

Published in final edited form as:

*Methods Mol Biol.* 2019 ; 1871: 421–435. doi:10.1007/978-1-4939-8814-3\_24.

## Protein-specific analysis of invertebrate glycoproteins

Alba Hykollari, Daniel Malzl, Iain B. H. Wilson, and Katharina Paschinger

Department für Chemie, Universität für Bodenkultur, Muthgasse 18, A-1190, Vienna, Austria

### Summary

N-glycans are post-translational modifications of proteins attached to the amide side chains of asparagine residues, with possible heterogeneity due to different structures being possible at the same glycosylation site. In contrast to the mammalian systems, invertebrate N-glycosylation presents a challenge in analysis as there exist unfamiliar epitopes and a high degree of structural and isomeric variation between different species. A simple analytical approach to analyse N-glycans on specific glycoproteins is presented, which involves a combination of tryptic peptide mass spectrometry and “off-line” RP-HPLC MALDI-TOF MS/MS complemented by blotting to recognize specific epitopes. An additional N-glycan enrichment and labelling step can facilitate the analysis of single structures and even provide isomeric separation of N-glycans from specific proteins.

### Keywords

glycosylation; glycoproteomics; mass spectrometry; “off-line” MALDI-TOF MS(MS)

## 1 Introduction

Amongst the post-translational modifications, the various forms of glycosylation represent a big challenge in terms of analysis. Most commonly, glycans are N- or O-linked to amino- or hydroxyl-side chains of proteins, although C- and S-linkages are also known (1). Probably N-glycans are the most studied forms and are known from bacteria, archaea and almost all eukaryotes; in the latter case, asparagine residues are modified with an oligosaccharide via a core *N*-acetylglucosamine residue (2). Certainly, N-glycans from mammals are quite well studied, whereas for invertebrate organisms the N-glycan structures and their functions remain rather unknown. In contrast to their simple body shape or size, recent studies on N-glycosylation have proven that invertebrate organisms synthesize complicated N-glycomes, which compete in terms of complexity with those of vertebrates (3–6).

Due to the high variety of N-glycan core and antennal modifications in invertebrates, their analysis is challenging and more time consuming, as most of the available bioinformatics tools are based on mammalian structures and so have limited utility when considering invertebrate glycopeptide and N-glycan data. On the other hand, invertebrate glycoproteins are biomedically relevant either due to their immunogenicity (e.g., in venoms),

immunomodulatory activity and relevance as vaccine targets (e.g., parasite glycoproteins) or in the use of invertebrate cell lines (e.g., for baculovirus-based systems) for production of recombinant biopharmaceuticals (7–9). Thus, there is a need to adequately determine invertebrate N-glycan structures at the glycomic and glycoprotein levels. Clearly, glycan annotations on the basis of  $m/z$  alone are insufficient and are often misleading. Orthogonal proofs are therefore necessary, including the use of specific detection reagents, MS/MS fragmentation, chemical or exoglycosidase treatments or reference to in-depth glycomic analyses from the same organism (10,11).

The analysis of N-glycans is often required from either purified single proteins or low amounts of biological material. Here we describe procedures for protein-specific N-glycan purification, enrichment and analysis, successfully used with invertebrate glycoproteins. The protein is first screened for N-glycan epitopes with Western blot, whereby the affinity of specific antibodies, lectins and pentraxins, can give first impressions as to the modifications of the oligosaccharides attached to proteins (e.g., core fucose, terminal galactose, phosphorylcholine, phosphoethanolamine). The tryptic digest of the protein and subsequent peptide mass spectrometry fingerprinting help to identify the protein *in silico* and provide (glyco)peptides for further analysis. Peptide:N-glycosidases can then be used to cleave the N-glycans from the peptides prior to mass spectrometric analyses; as amounts allow, these can be fluorescently labeled and subject to HPLC and MS/MS analyses. Thus, we can go beyond the typical glycoproteomic procedures in order to more firmly define N-glycan structures on specific glycoproteins.

## 2 Materials

### 2.1 Equipment

1. Probe sonifier, e.g., Branson Sonifier 250
2. Porcelain mortar with pestle; tight fitting glass homogenizer (Wheaton; customise as required)
3. Vacuum centrifuge (e.g., Speedvac, Thermo)
4. Micro-centrifuge (e.g., Heraeus, Thermo)
5. Lyophilizer (Labconco)
6. Mini Protean® Tetra cell and Power Pac power supply (Bio-Rad)
7. Trans blot SD semi dry electrophoretic transfer cell (Bio-Rad)
8. Glass columns of 1 cm diameter and 50 cm length (Bio-Rad)
9. Multifunctional microtitre plate reader (e.g., Infinite M200, Tecan); black 96-well microtitre plates, e.g., Microfluor™ 1 or LumiNunc (Thermo).
10. HPLC liquid chromatograph with fluorescence detector (e.g., Shimadzu Nexera); reverse phase chromatography column, e.g., Ascentis® Express RP-Amide (150 mm × 46 mm, 2.7 μm, Supelco)

11. MALDI-TOF-TOF-MS: Autoflex Speed or UltrafleXtreme MALDI-TOF-TOF; appropriate MALDI polished or ground steel target plate (Bruker Daltonics, Billerica, MA). Alternatives are available commercially from Shimadzu or Applied Biosystems.

