is accounted for by several factors: carbohydrates can exist in a number of epimeric forms with variations occurring in anomeric configuration, ring size, core and branching sequences and the linking, branching and reducing terminal attachments can also differ [5].

*Address correspondence to this author at the California Institute of Technology, Beckman Institute, Proteome Exploration Laboratory, 1200 E California Blvd, MC139-74, Pasadena, CA 91125, USA;
Tel: (001) 626-395-2339; Fax: (001)626-449-4159;
E-mail: shess@caltech.edu

In biological systems these sugar moieties are most commonly covalently attached to the polypeptide backbone of proteins and are amongst the most abundant co- and post-translational modifications observed in nature, with as much as half of all proteins being predicted to be glycosylated [1, 6-8].

Glycosylation events give rise to the production of a large number of glycoproteins. These differ by the amino acid to which the sugar is connected as well as the composition of the attached carbohydrate units. Diversity is also increased when taking into account the position and linkages between the sugars and also the monomeric sequences and their anomeric forms. Separate populations of glycoproteins also exist that have identical amino acid sequences but differ only in the compositions of glycans attached to the sites of glycosylation, these different 'glycoforms' of the same protein produce a high degree of micro-heterogeneity [1, 2, 7].

Due to their prominent positions on the cell surface, glycoproteins are often the first cellular components encountered by pathogens. After cell-pathogen contact, glycoproteins are ideally suited to modulate a myriad of signaling processes [9]. Almost all of the key components of the immune system are glycoproteins. In fact, all of the immunoglobulins are glycosylated; here the sugar moieties are responsible for imparting stability and are also involved in the production of specific recognition epitopes that facilitate binding to receptors for the initiation of signaling events [10].

In a number of human pathogens glycans have been found attached to the surfaces of pilin, flagellin and other surface and secreted proteins where it has been demonstrated that they play a role in adhesion, colonization, pathogenicity, immune evasion and virulence of these bacteria [11-14].

1570-1646/10 $55.00+.00

Given the ubiquitous distribution of these glycoproteins the study of the 'glycome', i.e.; all the sets of glycans/glycoproteins expressed within a biological system, must be an intrinsic part of any research into the understanding of biological processes [15].

THE ROLE OF MASS SPECTROMETRY (MS) IN GLYCOPROTEOMICS

A dizzying array of glycan structures can be expressed in an organism at any specific time. Given the physiological importance of this co/post-translational process numerous research groups have dedicated significant time and effort to the analysis and characterization of specific glycan structures and the proteins to which these glycans are attached. This work has given rise to the fields of 'glycomics', which attempts to study the significance and characterize the total N- and O-glycosylation expression within a biological system, and 'glycoproteomics', which is concerned with glycopeptide analysis [16].

Since the first successful structural analysis of a complex naturally occurring glycopeptide, leading to the elucidation of the full primary structure of the antifreeze glycoprotein from the Antarctic Fish *Trematomus borchgrevinki*, [17] mass spectrometry has been a central and irreplaceable enabling technology in the analysis and elucidation of glycoproteins [18].

For a number of years characterization of glycan structures has almost exclusively been the domain of mass spectrometry [19] and in conjunction with nuclear magnetic resonance (NMR) has been used extensively to define novel and unusual glycan structures associated with bacterial proteins [2]. Mass spectrometry as an analytical technique requires that samples to be analyzed are converted into ions in the gaseous phase. The mass spectrometer separates these ions according to their mass-to-charge ratio (m/z) and records the relative abundance of each of the ions to obtain a mass spectrum [20]. For more detailed information on mass spectrometry principles and instrumentation the reader is directed to the reviews of Domon & Abersold [21], Abersold & Mann [22], Graham *et al.* [23] , Smith *et al.* [24, 25] and Yates *et al.* [26, 27].

The development of 'soft ionization' techniques in the 1980s allowed the ionization of large, polar, and thermally labile biomolecules including glycopeptides [28] that did not previously lend themselves to such analytical techniques [29, 30]. The two critical 'soft ionization' technologies are electrospray ionization mass spectrometry (ESI-MS) and matrix-assisted laser desorption ionization mass spectrometry (MALDI- MS).

Electrospray Ionization (ESI)

ESI has become a mainstream method for life science research as a result of its high sensitivity and broad applicability. ESI is typically carried out in conjunction with high performance liquid chromatography (HPLC) and can be used with a nano electrospray configuration [31]. There are two models of how ions are generated in ESI, the charged droplet model and the ion evaporation model. In the charged droplet model, ESI generates charged micro-droplets containing analyte ions. The sample of interest is dissolved in a solvent and then pumped through a thin capillary or needle that is raised to a high potential that may be positive or negative. Due to the application of the electric field the solution exits the tip of the capillary in the shape of a cone, known as the Taylor cone [32]. Charged droplets are sprayed from the apex of the cone when the electrostatic repulsion of the charged molecules approaches the surface tension of the solution. These small charged droplets travel down a pressure and potential gradient towards an orifice in the mass spectrometer. As the droplets traverse this path they become desolvated and reduced in size, however, their charge remains constant [33]. Electrostatic stress near the surface of the droplet increases as the droplet shrinks. When the force of electrostatic repulsion between like charges becomes equal to the surface tension of the droplet, it undergoes Coulombic fission leading to the production of smaller droplets [34]. In the charged residue model these fissions ultimately produce droplets small enough to contain only one analyte molecule. As the last of the solvent evaporates from the droplet the analyte molecule retains some of the charge of the droplet to become an ion [35]. In the ion evaporation model following Coulombic fissions a droplet does not have to be formed containing only one analyte ion. Instead solvent evaporation increases the surface charge density at the droplet surface. When the droplet becomes small enough, the small radius of curvature and high charge density are strong enough to overcome the solvation forces holding the ion to the droplet and it is lifted from the droplet surface into the ambient gas [36]. In the first model the solvent leaves the ion and in the second the ion leaves the solvent [37]. The gas phase ions then enter the mass spectrometer and are detected as a series of multiply charged ions (e.g. $[M+nH]^{n+}$). ESI as an ionization process is relatively gentle, resulting in little or no fragmentation of the glycans or glycoproteins/glycopeptides being studied [28]. The absence of matrix adducts also leads to cleaner mass spectra when compared to MALDI [38]. However, native glycans are not ionized well by this process and therefore information on them can be lost [39]. The sample being investigated must also be free of salts and other contaminants, so some form of sample clean up is usually required prior to ESI [38].

Matrix-Assisted Laser Desorption-Ionization (MALDI)

Matrix-assisted laser desorption-ionization (MALDI) is a method that uses a laser to desorb intact molecular ions of proteins and peptides that are co-crystallized in a photoactive matrix solution [40, 41]. Unlike ESI, generally only singly charged ions are observed for MALDI. In MALDI, the analyte of interest is co-crystallized with an excess of matrix, that is used as a diluent preventing the analyte from forming large aggregates that would otherwise be too large to desorb [42, 43]. The matrix also absorbs UV light from a laser thus facilitating analyte desorption and ionization [42]. There are a number of different matrices that may be used in MALDI-MS. Typically, matrices for the analysis of glycans are 2,5-dihydroxybenzoic acid and its derivatives [44]. Analyte and matrix are both spotted onto a metal target plate, which is then inserted into a high vacuum source region within the mass spectrometer [43]. The target plate is subjected to laser bombardment and analyte molecules are vaporized along

with the matrix molecules. The process of desorption and ionization in MALDI is not fully understood with several influencing factors such as, laser wavelength, pulse width and chemical properties of the matrix and analyte [45]. During laser irradiation a gas jet of matrix neutrals and surrounding analyte molecules is formed. The matrix molecules are strongly excited at this stage and analyte molecules are thought to be ionized as a result of multi-step interactions with the matrix resulting in proton transfer giving both protonated and deprotonated analyte ions [42, 43]. One of the advantages of this technique is that you can obtain spectra from unmodified glycans [39, 44, 46] and the technique also has a high tolerance for salts and other contaminants [38]. MALDI is, however, a harsher ionization process than ESI resulting in the fragmentation of acidic glycans particularly sialic acid (**9**) residues [39].

TANDEM MS

The use of mass spectrometry in glycoproteomics is primarily advantageous because it has the ability to produce mass spectra, which tend to be unique for a particular glycopeptide/glycoprotein and thus can be used in the characterization of the peptide/protein and the mapping of any modifications [18]. The investigation of glycoproteins may be carried out using one of two approaches either 'top down' or 'bottom up' [47]. The goal of the top down methodology is to identify intact glycoproteins using mass spectrometry, without the need for prior proteolytic digestion of the sample. Initially proteins are introduced into the mass spectrometer in the gas phase and are then fragmented [48-50]. This methodology has proved to be extremely useful in bacterial glycoproteomics in the elucidation of novel glycan structures such as those attached to the pilin in *Neisseria sp*. [51, 52] and *Mycobacterium tuberculosis* [53]. This methodology is not as widely used as the bottom up approach as it typically requires a high resolution mass spectrometer such as FT-ICR [47] or MALDI/TOF-TOF [50]. However, this technique is increasingly being used in conjunction with bottom up methodologies, to gain wider structural information for the identification of bacterial glycoprotein's and glycan structures [54]. The more widely used approach when investigating glycopeptides is the bottom up strategy. This approach refers to any methodology that identifies proteins from the analysis of peptides derived from the proteolytic digestions of those proteins [55].

Whilst the knowledge of the mass of a glycoprotein or glycan may provide initial information regarding composition, to acquire structural or peptide sequence information from glycopeptides, it is necessary to induce fragmentation of the glycopeptides of interest *via* tandem mass spectrometry. This is not possible with soft ionization techniques such as ESI and MALDI, however, the use of these techniques in conjunction with tandem mass spectrometry has allowed the structural elucidation of a wide range of glycopeptides and their glycan structures [18, 56].

Fragmentation of the glycopeptides can be achieved by several different methods, which can induce fragmentation of both the peptide backbone and the glycan moiety. The most common fragmentation technique is collisional induced dissociation (CID), sometimes also referred to as collisional activated dissociation (CAD), but other complimentary techniques include infrared multiphoton dissociation (IRMPD) [57-59], electron capture dissociation (ECD) [58, 60] and electron transfer dissociation (ETD) [61]. It has been demonstrated that in tandem MS of glycans those that are adducted with a Na^+ or other metal ion rather than merely being protonated provide a larger number of structurally informative fragments [62].

Collisional Induced Dissociation (CID)

In tandem MS peptides are individually ionized in the source region using ESI or MALDI. These peptides are then further separated, based on their *m*/*z* ratio. The selected ions are allowed into a collision cell, which is filled with an inert gas such as xenon, argon or nitrogen, collisions then occur between the precursor ion and inert-gas atoms (molecules). In these collisions part of the precursor ion's translational energy can be converted into internal energy, and as a result of single or multiple collisions an unstable excited state is populated. Excited precursor ions decompose to produce product ions during collision-induced dissociation [34]. The types of fragment ions observed in an MS/MS spectrum depend on many factors that include the primary sequence of the peptide, the amount of internal energy and the charge state [63].

The main types of ions observed in the fragmentation of protonated peptides are well established noting that fragments can only be detected if they carry a charge. If the charge is retained on the N-terminal fragment the ion is classified as a, b, or c and x, y, or z if the charge is carried on the C-terminus. The nomenclature for fragmentation ions was described by Roepstorff [64] and subsequently modified by Johnson *et al.* [63] Fig. (**1A**).

The glycan moiety generally fragments in one of two ways, either by fragmentation of the bond between two adjacent sugar groups or by cross-ring cleavages occurring *via* rupture of two bonds within the same sugar unit [19, 46]. The nomenclature for fragmentation of carbohydrates is also well established, ions retaining the charge on the non-reducing terminus are classified as A, B and C, those that retain the charge on the reducing terminus are X, Y and Z. A and X are the result of cross-ring cleavages, whereas B, C, Y, and Z are the result of glycosidic bond cleavages [65] (Fig. **1B**).

CID for the most part produces fragments from glycopeptides that arise from glycosidic bond cleavage providing invaluable information on the sequence and connectivity of glycopeptides [57, 61]. However, CID has inherent disadvantages, the efficiency of fragmentation and also the degree of fragmentation decreases with increasing mass of the oligosaccharide. Also a number of successive tandem mass spectrometry steps may be required to gain insight into the structure of the glycan, and this is tempered with the loss of material at each step in CID making interpretation of spectra at later MS^n more difficult.

Infrared Multiphoton Dissociation (IRMPD)

In IRMPD the ion of interest is isolated and irradiated with an infrared laser beam, as the ions absorb energy from

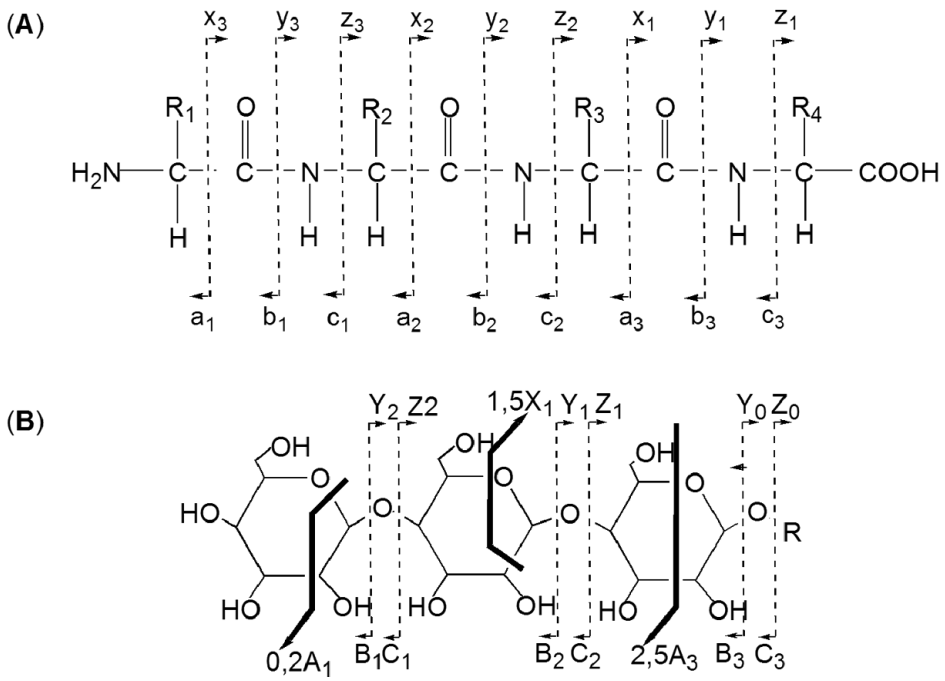

Fig. (1). (**A**) Ions generated from the fragmentation of peptides/proteins. (**B**) Ions generated from the fragmentation of carbohydrates/glycan chains.

the beam they become vibrationally excited, leading to bond scission. IRMPD produce fragmentation ions that are, to all intense purposes, the same as those for CID [57]. There are, however, some advantages, since this process does not depend on resonant energy transfer, the energy imparted to the glycopeptide can be more easily controlled, by adjusting the laser intensity and the duration of exposure, and since both the precursor and fragment ions can absorb the infrared photons the fragment ions can undergo secondary fragmentation providing further structural information without the need for further MS^n [59, 66]. There is no need for the introduction of collision gas, and thus no subsequent perturbation of vacuum and loss of resolution. Therefore, IRMPD has a higher duty cycle and can do more scans in a given time, resulting in more structural information [57, 58]. It has also been demonstrated, in this process, that the efficiency of fragmentation of oligosaccharides is enhanced as the size of the oligosaccharide increases [67]. This process has also been used to obtain structural information from N-linked glycans *via* cross-ring cleavage, generating information, which is not readily obtainable using CID [66].

Electron Capture Dissociation (ECD)

During ECD multiply protonated precursor ions are confined and are irradiated with low energy electrons (<0.2 eV) from an electron gun. This causes the peptide to fragment, producing $[M + nH]^{n+} + e^- \rightarrow [M + (n-1)H]^{(n-1)+}$ as the major product ion. This species undergoes molecular rearrangement and dissociates into sequence specific products with cleavages observed at the polypeptide backbone producing c and $z^{\cdot}$ product ions [68, 69]. Within this technique even though the backbone of the peptide is fragmented the more labile modifications like glycosylation remain intact on the polypeptide backbone, enhancing the ability to identify and localize the sites of glycosylation. This has great promise particularly for the identification of O-linked glycosylation sites, which have no consensus sequence [69, 70].

Electron Transfer Dissociation (ETD)

In ETD a singly charged anion, originally anthracene and eventually fluoranthracene is used to transfer an electron to a multiply charged peptide or protein. The resulting odd electron species undergoes subsequent recombination. This results in the production of c and z fragment ions from the peptide backbone. Thus, labile post-translational modifications are conserved, so again, this represents a potential tool for localizing sites of glycan attachment [61, 70, 71].

Glycosylation in Eukaryota

Attachment of glycans to the polypeptide backbone of proteins occurs *via* several distinct linkages including N-linked, O-linked, glycosylphosphatidyl inositol anchors and C-glycosylation [72]. The common monosaccharides involved in glycosylation can be seen in Fig. (**2**).

N-linked glycosylation is a highly specific co-translational reaction that is to say that it occurs during protein synthesis before the protein is folded into its final conformation [73]. In eukaryotes, N-acetyl glucosamine (GlcNAc) (**5**) is linked to the amide nitrogen of an asparagine residue, within the highly conserved consensus sequence Asn-X-Ser/Thr, where X may be any amino acid apart form proline [14, 38, 72-74]. The β-glycosamine linkage between GlcNAc and Asn is the most widely distributed carbohydrate-peptide bond in nature [75]. All eukaryotic N-glycans contain a conserved pentasaccharide core, the trimannosyl chitobiose core, made up of two GlcNAc (**5**) and three mannose (**3**) units shown in Fig. (**3**). The different permutations of N-glycans in eukaryotes are all manufactured from a common precursor, which is effectively the

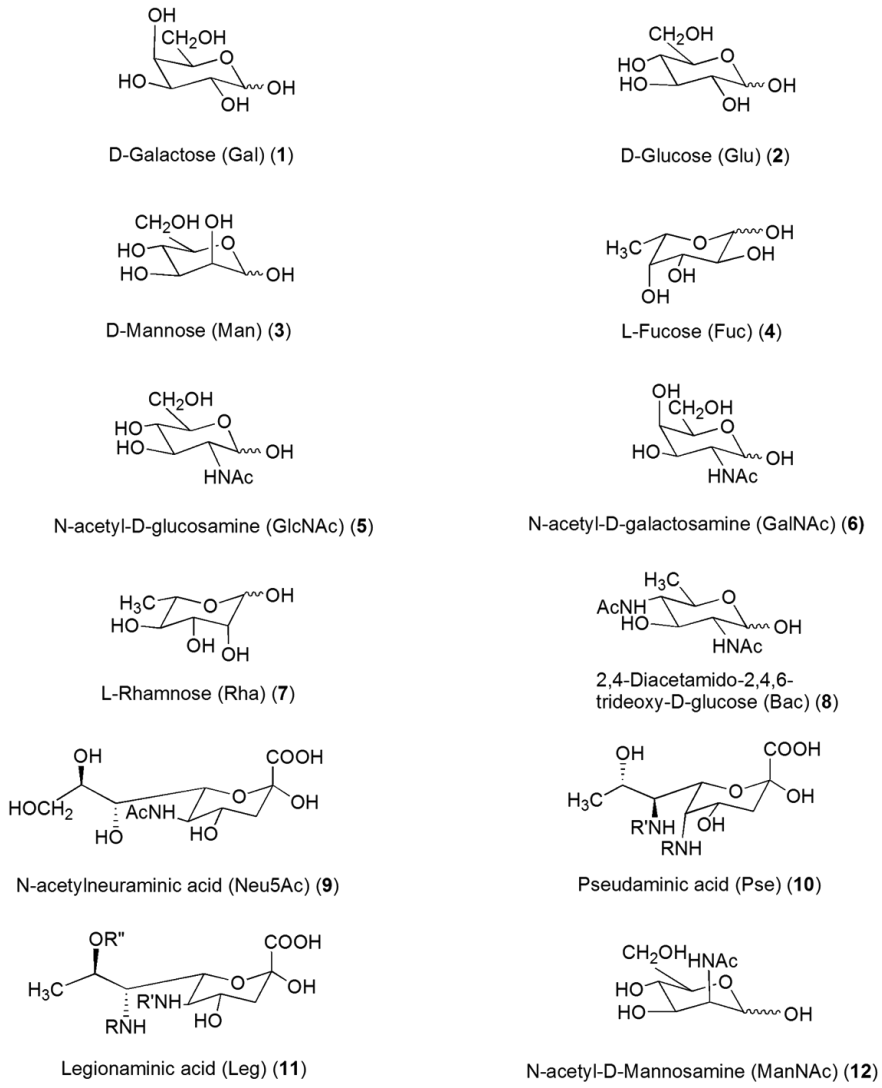

Fig. (2). The structures of common monosaccharides utilized in glycosylation by eukaryotes and prokaryotes. R, R' and R" represent positions where different substituents may be attached to give derivatives of the parent monosaccharide.

trimannosyl chitobiose core with one or more antennae attached to each of the two outer mannose residues. A number of enzymes are then involved in the processing of this structure through cycles of trimming and addition of carbohydrate residues to produce the three main classes of N-glycans shown in Fig. (**3**):

(1) High mannose – where only mannose residues are attached to the chitobiose core.

