Ultrasensitive profiling and sequencing of N-linked oligosaccharides using standard DNA-sequencing equipment

Nico Callewaert, Steven Geysens, Francis Molemans, and Roland Contreras

Unit of Fundamental and Applied Molecular Biology, Department of Molecular Biology, Ghent University and Flanders Interuniversity Institute for Biotechnology, K.L.-Ledeganckstraat 35, 9000 Ghent, Belgium

Received on July 21, 2000; revised on November 16, 2000; accepted on November 16, 2000

The analysis of protein-linked glycans is of increasing importance, both in basic glycobiological research and during the production process of glycoprotein pharmaceuticals. In many cases, the amount of glycoprotein available for typing the glycans is very low. This, combined with the high branching complexity typical for this class of compounds, makes glycan typing a challenging task. We present here methodology allowing the medium-throughput analysis of N-glycans derived from low picomole amounts of glycoproteins using the standard DNA-sequencing equipment available in any life sciences laboratory. The high sensitivity of the overall analytical process (from glycoprotein to results) is obtained using state-of-the-art deglycosylation procedures combined with a highly efficient and reproducible novel postderivatization cleanup step involving Sephadex G10 packed 96-well filterplates. All sample preparation steps (enzymatic deglycosylation with PNGase F, desalting, derivatization with 8-amino-1,3,6-pyrenetrisulfonic acid, and postderivatization cleanup) are performed using 96-well-based plates. This integrated sample preparation scheme is also compatible with capillary electrophoresis and MALDI-TOF-MS platforms already in use in some glycobiology labs and anticipates the higher throughput that will be offered by the capillary-array-based DNA sequencers currently penetrating the market. The described technology should bring high-performance glycosylation analysis within reach of each life sciences lab and thus help expedite the pace of discovery in the field of glycobiology.

Key words: α,3-acid glycoprotein/APTS/DNA sequencer/MALDI-TOF-MS/N-glycan.

Introduction

Protein-linked carbohydrates can be analyzed using a variety of high-resolution techniques, such as high-performance liquid chromatography (Guile et al., 1996), capillary electrophoresis (Suzuki and Honda, 1998), and mass spectrometry (MS) (Rudd and Dwek, 1997), especially matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) MS.

For ultrasensitive detection in chromatographic and electrophoretic analytical schemes, it is necessary to derivatize the glycans with a fluorophore or a chromophore, the properties of which can be tailored for the specific application. Almost without exception, the tag is introduced at the reducing terminus of the glycan via reductive amination (Honda, 1996; Paulus and Klockow, 1996). However, it is invariably necessary to use a large excess of fluorescent tag for quantitative derivatization to occur. This not only requires high purity of the fluorescent tag preparation but also implies that this excess tag be efficiently removed if trace amounts of derivatized glycans are to be detected. Many authors have circumvented this difficult problem by derivatizing large amounts of glycans and diluting the derivatization mixture prior to analysis (Liu et al., 1991; Guttman et al., 1996a). This may result in high theoretical sensitivities, but it is not very useful for real-world high-sensitivity glycosylation analysis, where the amounts of starting material are low.

For MALDI-TOF-MS of oligosaccharides, no derivatization is necessary (Kuster et al., 1997), but here highly efficient desalting procedures are required, and this also gets more complicated in trace analysis. However, good-quality MALDI-TOF spectra have been obtained from the N-glycans derived from submicrogram amounts of tissue plasminogen activator by Papac and colleagues (Papac et al., 1998); we are using the deglycosylation procedure described in the aforementioned study and a 96-well-format desalting step adapted from it in our integrated 96-well sample preparation strategy (see Figure 1). A caveat relating to MALDI-TOF-MS of glycans is that isobaric (regio- and stereoisomeric) structures, so prevalent in the field of carbohydrates, are not resolved (Kuster et al., 1997).