## 2.2 Reagents, Buffers and Columns (see Notes 1 and 2)

**2.2.1 Disruption of biological material and SDS-PAGE sample preparation—1.** 2 x SDS-PAGE reducing sample buffer containing 200 mg SDS, 154 mg DTT, 5 ml stacking gel buffer, 3.6 ml 87 % glycerol (make up to 10 ml with water, then add a few crystals of bromophenol blue)

### 2.2.2 SDS-PAGE and Western blotting

1. 12% SDS-PAGE gel (using 40% acrylamide stock from Bio-Rad, diluted with either stacking gel buffer with 0.5 M Tris/HCl pH 6.8 or separation gel buffer with 1.5 M Tris/HCl pH 8.8)
2. SDS-PAGE running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS; components from either VWR or Roth)
3. Western blotting transfer buffer (25 mM Tris, 192 mM glycine, 10% methanol; VWR or Roth)
4. SDS-PAGE protein standard ladder (e.g., Thermo PageRuler™)
5. Nitrocellulose membrane (BioTrace™ NT from Pall Life Science)
6. Extra thick blotting paper (Bio-Rad)
7. 0.5 % (w/v) Ponceau S (Sigma) in 1 % (v/v) acetic acid solution
8. Membrane washing buffer: Tris buffered saline (TBS, i.e., 0.1 M Tris/HCl, pH 7.4, 0.1 M NaCl; typically made as a 10-fold concentrated stock) with 0.05% Tween (Sigma)
9. Membrane blocking and antibody/lectin dilution buffer: Tris buffered saline with 0.05% Tween and 0.5% BSA (Roth)
10. Primary, secondary antibodies, lectins and pentraxins (Sigma or Vector Laboratories) (see Table 1)
11. SigmaFAST BCIP/NBT or SigmaFAST 3,3'-diaminobenzidine tetrahydrochloride tablets (Sigma), dissolved in 10ml and 5ml respectively.

### 2.2.3 Tryptic peptide mapping

1. Colloidal Coomassie Blue staining solution: 0.02% (w/v) Coomassie Brilliant Blue G-250 (Bio-Rad), 5% aluminium sulphate-(14-18)-hydrate [ $\text{Al}_2(\text{SO}_4)_3 \cdot 16\text{H}_2\text{O}$  (Roth)], ethanol 96% (VWR), phosphoric acid 85% (Roth). Weigh in 100 g of aluminium sulphate and dissolve it in 1500 ml of water; add 200 ml ethanol and mix well; add 0.4 g of Coomassie Brilliant Blue G-250 and

mix well for at least 30 minutes; add slowly 47 ml of phosphoric acid and mix well; make up to 2000 ml with water (see Note 3).

2. Acetonitrile LC-MS grade (VWR), ammonium bicarbonate (Roth), water HPLC super gradient grade (VWR), iodoacetamide (Sigma), dithiothreitol (Roth), trifluoroacetic acid (Fluka)
3. Sequencing grade modified trypsin dissolved to 0.1 mg/ml in 50 mM acetic acid (Promega); typically cleaves after Arg and Lys residues.

#### 2.2.4 N-glycome release and analysis

1. Peptide:N-glycosidase F (PNGase F, recombinant from *Flavobacterium meningosepticum*, Sigma)
2. Peptide:N-glycosidase A (either purified native from almond meal, PNGase A from Sigma, or recombinant Endo H-treated from *Oryza sativa* and expressed in *Pichia pastoris*, PNGase Ar from NEB)
3. For PNGase F: 100 mM McIlvaine phosphate/citrate buffer (pH 7.5) or 50 mM ammonium hydrogen carbonate (pH 8; mixture of ammonium carbonate and ammonium hydrogen carbonate)
4. For PNGase A: 20 mM ammonium acetate (pH 5)
5. 1-3 ml solid-phase extraction column and frits (Supelco)
6. Acetonitrile, isopropanol, acetic acid, water (as above)
7. Dowex AG® 50W-X8 200-400 mesh H<sup>+</sup> form (Bio-Rad, biotechnology grade; washed serially with 0.1 M NaOH, water, 0.1 M HCl, water, 1 M ammonium acetate and water) and pre-equilibrated with 2% acetic acid prior to usage; C18 material (Lichroprep, Merck); non-porous graphitized carbon material (NPGC; e.g., Supelco ENVICarb™)
8. MALDI matrices:  $\alpha$ -cyanocinnamic acid (ACH, Sigma; 10 mg/ml in 0.1% trifluoroacetic acid/50% acetonitrile) or 6-aza-thiothymine (ATT, Sigma; 3 mg/ml ATT dissolved in 50% ethanol)
9. Glycan labelling: 2-aminopyridine (PA, >99% purity, Sigma), sodium cyanoborohydride (95% purity, Sigma), hydrochloric acid (37% HCl, Roth).
10. Gel filtration: Sephadex G-15 and G-25 coarse (GE Healthcare).

#### 2.2.5 Proteome and N-glycan data analysis

##### 2.2.5.1 Database search programs

1. Matrix Science web server ([www.matrixscience.com/cgi](http://www.matrixscience.com/cgi))
2. ProteinProspector ([prospector.ucsf.edu/prospector/cgi-bin](http://prospector.ucsf.edu/prospector/cgi-bin))

##### 2.2.5.2 Peptide/Protein utility program

1. Theoretical peptide mass calculator ([www.expasy.org](http://www.expasy.org))

### 2.2.5.3 N-glycan analysis

1. Glycoworkbench ([www.glycoworkbench.org](http://www.glycoworkbench.org))
2. FlexAnalysis Bruker software

## 3 Methods (see flow chart and example data in Figures 1 and 2)

### 3.1 Sample preparation and glycoepitope recognition

**3.1.1 Sample preparation for glycoprotein analysis (See Notes 1 and 2)**—The purification procedure of the (glyco)protein of interest depends on the biological material which can be whole organisms, cells, tissues, cyst fluid, semi-purified proteins or secreted (glyco)proteins in culture media or buffer.