(2) Complex N-glycans - these have had all mannose residues removed and can have two (biantennary), three (triantennary) or four (tetraantennary) antennae.

(3) Hybrid N-glycans - where one of the antennae consists of only mannose residues [38, 39, 72, 73].

O-linked glycosylation in eukaryotes is a simpler process than N-glycosylation and is a post-translational event [72, 73]. There is no known consensus sequence for O-linked glycosylation; a variety of monosaccharides can be attached to any amino acid with a hydroxyl functional group namely serine, threonine, tyrosine, hydroxyproline and hydroxylysine [7, 75, 76].

Although there is no consensus sequence, β-linkage of N-acetyl-galactosamine (GalNAc) (**6**) to Ser/Thr residues is widespread in eukaryotes, with eight associated O-glycan core structures [77]. The GalNAc (**6**) to Ser/Thr linkage is found in many cytoplasmic and nuclear proteins [75]. Mucins are also associated with these Ser/Thr O-glycans found in abundance in tandem repeat domains rich in proline [7, 16, 28, 72].

Glycosylphoshatidyl anchors are mainly distributed among biologically important cell surface glycoproteins. This process involves the linkage of phosphoethanolamine to mannose (**3**), which is then attached to the carboxyl terminal of the target protein [72, 75].

C-glycosylation is the only form of glycosylation that does not use an amino acid functional group in its bonding. Instead the linkage occurs *via* the anomeric carbon of an α-

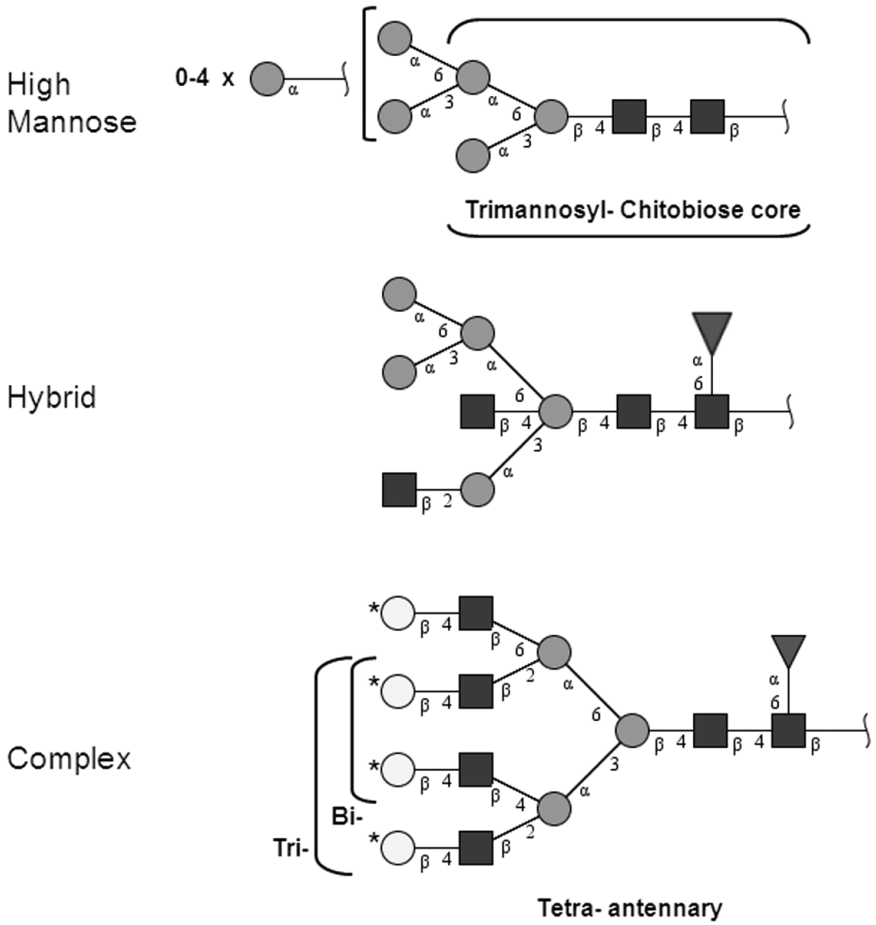

Fig. (3). The three main classes of N-glycans in Eukaryotae: high mannose, hybrid and complex type. ● Mannose residues (**3**), ○ Galactose residues (**1**), ■ N-acetyl-glucosamine (**5**), ▲ Fucose residues (**4**), *a number of varying carbohydrate constituents can be added to these antennae.

mannose residue (**3**) binding to the C_2 atom of the tryptophan side chain [75, 78, 79].

Glycosylation in Prokaryota

Glycosylation was once thought to be the sole dominion of the Eukaryota. It was with the discovery of N- and O-linked carbohydrates attached to a surface layer protein of the halophilic archaeon *Halobacterium salinarium* in 1976, that this long held misconception was dispelled [2, 80-82]. Over the following years, a plethora of surface layer and non-surface layer glycoproteins have been identified in Archaea and Bacteria [2, 80, 81, 83] including such milestones as the first structural evidence for glycosylation of pili *via* an O-linked trisaccharide in *Neisseria meningitidis* [84] and the first evidence for N-linked glycosylation in bacteria, with the identification of an N-linked heptasaccharide on protein PEB3 from *Campylobacter jejuni* [85].

N- and O-linked glycosylation events have been found to be widespread in Prokaryota. There are important distinctions between eukaryotic and prokaryotic glycosylation events. In prokaryotic glycosylation, the common structural units such as the chitobiose core and antennae are not necessarily present [86]. This is aptly demonstrated by a number of Archae that directly link their glycan chains to S-layer proteins by N-glycosidic linkages for example Glc (**2**) to Asn [87] and GalNAc (**6**) to Asn in the case of *Halobacterium salinarium* [82] and Rha (**7**) to Asn for the Archae *Methanotrix soehngenil* [88] and this process has also been observed in bacteria in *Bacillus stearothermophilus* [87, 89]. Whilst a number of prokaryotes have the eukaryotic N-glycosylation consensus sequence Asn-X-Ser/Thr, in *Campylobacter jejuni* this has been found to be extended to Asp/Glu-Y-Asn-X-Ser/Thr, where X and Y can be any amino acid but proline [90]. In fact, the optimal sequence was found to be Asp-Gln-Asn-Ala-Thr [91].

Also prokaryotic N-linked glycosylation is not always *via* GlcNAc (**5**) to Asn. In the case of *Campylobacter jejuni* the linkage to the Asn is *via* a unique bacterial sugar Bac (**8**), a derivative of bacillosamine [92].

Whilst a number of prokaryotes conform to the common eukaryotic GalNAc (**6**) O-linked glycosylation at Ser/Thr [93], there are instances, where different consensus sequence sites such as mannose (**3**) linked D-S* and D-T*-T in *Chryseobacterium meningosepticum* [94] and V-Y* in *Thermoanaerobacter kivui* (* denotes site of attachment) have been identified [95].

Prokaryotic protein glycans are composed of far more unusual and varied structures than those found in eukaryotic

systems (**8-11)** Fig. (**2**) [2, 96]. This significantly complicates carbohydrate structure determination in prokaryotes. Despite this restriction prokaryotic 'glycomics' has literally exploded in the last decade [96]. This heightened interest is due in part to the large number of glycosylation events that have been identified and associated with human pathogens, such as *Campylobacter jejuni*, *Helicobacter pylori*, *Nesseiria gonorrhoea*, *Pseudomonas aeruginosa* and *Mycobacterium tuberculosis* to name but a few [2, 14, 81, 96, 97]. It has been shown that these glycosylated proteins are involved in pathogenicity, adhesion and immune evasion of these bacteria [11, 98-100].

Although Prokaryotes lack cellular organelles similar processes for glycan biosynthesis are conserved between them and the Eukaryotes. For example the separation of the respective pathways by cellular membranes and the use of nucleotide activated carbohydrates and membrane bound intermediates as well as the transfer of the glycans to similar consensus sequences on target proteins [97]. However, there are also a number of differences between the prokaryotic and eukaryotic glycosylation systems including additional enzymatic processing steps that take place within the Endoplasmic Reticulum and Golgi Apparatus of the Eukaryotes. This review will focus on prokaryotic glycosylation. For more detailed information on eukaryotic glycosylation the reader is directed to the reviews of Reuter and Gabius [1] and Weerapana and Imperiali [101].

ANALYTICAL ENRICHMENT TECHNOLOGIES

Post-translational modifications are often found at substoichiometric levels making their identification and analysis particularly difficult [102]. This is especially true of glycoproteins and glycopeptides, which often are in low abundance with their non-glycosylated counterparts in complex mixtures. Therefore, some sort of enrichment techniques need to be employed to purify and concentrate these glycoforms prior to mass spectrometry analysis [103].

Lectins

Lectins are a widely used tool in research into carbohydrate chemistry, traditionally they have been used for histochemical detection of sugar chains on cell surfaces [104, 105]. These multivalent proteins are becoming increasingly prevalent as analytical tools in glycoproteomics due to their unique affinity towards carbohydrate epitopes by recognizing not only specific residues but also being able to distinguish between linkage and terminal modifications of complex oligosaccharide chains [106-108]. There are large arrays of lectins used in the analysis of glycoproteins and peptides, for the capture of N-linked glycans, two with the broadest applicability are Concanavalin A (ConA) and Wheat Germ agglutinin (WGA), whilst for the capture of O-linked glycans Jacalin is widely used.

Con A was first isolated by Sumner from the Jack bean (*Canavalia ensiformis*) in 1919 [109]. It has been shown to bind predominantly to α-D-mannose (**3**) [110], which is prevalent in eukaryotic N-linked glycans. It has a marked preference for the α-configuration at the C-1 position of the D-pyranose ring. It has a general propensity for α-D-mannosyl (**3**) and α-D-glucosyl (**2**) residues at the non-reducing termini of oligosaccharides although certain non-terminal α-D-mannosyl (**3**) and α-D-glucosyl (**2**) residues have also been seen to bind. Con A has been widely studied and has been shown to exist as a dimer below pH 5.5 and a tetramer with a total of four binding sites above pH 5.5 [111].

WGA was identified during work carried out on tumor cells by Aub and colleagues [112]. It is a dimeric protein with four binding sites that has specificity for GlcNAc (**5**) and its β-(1→4) linked oligosaccharides and analogous sugars [113]. It has also been shown to have affinity for sialic acid (neuraminic acid) (**9**) [111], which is present in a wide variety of glycans as an end cap bonded to Gal (**1**) or GalNAc (**6**) [105].

Jacalin is another lectin of interest that was isolated from seeds of the Jackfruit (*Artocarpus integrifolia*) [114]. Jacilin is a tetrameric protein that has specificity towards α-D-Galactose (**1**) and α-D-Galactose (**1**) attached to GalNAc (**6**) [115] and has been found to bind to Gal (**1**) β-(1→3) GalNAc (**6**)-α Ser [116] and as such is used in the capture of O-linked glycans.

These (and other) lectins can be used separately, or as was first demonstrated by Yang and Hancock [111] can be combined together to produce multi-lectin affinity columns, which can be used to enrich for a number of O- and N-linked glycoproteins present in a sample prior to MS analyses [117]. Madera *et al.* [107] have also attempted to take this lectin binding technology for the analyses of glycopeptides a step further by coupling it directly online to the mass spectrometer by producing lectin microcolumns that can be directly incorporated into MS front-end configurations. Work is even ongoing on the production of multiple lectin 'microarrays' to try and carry out large scale capture of as wide a proportion of the glycome, as possible, in a single experiment [118].

Graphite Carbon Column Chromatography

Charcoal columns have been used for the separation of oligosaccharides since the 1950's [119]. These, however, were not robust enough to be applied to HPLC techniques. The advent of Graphite Carbon Chromatography columns (GCC) enabled Koizumi and co-workers to use these columns for the resolution of carbohydrates including mono- and disaccharides as well as isomeric and even different anomeric forms [119]. The column retention appears to be based on both the carbohydrate and peptide portions present in glycoproteins and glycopeptides [120]. These columns are now widely used in glycomic/glycoproteomic research interfaced directly online to the mass spectrometer [121] for the analysis of N- and O-linked glycans [122] and have even recently been adapted in a Chip format for the Agilent mass spectrometers [123].

Hydrophilic Interaction Chromatography

Hydrophilic interaction chromatography (HILIC) has been used for the analysis of sugars and oligosaccharides since the 1970's [124] and efficient high resolution separation of glycans and glycopeptides can be achieved using amine and amide based columns [77, 125]. The glycans in-

teract with the hydrophilic (polar) stationary phase through H-bonding, ionic interactions and dipole-dipole interactions [125] and are eluted from the column in a solution of organic modifier and water, where increasing concentrations of water are used to remove the glycans [125]. Guile *et al.* [126] demonstrated that by labeling the reducing ends of glycans by reductive amination followed by separation on a HILIC column, it was possible to simultaneously resolve subpicomolar mixtures of neutral and acidic glycans. For structural characterization of the glycans their elution positions from the HILIC column were calculated in glucose units with reference to a dextran ladder [126]. HILIC of underivatized glycans has also been directly interfaced online in the nanoLC format enabling the detection of oligosaccharide mixtures at low-femtomole sensitivity [127]. The same group also analyzed glycopeptides by this methodology [128]. HILIC micro columns have also been used for the capture of glycopeptides from N-glycosylated proteins with subsequent direct release onto MALDI target plates for direct MS analysis [129].

Reversed Phase Chromatography

Native carbohydrates are not particularly well retained on reversed phase (RP) HPLC columns, however, derivatization of the glycans by either permethylation or reductive amination renders them more hydrophobic and, therefore, amenable to RP separation, with the added advantage that the RP column can be directly coupled to the mass spectrometer [39, 46]. The separation of the glycans is based mainly on the hydrophobicity of the tag. As the size of the glycan increases, so does the hydrophilicity of the tag, thus the larger the tagged glycan the less effectively it binds to the column [77, 130].

Periodate Oxidation Coupled to Hydrazide Resin Capture

Periodate oxidation allows for the identification of sites of N-linked or O-Linked glycosylation. As illustrated in Fig. (**4**), in this process the vicinal cis-diol groups of carbohydrates in a glycan chain are converted to aldehydes, these in turn are reacted with immobilized hydrazide groups on a solid support forming a covalent bond. The immobilized glycoprotein undergoes proteolysis on the support and non-glycosylated peptides are removed. At this point the attached glycopeptide can be isotopically labeled for quantitative analysis and ε-amino groups of lysines are guanidinated to form homoarginines, facilitating detection by MS. The N-linked glycopeptides are then released from the column using PNGase F and are identified by mass spectrometry. O-linked glycopeptides can be released by β-elimination [131] or with a number of unspecific enzymes [132, 133].

β-Elimination Followed by Michael-Addition (BEMAD)

A number of groups have used β-elimination followed by Michael-addition to introduce functional groups into the sites of phosphorylation and O-linked glycosylation in peptides. This enables both the labeling and separation of the peptides in order to aid in the mapping and identification of these sites [134, 135]. In an interesting extension of this process Wells *et al.* [102] developed a methodology that was directed specifically towards O-linked glycan labeling and not phosphorylation site labeling. In this technique the O-linked glycans undergo mild β-elimination followed by Michael addition with dithiothreitol effectively adding an affinity tag to the site of O-glycosylation allowing the labeled peptide to be captured and the site of glycosylation accurately mapped.

GLYCAN RELEASE

Given the different linkage chemistries involved in the attachment of O- and N-linked glycans, different approaches, either enzymatic or chemical, must be taken to release the intact oligosaccharide chains [72].

N-linked Glycan Release

From its identification [136] to its purification [137] the most widely used enzyme for release of N-linked glycans has been peptide-*N*-glycosidase-F (PNGase F). This enzyme cleaves the amide bond between the asparagine and the glycosyl amine, releasing the glycosyl amine, which is readily converted to a glycan with a reducing end (aldehyde functional group) [77, 137]. Cleavage by PNGase F leads to the conversion of asparagine to aspartic acid, therefore effectively labeling the sites that had been occupied by the glycan structure [44, 72]. However, care must be taken when interpreting these data since hydrolysis of aspartic acid is a common side reaction leading to asparagine and thus potential misinterpretations [138, 139]. PNGase F releases most Asn N-linked glycans from their proteins, unless there is α(1→3) fucosylation (**4)** of the inner GlcNAc (**5**). Also as is found in

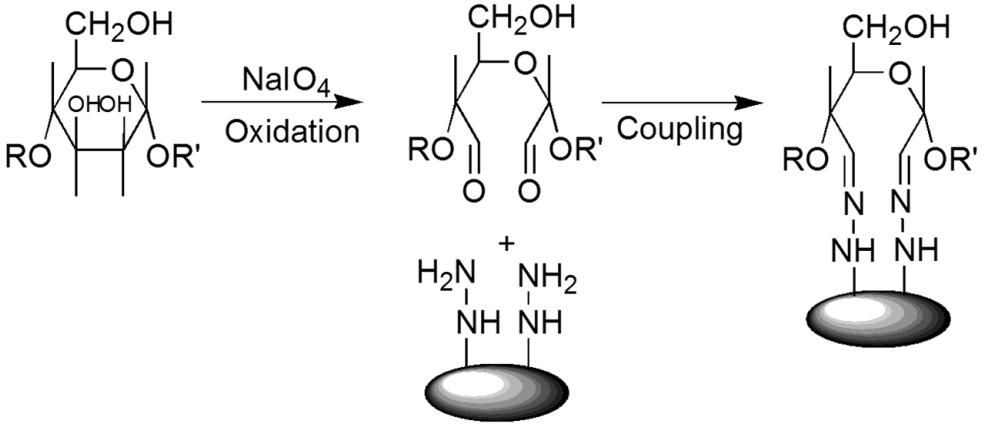

Fig. (4). Periodate oxidation and hydrazide resin capture of glycopeptides. R and R´ groups represent glycan chains and polypeptide chains.