1. Heat inactivate the biological material in boiling water for 10 minutes. After cooling the sample (cells, worms, royal jelly, etc.), homogenize the material using a sonifier or mortar and pestle or tight fitting glass homogenizer; for tissues or fungal mycelium, lyophilize after heat inactivation and ground to a fine powder in liquid nitrogen. Large volumes of proteinaceous liquid samples are concentrated with an additional precipitation step (e.g., methanol). Alternatively, after homogenization, purify a subset of glycoproteins by an additional enrichment step using affinity chromatography (if available) e.g., agarose immobilized monoclonal antibodies (12).
2. Prior to SDS-PAGE, precipitate an aliquot of the samples with a five-fold excess volume of methanol, incubate at -80°C for one hour and centrifuge for 10 minutes at 4°C, 21000 g. Dry the protein pellet at 65°C for several minutes to evaporate excessive methanol and re-dissolve the pellet in 20 µl SDS-PAGE sample buffer. In addition, heat treat the mixture for 10 minutes at 95°C and after cooling, centrifuge again for 5 minutes at room temperature, 21 000g.

**3.1.2 SDS-PAGE and Western Blotting**—For initial screening of the N-glycan epitopes, approx. 2 µg of proteins are subject to SDS-PAGE under reducing conditions, followed by protein transfer to a nitrocellulose membrane (Western blotting).

1. Check the quality of the successful transfer by incubating the membrane with Ponceau S staining solution for 1 minute. After de-staining with water (protein bands will stain red), block the membrane with Tris buffered saline containing 0.05% Tween and 0.5% BSA for 1h at room temperature under smooth shaking.
2. Wash the membrane three times using Tris buffered saline with 0.05% Tween (washing buffer).
3. Incubate with biotinylated lectins, pentraxins or primary antibodies in blocking/dilution buffer for 60 minutes (see Table 1 and Note 4).
4. Wash the membrane again thrice as above and incubate with the relevant peroxidase or alkaline phosphatase conjugated secondary antibodies in blocking/dilution buffer for 60 minutes.

5. Again wash the membrane three times as above.
6. Develop the Western blots respectively for peroxidase or phosphatase conjugates with either SigmaFAST 3,3'-diaminobenzidine tetrahydrochloride or SigmaFAST BCIP/NBT (dissolve tablets first in water). Chemiluminescence or other detection methods can also be used.

### 3.2 Tryptic peptide mapping (see Note 5)

1. For the peptide mass fingerprinting identification of proteins with MALDI-TOF MS, apply 10 µg of protein to the SDS-PAGE and stain with Coomassie Blue.
2. After de-staining the gel with water, excise the protein bands in small pieces on glass plates using a clean scalpel.
3. Wash/destain the gel pieces twice with 50% acetonitrile in water and successively once with 1:1 0.1 M ammonium bicarbonate/acetonitrile and 100% acetonitrile only, prior to drying in a Speedvac.
4. In addition, reduce the gel pieces with 10 mM DTT for one hour at 56°C and alkylate for 45 minutes at room temperature with iodoacetamide (55 mM in 0.1 M ammonium bicarbonate) in the dark. Subject the gel pieces to a second round of serial washing (twice 50% acetonitrile, 1:1 0.1 M ammonium bicarbonate/acetonitrile and 100% acetonitrile only) and drying in a Speedvac.
5. For proteolytic digestion, cover the gel pieces with a 1:2 mixture of 0.1 M ammonium bicarbonate/trypsin (100 ng/µl) and incubate overnight at 37°C.
6. Extract the peptides at room temperature three times using acetonitrile/water/trifluoroacetic acid in a ratio of 660:330:1 (v/v/v). Dry the enriched glycopeptides using a vacuum centrifuge and redissolve them in 5 µl water, prior to spotting on a target plate for MALDI-TOF MS analysis.
7. Spot 0.5 µl of the peptides before applying the matrix (either α-cyanocinnamic acid (ACH) or 6-aza thiothymine (ATT)). The peptides are typically measured in the positive ion mode and 2000 shots are summed for MS and 4000 for MSMS. The spectra are processed with the manufacturer's software; for the Bruker Flexanalysis software, this includes the SNAP algorithms with corresponding signal to noise thresholds.

### 6.3 (Glyco)peptide analysis (see Note 6)

1. Predict/identify the corresponding proteins with e.g., the MASCOT program (Matrix Science web server) or MS-Fit (ProteinProspector server) using the peptide masses obtained from tryptic digest and MALDI-TOF MS results. Use one of the sequence databases available online such as Swissprot or Uniprot. In parallel, the list of theoretical peptide masses can be generated by online software (e.g., MS-digest at [prospector.ucsf.edu](http://prospector.ucsf.edu) or [web.expasy.org/peptide-mass](http://web.expasy.org/peptide-mass)).

2. Verify the selected 'peptide-hits' with the sequences of the single masses when subject to MALDI-TOF MS/MS. In order to obtain optimal sequence coverage, allow a mass tolerance of 0.5 Da, one missed cleavage site and consider all fixed modifications (e.g. carbamidomethylation of Cys residues if alkylated) and potentially known contaminants for mass spectrometric fingerprint analysis (e.g., human keratin). Include in the results the protein accession number, number of successfully assigned peptides and the percentage of sequence coverage, the software version, number of database entries and number of species selected for the software search. Glycosylated peptides will not be identified unless subject to PNGase digestion, whereby Asn residues will be converted to Asp ( $m/z = +1$  Da); controls or digestion in  $^{18}\text{O}\text{-H}_2\text{O}$  may be necessary to assess for non-PNGase-mediated deamination.

### 3.4 N-glycome release and analysis

#### 3.4.1 N-glycome release from intact glycoproteins (see Note 7)

1. The peptide:N-glycosidase F (PNGase F) can release N-glycans from undigested proteins. Denature approximately 8  $\mu\text{g}$  of protein first in 10  $\mu\text{l}$  0.5% SDS in water for 5 minutes at 95°C. Alternatively, recombinant PNGase Ar from rice may be partly effective and release also any core  $\alpha$ 1,3-fucosylated structures, which commonly occur in invertebrates and plants.
2. After cooling, add 3  $\mu\text{l}$  of 100 mM McIlvaine phosphate/citrate buffer pH 7.5 and 2  $\mu\text{l}$  of PNGase F to the sample prior to incubation for 2 days at 37°C.
3. Mix approximately 2  $\mu\text{g}$  of either glycosylated or deglycosylated protein with 2 x SDS-PAGE buffer and after heat denaturation and a short centrifugation step, apply both samples to SDS-PAGE and Western blotting to estimate the degree of deglycosylation and N-glycan epitope removal.

#### 3.4.2 N-glycome release from glycopeptides (see Notes 7, 8 and 9)

1. After protein identification, heat treat the glycopeptides to inactivate the protease and incubate 90% of the sample with either PNGase F or PNGase Ar (see Note 7).
2. Optimal conditions for the PNGase Ar activity (use approximately 5 U/reaction) are 20 mM ammonium acetate buffer, pH 5 for 2 days at 37°C.
3. Purify the released N-glycans using two different columns packed with Lichrorep C18/Dowex AG 50 and non-porous graphitized carbon/Lichrorep C18 (see Note 8). Wash first the Lichrorep C18/Dowex AG 50 column with 2 % acetic acid and 60% isopropanol and equilibrate with 2 % acetic acid. Apply the glycopeptide sample after acidifying with 10% acetic acid and collect immediately the unbound released N-glycans in the flow-through and wash fractions (three column volumes of 2% acetic acid).
4. Apply the flow-through/wash from the Lichrorep/Dowex column directly to a non-porous graphitized carbon/Lichrorep C18 column (prewashed and pre-

equilibrated with first 100% acetonitrile then water). After sample application, wash the column with water and elute the N-glycans with 40 % acetonitrile containing 0.1 % trifluoroacetic acid. Due to the presence of TFA, this sample contains a mixed pool of the neutral and anionic N-glycans.

5. Lyophilise the purified N-glycans overnight and after dissolving them in water, spot an aliquot for MALDI-TOF MS/MS analysis with 6-azathiothymine (ATT); regarding acquisition and interpretation of mass spectra, refer to Notes 6 and 9. In comparison to peptides, higher laser power and detector gain settings are necessary to detect glycans. For a more detailed analysis, label the N-glycans by reductive amination using 2-aminopyridine and in addition subject them to HPLC and MALDI-TOF MS analysis as described below.
6. Fluorescent labelling is performed as follows: dissolve 100 mg 2-aminopyridine in 76  $\mu$ l concentrated HCl and 152  $\mu$ l water; add 80  $\mu$ l of this solution to the dried glycan sample, prior to incubation in boiling water for 15 minutes. Then prepare a solution of 4.4 mg of sodium cyanoborohydride in a mixture of 9  $\mu$ l of the aforementioned 2-aminopyridine solution and 13  $\mu$ l water; add 4  $\mu$ l of this cyanoborohydride-aminopyridine solution to the sample and continue the incubation overnight at 90 °C.
7. Removal of excess labelling reagent is performed immediately the following day by gel filtration. Dilute the sample in 1.5 ml of 0.5% acetic acid (i.e., no more than 5% of the gel filtration column volume), apply to a 30 ml Sephadex G-15 column (1  $\times$  40 cm) equilibrated in 0.5% acetic acid, and collect 1.5 ml fractions. Transfer aliquots of fractions (80  $\mu$ l) to a 96 F black plate and detect fluorescence in a microtiter plate reader (excitation/emission: 320/400nm). Pool fluorescent glycans eluting before the excess labeling reagent and lyophilise.
8. Dissolve dried sample by washing the flask four-times with 20  $\mu$ L of water and transfer to a microcentrifuge tube; re-lyophilise as required and analyse an aliquot by MALDI-TOF MS.
9. Inject the major portion of sample onto an Ascentis® Express RP-Amide column pre-equilibrated with 100 mM ammonium acetate (pH 4; buffer A); elute at 0.8 ml/min using a linear gradient of 30% (v/v) MeOH (buffer B) from 0% B up to 35% B over 35 minutes (higher percentages of B generate higher pressure). The glycans are detected by fluorescence using excitation/emission wavelengths of 320/400 nm and the column is calibrated in terms of glucose units with a fluorescently-labelled oligoglucose standard (partial dextran hydrolysate). Collect fractions based on fluorescence intensity and lyophilize prior to another round of MALDI-TOF MS and MS/MS to identify the glycans in the fractions (for example data, refer to Figure 2). Normal phase or non-fused core reversed phase columns can also be used (15).