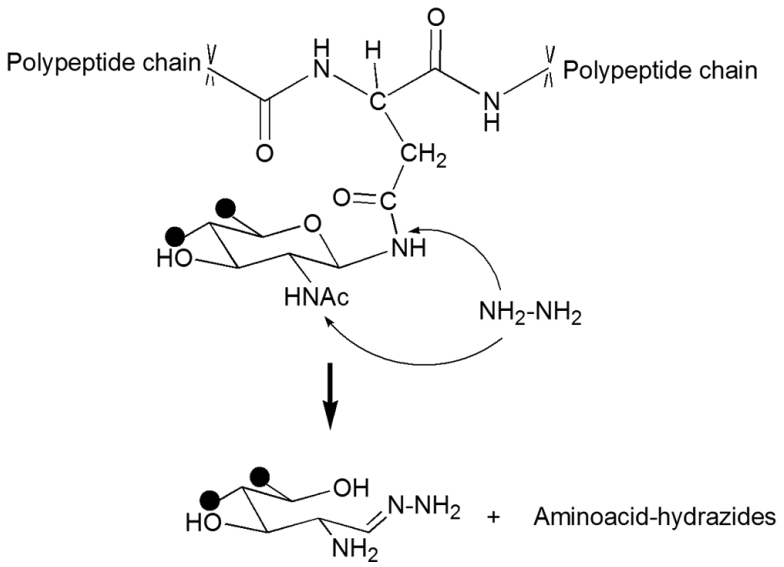

Fig. (5). Hydrazinolysis of glycopeptides/proteins. ● Glycan chains.

a number of prokaryotes the N-linkage, rather than occurring through GlcNAc (**5**), is *via* an unusual sugar moiety. This is the case in *Campylobacter jejuni* where the N-linkage is through di-acetyl bacillosamine (**8**) and PNGase F has no effect as it does not recognize this sugar as a substrate [140]. In this case Liu *et al*. [141] have developed a strategy using the enzyme Pronase E that results in the production of Asn linked glycans that can be further derivatized and studied.

A chemical method for the release of N and O-linked glycans is by hydrazinolysis shown in Fig. (**5**). This method releases the glycans, but the peptide bonds are destroyed leaving no information on the protein or peptide to which the glycan was attached. Also information on *N*- and *O*- acetyl and *N*-glycosyl substituents of sialic acid (**9**) modifications can be lost [39, 142, 143].

O-linked Glycan Release

Unlike PNGase F for N-linked glycan analysis, there is no broad specificity ubiquitous enzyme for the release of O-linked glycans. A number of exoglycosidases need to be used to reduce the glycan chain to the Gal (**1**) β(1→3)GalNAc (**6**) core attached to Ser/Thr, that can then be released by O-glycosidase [131-133]. Not all O-linked glycans, particularly those present in prokaryotes, contain this core structure. Therefore, O-linked glycans to Ser/Thr are usually released by chemical methods, the most common of which involves β-elimination under alkaline conditions in the presence of excess sodium borohydride ($NaBH_4$), also termed 'reductive' β-elimination.

The β-elimination converts Ser/Thr glycan linked amino acids to unsaturated hydroxyamino acids, however, the presence of the $NaBH_4$ ultimately reduces the unsaturated hydroxyamino acids to alanine and 2-aminobutyric acid. Degradation of peptides is common in β-elimination and obtaining sequence information can therefore be difficult [144]. The common addition of $NaBH_4$ also prevents 'peeling' of the released glycans by conversion of the terminal GalNAc (**6**) residue to its alditol (see Fig. (**6**)) [7, 39, 46]. During 'peeling' single monosaccharides are removed sequentially from the reducing end of the glycan chain. This is due to rearrangement under alkaline conditions of the reducing terminal to the keto form, followed by glycosidic bond scission [145]. Due to the conversion of the glycan to the alditol it cannot be further derivatized using reductive amination [77], which is commonly used for further glycan characterization. Using this β-elimination methodology - due to the presence of $NaBH_4$ - there can also be a significant release of N-glycans [146].

GLYCAN DERIVATIZATION

Glycans that have been released from their peptide or protein by enzymatic or chemical processes will contain a reducing terminus. One of the problems in the mass spectrometric analysis of these glycans is that due to the diversity of their composition and the presence, in many cases, of acetyl and acidic residues, the glycan moieties do not readily ionize [147, 148]. Thus derivatization of the glycan moieties is often performed prior to MS analyses.

Permethylation

One of the most commonly used techniques is to derivatize the carbohydrate to form acid stable methyl esters [149]. The most widely used method for the complete methylation (per-methylation) of carbohydrates was initially developed by Hakomori in 1964 [150]. This permethylation technique replaced previous techniques as it could be carried out quickly and efficiently. In this method the methylsulfinyl carbanion, produced from NaH/NaOH and DMSO, acts as a strong base, deprotonating all the labile sites on the carbohydrate, -OH, $-NH_2$ and -COOH. Methylation of these sites then occurs rapidly in the presence of methyl iodide shown in Fig. (**7**). All hydroxyl, carbonyl and amino groups are converted to methyl esters and all N-acyl groups are N-methylated [149, 151].

There have been a number of improvements to the original method, particularly the simplified approach of Ciucanu and Kerek [152] resulting in the production of methylated

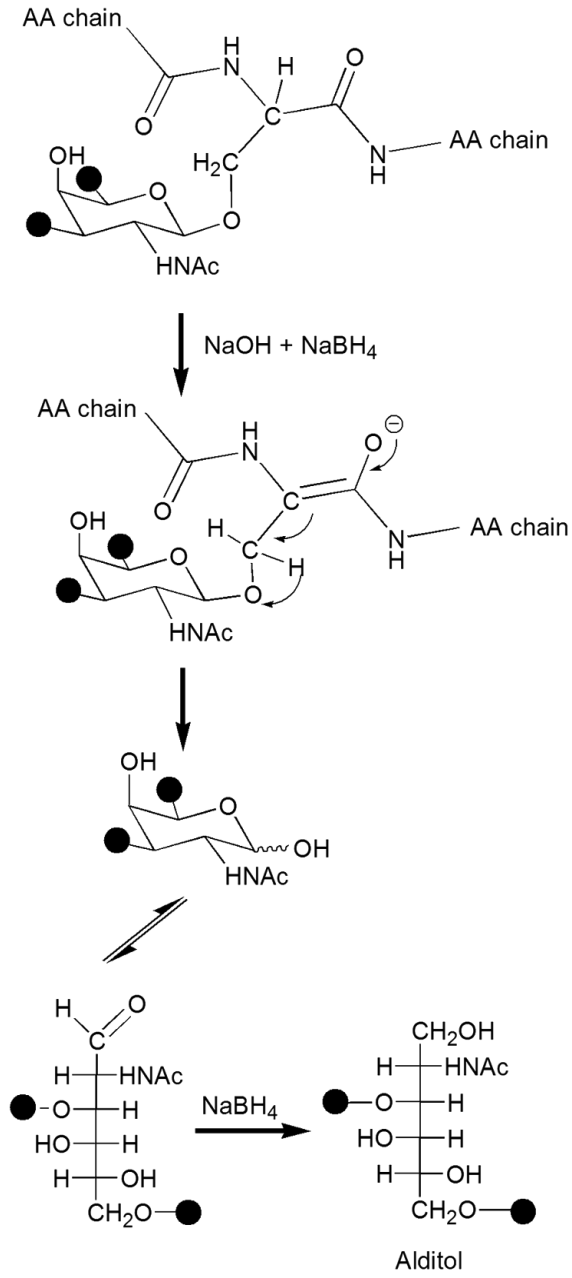

Fig. (6). Reductive β-elimination of O-linked glycans. ● Glycan chains.

glycans that ionize more efficiently and with significantly improved sensitivity, than their underivatized native forms. The improvement in glycan ionization is due to a combination of the highly polar hydroxyl and carboxyl groups being converted to non-polar derivatives and sialic acid (**9**) residues also being stabilized by the removal of the highly acidic proton on their carboxyl groups [46, 148, 153, 154]. The permethylation process also aids in the sample clean-up of these derivatized glycans prior to MS as the addition of the methyl groups make the glycans more hydrophobic and, therefore, amenable to separation and desalting *via* RP separation [46]. Another advantage of this derivatization is that the fragmentation pathways are much more informative and predictable leading to ions that can be unambiguously assigned to unique features [154].

This technique is also used widely in the establishment of the position of glycosidic linkages between sugar residues [149, 155]. The advantage here is that the glycosidic linkage oxygen will not be methylated. The fragmentation ions can be analyzed by tandem MS and the linkage site, where the glycosidic bond was, lacks the mass of the methyl group and, therefore, the linkage position can be established [62]. Another approach is to reduce the glycan to its alditol using $NaBH_4$, then permethylate it and the linkage to the reduced sugar can then be determined from the fragmentation of the alditol moiety [38, 155].

Reductive Amination

Reductive amination is another method used for the derivatization of glycans. In this technique the derivatives are

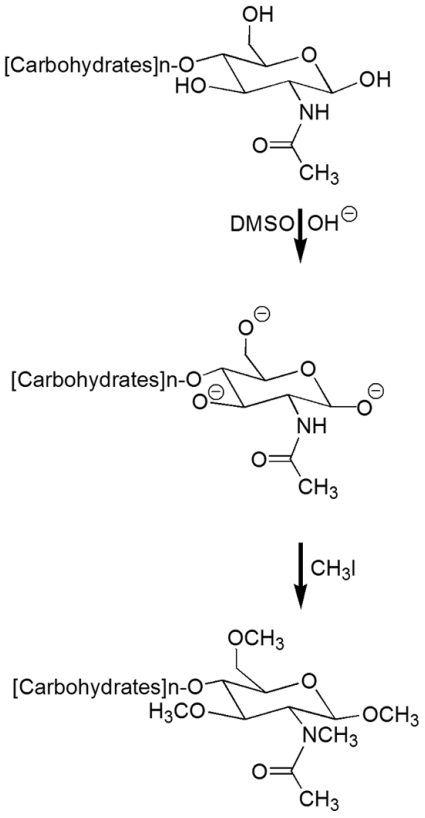

Fig. (7). Permethylation of glycan chains.

added to the reducing end of the glycan by reductive amination with an aromatic amine, such as the natural amino acid analogue 2-aminobenzoic acid. These aromatic amines can be fluorophores or chromophores and their addition leads to increased sensitivity for these derivatives and also enhanced MS signal intensity [39]. Fig. (**8**) illustrates the widely used scheme for reductive amination *via* Schiff base. The Schiff base formation is facilitated by the presence of an acid and the amine acts as a free base during the protonation of the carbonyl oxygen, once the Schiff base is formed it is then reduced, usually with cyanoborohydride ($NaBH_3CN$) to obtain the stable derivative [77]. This method works well for enzymatically released N-linked glycans and glycans that retain a reducing end but it is not amenable to O-linked glycans that have been released *via* reductive β-elimination, as the aldehyde group will have been reduced to an alcohol [38].

GLYCOINFORMATICS

One of the goals of glycoproteomics is to have the ability to analyze intact glycopeptides and to provide both peptide sequence and structural information on the attached glycan [156]. However, glycoproteomics places a huge burden on the researcher, generating large amounts of mass spectral data [16]. A major restriction here is the lack of rapid accurate automated tools for the retrieval of structural information from this MS data [18]. Without such tools each mass spectrum must be manually annotated by an expert, a task that is not too cumbersome with one spectrum but which becomes colossal in magnitude when applied to a full LC-MS/MS run [6]. For an in-depth analysis of these problems and the approaches being used to tackle them, the reader is directed to the reviews of Pérez & Mulloy [6, 157] and Aoki-Kinoshita [6].

In order to aid research in glycomics and glycoproteomics, several databases have been developed that contain structural data on complex glycans: Consortium for Functional Glycomics (http://www.functionalglycomics.org/glycomics/common/jsp/firstpage.jsp), the Kyoto Encylopedia of Genes and Genomes glycome information resource (htpp://www.genome.jp/keg/glycan/), the Japanese Consortium for Glycobiology and Glycotechnology Databases

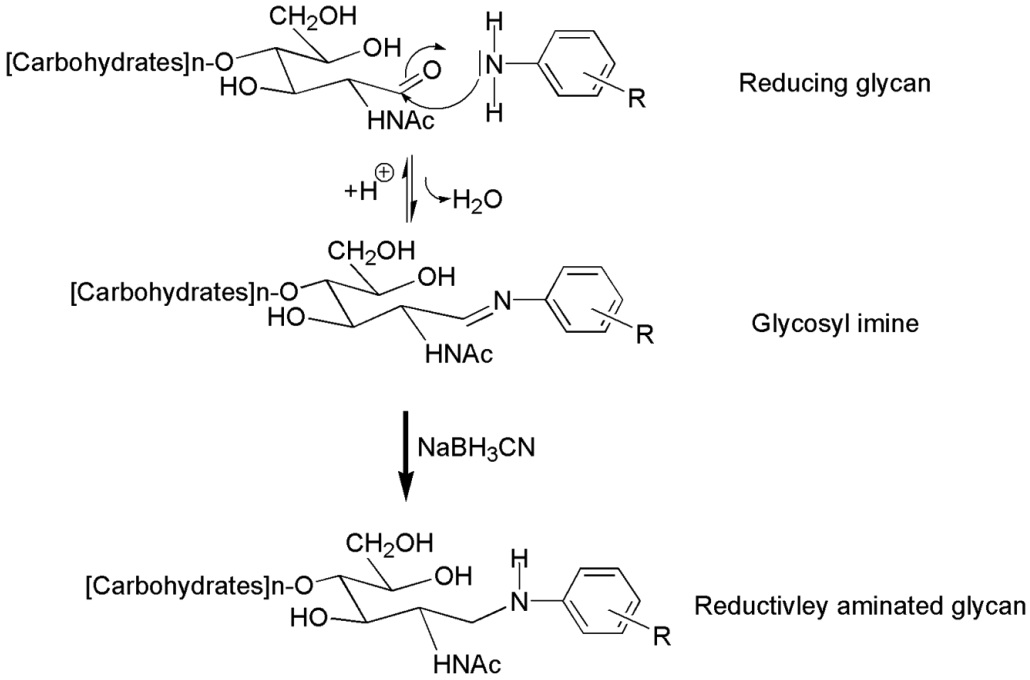

Fig. (8). Labeling of glycan chains by reductive amination. R group can be a number of different substituents.

(htpp://jcggdb.jp/index_en.html), Glycosciences.de (http://-www.dkfz.de/spec/glycosciences.de/sweetdb/index.php) and EuroCarbDB (http://www.ebi.ac.uk/eurocarb/home.action) [16].

There has also been the production of databases of theoretical mass spectra corresponding to known glycan structures that can be used for comparison with generated mass spectra such as Glyco-Fragment and GlycoSearchMS [158].

The production of GlycoWorkBench [159] to provide support for the routine semi-automated interpretation and assignment of glycomic MS spectra has been of immense aid to researchers in this field. This software provides a free easy to use graphical interface, and visual editor based on GlycanBuilder [160] for the interpretation of MS data.

The development of Cartoonist allowed for the first comprehensive attempt to automate the analysis of intact permethylated glycans. The system selects plausible annotations from a library, the system then determines the precision and calibration of the mass spectrometer used to generate the data and then assigns a confidence score to the most likely glycan structures and annotates the appropriate peaks in the mass spectrum [161]. The problem with these databases is that they are only applicable to the identification of glycans after release from and generally with prior knowledge of the peptide it was attached to.

The latest generation of bioinformatics programs, Peptoonist an extension of the Cartoon program [162], GlycoMiner [163] and Branch-and-Bound [156] are now capable of providing information on the sequence of N-glycopeptides and the structure of its glycan attachments.

ROLE OF GLYCOSYLATION IN *Campylobacter sp.*

Campylobacter jejuni and *Campylobacter coli* are recognized as the most common causative agents in human enterocolitis and are one of the main causes of severe bacterial diarrheal disease in humans. They account for more cases worldwide than those due to *Salmonella sp., Shigella sp.,* or *Escheria coli.* In the United States over 99% of reported infection with *Campylobacter* are from *C. jejuni.* In developing countries the incidences of *Campylobacter* infection are estimated at about 800 notifications per 100,000 population. In tropical developing countries *Campylobacter* infection is hyper-endemic among young people causing high mortality rates in children and infants. Several studies estimate infection rates in the under fives to be as much as 60,000 per 100,000 population [100, 164-167]. This bacterium is commensal in both poultry and cattle, with birds acting as its natural reservoir. Human infection generally occurs from the consumption of contaminated food or water [167, 168].

The flagellin of *Campylobacter* are the most heavily glycated bacterial proteins described to date and these species also produce a number of other glycated proteins that are destined for deposition on the cell surface or to be secreted [13]. These glycosylated proteins have been demonstrated to be highly immunogenic in nature [12]. They have been shown to play pivotal roles in virulence, microcolony formation and resistance to acid phagocytosis [13] as well as mediating the ability of *C. jejuni* to adhere to and invade cells *in vitro* and to colonize the intestinal tract of mammals [11, 12].

O-Linked (Flagellin) Glycosylation: Pseudaminic and Legionaminic Acid

Flagellins of *Campylobacter* bacteria are the immunodominant antigen during infection and are involved in the colonization of the GI-tract. Thus, investigation of these flagellins may provide insights into the capability of these organism to cause disease and may also lead to the elucidation of mechanism to inhibit or inactivate their ability to colonize and cause disease in hosts [169].

In *C. coli* VC167 two variants were identified that produce different flagellins. These were identified as T1 and T2 based on reactions with specific polyclonal antibodies. The genes for both were cloned and consist of two flagellin genes, *flaA* and *flaB.* Both gene products of *fla*A and *fla*B must be expressed in *Campylobacter sp*. to produce a fully active flagellum [169]. The predicted masses for these gene products were all between 58-59 kDa. Migration on SDS-PAGE gels showed that the masses were, in fact, much higher, suggesting that they were post-translationally modified.

Further evidence for the post-translational modification of these flagellins was the observation that the antigenic nature of the flagellin is based on the genetic background of the host strain and not the specific flagellin gene. Also, of the genes required for the synthesis of this post-translational modification, one had high sequence homology to *N*-acetyl neuraminic acid (Neu5Ac) synthetase in *Neisseria* and *Escherichia coli,* which is involved in the synthesis of capsular sialic acid (Neu5Ac) (**9**) [170].

Doig *et al.* [165] examined the flagellin of *C. coli*, *C. jejuni* and *C. fetus*. The first evidence that this post-translational modification was glycosylation was shown by the fact that the flagellin in all species underwent mild periodate oxidation producing free aldehydes, which were then able to be hydrazine biotinylated. The presence of terminal sialic acid (**9**) residues was also confirmed, using an adaptation of this method, where binding of flagellin only occurred in the presence of sialic acid (**9**) binding LFA (Limax flavus agglutinin) lectin. The inability of LFA to completely remove antigenicity, however, indicated that other molecules were also involved in sero-specificity; the fact that tunicamycin, an N-linked glycosylation inhibitor failed to prevent flagellin glycosylation suggested that this was an O-linked process [165].