## 4 Notes

1. The quality of water and other reagents (acetonitrile, methanol, isopropanol) used for analytical purposes should be high and free of ionic and microbial contaminants.
2. In general, contaminants should be avoided; to prevent analysis of ‘foreign’ components from the food/nutrition source or media (e.g. foetal calf serum), the material (whole organisms or cells) should be washed several times before the heat treatment and homogenization. After collection, the biological material should be stored at  $-80^{\circ}\text{C}$ , if not immediately homogenized. To prevent hydrolysis of the anionic or zwitterionic residues (e.g., phosphate, sialic acid, PC or PE), the samples should be heat treated only in water and not in acidic buffers; however, heat inactivation is necessary to prevent degradation of the glycans by endogenous glycosidases. For small amounts of biological samples, also a lysis buffer supplemented with protease inhibitor cocktail (Sigma) can be used prior to SDS-PAGE. The methanol precipitation step after cell lysis helps to desalt the sample and so avoid smearing upon electrophoresis.
3. Colloidal Coomassie aggregates and tiny blue dots are visible. Make sure that the staining solution is mixed well (e.g., with a magnetic mixer) before each use.
4. Results obtained from antibody or lectin binding are no structural proof of the N-glycans on the glycoprotein as their specificities are sometimes wide or not fully determined. Positive and negative controls and pull-downs to ‘pre-clear’ endogenous biotinylated proteins, as well as Western blots with and without lectins/antibodies (i.e., just secondary reagents) or after glycosidase digestions should be considered for data interpretation. The ‘mini-description’ of the epitopes in Table 1 is based on determination of binding of the antibodies, lectins or pentraxins to standard ligands; these determinations are by no means exhaustive as invertebrate standards are rarely tested (13,16). Nevertheless, anti-horseradish peroxidase is valuable for screening of core  $\beta$ 1,2-xylose and core  $\alpha$ 1,3-fucose (17), but the anti-xylose and anti-fucose components of the antisera are difficult to properly separate. Phosphorylcholine (PC) epitopes can be detected with either the TEPC-15 antibody or by human C-reactive protein (12).
5. The peptides measured “off-line” by MALDI-TOF MS can be sometimes suppressed by contaminant ions generating from the protease itself (e.g., trypsin), which is in part autohydrolysed. It is recommended to generate online the theoretical peptide masses of the protease as well as the target protein using the MS-digest software.
6. The method described here is a simple and initial procedure for glycoprotein identification by off-line peptide mass fingerprinting and (glyco)peptide analysis of the selected proteins before and after PNGase F or A digestion followed by MALDI-TOF MS. For qualitative/quantitative peptide studies several “on-line” methods such as LC-ESI MS/MS can also be employed. Invertebrate N-glycomes can dramatically differ from those of mammalian systems, so the N-

glycan assignment on the defined glycopeptides should be based at least on MS/MS data analysis as compositions based on mass alone can be misleading: for instance, a difference of 324 Da can either correspond to two hexoses or one methylaminoethylphosphonate-modified HexNAc as seen, e.g., in molluscs. Also, a difference of 176 Da may be either a methylated hexose or a glucuronic acid (11). Nevertheless, mass differences of 146, 162 or 203 can suggest the presence of fucose, hexose and *N*-acetylhexosamine residues. There are various bioinformatics tools for the automated glycopeptide and glycan identification and the following software can be applied for glycopeptide MS: “GlycoMod”, “GlycoX”, “GlycopepDB” “Massy tools” and “GlycoSpectrumScan” and MSMS “GlycoMiner”, “Protein Prospector”, “GlycopepID”, “GlycoMasterDB” and many more.(18) As these are generally applied to mammalian glycomes and glycoproteomes, caution is required when using search engines to annotate invertebrate glycans. For publication, consider the MIRAGE guidelines for presentation of glycomic data (19) as well as use of the diagrammatic Symbol Nomenclature for Glycans (20).

7. PNGase F can release N-glycans from both glycoproteins and glycopeptides, whereas recombinant PNGase Ar still works best on peptides. PNGase F does not release N-glycans with core  $\alpha$ 1,3-fucose modification (but does release core  $\alpha$ 1,6-fucosylated or  $\beta$ 1,3-mannosylated structures), while recombinant PNGase Ar can release substituted core  $\alpha$ 1,3-fucosylated glycans (6). The degree of protein deglycosylation can be monitored with SDS PAGE (reduced size of the protein after deglycosylation) and Western blotting (reduced or abolished N-glycan epitope binding).
8. After PNGase F or A digest of the glycopeptides one aliquot of the sample should be analysed with MALDI-TOF MS to verify deglycosylated peptides and potential “occupation” of an N-glycosylation site of the protein. The released glycopeptides should be acidified with 10 % acetic acid before Dowex cation exchange chromatography. For N-glycan recovery protocol after PNGase F or A release, refer to our recent protocol on “Analysis of invertebrate and protist N-glycans” (15). For O-glycosylation, there is no single universal de-O-glycosylation enzyme available; O-glycanase has a restricted substrate specificity and will not remove most extended GalNAc-Ser/Thr (mucin-type) or other O-glycan structures.
9. The released N-glycans should be measured in positive and negative ion mode for the identification of potential anionic residues as sulphate (+80 Da), phosphate (+80 Da), glucuronic acid (+176 Da), phosphoethanolamine (+123 Da) and aminoethylphosphonate (+107 Da; +121 Da if methylated) (11). Sialic acids are rare in invertebrates and have been only convincingly proven in *Drosophila* or in echinoderms (21,22), but are absent, e.g., from nematodes.

## Acknowledgments

This work was supported by the Austrian Fonds zur Förderung der wissenschaftlichen Forschung (FWF; grants P26662, P25058 and P23922 to A.H., K.P and I.B.H.W.).