Fry *et al.* cloned a number of genes that had significant homology to enzymes involved in lipopolysaccharide (LPS) biosynthesis, from *C. jejuni* 81116 into *E. coli.* Lipopolysaccharides are an abundant component of the outer membrane of Gram-negative bacteria that are involved in modulation of host immune response [100]. However, a further study using site specific mutations of these genes from *C. jejuni* 81-176 demonstrated that at least in *Campylobacter* they did not play any role in LPS biosynthesis. Rather, based on sequence homology of the genes in 81-176, they were classified as protein glycosylation *pgl* genes. *Pgl*E&F showed significant homology to flagellin modification *Flm* genes *FlmB&A,* which were involved in glycosylation of flagellin in other bacteria. *Pgl*B had significant homology to an oligosaccharide transferase involved in protein glycosylation in yeast. *Pgl*A is homologous to a galactosyl transferase in the

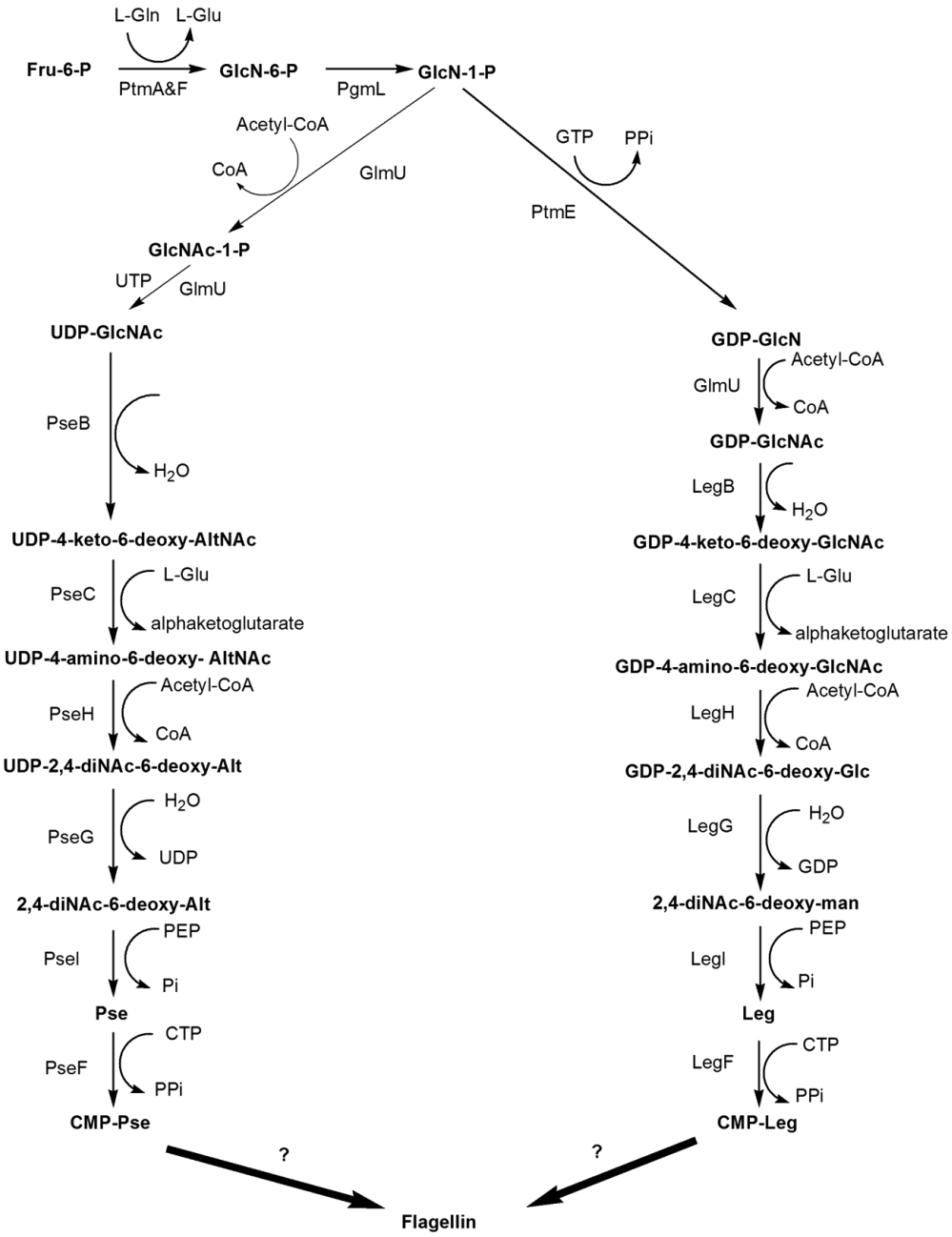

Fig. (9). The CMP-Pseudaminic and CMP-Legionaminic pathway. The pathways begin with PtmF&A acting in concert to convert fructose-6-P to glucosamine-6-P. PglM (Cj1407c) acts as a phosphoglucomutase converting glucose-6-P to glucosamine-1-P.

CMP –Pse biosynthetic pathway: The full process starts with GlmU proposed to produce UDP-GlcNAc from glucosamine-1-P, the C-4,6 dehydration and C-5 epimerization of UDP-GlcNAc by PseB forming UDP-4-keto-6-deoxy-AltNAc, C-4 aminotransferase of this by PseC forms UDP-4-amino-6-deoxy- AltNAc, N-4 deacytylation of this by PseH forms UDP-2,4-diNAc-6-deoxy-Alt, removal of UDP from C-1 of this by PseG produces 2,4-diNAc-6-deoxy-Alt condensation of this with pyruvate by PseI forms pseudaminic acid (**13**) and finally activation of this with CMP by PseF forms CMP-pseudaminic acid, ready for attachment to Ser/Thr groups at the appropriate surface exposed regions on flagellin in the cytoplasm.

CMP-Leg biosynthetic pathway: PtmE acts as the nucleotidyl transferase adding GTP to glucosamine-1-P producing GDP-GlcN. With the initial GDP sugar produced the acetylation of this is the next step, here GlmU is proposed to produce GDP-GlcNAc, but other enzymes may be involved. LegB acts as a dehydratase forming GDP-4-keto-6-deoxy-GlcNAc. LegC catalyzes the amino transfer of this to produce GDP-4-amino-6-deoxy-GlcNAc. LegH (Cj1298) acts as the *N*-acetyltransferase converting this to GDP-2,4-diNAc-6-deoxy-Glc. LegG then performs a C_2 epimerization resulting in the removal of the GDP to form 2,4-diNAc-6-deoxy-Man. LegI (Cj1327) catalyzed the condensation of this with pyruvate to form legionaminic acid (**16**). LegF (Cj1331) then catalyzed the addition of CMP to give the activated legionaminic acid for attachment to flagellin.

trisaccharide modification of pilin in *N. meningitidis*. This research also showed that the carbohydrate components attached to proteins were immunogenic in nature, including those of flagellin proteins [12].

Elucidation of the Pseudaminic Acid Biosynthetic Pathway

The proposed biosynthetic pathway for the production of cytidine monophosphate (CMP)-pseudaminic acid is shown in Fig. (**9**).

Although *C. jejuni* enteritis is somewhat debilitating in healthy individuals it also tends to be self-limiting. More serious conditions that can arise from *C. jejuni* infections include Guillain-Barré syndrome, which is the most common cause of acute flaccid paralysis in the United States [171] and Miller-Fisher syndrome. These diseases result in damage to the nerve gangliosides by antibody mediated axonal cytoskeletal breakdown and peri-synaptic Schwann cell damage. They are considered to be true cases of molecular mimicry mediated diseases, particularly related to antecedent *Campylobacter jejuni* infection. The link between these neuropathies and *C. jejuni* is thought to come about due to the bodies' production of antibodies against *C. jejuni* lipopolysaccharides (in *C. jejuni* these are analogous to lipo-oligosaccharides of mucosal pathogens), which cross-react with gangliosides. It is believed that lectin-binding proteins within the immune system recognize glycan binding sites on pathogens and activate the complement system thus producing these cross-reactive antibodies [171, 172]. This cross-reactivity is thought in part to be due to the presence of *N*-acetyl neuraminic acid (Neu5Ac) (**9**) that is displayed in prominent positions on vertebral cells and is critical for a number of physiological processes including neuronal plasticity [173], and is also found on the oligosaccharide portion of *C. jejuni* cell surface proteins. Investigation of glycoproteins and their associated glycans may help elucidate the mechanisms of this cross-reactivity and provide leads for new therapeutic intervention for *C. jejuni* infections.

Three putative Neu5Ac (**9**) synthetase genes were predicted from analysis of the *C. jejuni* NCTC 11168 genome sequence. *Neu*B encodes Neu5Ac (**9**) synthetase, which catalyses the formation of Neu5Ac (**9**) from *N*-acetyl-D-mannosamine (**12**) and phosphoenolpyruvate. *Neu*B2 (Cj1327) and *neu*B3 (Cj1317) were found to be part of a nineteen gene cluster from Cj1317 to Cj1337. Adjacent to this cluster were found the flagellin structural genes *fla*A&B. These genes were thus hypothesized to play a role in the post-translational modification of the flagellin subunit [168].

*Neu*B1 (Cj1141) was identified as the gene responsible for lipo-oligosaccharides sialylation within *C. jejuni* NCTC 11168. GC-MS analysis of purified lipo-oligosaccharides confirmed that a peak corresponding to Neu5Ac (**9**) was present at 23 minutes in the wild type trace but absent in the *neu*B1 mutant. Also FAB-MS of permethylated lipo-oligosaccharides showed that the carbohydrate portions of wild type and mutant lipo-oligosaccharides differed by the equivalent mass of a trisaccharide comprising Neu5Ac (**9**)-Hex-HexNAc with the hexose likely to be galactose (1) [174]. The sequencing of the *C. jejuni* NCTC 11168 genome in 2000 showed two large regions lower in G+C content that encompass Cj1135-Cj1148 and Cj1421-Cj1442 the lipo-oligosaccharides and extracellular (capsular) protein biosynthesis gene clusters [168]. There were also found to be variations in the lengths in polyG:C (homopolymeric) tracts in certain parts of the genome that lead to slipped-strand mispairing during replication, which can affect translation and has been shown to be responsible for phase variation of surface proteins and antigenicity. Interestingly, the majority of the hypervariable regions were to be found in the gene clusters responsible for lipo-oligosaccharide biosynthesis, extracellular/capsular protein biosynthesis and flagellar modifications. Some of these variable genes could be assigned possible functions, such as Cj1139 encoding a β-(1→3)-glycosyltransferases. A number of others were found to be associated with two families, 617 and 1318. The 617 family consists of five members Cj617/618, Cj1305c, Cj1306c, Cj1310c and Cj1342c, four of which are located in the flagellin glycosylation gene locus. There are no homologues of the 617 found outside *C. jejuni*. However, the 1318 family, at least in enterobacteriaceae are found within lipo-oligosaccharide gene clusters [98, 168]. These seven genes were found to be involved in variation of motility and were termed motility accessory factors (*maf*): Cj1318 (*maf*1), Cj1333 (*maf*2), Cj1334 (*maf*3), CJ1335/6 (*maf*4), Cj1337 (*maf*5), Cj1341 (*maf*6) and Cj1342 (*maf*7) [169].

The use of mass spectrometry in the analysis of O-linked glycosylation in *Campylobacter sp.* was indispensable in the extensive analysis of *Campylobacter sp.* flagellin carried out by Thibault and co-workers [175]. In this study flagellin were isolated from three *C. jejuni* species and one *C. coli* species. The intact flagellins were subjected to MS analyses, whilst flagellin glycopeptides were subjected to MS and tandem MS analyses before and after β-elimination. Analysis of the intact *Campylobacter* flagellin demonstrated that at least 10% of the mass of the flagellins was due to the attachment of glycan moieties. Analysis of the tryptic peptides derived from the purified flagellin identified three unexpected modifications of the peptides relating to species of masses 317 m/z (316 Da), 316 m/z (315 Da) and 409 m/z (408 Da). Accurate mass measurements were obtained for the oxonium ion relating to the 316 Da modification; the exact mass was found to be 316.122 ± 0.004 Da. The tandem MS spectrum of this moiety was characterized by the loss of neutral groups such as water, ketene and formic acid. These experiments indicate that the unusual glycan was a diamino sugar containing an acid group with two *N*-acetyl functionalities and a modified C_7 side chain. This was identified as pseudaminic acid (Pse5Ac7Ac) (**13**); a unique nine carbon bacterial carbohydrate shown in Fig. (**10**); and not, as had been previously thought sialic acid (**9**) [165].

The 316 m/z ion corresponding tandem MS spectrum showed losses consistent with the substitution of one of the two acetamido groups by an acetoamidino group, which resulted in a glycan moiety 1 Da lower than that for pseudaminic acid (Pse5Ac7Ac) (**13**) and was termed 5-acetamidino-7-acetamido-Pse (Pse5Am7Ac) (**15**). The 408 moiety was shown to have a mass difference of 92.027 Da to the pseudaminic acid residue and was consistent with the 2 *N*-acetyl groups being substituted for 2 *N*-2,3-dihyrdroxypropionyl groups (Pse5Pr7Pr) (**14**) (see Fig. (**10**)).

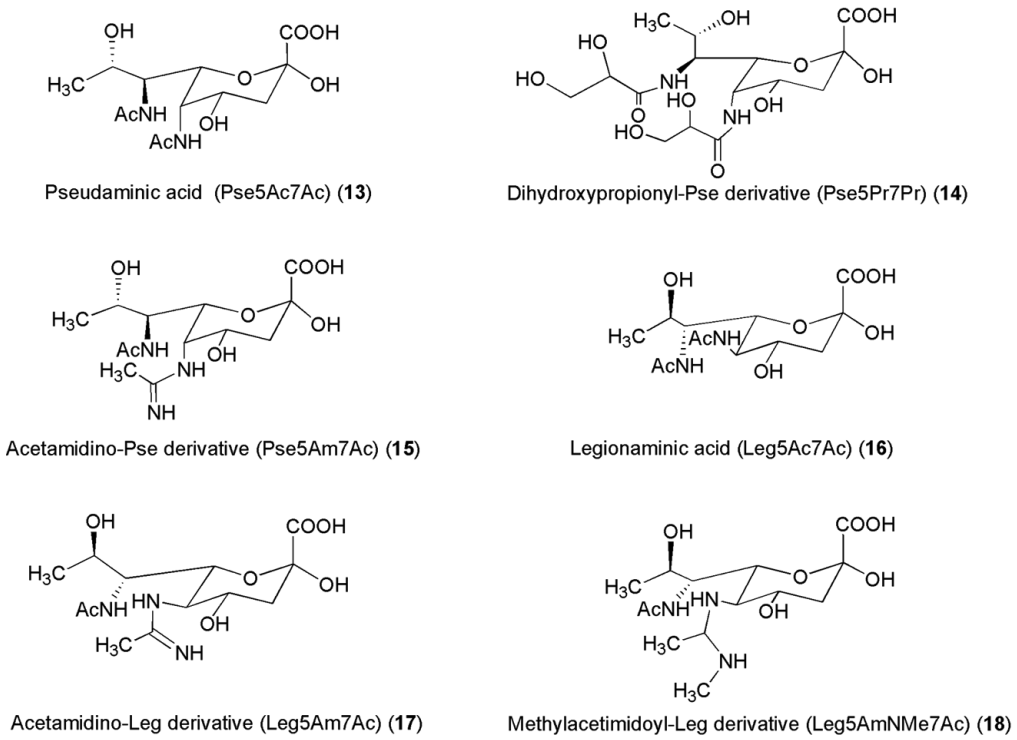

Fig. (10). The structures of Pseudaminic Acid (**13**), Legionaminic acid (**16**) and their discussed derivatives.

To determine the sites of O-linked glycosylation, the purified glycopeptides were subjected to β-elimination. Ten sites were identified as glycated by MS and a further nine were assigned using Edman sequencing. Most of these sites were localized in a region of the flagellin that was surface exposed. This work also identified the previously unknown functionality of the gene product Cj1316c (PseA) as responsible for the conversion of pseudaminic acid to 5-acetamidino-7-acetamido-Pse (**15**) [175]. The gene Cj1293 (*pseB*) was found to be central in flagellin glycosylation with pseudaminic acid (Pse5Ac7Ac) (**13**) and in flagellum assembly. *PseB* mutants were found to be non-motile, with accumulation of unglycosylated flagellin intracellularly. Flagellin taken from wild type and mutant were digested and subjected to tandem MS, only in the wild type was the characteristic 317 m/z pseudaminic acid oxonioum ion observed. Complementation with *pseB* restored motility and glycosylation patterns. *PseB* was found to have a high degree of homology to a *H. pylori* Uridiidine diphosphate (UDP)-GlcNAc C_6-dehydratase. When this protein was introduced into the *pseB* mutant, tandem MS analysis revealed the presence of the 317 m/z oxonium ion and 66% restoration of glycosylation, suggesting that *pseB* is a UDP-GlcNAc C_6-dehydratase involved in the production of pseudaminic acid. [167, 176]. Further studies proved that *pseB* in fact produced UDP-4-keto-6-deoxy-N-Acetyl-altrosamine (AltNAc), from UDP-GlcNAc, the first precursor in the production of pseudaminic acid (**13**) [177]. *PseC* (Cj1294c) was shown to use UDP-4-keto-6-deoxy-AltNAc as a substrate to produce UDP-4-amino-6-deoxy-AltNAc [167]. These findings were corroborated by a separate study that same year carried out on dehydratase and aminotransferase proteins from *Campylobacter* and homologous proteins in *Helicobacter* [178].

In an effort to further enhance knowledge of O-linked pseudaminic acid glycosylation Guerry *et al.* [13] selectively created knockout mutants of the flagellin glycosylation locus in *C. jejuni* 81-76. This species has only 21 of the 50 proposed genes involved in the bacterial glycosylation locus of *C. jejuni* NCTC 11168. Only nine mutations affected modification of the flagellum with pseudaminic acid (Pse5Ac7Ac) (**13**) or its derivatives. Mutation of *pseB, pseC*, *pseF*, *pseI/neuB3* (involved in pseudaminic acid (**13**) biosynthesis) and *pseG*, *pseH* and *pseE* (whose function were unknown at the time) all resulted in non-motile phenotypes. Based on homology to the sialic acid (**9**) biosynthetic pathway *pseF* was proposed to attach pseudaminic acid (**13**) to CMP converting it to its activated form for attachment to polypeptide backbones.

There was also a loss of the acetamidino derivative of pseudaminic acid (**15**) oxonioum ion (316 m/z) in *pseA* and *pseD* mutants, whilst still being evidence for the presence of pseudaminic acid (**13**), suggesting that both genes are involved in biosynthesis of acetamidino derivatives of pseudaminic acid (**15**) from pseudaminic acid (**13**) and its subsequent transfer to flagellin. Of the five *maf* homologs only *pseD&E* had an observable effect on flagellin. All the mutants resulting in the loss of acetamidino derivatives of pseudaminic acid (**15**) from the flagellin failed to autoagglutinate, a critical process for virulence, microcolony formation, and resistance to acid phagocytosis. This suggests that glycans on flagellin surface play a critical role in virulence [13].

Studies involving mass spectrometry as a key technique lead to the elucidation of all of the components involved in the biosynthesis of CMP-pseudaminic acid. These studies demonstrated that *pseD&E* were not involved in this biosyn-

thetic pathway and that *pseG* acted as a hydrolase removing UDP from one of the sugar intermediates [34, 179, 180] to produce the substrate for *pseI*, which could then produce pseudaminic acid (**13**) [181].

Elucidation of the Legionaminic Acid Biosynthetic Pathway

The proposed pathway can be seen in Fig. (**9**). It proceeds in two parts, the first part produces the Guanosine diphosphate (GDP) sugar and the second produces the final CMP-linked legionaminic acid (**16)**.

Whilst studying O-linked flagellar glycosylation in *Campylobacter*, Logan and co-workers investigated *C. coli* VC167 [182], which lacks a full set of *pgl* genes but contains another set of genes *ptmA-F*, which they believed were responsible for the production of pseudaminic acid (**13**) in this organism. Most other *Campylobacter* have this gene set, although not the sub-species *C. jejuni* 81-76. The homologues in *C. jejuni* NTC11168 are, *ptmC* (Cj1327), *ptmD* (Cj1328), *ptmE* (Cj1329), *ptmF* (Cj1330), *ptmB* (Cj1331) and *ptmA* (Cj1332).

Comparisons of flagellins from both *C. coli* VC167 and *C. jejuni* 81-76 showed that both contained the acetamidino form of pseudaminic acid (**15**). However, tandem MS of peptides from *C. coli* identified four carbohydrate modifications with neutral losses of 316 m/z, 317 m/z, 432 m/z and 433 m/z. Fragmentation data showed that the profile for the 317 m/z was consistent with that for pseudaminic acid (**13**). The spectrum for the 316 m/z, whilst expected to be the acetamidino form of pseudaminic acid (**15**), turned out not to be as it had a different fragmentation profile and remained unidentified. The profiles of both the 432/433 m/z moieties showed a loss of 116 m/z, which could be accounted for by the addition of a deoxypentose to the 316/317 m/z. Apart from the 116 m/z ion the 432 spectrum resembled that for the unidentified 316 m/z and the 433 resembled that of the 317 m/z (pseudaminic acid) (**13**). Mutations of the *ptm* genes in *C. coli* VC167 led to tryptic peptides that had previously carried the 316 m/z and 432 m/z being replaced with 317 m/z and 433 m/z and a resultant loss of antigenicity of the *C. coli* flagellin. Thus, there appeared to be two distinct pseudaminic acid (**13**) structures on the flagellin of the two strains [182].

Building on previous work, [180] McNally carried out a targeted metabolomic analysis of *C. coli* VC167, using NMR in conjunction with top down analysis and HILIC MS on the flagellin and its associated tryptic fragments [54]. This work lead to the identification of two structurally distinct CMP-linked carbohydrate derivatives of legionaminic acid (Leg5Ac7Ac) (**16**), an acetamidino (Leg5Am7Ac) (**17**) and *N*-methylacetimidoyl (Leg5AmNMe7Ac) (**18**) form, shown in Fig. (**10**). Thus, the previously unidentified 316 m/z was, in fact, legionaminic acid (**16**). Further work with knockout mutants identified that the *ptm* genes were thus responsible for the production of legionaminic acid (**16**) and its CMP-activated derivatives and that this is a completely separate pathway to that for pseudaminic acid (**13**). McNally also identified a further two genes involved in the pathway *ptmG&H* and with his dataset made the first attempt to predict a full biosynthetic pathway for legionaminic acid [54].

However, the entire legionaminic pathway in *C. jejuni* has only recently been elucidated by Schoenhofen *et al.* [183] using a systems biology approach involving bioinformatics, comparative genomics, metabolomics and functional assays. This considerable and impressive piece of work has identified eleven enzymes involved in the production of CMP-activated legionaminic acid. Interestingly unlike the di-*N*-acetyl bacillosamine (**8**) and pseudaminic (**13**) pathways, this pathway acts through GDP as opposed to UDP intermediates; this may provide a mechanism for the pathways to stay segregated.

N-linked Glycosylation: Di-*N*-Acetyl Bacillosamine Biosynthetic Pathway

The proposed mechanism for N-linked glycosylation in *C. jejuni* is shown in Fig. (**11**).

Examination of the full genome sequence of *C. jejuni* NCTC 11168 showed that there are effectively two glycosylation loci. The first being the flagellin biosynthesis and modification locus containing about 50 genes in the region Cj1293-Cj1344 encompassing the gene loci for the pseudaminic acid (**13**) and legionaminic acid (**16**) biosynthesis pathways [168, 169, 174].

The second region from Cj1119-Cj1152 encoded for lipo-oligosaccharide and general glycosylation, including flagellin. Contained within this region is the *pgl/wla* gene locus Cj1119-Cj1131, which is highly conserved in *Campylobacter* and does not contain phase variable genes with homopolymeric sequences [12, 98, 100, 168].

When investigating this gene locus Linton *et al.* [184] identified two glycosylated proteins from *C. jejuni* NCTC 11168. Lectin capture of the proteins using Soya Bean Agglutinin (SBA) demonstrated that the proteins contained terminal GalNAc (**6**) residues, treatment with α-*N*-acetylgalactosamidase removed their ability to bind to SBA, demonstrating that the terminal GalNAc (**6**) residues were attached *via* an α(1→3) linkage. Disruption of the *C. jejuni* NCTC 11168 pgl genes: *pglH, pglI, pglJ, pglA, pglD and pglE* resulted in the loss of binding of these proteins to SBA, thus demonstrating that the *pgl* locus is responsible for the observed glycosylation. After lectin capture the glycoproteins were separated on an SDS gel, bands were extracted digested and subjected to MALDI-TOF analyses with peptide mass fingerprinting and identified as PEB3 (Cj0289c), a highly immunogenic protein and a putative periplasmic protein (Cj1670c). Using computer based predictions Linton *et al.* suggested that the gene products of *pglH, pglI, pglJ, pglA,* and *pglC* acted as putative glycosyltransferases; *pglB, wlaJ and pglG* were integral membrane proteins; *pglD, pglE* and *pglF* were involved in sugar biosynthesis; *walB* encoded an ABC transporter and *galE*, which encodes a UDP-glucose 4-epimerase (conversion of UDP-glucose to UDP-galactose) [166] are all involved in this protein glycosylation process [184].

Young *et al.* [85] carried out an extensive analysis of the *C. jejuni* NCTC 11168 glycoprotein complement identifying more than thirty potential glycoproteins. They specifically

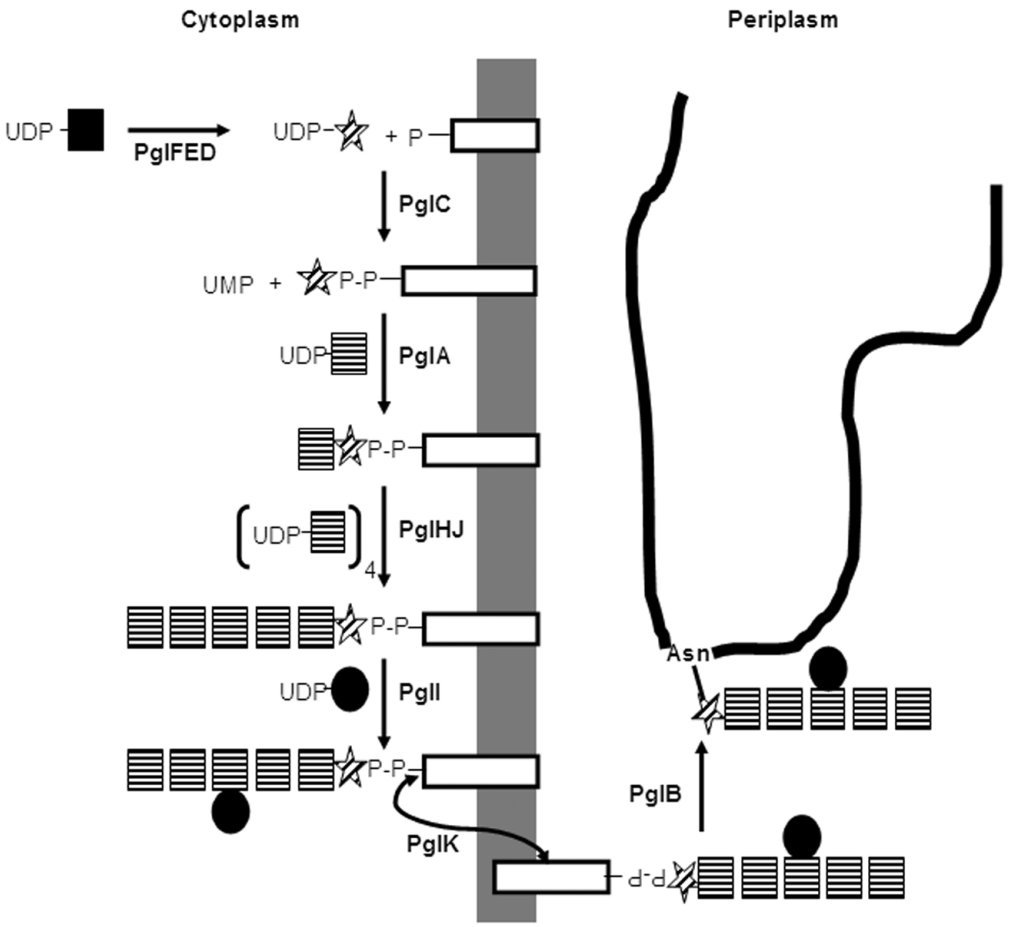

Fig. (11). N-linked glycosylation in *C. jejuni*: The process is as follows: UDP-Bac is produced by the sequential action of PglF, a dehydratase, PglE, an aminotransferase and PglD, an acetyl transferase, on UDP-GlcNAc. The UDP-Bac is transferred to a lipid carrier, undecaprenyl phosphate, attached to the cytoplasmic face of the cell membrane, by the action of PglC. PglA adds the first GalNAc (**6**) residue to the Bac, the next residue is added by PglH, PglJ until all four residues have been added. PglI attaches the glucose (**2**) residue to the side chain of the third GalNAc (**6**), completing the heptasaccharide chain. PglK flips this entire chain across the inner membrane to the periplasmic space where pglB attaches it to the N-glycosylation sites on proteins. UDP Uridine diphosphate, ■ UDP-GlcNAc, P-▭ undecaprenyl phosphate, ● Glucose (**2**), ▤ GalNAc (**6**), ✡ diacetyl Bacillosamine (**8**).

isolated and analyzed the PEB3 (Cj0289c) protein. MS analysis of the PEB3 showed an expected MS peak at 25,454 Da and an unexplained peak at 26,861 Da. The fraction containing the 26,861 Da peak was tryptically digested and all peptides could be assigned to PEB3 bar one. This tryptic peptide was identified as DFNVSK and tandem MS confirmed that it was a glycopeptide; the mass of the oligosaccharide was 1406 Da with a fragmentation profile equivalent to the loss of 5 x HexNAc (203 Da), a single Hex (162 Da) and an unidentified 228 Da moiety. To determine whether the oligosaccharide was O-linked, β-elimination was attempted, however, this failed to remove the oligosaccharide from the peptide.

Further tandem MS and examination of the b_3 +228 ion fragment showed that the glycan structure was attached to the peptide backbone *via* the 228 moiety being bound to an Asn. This was thus the first instance of an N-linked glycosylation event in *C. jejuni*, further evidenced by the fact that the Asn to which the sugar was attached was within the eukaryotic N-linked consensus sequence Asn-X-Ser. Isolation of other glycoproteins in this study using SBA revealed a number of other N-linked glycoproteins with the same glycan structure attachment and linkage to the eukaryotic sequon. NMR studies elucidated the heptasaccharide structure of the glycan to be GalNAc(**6**)-α(1→4)-GalNAc(**6**)-α(1→4)–[Glc(**2**)-β-(1→3)]GalNAc(**6**)-α(1→4)-GalNAc (**6**)-α(1→4)-GalNAc (**6**)-α(1→3)–Bac (**8**)-β–*N*-Asn (see Fig. (**12**)). Bac is 2,4-diacetamido-2,4,6-trideoxyglucopyranose, a di-*N*-acetyl derivative of bacillosamine (**8**) an unusual carbohydrate specific to certain bacteria [85].

Further investigations of this gene set demonstrated that *pglB* has high sequence homology to yeast STT3, which has been shown to be an essential subunit of the N-linked oligosaccharyltransferase complex. The PglB protein also contains the highly conserved WWDYGY motif, which is found in a number of STT3 oligosaccharyltransferase homologues, this is located on the hydrophilic C-terminal of the protein orientated towards the periplasmic space [185]. Wacker *et al.* [185] demonstrated that the *pglB* gene product functions as an oligosaccharide transferase and is responsible for the N-

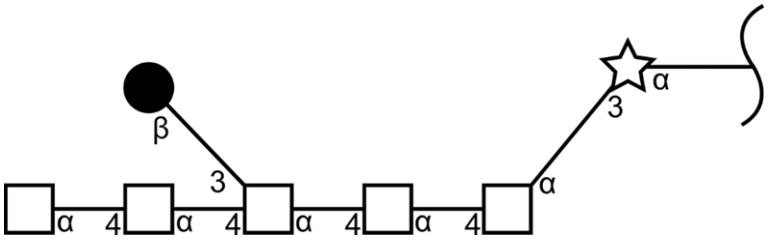

Fig. (12). The structure of the heptasaccharide chain involved in *C. jejuni* N-linked glycosylation. ● Glucose (**2**), □ GalNAc (**6**), ☆ diacetyl Bacillosamine (**8**).

linkage of the heptasaccharide sugar to the polypeptide backbone of the glycoprotein. Mutants without this gene produced proteins without immunoreactivity or the ability to bind SBA. When PEB3 was purified from this mutant MS analysis showed that it completely lacked the glycan attachment. It was, therefore, postulated that in *C. jejuni* the heptasaccharide is assembled in the cytoplasm on a lipid carrier, probably a bactoprenyl pyrophosphate, translocated to the periplamic side of the membrane and then transferred to the protein by PglB [185].

Mutation of the *pgl* genes was also shown to have no effect on lipo-oligosaccharide or capsular oligosaccharide proteins demonstrating that this gene locus is not involved in these pathways but in the glycosylation of glycoproteins [185].

All the proteins identified in this study containing the Bac (**8**) heptasaccharide were classified as periplasmic, suggesting that the periplasmic space may function in a similar manner to the endoplasmic reticulum in eukaryotes [185]. In fact, all N-linked proteins identified to date contain predicted signal peptides and would, therefore, have to transverse the periplasmic space to move extracellular [186].

This study showed that multiple glycoproteins contain the same glycan, demonstrating that several biological functions could be affected by alteration of the *pgl* gene locus [184]. In fact, previous studies on the *pgl* gene locus had already shown it to play a role in immunogenicity, with deletion of genes within this set leading to a reduction in the ability of *C. jejuni* to adhere to and invade cells *in vitro* and to colonize the intestinal tract of mice [11, 12].

Further elucidation of the mechanism of N-linked glycosylation was provided by the introduction of the *C. jejuni pgl* locus into *E. coli* with concomitant knocking out of each of the functional genes. The effect of the knock out of each of these genes was studied by tandem MS analysis of the glycosylation of PEB3 glycopeptides [187] and provided fundamental insights. Inactivation of *pglB* and *galE* resulted in peptide fragmentation patterns lacking any glycan structures. *pglB* has thus a role in the protein glycosylation reaction, presumably as an oligosaccharyl transferase. Given the structure of the heptasaccharide it is likely that *galE* produces a protein that is involved in the biosynthesis of the GalNAc (**6**) sugars [187]. Inactivation of *pglI* led to the production of glycopeptides with fragmentation patterns consistent with a hexasaccharide sequence and when compared to the wild type heptasaccharide tandem MS profile was shown to be missing a single hexose residue. It, therefore, seems likely that *pglI* encodes a $\beta(1\rightarrow3)$ glucotransferase that links the side chain glucose to the glycan chain once the hexasaccharide core has already been formed.

Inactivation of *pglA* resulted in a glycopeptide with a tandem MS profile consistent with one Bac residue attached to the peptide suggesting that it produces an $\alpha(1\rightarrow3)$-*N*-acetylgalactosaminyltransferase, attaching the first GalNAc (**6**) to the Bac (**8**) sugar. Similarly, inactivation of *pglJ* led to tandem MS fragments consistent with the GalNAc (**6**)-Bac (**8**) disaccharide being present indicating its role as a $\alpha(1\rightarrow4)$-*N*-acetylgalactosaminyltransferase in the addition of the second GalNAc (**6**) to the heptasaccharide glycan. Likewise, *pglH* is responsible for the addition of the third GalNAc (**6**) and *pglJ* for the fourth; it is currently unsure how the fifth and sixth are added but it could be carried out by further additions from *pglJ/H*.

The final putative transferase *pglC* is thought to be involved in the transfer of the first sugar residue onto a lipid carrier (undecaprenyl pyrophosphate), there is some evidence based on its sequence homology to a *Salmonella* protein that carries out a similar function. However, direct evidence for its role came from Glover *et al.*, who demonstrated that PglC transfers UDP-Bac directly to undecaprenyl phosphate lipid to form Und-PP-Bac [188]. G*alD, E&F* were also suggested to be involved in Bac (**8**) biosynthesis [187].

Studies on the characterization of this pathway *in vivo* were performed by the use of high-resolution magic angle spinning nuclear magnetic resonance (HR-MAS NMR), with MS analysis being an integral part of the confirmatory process for the mechanisms of this pathway. Interestingly, after knockout of the *wlaB* gene, there was a loss of N-linked glycan detection in both NMR & MS and also removal of SBA reactivity and loss of colonization ability of the *C. jejuni*. This demonstrated that the ABC transporter WlaB (renamed *pglK*) is essential in N-linked glycosylation [189].

A number of studies, using MS as core technology, identified the enzymatic activities of PglF [177, 178], PglE [167] and PglD [92]. The biosynthetic pathway was finally fully characterized by Oliver *et al*. [92]. Astonishingly, this group was able to produce the full heptasaccharide in 'one shot' by mixing the various proteins from the *pgl* gene loci with co-factors in two test tubes [92].

ROLE OF GLYCOSYLATION IN *Mycobacterium tuberculosis*

Infections with *Mycobacterium tuberculosis (Mtb)* cause tuberculosis, one of the most widespread and deadliest infectious diseases worldwide. Once thought to be defeated, a resurgence of *Mtb* started to occur in the 80s due to the de-

velopment of multidrug resistant (MDR) and extensively drug resistance (XDR) strains and due to HIV co-infection [190]. Although tuberculosis primarily attacks the lungs, other organs such as the central nervous system, the gastrointestinal system etc. can also be affected [191]. A number of surface proteins of *Mtb* are involved in cell-cell recognition and thus mediate immunological responses [192-194].

Particularly, host-pathogen interactions of mycobacterial surface molecules with the mannose receptor (MR) on the host's macrophages [195-198], the dendritic cell-specific ICAM-3 grabbing nonintegrin (DC-SIGN) on dendritic cells and monocyte-derived macrophages [199, 200] and Dectin-1 on macrophages [201] have been documented. Evidently, carbohydrates play a pivotal role in these interactions and thus in the mycobacterial infestation of the hosts. Since *Mtb* is phagocytosed but not digested, mycobacterial intracellular survival seems to be dependent on the bacterium's strategy to escape immune responses in macrophages.

Next to the abundant cell wall lipoglycans (lipoarabinomannans (LAMs), lipomannans (LMs) and phosphatidylinositol mannosides (PIMs)) glycoproteins emerge as virulence factors. In fact, glycosylated *Mtb* proteins were already pursued in the 80s for their antigeneic properties [202, 203]. However, until the 90s, most glycoproteins in *Mtb* were identified solely on their ability to bind the lectin ConA [204-205]. In these initial studies, a 19kDa glycoprotein [205], a 45/47 glycoprotein [206] and a 38 KDa and 50/55 kDa glycoprotein [204] were detected. The 45/47 glycoprotein is homologous to 50/55 kDa glycoprotein [206]. Due to the ubiquitous presence of cell wall lipoglycans (LAMs, LMs and PIMs) that may easily contaminate proteins and thus lead to false positive glycoprotein identifications, a real breakthrough was achieved, when glycosyl units were found to be covalently bound to peptides by MS [207]. Four threonine residues in 45/47 kDa glycoprotein Apa (Rv1860) were found to be linked to mannose (3), mannobiose and manotriose [207]. Horn showed later that, in fact, up to 9 mannose residues can be attached to the threonines of Apa [208].

Through site-directed mutagenesis and ConA binding, two threonines were shown to be glycosylated in the 19 kDa antigen LpqH lipoprotein and this glycosylation protected the proteins from proteolysis [209]. Our MS data suggest further glycosylation sites in this region (Bell *et al.* unpublished results). LpqH has been described as adhesin, binding to mannose receptors and thus inducing phagocytosis of monocytic cells [195]. In macrophages, LpqH induces a TLR-2 dependent bactericidal response [210], but at the same time cytokine production [211] antigen-processing and MHC II expression is reduced [211-213], thus successfully evading the immune response of its host. The B-cell antigen Superoxide dismutase C, SodC (Rv0432), is another putative lipoprotein that was recently shown to be glycosylated by MS. Of the six glycosylation sites, three were found at threonines (T 45, T46 and T51) and three were found at serines (S48, S53 and S58) [214].

Secreted glycosylated and nonglycosylated proteins of *Mtb* were also characterized in a top-down fashion using ECD after ConA column enrichment of cell culture filtrate [215]. Of the 689 components that were characterized by mass spectrometry, ten proteins were further characterized. Consecutive losses of 162 Da indicative of hexoses were shown for a fragment of the ESAT-6 like protein esxB (CFP-10) and other unnamed proteins [215]. Together with ESAT-6, CFP-10 forms a heterodimer that induces a strong T-cell mediated immune response [216]. The role that glycosylation plays in this process has not yet been explored.