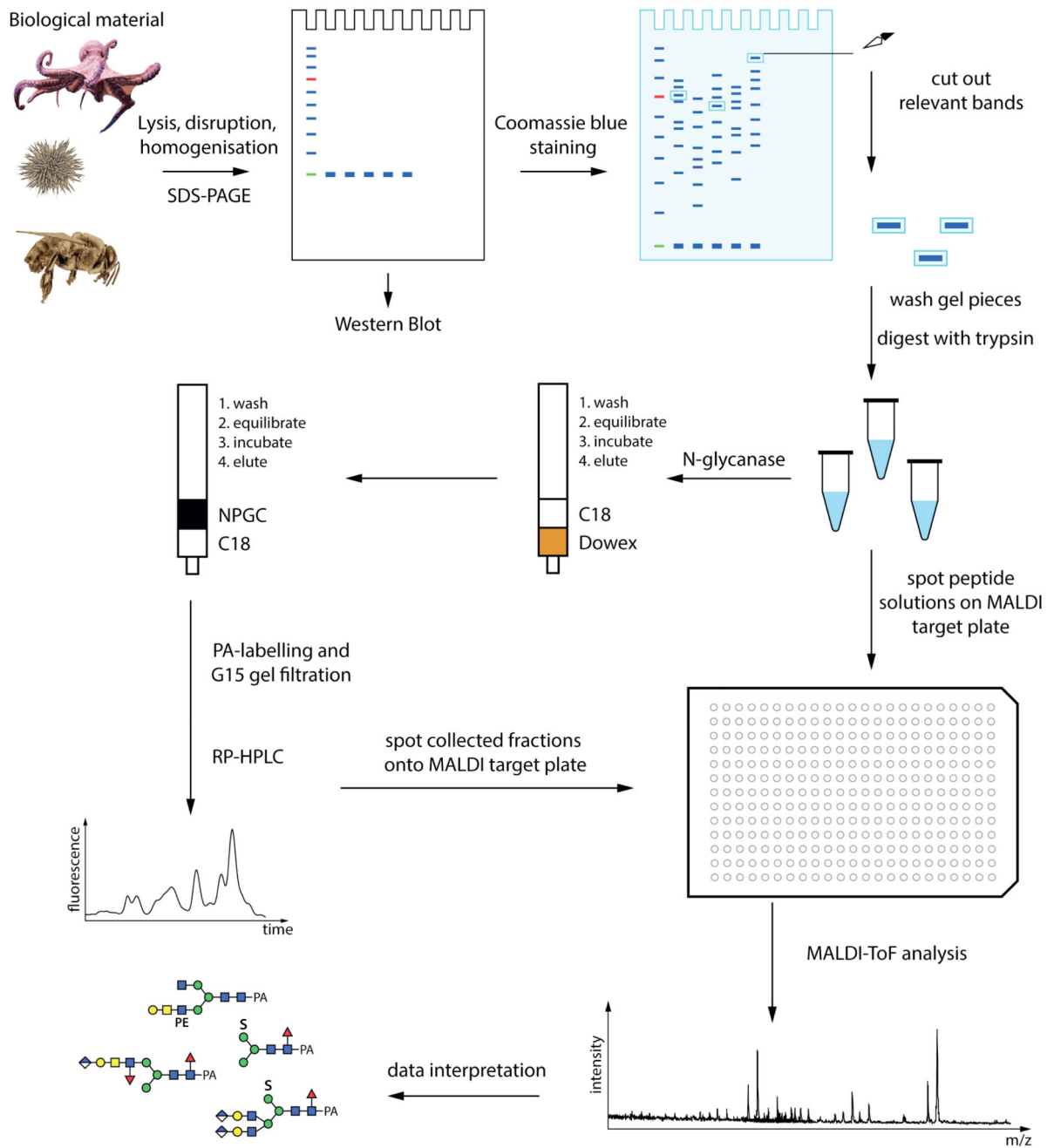
## Abbreviations

<b>DTT</b>	dithiothreitol
<b>HRP</b>	horseradish peroxidase
<b>MALDI-TOF MS</b>	matrix-assisted laser-desorption/ionisation time-of-flight mass spectrometry
<b>NPGC</b>	non-porous graphitized carbon
<b>PC</b>	phosphorylcholine
<b>PE</b>	phosphoethanolamine
<b>RP-HPLC</b>	reversed phase high pressure liquid chromatography
<b>SDS-PAGE</b>	sodium dodecyl sulphate polyacrylamide gel electrophoresis

## References

1. Spiro RG. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*. 2002; 12:43R–56R.
2. Aebi M. N-linked protein glycosylation in the ER. *Biochim Biophys Acta*. 2013; 1833:2430–2437. [PubMed: 23583305]
3. Schiller B, Hykollari A, Yan S, Paschinger K, Wilson IBH. Complicated N-linked glycans in simple organisms. *Biol Chem (Hoppe Seyler)*. 2012; 393:661–673.
4. Eckmair B, Jin C, Abed-Navandi D, Paschinger K. Multi-step fractionation and mass spectrometry reveals zwitterionic and anionic modifications of the N- and O-glycans of a marine snail. *Mol Cell Proteomics*. 2016; 15:573–597. [PubMed: 26598642]
5. Stanton R, Hykollari A, Eckmair B, Malzl D, Dragosits M, Palmberger D, Wang P, Wilson IBH, Paschinger K. The underestimated N-glycomes of lepidopteran species. *Biochim Biophys Acta*. 2017; 1861:699–714.
6. Yan S, Vanbeselaere J, Jin C, Blaukopf M, Wols F, Wilson IBH, Paschinger K. Core Richness of N-Glycans of *Caenorhabditis elegans*: A Case Study on Chemical and Enzymatic Release. *Anal Chem*. 2018; 90:928–935. [PubMed: 29182268]
7. Tretter V, Altmann F, Kubelka V, März L, Becker WM. Fucose  $\alpha$ 1,3-linked to the core region of glycoprotein N-glycans creates an important epitope for IgE from honeybee venom allergic individuals. *Int Arch Allergy Immunol*. 1993; 102:259–266. [PubMed: 7693094]
8. Prasanphanich NS, Mickum ML, Heimburg-Molinaro J, Cummings RD. Glycoconjugates in host-helminth interactions. *Front Immunol*. 2013; 4:240. [PubMed: 24009607]
9. Geisler C, Mabashi-Asazuma H, Jarvis DL. An Overview and History of Glyco-Engineering in Insect Expression Systems. *Methods Mol Biol*. 2015; 1321:131–152. [PubMed: 26082220]
10. Hykollari A, Malzl D, Yan S, Wilson IBH, Paschinger K. Hydrophilic interaction anion exchange for separation of multiply modified neutral and anionic Dictyostelium N-glycans. *Electrophoresis*. 2017; 38:2175–2183. [PubMed: 28556908]
11. Paschinger K, Wilson IBH. Analysis of zwitterionic and anionic N-linked glycans from invertebrates and protists by mass spectrometry. *Glycoconj J*. 2016; 33:273–283. [PubMed: 26899268]
12. Paschinger K, Gonzalez-Sapienza GG, Wilson IBH. Mass spectrometric analysis of the immunodominant glycan epitope of *Echinococcus granulosus* antigen Ag5. *Int J Parasitol*. 2012; 42:279–285. [PubMed: 22342524]