Espitia *et al.* [204] were the first to use bottom-up glycoproteomics to globally characterize glycoproteins in cell culture filtrates [217]. Using ConA lectin affinity capture, followed by 2D-GE combined with LC-MS/MS, 41 putative mannosylated proteins were identified, with many belonging to the putative lipoproteins Lpp. While a number of prediction tools such as SignalP predictions (http://www.cbs.dtu.dk/services/SignalP), NetOglyc predictions (http://www.cbs.dtu.dk/services/NetOGlyc) and LipoP algorithms (http://www.cbs.dtu.dk/services/LipoP) were used in this study, no attempt was made to characterize individual glycosylation sites by MS.

All of the glycoproteins characterized in *Mtb* so far are O-glycosylation with mannose as the most prominent sugar. It is generally believed that *Mtb* does not have N-glycosylated proteins since it is lacking the analogon to PglA. *Mtb*, however, possesses a homolog to PglB, namely Rv1505c, a currently uncharacterized protein [218]. PglB is the protein that attaches the activated sugar to the consensus protein sequence site in the periplasm. It may therefore be hypothesized that the homolog Rv1505c could have the same function in *Mtb*. In addition, there is also a homolog to *PglE,* namely Rv1504c [218]. In both, Neisseria sp. and *Mtb*, the genes for these proteins are located adjacent to each other on Chromosome 1. Whether or not N-glycosylated proteins exist in *Mtb* remains to be determined.

CONCLUSIONS

Much progress has been made when comparing the earlier dogma that only eukaryotes possess glycoproteins to the plethora of O- and N-glycosylated proteins now known to exist in prokaryotes. Due to the advances in MS technology in terms of ionization efficiency and sensitivity, MS has often been an enabling part showing definite proof of glycosylation on the molecular level. At the same time, more research is necessary to elucidate the synthesis, structure and complex interactions of bacterial glycoproteins with host cells. A better understanding of their role in host-pathogen interactions will ultimately lead to better cures.

ACKNOWLEDGEMENTS

We thank the Beckman Institute and the Gordon and Betty Moore Foundation for financial support.

ABBREVIATIONS

ESI-MS = Electrospray Ionization Mass Spectrometry

MALDI- MS = Matrix-Assisted Laser Desorption Ionization Mass Spectrometry

HPLC = High Performance Liquid Chromatography

CID = Collisional Induced Dissociation

CAD = Collisional Activated Dissociation

IRMPD = Infrared Multiphoton Dissociation

ECD = Electron Capture Dissociation

ETD = Electron Transfer Dissociation

GlcNAc = N-Acetyl Glucosamine

GalNAc = N-Acetyl-Galactosamine

ConA = Concanavalin A

WGA = Wheat Germ Agglutinin

GCC = Graphite Carbon Chromatography Columns

HILIC = Hydrophilic Interaction Chromatography

RP = Reversed Phase

PNGase F = Peptide-*N*-Glycosidase-F

$NaBH_4$ = Sodium Borohydride

$NaBH_3CN$ = Cyanoborohydride

Neu5Ac = *N*-Acetyl Neuraminic Acid

CMP = Cytidine Monophosphate

Pse5Ac7Ac = Pseudaminic Acid

Pse5Am7Ac = 5-Acetamidino-7-Acetamido-Pseudaminic Acid

Pse5Pr7Pr = *N*-2,3-Dihyrdroxypropionyl Pseudaminic Acid

UDP = Uridine Diphosphate

AltNAc = N-Acetyl-Altrosamine

GDP = Guanosine Diphosphate

Leg5Ac7Ac = Legionaminic Acid

Leg5Am7Ac = Acetamidino Derivative of Legionaminic Acid

Leg5Am-NMe7Ac = *N*-Methylacetimidoyl Derivative of Legionaminic Acid

Bac = Di-*N*-Acetyl Bacillosamine

SDS = Sodium Dodecyl Sulphate

Und-PP = Undecaprenyl Pyrophosphate

HR-MAS NMR = High-Resolution Magic Angle Spinning nuclear Magnetic Resonance

SBA = Soyabean Agglutinin

MDR = Multidrug Resistant

LAMs = Lipoarabinomannans

LMs = Lipomannans

PIMs = Phosphatidylinositol Mannosides

REFERENCES

[1] Reuter, G. and Gabius, H.J. Eukaryotic glycosylation: Whim of nature or multipurpose tool? *Mol. Cell Prot.*, **1999**, *55*, 368-422.

[2] Hitchen, P.G. and Dell, A. Bacterial glycoproteomics. *Microbiol.*, **2006**, *152*, 1575-1580.

[3] Heiskanen, A.; Hirvonen, T.; Salo, H.; Impola, U.; Olonen, A.; Laitinen, A.; Tiitinen, S.; Natunen, S.; Aitio, O.; Miller-Podraza, H.; Wuhrer, M.; Deelder, A.M.; Natunen, J.; Laine, J.; Lehenkari, P.; Saarinen, J.; Satomaa, T. and Valmu, L. Glycomics of bone marrow-derived mesenchymal stem cells can be used to evaluate their cellular differentiation stage. *Glycoconj. J.*, **2009**, *26*, 367-384.

[4] Hellerqvist, C.G. Linkage analysis using lindberg method. *Meth. Enzymol.*, **1990**, *193*, 554-573.

[5] Laine, R.A. A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05x10(12) structures for a reducing hexasaccharide - the isomer-barrier to development of single-method saccharide sequencing or synthesis systems. *Glycobiology*, **1994**, *4*, 759-767.

[6] Aoki-Kinoshita, K.F. An introduction to bioinformatics for glycomics research. *PLoS Comput. Biol.*, **2008**, *4*, 1-7.

[7] Reinders, J.; Lewandrowski, U.; Jan, M.; Yvonne, W. and Albert, S. Challenges in mass spectrometry-based proteomics. *Proteomics*, **2004**, *4*, 3686-3703.

[8] Apweiler, R.; Hermjakob, H. and Sharon, N. On the frequency of protein glycosylation, as deduced from analysis of the swiss-prot database. *Biochim. Biophys. Acta -General Subjects*, **1999**, *1473*, 4-8.

[9] Lanctot, P.M.; Gage, F.H. and Varki, A.P. The glycans of stem cells. *Curr. Opin. Chem. Biol.*, **2007**, *11*, 373-380.

[10] Rudd, P.M.; Elliott, T.; Cresswell, P.; Wilson, I.A. and Dwek, R.A. Glycosylation and the immune system. *Science*, **2001**, *291*, 2370-2376.

[11] Szymanski, C.M.; Burr, D.H. and Guerry, P. Campylobacter protein glycosylation affects host cell interactions. *Infect. Immun.*, **2002**, *70*, 2242-2244.

[12] Szymanski, C.M.; Yao, R.J.; Ewing, C.P.; Trust, T.J. and Guerry, P. Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Mol. Microbiol.*, **1999**, *32*, 1022-1030.

[13] Guerry, P.; Ewing, C.P.; Schirm, M.; Lorenzo, M.; Kelly, J.; Pattarini, D.; Majam, G.; Thibault, P. and Logan, S. Changes in flagellin glycosylation affect *Campylobacter* autoagglutination and virulence. *Mol. Microbiol.*, **2006**, *60*, 299-311.

[14] Benz, I. and Schmidt, M.A. Never say never again: Protein glycosylation in pathogenic bacteria. *Mol. Microbiol.*, **2002**, *45*, 267-276.

[15] Hirabayashi, J.; Arata, Y. and Kasai, K. Glycome project: Concept, strategy and preliminary application to *Caenorhabditis elegans*. *Proteomics*, **2001**, *1*, 295-303.

[16] North, S.J.; Hitchen, P.G.; Haslam, S.M. and Dell, A. Mass spectrometry in the analysis of n-linked and o-linked glycans. *Curr. Opin. Struct. Biol.*, **2009**, *19*, 1-9.

[17] Morris, H.R.; Thompson, M.R.; Osuga, D.T.; Ahmed, A.I.; Chan, S.M.; Vandenheede, J.R. and Feeney, R.E. Antifreeze glycoproteins from blood of an antarctic fish - structure of proline-containing glycopeptides. *J. Biol. Chem.*, **1978**, *253*, 5155-5162.

[18] Tissot, B.; North, S.J.; Ceroni, A.; Pang, P.-C.; Panico, M.; Rosati, F.; Capone, A.; Haslam, S.M.; Dell, A. and Morris, H.R. Glycoproteomics: Past, present and future. *FEBS Lett.*, **2009**, *583*, 1728.

[19] Devakumar, A.; Mechref, Y.; Kang, P.; Novotny, M.V. and Reilly, J.P. Identification of isomeric N-glycan structures by mass spectrometry with 157 nm laser-induced photofragmentation. *J. Am. Soc. Mass Spectrom.*, **2008**, *19*, 1027-1040.

[20] Bonner, P.L.R.; Lill, J.R.; Hill, S.; Creaser, C.S. and Rees, R.C. Electrospray mass spectrometry for the identification of MHC class I-associated peptides expressed on cancer cells. *J. Immunol. Methods*, **2002**, *262*, 5-19.

[21] Domon, B. and Aebersold, R. Mass spectrometry and protein analysis. *Science*, **2006**, *312*, 212-217.

[22] Aebersold, R. and Mann, M. Mass spectrometry-based proteomics. *Nature*, **2003**, *422*, 198-207.

[23] Graham, R.; Graham, C. and McMullan, G. Microbial proteomics: A mass spectrometry primer for biologists. *Microb. Cell Fact.*, **2007**, *6*, 26.

[24] Liu, T.; Belov, M.E.; Jaitly, N.; Qian, W.J. and Smith, R.D. Accurate mass measurements in proteomics. *Chem. Rev.*, **2007**, *107*, 3621-3653.

[25] Smith, R.D. Future directions for electrospray ionization for biological analysis using mass spectrometry. *BioTechniques*, **2006**, *41*, 147-148.

[26] Yates, J.R. Mass spectral analysis in proteomics. *1056-8700*, **2004**, *33*, 297-316.
[27] Han, X.M.; Aslanian, A. and Yates, J.R. Mass spectrometry for proteomics. *Curr. Opin. Chem. Biol.*, **2008**, *12*, 483-490.
[28] Dell, A. and Morris, H.R. Glycoprotein structure determination mass spectrometry. *Science*, **2001**, *291*, 2351-2356.
[29] Poland, G.A.; Ovsyannikova, I.G.; Johnson, K.L. and Naylor, S. The role of mass spectrometry in vaccine development. *Vaccine*, **2001**, *19*, 2692-2700.
[30] Fenn, J.B.; Mann, M.; Meng, C.K.; Wong, S.F. and Whitehouse, C.M. Electrospray ionization for mass spectrometry of large biomolecules. *Science*, **1989**, *246*, 64-71.
[31] Mann, M.; Hendrickson, R.C. and Pandey, A. Analysis of proteins and proteomes by mass spectrometry. *Annu. Rev. Biochem.*, **2001**, *70*, 437-473.
[32] Manisali, I.; Chen, D.D.Y. and Schneider, B.B. Electrospray ionization source geometry for mass spectrometry: Past, present, and future. *Trends Analyt. Chem.*, **2006**, *25*, 243-256.
[33] Cole, R.B. Some tenets pertaining to electrospray ionization mass spectrometry. *J. Mass Spectrom.*, **2000**, *35*, 763-772.
[34] Griffiths, W.J.; Jonsson, A.P.; Liu, S.; Rai, D.K. and Wang, Y. Electrospray and tandem mass spectrometry in biochemistry. *Biochem. J.*, **2001**, *355*, 545-561.
[35] Dole, M.; Mack, L.L. and Hines, R.L. Molecular beams of macroions. *J. Chem. Phys.*, **1968**, *49*, 2240-&.
[36] Iribarne, J.V. and Thomson, B.A. Evaporation of small ions from charged droplets. *J. Chem. Phys.*, **1976**, *64*, 2287-2294.
[37] Desiderio, D.M. Mass spectrometry: Clinical and biomedical applications; Plenum Press: New York, 1992.
[38] Zaia, J. Mass spectrometry and the emerging field of glycomics. *Chem. Biol.*, **2008**, *15*, 881-892.
[39] Harvey, D.J. Identification of protein-bound carbohydrates by mass spectrometry. *Proteomics*, **2001**, *1*, 311-328.
[40] Karas, M. and Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10000 daltons. *Anal. Chem.*, **1988**, *60*, 2299-2301.
[41] Hillenkamp, F.; Karas, M.; Beavis, R.C. and Chait, B.T. Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Anal. Chem.*, **1991**, *63*, 1193A-1203A.
[42] Karas, M.; Glückmann, M. and Schäfer, J. Ionization in matrix-assisted laser desorption/ionization: Singly charged molecular ions are the lucky survivors. *J. Mass Spectrom.*, **2000**, *35*, 1-12.
[43] Kicman, A.T.; Parkin, M.C. and Iles, R.K. An introduction to mass spectrometry based proteomics--detection and characterization of gonadotropins and related molecules. *Mol. Cell. Endocrinol.*, **2007**, *260-262*, 212-227.
[44] Harvey, D.J. Analysis of carbohydrates and glycoconjugates by matrix-assisted laser desorption/ionization mass spectrometry: An update for 2003-2004. *Mass Spectrom. Rev.*, **2009**, *28*, 273-361.
[45] Dreisewerd, K.; Schürenberg, M.; Karas, M. and Hillenkamp, F. Influence of the laser intensity and spot size on the desorption of molecules and ions in matrix-assisted laser desorption/ionization with a uniform beam profile. *Int. J. Mass Spectrom. Ion Process.*, **1995**, *141*, 127-148.
[46] Morelle, W. and Michalski, J.C. Analysis of protein glycosylation by mass spectrometry. *Nat. Protoc.*, **2007**, *2*, 1585-1602.
[47] Loo, J.A.; Quinn, J.P.; Ryu, S.I.; Henry, K.D.; Senko, M.W. and McLafferty, F.W. High-resolution tandem mass-spectrometry of large biomolecules. *Proc. Natl. Acad. Sci. USA*, **1992**, *89*, 286-289.
[48] Ginter, J.M.; Zhou, F. and Johnston, M.V. Generating protein sequence tags by combining cone and conventional collision induced dissociation in a quadrupole time-of-flight mass spectrometer. *J. Am. Soc. Mass Spectrom.*, **2004**, *15*, 1478-1486.
[49] Bogdan, B. and Richard, D.S. Proteomics by fticr mass spectrometry: Top down and bottom up. *Mass Spectrom. Rev.*, **2005**, *24*, 168-200.
[50] Demirev, P.A.; Feldman, A.B.; Kowalski, P. and Lin, J.S. Top-down proteomics for rapid identification of intact microorganisms. *Anal. Chem.*, **2005**, *77*, 7455-7461.
[51] Chamot-Rooke, J.; Rousseau, B.; Lanternier, F.; Mikaty, G.; Mairey, E.; Malosse, C.; Bouchoux, G.; Pelicic, V.; Camoin, L.; Nassif, X. and Dumenil, G. Alternative *Neisseria spp*. Type iv pilin glycosylation with a glyceramido acetamido trideoxyhexose residue. *Proc. Natl. Acad. Sci. USA*, **2007**, *104*, 14783-14788.
[52] Aas, F.E.; Vik, A.; Vedde, J.; Koomey, M. and Egge-Jacobsen, W. Neisseria gonorrhoeae O-linked pilin glycosylation: Functional analyses define both the biosynthetic pathway and glycan structure. *Mol. Microbiol.*, **2007**, *65*, 607-624.
[53] Ge, Y.; ElNaggar, M.; Sze, S.K.; Oh, H.B.; Begley, T.P.; McLafferty, F.W.; Boshoff, H. and Barry, C.E. Top down characterization of secreted proteins from *Mycobacterium tuberculosis* by electron capture dissociation mass spectrometry. *J. Am. Soc. Mass Spectrom.*, **2003**, *14*, 253-261.
[54] McNally, D.J.; Aubry, A.J.; Hui, J.P.M.; Khieu, N.H.; Whitfield, D.; Ewing, C.P.; Guerry, P.; Brisson, J.R.; Logan, S.M. and Soo, E.C. Targeted metabolomics analysis of *Campylobacter coli* VC167 reveals legionaminic acid derivatives as novel flagellar glycans. *J. Biol. Chem.*, **2007**, *282*, 14463-14475.
[55] Lin, D.; Tabb, D.L. and Yates, J.R. Large-scale protein identification using mass spectrometry. *Biochim. Biophys. Acta - Proteins Proteomics*, **2003**, *1646*, 1-10.
[56] Vachet, R.W.; Ray, K.L. and Glish, G.L. Origin of product ions in the ms/ms spectra of peptides in a quadrupole ion trap. *J. Am. Soc. Mass Spectrom.*, **1998**, *9*, 341-344.
[57] Xie, Y.M. and Lebrilla, C.B. Infrared multiphoton dissociation of alkali metal-coordinated oligosaccharides. *Anal. Chem.*, **2003**, *75*, 1590-1598.
[58] Hakansson, K.; Cooper, H.J.; Emmett, M.R.; Costello, C.E.; Marshall, A.G. and Nilsson, C.L. Electron capture dissociation and infrared multiphoton dissociation ms/ms of an N-glycosylated tryptic peptide to yield complementary sequence information. *Anal. Chem.*, **2001**, *73*, 4530-4536.
[59] Pikulski, M.; Hargrove, A.; Shabbir, S.H.; Anslyn, E.V. and Brodbelt, J.S. Sequencing and characterization of oligosaccharides using infrared multiphoton dissociation and boronic acid derivatization in a quadrupole ion trap. *J. Am. Soc. Mass Spectrom.*, **2007**, *18*, 2094-2106.
[60] Adamson, J.T. and Hakansson, K. Electron capture dissociation of oligosaccharides ionized with alkali, alkaline earth, and transition metals. *Anal. Chem.*, **2007**, *79*, 2901-2910.
[61] Perdivara, I.; Petrovich, R.; Alliquant, B.; Deterding, L.J.; Tomer, K.B. and Przybylski, M. Elucidation of O-glycosylation structures of the beta-amyloid precursor protein by liquid chromatography-mass spectrometry using electron transfer dissociation and collision induced dissociation. *J. Proteome Res.*, **2009**, *8*, 631-642.
[62] Ashline, D.; Singh, S.; Hanneman, A. and Reinhold, V. Congruent strategies for carbohydrate sequencing. 1. Mining structural details by msn. *Anal. Chem.*, **2005**, *77*, 6250.
[63] Johnson, R.S.; Martin, S.A.; Biemann, K.; Stults, J.T. and Watson, J.T. Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: Differentiation of leucine and isoleucine. *Anal. Chem.*, **2002**, *59*, 2621-2625.
[64] Roepstorff, P. and Fohlman, J. Letter to the editors. *Biol. Mass Spectrom.*, **1984**, *11*, 601.
[65] Domon, B. and Costello, C.E. A systematic nomenclature for carbohydrate fragmentations in Fab-ms ms spectra of glycoconjugates. *Glycoconj. J.*, **1988**, *5*, 397-409.
[66] Lancaster, K.S.; An, H.J.; Li, B.S. and Lebrilla, C.B. Interrogation of n-linked oligosaccharides using infrared multiphoton dissociation in ft-icr mass spectrometry. *Anal. Chem.*, **2006**, *78*, 4990-4997.
[67] Zhang, J.H.; Schubothe, K.; Li, B.S.; Russell, S. and Lebrilla, C.B. Infrared multiphoton dissociation of O-linked mucin-type oligosaccharides. *Anal. Chem.*, **2005**, *77*, 208-214.
[68] Cooper, H.J.; Hakansson, K. and Marshall, A.G. The role of electron capture dissociation in biomolecular analysis. *Mass Spectrom. Rev.*, **2005**, *24*, 201-222.
[69] Bakhtiar, R. and Guan, Z.Q. Electron capture dissociation mass spectrometry in characterization of peptides and proteins. *Biotechnol. Lett.*, **2006**, *28*, 1047-1059.
[70] Syka, J.E.P.; Coon, J.J.; Schroeder, M.J.; Shabanowitz, J. and Hunt, D.F. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. USA*, **2004**, *101*, 9528-9533.
[71] Coon, J.J.; Syka, J.E.P.; Schwartz, J.C.; Shabanowitz, J. and Hunt, D.F. Anion dependence in the partitioning between proton and electron transfer in ion/ion reactions. *Int. J. Mass Spectrom.*, **2004**, *236*, 33-42.
[72] Morelle, W.; Canis, K.; Chirat, F.; Faid, V. and Jean-Claude, M. The use of mass spectrometry for the proteomic analysis of glycosylation. *Proteomics*, **2006**, *6*, 3993-4015.
[73] Khmelnitsky, Y.L. Current strategies for *in vitro* protein glycosylation. *J. Mol. Catal., B Enzym.*, **2004**, *31*, 73-81.