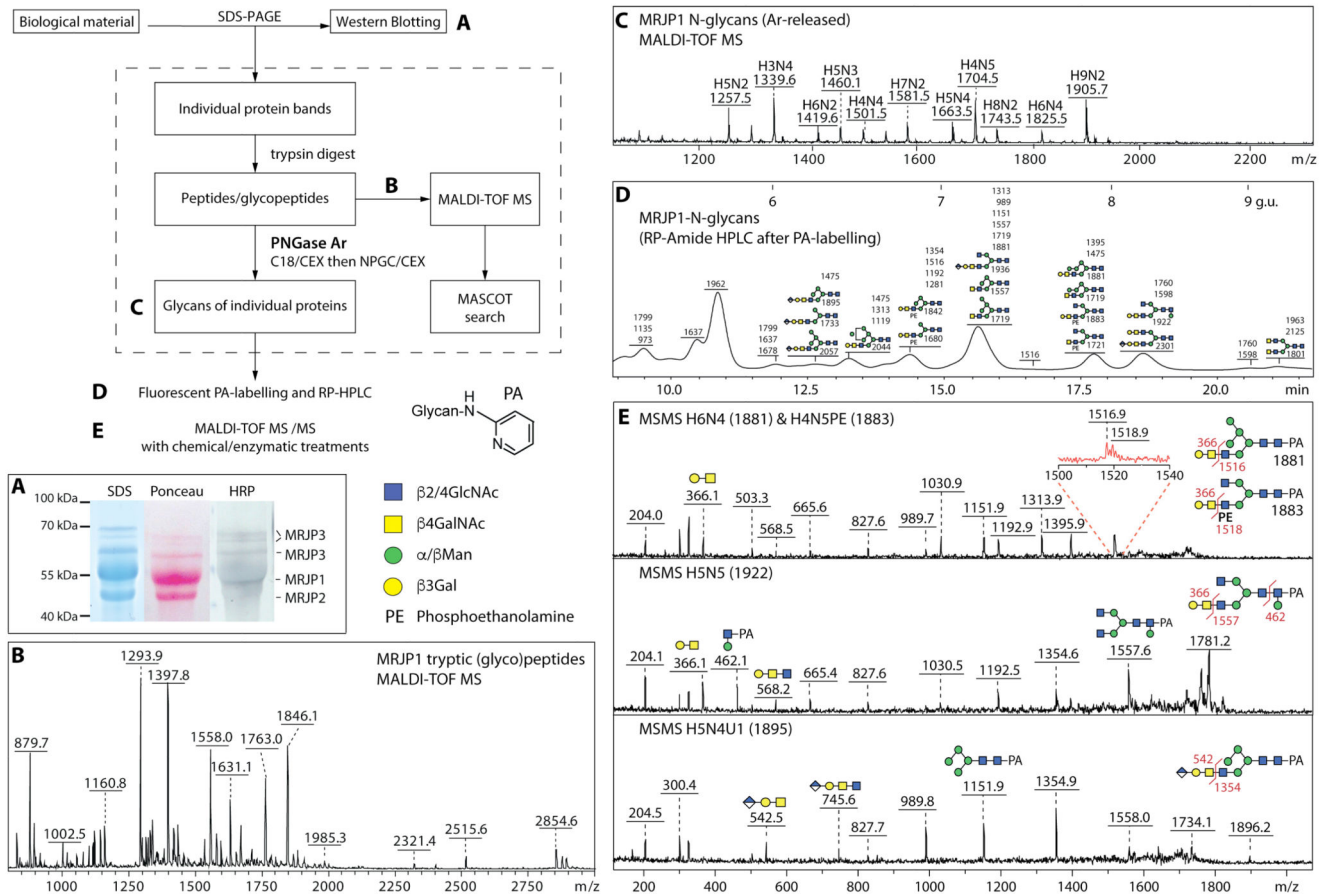
13. Iskratsch T, Braun A, Paschinger K, Wilson IBH. Specificity analysis of lectins and antibodies using remodeled glycoproteins. *Anal Biochem.* 2009; 386:133–146. [PubMed: 19123999]
14. Mikolajek H, Kolstoe SE, Pye VE, Mangione P, Pepys MB, Wood SP. Structural basis of ligand specificity in the human pentraxins, C-reactive protein and serum amyloid P component. *J Mol Recognit.* 2011; 24:371–377. [PubMed: 21360619]
15. Hykollari A, Paschinger K, Eckmair B, Wilson IBH. Analysis of Invertebrate and Protist N-Glycans. *Methods Mol Biol.* 2017; 1503:167–184. [PubMed: 27743366]
16. Purohit S, Li T, Guan W, Song X, Song J, Tian Y, Li L, Sharma A, Dun B, Mysona D, Ghamande S, et al. Multiplex glycan bead array for high throughput and high content analyses of glycan binding proteins. *Nat Commun.* 2018; 9:258. [PubMed: 29343722]
17. Paschinger K, Rendi D, Wilson IBH. Revealing the anti-HRP epitope in *Drosophila* and *Caenorhabditis*. *Glycoconj J.* 2009; 26:385–395. [PubMed: 18726691]
18. Tsai PL, Chen SF. A Brief Review of Bioinformatics Tools for Glycosylation Analysis by Mass Spectrometry. *Mass Spectrom (Tokyo).* 2017; 6:S0064. [PubMed: 28337402]
19. York WS, Agravat S, Aoki-Kinoshita KF, McBride R, Campbell MP, Costello CE, Dell A, Feizi T, Haslam SM, Karlsson N, Khoo KH, et al. MIRAGE: the minimum information required for a glycomics experiment. *Glycobiology.* 2014; 24:402–406. [PubMed: 24653214]
20. Varki A, Cummings RD, Aebi M, Packer NH, Seeberger PH, Esko JD, Stanley P, Hart G, Darvill A, Kinoshita T, Prestegard JJ, et al. Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology.* 2015; 25:1323–1324. [PubMed: 26543186]
21. Aoki K, Perlman M, Lim JM, Cantu R, Wells L, Tiemeyer M. Dynamic developmental elaboration of N-linked glycan complexity in the *Drosophila melanogaster* embryo. *J Biol Chem.* 2007; 282:9127–9142. [PubMed: 17264077]
22. Miyata S, Sato C, Kumita H, Toriyama M, Vacquier VD, Kitajima K. Flagelliasialin: a novel sulfated  $\alpha$ 2,9-linked polysialic acid glycoprotein of sea urchin sperm flagella. *Glycobiology.* 2006; 16:1229–1241. [PubMed: 16887955]