[74] Harvey, D.J. Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates. *Mass Spectrom. Rev.*, **1999**, *18*, 349-450.

[75] Spiro, R.G. Protein glycosylation: Nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*, **2002**, *12*, 43R-56R.

[76] Morris, H.R.; Chalabi, S.; Panico, M.; Sutton-Smith, M.; Clark, G.F.; Goldberg, D. and Dell, A. Glycoproteomics: Past, present and future. *Int. J. Mass Spectrom.*, **2007**, *259*, 16.

[77] Anumula, K.R. Advances in fluorescence derivatization methods for high-performance liquid chromatographic analysis of glycoprotein carbohydrates. *Anal. Biochem.*, **2006**, *350*, 1-23.

[78] Hofsteenge, J.; Muller, D.R.; Debeer, T.; Loffler, A.; Richter, W.J. and Vliegenthart, J.F.G. New-type of linkage between a carbohydrate and a protein - C-glycosylation of a specific tryptophan residue in human RNnase US. *Biochemistry*, **1994**, *33*, 13524-13530.

[79] de Beer, T.; Vliegenthart, J.F.G.; Loeffler, A. and Hofsteenge, J. The hexopyranosyl residue that is C-glycosidically linked to the side chain of tryptophan-7 in human RNase US is Alpha.-mannopyranose. *Biochemistry*, **2002**, *34*, 11785.

[80] Yurist-Doutsch, S.; Chaban, B.; VanDyke, D.J.; Jarrell, K.F. and Eichler, J. Sweet to the extreme: Protein glycosylation in archaea. *Mol. Microbiol.*, **2008**, *68*, 1079-1084.

[81] Abu-Qarn, M.; Eichler, J. and Sharon, N. Not just for eukarya anymore: Protein glycosylation in bacteria and archaea. *Curr. Opin. Struct. Biol.*, **2008**, *18*, 544.

[82] Mescher, M.F. and Strominger, J.L. Purification and characterization of a prokaryotic glycoprotein from cell-envelope of halobacterium-salinarium. *J. Biol. Chem.*, **1976**, *251*, 2005-2014.

[83] Schaffer, C. and Messner, P. Surface-layer glycoproteins: An example for the diversity of bacterial glycosylation with promising impacts on nanobiotechnology. *Glycobiology*, **2004**, *14*, 31R-42R.

[84] Stimson, E.; Virji, M.; Makepeace, K.; Dell, A.; Morris, H.R.; Payne, G.; Saunders, J.R.; Jennings, M.P.; Barker, S.; Panico, M.; Blench, I. and Moxon, E.R. Meningococcal pilin - a glycoprotein substituted with digalactosyl 2,4-diacetamido-2,4,6-trideoxyhexose. *Mol. Microbiol.*, **1995**, *17*, 1201-1214.

[85] Young, N.M.; Brisson, J.R.; Kelly, J.; Watson, D.C.; Tessier, L.; Lanthier, P.H.; Jarrell, H.C.; Cadotte, N.; Michael, F.S.; Aberg, E. and Szymanski, C.M. Structure of the N-linked glycan present on multiple glycoproteins in the gram-negative bacterium, campylobacter jejuni. *J. Biol.Chem.*, **2002**, *277*, 42530-42539.

[86] Schaffer, C.; Graninger, M. and Messner, P. Prokaryotic glycosylation. *Proteomics*, **2001**, *1*, 248-261.

[87] Wieland, F.; Heitzer, R. and Schaefer, W. Asparaginylglucose - novel type of carbohydrate linkage. *Proc. Natl. Acad. Sci. USA. Biological Sciences*, **1983**, *80*, 5470-5474.

[88] Pellerin, P.; Fournet, B. and Debeire, P. Evidence for the glycoprotein nature of the cell sheath of methanosaeta-like cells in the culture of *Methanothrix-soehngenii* strain Fe. *Can. J. Microbiol.*, **1990**, *36*, 631-636.

[89] Messner, P. and Sleytr, U.B. Asparaginyl-rhamnose - a novel type of protein-carbohydrate linkage in a eubacterial surface-layer glycoprotein. *FEBS Lett.*, **1988**, *228*, 317-320.

[90] Kowarik, M.; Young, N.M.; Numao, S.; Schulz, B.L.; Hug, I.; Callewaert, N.; Mills, D.C.; Watson, D.C.; Hernandez, M.; Kelly, J.F.; Wacker, M. and Aebi, M. Definition of the bacterial N-glycosylation site consensus sequence. *EMBO J.*, **2006**, *25*, 1957-1966.

[91] Chen, M.M.; Glover, K.J. and Imperiali, B. From peptide to protein: Comparative analysis of the substrate specificity of N-linked glycosylation in *C-jejuni*. *Biochemistry*, **2007**, *46*, 5579-5585.

[92] Olivier, N.B.; Chen, M.M.; Behr, J.R. and Imperiali, B. *In vitro* biosynthesis of UDP-N,N'-diacetylbacillosamine by enzymes of the *Campylobacter jejuni* general protein glycosylation system. *Biochemistry*, **2006**, *45*, 13659-13669.

[93] Moens, S. and Vanderleyden, J. Glycoproteins in prokaryotes. *Arch. Microbiol.*, **1997**, *168*, 169-175.

[94] Plummer, T.H.; Tarentino, A.L. and Hauer, C.R. Novel, specific O-glycosylation of secreted *Flavobacterium-meningosepticum* proteins - asp-ser-asterisk and asp-thr-asterisk-thr consensus sites. *J. Biol. Chem.*, **1995**, *270*, 13192-13196.

[95] Peters, J.; Rudolf, S.; Oschkinat, H.; Mengele, R.; Sumper, M.; Kellermann, J.; Lottspeich, F. and Baumeister, W. Evidence for tyrosine-linked glycosaminoglycan in a bacterial surface protein. *Biol. Chem. Hoppe-Seyler*, **1992**, *373*, 171-176.

[96] Schmidt, M.A.; Riley, L.W. and Benz, I. Sweet new world: Glycoproteins in bacterial pathogens. *Trends Microbiol.*, **2003**, *11*, 554-561.

[97] Szymanski, C.M. and Wren, B.W. Protein glycosylation in bacterial mucosal pathogens. *Nat. Rev. Microbiol.*, **2005**, *3*, 225-237.

[98] Linton, D.; Karlyshev, A.V. and Wren, B.W. Deciphering *Campylobacter jejuni* cell surface interactions from the genome sequence. *Curr. Opin. Microbiol.*, **2001**, *4*, 35-40.

[99] Brimer, C.D. and Montie, T.C. Cloning and comparison of flic genes and identification of glycosylation in the flagellin of *Pseudomonas aeruginosa* a-type strains. *J. Bacteriol.*, **1998**, *180*, 3209-3217.

[100] Fry, B.N.; Korolik, V.; ten Brinke, J.A.; Pennings, M.T.T.; Zalm, R.; Teunis, B.J.J.; Coloe, P.J. and van der Zeijst, B.A.M. The lipopolysaccharide biosynthesis locus of *Campylobacter jejuni* 81116. *Microbiol.*, **1998**, *144*, 2049-2061.

[101] Weerapana, E. and Imperiali, B. Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. *Glycobiology*, **2006**, *16*, 91R-101R.

[102] Wells, L.; Vosseller, K.; Cole, R.N.; Cronshaw, J.M.; Matunis, M.J. and Hart, G.W. Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications. *Mol. Cell. Prot.*, **2002**, *1*, 791-804.

[103] Amon, S.; Zamfir, A.D. and Rizzi, A. Glycosylation analysis of glycoproteins and proteoglycans using capillary electrophoresis-mass spectrometry strategies. *Electrophoresis*, **2008**, *29*, 2485-2507.

[104] Ohyama, Y.; Kasai, K.I.; Nomoto, H. and Inoue, Y. Frontal affinity-chromatography of ovalbumin glycoasparagines on a Concanavalin a-sepharose column - a quantitative study of the binding-specificity of the lectin. *J. Biol. Chem.*, **1985**, *260*, 6882-6887.

[105] Qiu, R.Q. and Regnier, F.E. Comparative glycoproteomics of N-linked complex-type glycoforms containing sialic acid in human serum. *Anal. Chem.*, **2005**, *77*, 7225-7231.

[106] Cummings, R.D. Use of lectins in analysis of glycoconjugates. *Guide to Techniques in Glycobiology*, **1994**, *230*, 66-86.

[107] Madera, M.; Mechref, Y. and Novotny, M.V. Combining lectin microcolumns with high-resolution separation techniques for enrichment of glycoproteins and glycopeptides. *Anal. Chem.*, **2005**, *77*, 4081-4090.

[108] Taketa, K. Characterization of sugar chain structures of human alpha-fetoprotein by lectin affinity electrophoresis. *Electrophoresis*, **1998**, *19*, 2595-2602.

[109] Sumner, J.B. The globulins of the jack bean, *Canavalia ensiformis*. *J. Biol. Chem.*, **1919**, *37*, 137-U131.

[110] Goldstein.I.J; Reichert, C.M. and Misaki, A. Interaction of Concanavalin-A with model substrates. *Ann. N. Y. Acad. Sci.*, **1974**, *234*, 283-296.

[111] Yang, Z.P. and Hancock, W.S. Approach to the comprehensive analysis of glycoproteins isolated from human serum using a multilectin affinity column. *J. Chromatogr. A.*, **2004**, *1053*, 79-88.

[112] Aub, J.C.; Lankester, A. and Tieslau, C. Reactions of normal and tumor cell surfaces to enzymes, .1. Wheat-germ lipase and associated mucopolysaccharides. *Proc. Natl. Acad. Sci. USA*, **1963**, *50*, 613-619.

[113] Miller, R.C. and Bowles, D.J. A comparative-study of the localization of wheat-germ-agglutinin and its potential receptors in wheat grains. *Biochem. J.*, **1982**, *206*, 571-576.

[114] Bunnmoreno, M.M. and Camposneto, A. Lectin(s) extracted from seeds of *Artocarpus-integrifolia* (jackfruit) - potent and selective stimulator(s) of distinct human t-cell and b-cell functions. *J. Immunol.*, **1981**, *127*, 427-429.

[115] Roquebarreira, M.C. and Camposneto, A. Jacalin - an IGA-binding lectin. *J. Immunol.*, **1985**, *134*, 1740-1743.

[116] Mahanta, S.K.; Sastry, M.V.K. and Surolia, A. Topography of the combining region of a thomsen-friedenreich-antigen-specific lectin jacalin (*Artocarpus-integrifolia* agglutinin) - a thermodynamic and circular-dichroism spectroscopic study. *Biochem. J.*, **1990**, *265*, 831-840.

[117] Yang, Z.P.; Hancock, W.S.; Chew, T.R. and Bonilla, L. A study of glycoproteins in human serum and plasma reference standards (hupo) using multilectin affinity chromatography coupled with RPLC-ms/ms. *Proteomics*, **2005**, *5*, 3353-3366.

[118] Hirabayashi, J. Oligosaccharide microarrays for glycomics. *Trends Biotechnol.*, **2003**, *21*, 141-143.

[119] Koizumi, K. High-performance liquid chromatographic separation of carbohydrates on graphitized carbon columns. *J. Chromatogr. A.*, **1996**, *720*, 119-126.

[120] Fan, J.Q.; Kondo, A.; Kato, I. and Lee, Y.C. High-performance liquid-chromatography of glycopeptides and oligosaccharides on graphitized carbon columns. *Anal. Biochem.*, **1994**, *219*, 224-229.

[121] Zhang, T.V.R.; Hao, Z. and Zabrouskov, V. Analysis of glycopeptides using porous graphite chromatography and ltq orbitrap xl etd hybrid ms, Thermo Scientific. Application Note 463, **2009**.

[122] Hashii, N.; Kawasaki, N.; Itoh, S.; Hyuga, M.; Kawanishi, T. and Hayakawa, T. Glycomic/glycoproteomic analysis by liquid chromatography/mass spectrometry: Analysis of glycan structure alternation in cells. *Proteomics*, **2005**, *5*, 4665-4672.

[123] Alley, W.R.; Mechref, Y. and Novotny, M.V. Use of activated graphitized carbon chips for liquid chromatography/mass spectrometric and tandem mass spectrometric analysis of tryptic glycopeptides. *Rapid Commun. Mass Spectrom.*, **2009**, *23*, 495-505.

[124] Hemstrom, P. and Irgum, K. Hydrophilic interaction chromatography. *J. Sep. Sci.*, **2006**, *29*, 1784-1821.

[125] Wuhrer, M.; de Boer, A.R. and Deelder, A.M. Structural glycomics using hydrophilic interaction chromatography (HILIC) with mass spectrometry. *Mass Spectrom. Rev.*, **2009**, *28*, 192-206.

[126] Guile, G.R.; Rudd, P.M.; Wing, D.R.; Prime, S.B. and Dwek, R.A. A rapid high-resolution high-performance liquid chromatographic method for separating glycan mixtures and analyzing oligosaccharide profiles. *Anal. Biochem.*, **1996**, *240*, 210-226.

[127] Wuhrer, M.; Koeleman, C.A.M.; Deelder, A.M. and Hokke, C.N. Normal-phase nanoscale liquid chromatography - mass spectrometry of underivatized oligosaccharides at low-femtomole sensitivity. *Anal. Chem.*, **2004**, *76*, 833-838.

[128] Wuhrer, M.; Koeleman, C.A.M.; Hokke, C.H. and Deelder, A.M. Protein glycosylation analyzed by normal-phase nano-liquid chromatography-mass spectrometry of glycopeptides. *Anal. Chem.*, **2005**, *77*, 886-894.

[129] Thaysen-Andersen, M.; Thogersen, I.B.; Nielsen, H.J.; Lademann, U.; Brunner, N.; Enghild, J.J. and Hojrup, P. Rapid and individual-specific glycoprofiling of the low abundance n-glycosylated protein tissue inhibitor of metalloproteinases. *Mol. Cell. Prot.*, **2007**, *6*, 638-647.

[130] Zaia, J. On-line separations combined with ms for analysis of glycosaminoglycans. *Mass Spectrom. Rev.*, **2009**, *28*, 254-272.

[131] Zhang, H.; Li, X.J.; Martin, D.B. and Aebersold, R. Identification and quantification of n-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.*, **2003**, *21*, 660-666.

[132] Burlingame, A.L. Characterization of protein glycosylation by mass spectrometry. *Curr. Opin. Biotechnol.*, **1996**, *7*, 4.

[133] Prime, S.; Dearnley, J.; Ventom, A.M.; Parekh, R.B. and Edge, C.J. Oligosaccharide sequencing based on exo- and endoglycosidase digestion and liquid chromatographic analysis of the products. *J. Chromatogr. A*, **1996**, *720*, 263.

[134] Rusnak, F.; Zhou, J. and Hathaway, G. Identification of phosphorylated and glycosylated sites in peptides by chemically targeted proteolysis. *J. Biomol. Tech.*, **2002**, *13*, 228-237.

[135] Steen, H. and Mann, M. A new derivatization strategy for the analysis of phosphopeptides by precursor ion scanning in positive ion mode. *J. Am. Soc. Mass Spectrom.*, **2003**, *14*, 83-83.

[136] Plummer, T.H.; Elder, J.H.; Alexander, S.; Phelan, A.W. and Tarentino, A.L. Demonstration of Peptide-N-Glycosidase-F activity in endo-beta-n-acetylglucosaminidase F preparations. *J. Biol. Chem.*, **1984**, *259*, 700-704.

[137] Tarentino, A.L.; Gomez, C.M. and Plummer, T.H. Deglycosylation of asparagine-linked glycans by Peptide - N-Glycosidase-F. *Biochemistry*, **1985**, *24*, 4665-4671.

[138] Geiger, T. and Clarke, S. Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides - succinimide-linked reactions that contribute to protein-degradation. *J. Biol. Chem.*, **1987**, *262*, 785-794.

[139] Catak, S.; Monard, G.; Aviyente, V. and Ruiz-Lopez, M.F. Deamidation of asparagine residues: Direct hydrolysis versus succinimide-mediated deamidation mechanisms. *J. Phys. Chem. A.*, **2009**, *113*, 1111-1120.

[140] Tretter, V.; Altmann, F. and Marz, L. Peptide-N4-(n-acetyl-beta-glucosaminyl) asparagine amidase-F cannot release glycans with fucose attached alpha-1-]3 to the asparagine-linked N-acetylglucosamine residue. *Eur. J. Biochem.*, **1991**, *199*, 647-652.

[141] Liu, X.; McNally, D.J.; Nothaft, H.; Szymanski, C.M.; Brisson, J.-R. and Li, J. Mass spectrometry-based glycomics strategy for exploring N-linked glycosylation in eukaryotes and bacteria. *Anal. Chem.*, **2006**, *78*, 6081.

[142] Takasaki, S.; Mizuochi, T. and Kobata, A. Hydrazinolysis of asparagine-linked sugar chains to produce free oligosaccharides. *Meth. Enzymol.*, **1982**, *83*, 263-268.

[143] Patel, T.; Bruce, J.; Merry, A.; Bigge, C.; Wormald, M.; Jaques, A. and Parekh, R. Use of hydrazine to release in intact and unreduced form both N-linked and O-linked oligosaccharides from glycoproteins. *Biochemistry*, **1993**, *32*, 679-693.

[144] Greis, K.D.; Hayes, B.K.; Comer, F.I.; Kirk, M.; Barnes, S.; Lowary, T.L. and Hart, G.W. Selective detection and site-analysis of O-GlcNAc-modified glycopeptides by beta-elimination and tandem electrospray mass spectrometry. *Anal. Biochem.*, **1996**, *234*, 38-49.

[145] Roberts, J.C. The chemistry of paper, ed.; the Royal Society of Chemistry: Cambridge, 1996.

[146] Cooper, C.A.; Packer, N.H. and Redmond, J.W. The elimination of O-linked glycans from glycoproteins under nonreducing conditions. *Glycoconj. J.*, **1994**, *11*, 163-167.

[147] Kang, P.; Mechref, Y.; Kyselova, Z.; Goetz, J.A. and Novotny, M.V. Comparative glycomic mapping through quantitative permethylation and stable-isotope labeling. *Anal. Chem.*, **2007**, *79*, 6064-6073.

[148] Alvarez-Manilla, G.; Warren, N.L.; Abney, T.; Atwood, J.; Azadi, P.; York, W.S.; Pierce, M. and Orlando, R. Tools for glycomics: Relative quantitation of glycans by isotopic permethylation using (CH3I)-C-13. *Glycobiology*, **2007**, *17*, 677-687.

[149] Price, N.P.J. Permethylation linkage analysis techniques for residual carbohydrates. *Appl. Biochem. Biotechnol.*, **2008**, *148*, 271-276.

[150] Hakomori, S.I. Rapid permethylation of glycolipid and polysaccharide catalyzed by methylsulfinyl carbanion in dimethyl sulfoxide. *J. Biochem.*, **1964**, *55*, 205-208.