**Figure 1. A potential glycome and glycoproteomic workflow.**

Starting from biological material, proteins can be separated by SDS-PAGE prior to Western blotting or peptide map fingerprinting. The peptides and glycopeptides are analysed directly by mass spectrometry; the glycans are released by an N-glycanase such as PNGase Ar and purified by two rounds of solid phase extraction prior to mass spectrometry and/or HPLC. Glycans (examples from honeybee royal jelly) are depicted according to the Symbol Nomenclature for Glycans, whereby circles, squares, triangles and diamonds respectively

represent hexose (here Man or Gal), *N*-acetylhexosamine (GalNAc or GlcNAc), deoxyhexose (Fuc) or hexuronic acids (GlcA); S, sulphate; PE, phosphoethanolamine.



**Figure 2. Example of a glycoproteomic study of honeybee glycoprotein MRJP1 found in royal jelly**

The flow chart scheme shows the glycoproteomic workflow for the biological sample; the letters refer to steps with exemplified by the following data. **(A)** N-glycan epitope detection by Western blotting of the royal jelly glycoproteins after incubation with anti-HRP antibodies. **(B)** Tryptic peptide mapping of one of the major royal jelly glycoproteins as measured with MALDI TOF MS in positive ion mode. **(C)** Free N-glycans from MRJP1 measured with MALDI-TOF MS after deglycosylation with PNGase Ar with the  $[M+Na]^+$  ions annotated with abbreviations of the form  $H_xN_y$ , where H is hexose and N is *N*-acetylhexosamine. **(D)** RP-HPLC chromatogram of reductively aminated N-glycans from royal jelly glycoprotein MRJP1 with fractions annotated with the detected glycan  $m/z$  values and calibrated in terms of glucose units. **(E)** MALDI-TOF MS/MS data of fractionated MRJP1 glycans, whereby  $[M+H]^+$  ions were fragmented and key B and Y ions are annotated. Note that the precursor ions for the H6N4 and H4N5PE structures cannot be separated ( $m/z = 2$ ), but the zoom shows that Y ions derived from both are present; the structure of the PA (pyridylamino) reducing-terminal label as well as the symbol nomenclature are shown beside the flowchart.

**Table 1**  
**List of selected antibodies, lectins and pentraxins for N-glycan epitope screening (Note 4)**

<b>Antibody (1<sup>st</sup>)</b>	<b>Dilution</b>	<b>Epitope(13,14)</b>	<b>Supplier</b>
Anti-HRP from rabbit, 10 mg/ml	1:10000	Core $\alpha$ 1,3-Fuc/ core $\beta$ 1,2-Xyl	Sigma
Anti-PC (TEPC-15 mouse IgA), 10 mg/ml	1:200	PC-Hex(NAc)	Sigma
<b>Antibody (2<sup>nd</sup>)</b>			
Anti-rabbit IgG from goat conjugated with alkaline phosphatase	1:2000		Vector Labs
Anti-mouse IgA from goat conjugated with alkaline phosphatase	1:10000		Sigma
<b>Pentraxin</b>			
C reactive Protein (CRP) from human plasma (CaCl <sub>2</sub> 2.5 mM added)	1:200	PC-Hex(NAc)	MP Biochemicals
Amyloid P component from human serum (SAP)	1:200	PE-Hex(NAc)	Sigma
<b>Pentraxin recognition</b>			
Anti-human C reactive protein from rabbit	1:1000		Dako
Anti-Amyloid P human serum component IgG from rabbit (anti SAP)	1:1000		Calbiochem
<b>Lectin</b>			
Biotinylated <i>Aleuria aurantia</i> lectin	1:1000	Core $\alpha$ 1,6-Fuc/Lex	Vector Labs
Biotinylated wheat germ agglutinin	1:1000	$\beta$ 1,4HexNAc/ $\alpha$ 2,3Sia	Vector Labs
Biotinylated peanut agglutinin	1:1000	Gal $\beta$ 1,3GalNAc	Vector Labs
<b>Lectin recognition</b>			
Anti-biotin from goat conjugated with alkaline phosphatase	1:10000		Sigma