[151] Large, D.G., Warren, C. D. Glycopeptides and related compounds: Synthesis, analysis and applications; Marcel Dekker, INC.: New York, 1997.

[152] Ciucanu, I. and Kerek, F. A simple and rapid method for the permethylation of carbohydrates. *Carbohydr. Res.*, **1984**, *131*, 209-217.

[153] Atwood, J.A.; Cheng, L.; Alvarez-Manilla, G.; Warren, N.L.; York, W.S. and Orlando, R. Quantitation by isobaric labeling: Applications to glycomics. *J. Prot. Res.*, **2008**, *7*, 367-374.

[154] Dell, A.; Reason, A.J.; Khoo, K.H.; Panico, M.; McDowell, R.A. and Morris, H.R. Mass-spectrometry of carbohydrate-containing biopolymers, in Guide to techniques in glycobiology. **1994**, 108-132.

[155] Tipton, R.S. and Horton, D. Advances in carbohydrate chemistry and biochemistry; Academic Press: New York, 1981.

[156] Peltoniemi, H.; Joenvaara, S. and Renkonen, R. De novo glycan structure search with the CID ms/ms spectra of native N-glycopeptides. *Glycobiology*, **2009**, *19*, 707-714.

[157] Perez, S. and Mulloy, B. Prospects for glycoinformatics. *Curr. Opin. Struct. Biol.*, **2005**, *15*, 517-524.

[158] Lohmann, K.K. and von der Lieth, C.W. Glycofragment and glycosearchms: Web tools to support the interpretation of mass spectra of complex carbohydrates. *Nucleic Acid Res.*, **2004**, *32*, W261-W266.

[159] Ceroni, A.; Maass, K.; Geyer, H.; Geyer, R.; Dell, A. and Haslam, S.M. Glycoworkbench: A tool for the computer-assisted annotation of mass spectra of glycans. *J. Prot. Res.*, **2008**, *7*, 1650-1659.

[160] Ceroni, A.; Dell, A. and Haslam, S. The glycanbuilder: A fast, intuitive and flexible software tool for building and displaying glycan structures. *Source Code for Biology and Medicine*, **2007**, *2*, 3.

[161] Goldberg, D.; Sutton-Smith, M.; Paulson, J. and Dell, A. Automatic annotation of matrix-assisted laser desorption/ionization N-glycan spectra. *Proteomics*, **2005**, *5*, 865-875.

[162] Goldberg, D.; Bern, M.; North, S.J.; Haslam, S.M. and Dell, A. Glycan family analysis for deducing N-glycan topology from single ms. *Bioinformatics*, **2009**, *25*, 365-371.

[163] Ozohanics, O.; Krenyacz, J.; Ludanyi, K.; Pollreisz, F.; Vekey, K. and Drahos, L. Glycominer: A new software tool to elucidate glycopeptide composition. *Rapid Comm. Mass Spectrom.*, **2008**, *22*, 3245-3254.

[164] Alfredson, D.A. and Korolik, V. Antibiotic resistance and resistance mechanisms in *Campylobacter jejuni* and *Campylobacter coli*. *FEMS Microbiol. Lett.*, **2007**, *277*, 123-132.

[165] Doig, P.; Kinsella, N.; Guerry, P. and Trust, T.J. Characterization of a post-translational modification of *Campylobacter* flagellin: Identification of a sero-specific glycosyl moiety. *Mol. Microbiol.*, **1996**, *19*, 379-387.

[166] Fry, B.N.; Feng, S.; Chen, Y.Y.; Newell, D.G.; Coloe, P.J. and Korolik, V. The galE gene of *Campylobacter jejuni* is involved in lipopolysaccharide synthesis and virulence. *Infect. Immun.*, **2000**, *68*, 2594-2601.

[167] Vijayakumar, S.; Merkx-Jacques, A.; Ratnayake, D.B.; Gryski, I.; Obhi, R.K.; Houle, S.; Dozois, C.M. and Creuzenet, C. Cj1121c, a novel UDP-4-keto-6-deoxy-GlcNAc C-4 aminotransferase essential for protein glycosylation and virulence in *Campylobacter jejuni*. *J. Biol. Chem.*, **2006**, *281*, 27733-27743.

[168] Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Davies, R.M.; Feltwell, T.; Holroyd, S.; Jagels, K.; Karlyshev, A.V.; Moule, S.; Pallen, M.J.; Penn, C.W.; Quail, M.A.; Rajandream, M.A.; Rutherford, K.M.; van Vliet, A.H.M.; Whitehead, S. and Barrell, B.G. The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature*, **2000**, *403*, 665-668.

[169] Karlyshev, A.V.; Linton, D.; Gregson, N.A. and Wren, B.W. A novel paralogous gene family involved in phase-variable flagella-mediated motility in *Campylobacter jejuni*. *Microbiol.*, **2002**, *148*, 473-480.

[170] Guerry, P.; Alm, R.A.; Power, M.E.; Logan, S.M. and Trust, T.J. Role of 2 flagellin genes in *Campylobacter* motility. *J. Bacteriol.*, **1991**, *173*, 4757-4764.

[171] Geleijns, K.; Roos, A.; Houwing-Duistermaat, J.J.; van Rijs, W.; Tio-Gillen, A.P.; Laman, J.D.; van Doorn, P.A. and Jacobs, B.C. Mannose-binding lectin contributes to the severity of Guillain-Barre syndrome. *J. Immunol.*, **2006**, *177*, 4211-4217.

[172] Lo, Y.L. Clinical and immunological spectrum of the Miller Fisher syndrome. *Muscle Nerve*, **2007**, *36*, 615-627.

[173] Lewis, A.L.; Desa, N.; Hansen, E.E.; Knirel, Y.A.; Gordon, J.I.; Gagneux, P.; Nizet, V. and Varki, A. Innovations in host and microbial sialic acid biosynthesis revealed by phylogenomic prediction of nonulosonic acid structure. *Proc. Natl. Acad. Sci. USA*, **2009**, *106*, 13552-13557.

[174] Linton, D.; Karlyshev, A.V.; Hitchen, P.G.; Morris, H.R.; Dell, A.; Gregson, N.A. and Wren, B.W. Multiple N-acetyl neuraminic acid synthetase (NeuB) genes in *Campylobacter jejuni*: Identification and characterization of the gene involved in sialylation of lipo-oligosaccharide. *Mol. Microbiol.*, **2000**, *35*, 1120-1134.

[175] Thibault, P.; Logan, S.M.; Kelly, J.F.; Brisson, J.R.; Ewing, C.P.; Trust, T.J. and Guerry, P. Identification of the carbohydrate moieties and glycosylation motifs in *Campylobacter jejuni* flagellin. *J. Biol. Chem.*, **2001**, *276*, 34862-34870.

[176] Goon, S.; Kelly, J.F.; Logan, S.M.; Ewing, C.P. and Guerry, P. Pseudaminic acid, the major modification on *Campylobacter* flagellin, is synthesized *via* the Cj1293 gene. *Mol. Microbiol*, **2003**, *50*, 659-671.

[177] Creuzenet, C. Characterization of Cj1293, a new UDP-GlcNAc C-6 dehydratase from *Campylobacter jejuni*. *FEBS Lett.*, **2004**, *559*, 136-140.

[178] Schoenhofen, I.C.; McNally, D.J.; Vinogradov, E.; Whitfield, D.; Young, N.M.; Dick, S.; Wakarchuk, W.W.; Brisson, J.R. and Logan, S.M. Functional characterization of dehydratase/aminotransferase pairs from *Helicobacter* and *Campylobacter* - enzymes distinguishing the pseudaminic acid and bacillosamine biosynthetic pathways. *J. Biol. Chem.*, **2006**, *281*, 723-732.

[179] Liu, F. and Tanner, M.E. Pseg of pseudaminic acid biosynthesis - a UDP-sugar hydrolase as a masked glycosyltransferase. *J. Biol. Chem.*, **2006**, *281*, 20902-20909.

[180] Schoenhofen, I.C.; McNally, D.J.; Brisson, J.R. and Logan, S.M. Elucidation of the CMP-pseudaminic acid pathway in *Helicobacter pylori*: Synthesis from udp-n-acetylglucosamine by a single enzymatic reaction. *Glycobiology*, **2006**, *16*, 8C-14C.

[181] Chou, W.K.; Dick, S.; Wakarchuk, W.W. and Tanner, M.E. Identification and characterization of NeuB3 from *Campylobacter jejuni* as a pseudaminic acid synthase. *J. Biol. Chem.*, **2005**, *280*, 35922-35928.

[182] Logan, S.M.; Kelly, J.F.; Thibault, P.; Ewing, C.P. and Guerry, P. Structural heterogeneity of carbohydrate modifications affects serospecificity of *Campylobacter* flagellins. *Mol. Microbiol.*, **2002**, *46*, 587-597.

[183] Schoenhofen, I.C.; Vinogradov, E.; Whitfield, D.M.; Brisson, J.R. and Logan, S.M. The CMP-legionaminic acid pathway in *Campylobacter*: Biosynthesis involving novel GDP-linked precursors. *Glycobiology*, **2009**, *19*, 715-725.

[184] Linton, D.; Allan, E.; Karlyshev, A.V.; Cronshaw, A.D. and Wren, B.W. Identification of N-acetylgalactosamine-containing glycoproteins PEB3 and CgpA in *Campylobacter jejuni*. *Mol. Microbiol.*, **2002**, *43*, 497-508.

[185] Wacker, M.; Linton, D.; Hitchen, P.G.; Nita-Lazar, M.; Haslam, S.M.; North, S.J.; Panico, M.; Morris, H.R.; Dell, A.; Wren, B.W. and Aebi, M. N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *e-coli*. *Science*, **2002**, *298*, 1790-1793.

[186] Szymanski, C.M.; Logan, S.M.; Linton, D. and Wren, B.W. *Campylobacter* - a tale of two protein glycosylation systems. *Trends Microbiol.*, **2003**, *11*, 233-238.

[187] Linton, D.; Dorrell, N.; Hitchen, P.G.; Amber, S.; Karlyshev, A.V.; Morris, H.R.; Dell, A.; Valvano, M.A.; Aebi, M. and Wren, B.W. Functional analysis of the *Campylobacter jejuni* N-linked protein glycosylation pathway. *Mol. Microbiol.*, **2005**, *55*, 1695-1703.

[188] Glover, K.J.; Weerapana, E.; Chen, M.M. and Imperiali, B. Direct biochemical evidence for the utilization of UDP-bacillosamine by pglC, an essential glycosyl-1-phosphate transferase in the *Campylobacter jejuni* N-linked glycosylation pathway. *Biochemistry*, **2006**, *45*, 5343-5350.

[189] Kelly, J.; Jarrell, H.; Millar, L.; Tessier, L.; Fiori, L.M.; Lau, P.C.; Allan, B. and Szymanski, C.M. Biosynthesis of the N-linked glycan in *Campylobacter jejuni* and addition onto protein through block transfer. *J. Bacteriol.*, **2006**, *188*, 2427-2434.

[190] Riccardi, G.; Pasca, M.R. and Buroni, S. *Mycobacterium tuberculosis*: Drug resistance and future perspectives. *Future Microbiol.*, **2009**, *4*, 597-614.

[191] Donoghue, H.D. and Holton, J. Intestinal tuberculosis. *Curr. Opin. Infect. Dis.*, **2009**, *22*, 490-496.

[192] Dietrich, J. and Doherty, T.M. Interaction of *Mycobacterium tuberculosis* with the host: Consequences for vaccine development. *Apmis*, **2009**, *117*, 440-457.

[193] Torrelles, J.B.; Azad, A.K.; Henning, L.N.; Carlson, T.K. and Schlesinger, L.S. Role of C-type lectins in mycobacterial infections. *Curr. Drug Targets*, **2008**, *9*, 102-112.

[194] Alteri, C.J.; Xicohtencatl-Cortes, J.; Hess, S.; Caballero-Olin, G.; Giron, J.A. and Friedman, R.L. *Mycobacterium tuberculosis* produces pili during human infection. *Proc. Natl. Acad. Sci. USA*, **2007**, *104*, 5145-5150.

[195] Diaz-Silvestre, H.; Espinosa-Cueto, P.; Sanchez-Gonzalez, A.; Esparza-Ceron, M.A.; Pereira-Suarez, A.L.; Bernal-Fernandez, G.; Espitia, C. and Mancilla, R. The 19-kda antigen of *Mycobacterium tuberculosis* is a major adhesin that binds the mannose receptor of thp-1 monocytic cells and promotes phagocytosis of mycobacteria. *Microb. Pathog.*, **2005**, *39*, 97-107.

[196] Schlesinger, L.S. Macrophage phagocytosis of virulent but not attenuated strains of *Mycobacterium-tuberculosis* is mediated by mannose receptors in addition to complement receptors. *J. Immunol.*, **1993**, *150*, 2920-2930.

[197] Villeneuve, C.; Gilleron, M.; Maridonneau-Parini, I.; Daffe, M.; Astarie-Dequeker, C. and Etienne, G. *Mycobacteria* use their surface-exposed glycolipids to infect human macrophages through a receptor-dependent process. *J. Lipid Res.*, **2005**, *46*, 475-483.

[198] Ortalomagne, A.; Dupont, M.A.; Lemassu, A.; Andersen, A.B.; Gounon, P. and Daffe, M. Molecular composition of the outermost capsular material of the tubercle bacillus. *Microbiol.*, **1995**, *141*, 1609-1620.

[199] van Kooyk, Y. and Geijtenbeek, T.B.H. DC-sign: Escape mechanism for pathogens. *Nat. Rev. Immunol.*, **2003**, *3*, 697-709.

[200] Tailleux, L.; Pham-Thi, N.; Bergeron-Lafaurie, A.; Herrmann, J.L.; Charles, P.; Schwartz, O.; Scheinmann, P.; Lagrange, P.H.; de Blic, J.; Tazi, A.; Gicquel, B. and Neyrolles, O. DC-sign induction in alveolar macrophages defines privileged target host cells for mycobacteria in patients with *tuberculosis*. *Plos Med.*, **2005**, *2*, 1269-1279.

[201] Yadav, M. and Schorey, J.S. The beta-glucan receptor dectin-1 functions together with tlr2 to mediate macrophage activation by *Mycobacteria*. *Blood*, **2006**, *108*, 3168-3175.

[202] Daniel, T.M.; Gonchoroff, N.J.; Katzmann, J.A. and Olds, G.R. Specificity of *Mycobacterium-tuberculosis* antigen-5 determined

with mouse monoclonal-antibodies. *Infect. Immun.*, **1984**, *45*, 52-55.

[203] Olds, G.R.; Sanson, A.J. and Daniel, T.M. Characterization of *Mycobacterium-tuberculosis* antigen-5 epitopes by using a panel of 19 monoclonal-antibodies. *J. Clin. Microbiol.*, **1987**, *25*, 471-475.

[204] Espitia, C. and Mancilla, R. Identification, isolation and partial characterization of *Mycobacterium-tuberculosis* glycoprotein antigens. *Clin. Exp. Immunol.*, **1989**, *77*, 378-383.

[205] Garbe, T.; Harris, D.; Vordermeier, M.; Lathigra, R.; Ivanyi, J. and Young, D. Expression of the *Mycobacterium-tuberculosis* 19-kilodalton antigen in *Mycobacterium-smegmatis* - immunological analysis and evidence of glycosylation. *Infect. immun.*, **1993**, *61*, 260-267.

[206] Dobos, K.M.; Swiderek, K.; Khoo, K.H.; Brennan, P.J. and Belisle, J.T. Evidence for glycosylation sites on the 45-kilodalton glycoprotein of *Mycobacterium-tuberculosis. Infect. Immun.*, **1995**, *63*, 2846-2853.

[207] Dobos, K.M.; Khoo, K.H.; Swiderek, K.M.; Brennan, P.J. and Belisle, J.T. Definition of the full extent of glycosylation of the 45-kilodalton glycoprotein of *Mycobacterium tuberculosis. J. Bacteriol.*, **1996**, *178*, 2498-2506.

[208] Horn, C.; Namane, A.; Pescher, P.; Riviere, M.; Romain, F.; Puzo, G.; Barzu, O. and Marchal, G. Decreased capacity of recombinant 45/47-kda molecules (apa) of *Mycobacterium tuberculosis* to stimulate T lymphocyte responses related to changes in their mannosylation pattern. *J. Biol. Chem.*, **1999**, *274*, 32023-32030.

[209] Herrmann, J.L.; Ogaora, P.; Gallagher, A.; Thole, J.E.R. and Young, D.B. Bacterial glycoproteins: A link between glycosylation and proteolytic cleavage of a 19 kda antigen from *Mycobacterium tuberculosis. EMBO J.*, **1996**, *15*, 3547-3554.

[210] Thoma-Uszynski, S.; Stenger, S.; Takeuchi, O.; Ochoa, M.T.; Engele, M.; Sieling, P.A.; Barnes, P.F.; Rollinghoff, M.; Bolcskei, P.L.; Wagner, M.; Akira, S.; Norgard, M.V.; Belisle, J.T.; Godowski, P.J.; Bloom, B.R. and Modlin, R.L. Induction of direct antimicrobial activity through mammalian toll-like receptors. *Science*, **2001**, *291*, 1544-1547.

[211] Pai, R.K.; Pennini, M.E.; Tobian, A.A.R.; Canaday, D.H.; Boom, W.H. and Harding, C.V. Prolonged Toll-like receptor signaling by *Mycobacterium tuberculosis* and its 19-kilodalton lipoprotein inhibits gamma interferon-induced regulation of selected genes in macrophages. *Infect. Immun.*, **2004**, *72*, 6603-6614.

[212] Fortune, S.M.; Solache, A.; Jaeger, A.; Hill, P.J.; Belisle, J.T.; Bloom, B.R.; Rubin, E.J. and Ernst, J.D. *Mycobacterium tuberculosis* inhibits macrophage responses to IFN-gamma through myeloid differentiation factor 88-dependent and -independent mechanisms. *J. Immunol.*, **2004**, *172*, 6272-6280.

[213] Noss, E.H.; Pai, R.K.; Sellati, T.J.; Radolf, J.D.; Belisle, J.; Golenbock, D.T.; Boom, W.H. and Harding, C.V. Toll-like receptor 2-dependent inhibition of macrophage class II MHC expression and antigen processing by 19-kda lipoprotein of *Mycobacterium tuberculosis. J. Immunol.*, **2001**, *167*, 910-918.

[214] Sartain, M.J. and Belisle, J.T. N-terminal clustering of the O-glycosylation sites in the *Mycobacterium tuberculosis* lipoprotein sodC. *Glycobiology*, **2009**, *19*, 38-51.

[215] Ge, Y.; ElNaggar, M.; Sze, S.K.; Bin Oh, H.; Begley, T.P.; McLafferty, F.W.; Boshoff, H. and Barry, C.E. Top down characterization of secreted proteins from *Mycobacterium tuberculosis* by electron capture dissociation mass spectrometry. *J. Am. Soc. Mass Spectrom.*, **2003**, *14*, 253-261.

[216] Majlessi, L.; Brodin, P.; Brosch, R.; Rojas, M.J.; Khun, H.; Huerre, M.; Cole, S.T. and Leclerc, C. Influence of ESAT-6 secretion system 1 (RD1) of mycobacterium tuberculosis on the interaction between mycobacteria and the host immune system. *J. Immunol.*, **2005**, *174*, 3570-3579.

[217] Gonzalez-Zamnorano, M.; Mendoza-Hernandez, G.; Xolalpa, W.; Parada, C.; Vallecillo, A.J.; Bigi, F. and Espitia, C. *Mycobacterium tuberculosis* glycoproteomics based on cona-lectin affinity capture of mannosylated proteins. *J. Proteome Res.*, **2009**, *8*, 721-733.

[218] Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W. and Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.*, **1990**, *215*, 403-410.

Received: September 30, 2009 Revised: December 04, 2009 Accepted: December 22, 2009