During numerous discussions with molecular biologists from diverse fields, it became apparent that many research possibilities concerning protein-linked glycans were not considered in these nonspecialized labs because of a lack of suitable equipment and expertise for glycan analysis and because of the appreciation that high-performance glycan analysis is highly complex and time-consuming. Therefore, we set out to develop a novel tool for the analysis of this class of biocompounds that had to be state-of-the-art in both resolution and sensitivity, easy to perform without the need for extra costly equipment, and able to handle a high throughput of samples.

The most widespread piece of equipment for the high-resolution electrophoretic separation and fluorometric detection of biomolecules is probably the Applied Biosystems series 377 DNA

---

1To whom correspondence should be addressed
sequencer, with thousands of machines in operation all over the world. These instruments use 12, 36, or 48 cm polyacrylamide-based gels as the separation matrix and contain an argon laser to excite the fluorescence of the analytes. As the laser scans past the bottom of the gel, all analytes have to pass through the same length of separation matrix, which helps maintain resolution over a very broad range of electrophoretic mobilities.

In the past, the ABI equipment has been used for the analysis of starch polymers (Morell et al., 1998), for which sensitivity of detection is not an issue because of the copious amounts of the analyte available. It has apparently not been realized so far that the same instruments could potentially be used for the analysis of the far more interesting protein-linked glycans. In this contribution, we show that the standard DNA-sequencing equipment present in the great majority of life science research environments can be adapted to a state-of-the-art glycan analysis tool. The integrated methodology presented here allows high-throughput fingerprinting and sequencing of N-glycans present on picomolar amounts of glycoproteins.

Results and discussion

Choice of the fluorescent label

The presence of an Ar laser in the ABI sequencers and the electrophoretic separation mechanism somewhat limit the choice of the fluorescent tag used to derivatize the glycans. The fluorescence of this tag must be excitable with 488 nm light, and it must contain several negative charges to obtain sufficient electrophoretic mobility of the (often neutral) carbohydrates toward the anode. In capillary electrophoresis and in the starch analysis studies using the ABI systems, 8-amino-1,3,6-pyrenetrisulfonic acid (Chen and Evangelista, 1995) has been used because it contains three sulfonic acid groups and has a λmax for fluorescence excitation of 434 nm after conjugation to a glycan (424 nm unconjugated), with a large tail of the absorption peak extending to 488 nm. Moreover, this label emits fluorescence with maximum intensity at 520 nm, close to the emission maximum of the green dyes used in DNA sequencing. This obviates the need for the creation of a new fluorescence-overlap correction matrix for the sequencer software, so that most machines will be useful for glycan analysis without any adaptation.

Postderivatization cleanup

Initial experiments indicated that it would be crucial to develop a highly efficient postderivatization cleanup step, as overloading of the gel was observed if more than 100 pmol of the free label was present in the sample (data not shown). The minimal concentration of APTS necessary to obtain quantitative derivatization in a reasonable time span (overnight incubation at 37°C) is about 10 nM (Evangelista et al., 1996). The derivatization reactions in this study were miniaturized to 1-µl volumes (performed in 250-µl ultraclean PCR tubes), which means that 10 nmol of the label is present. From these considerations, it is evident that a cleanup methodology is required that removes > 95% of the label to be able to load a significant fraction of the labeled glycans (± 20%) on a sequencing gel lane. For this purpose, we tried several approaches (paper chromatography, thin-layer chromatography, reaction with partially oxidized Sephadex G75 beads; Mort et al., 1998), but we finally turned to an approach using Sephadex G10 packed spin columns. This gel filtration resin has an exclusion limit of 700 Da for dextrans. APTS has a molecular mass of 454 Da, so addition of the chitobiose core sugar of N-glycans (424 Da) already leads to a molecular mass of 878 Da, exceeding the exclusion limit. We initially tested the cleanup efficiency of 1.2-cm beds of Sephadex G10, packed in micropin columns. Over 95% of the free APTS label could easily and reproducibly be removed, with a recovery of > 70% of labeled 14C-lactose (disaccharide) and 14C-sialyllactose (trisaccharide), the neutral and sialylated test compounds used. In a different control experiment, batches of a malt-oligosaccharide mixture containing 1 nmol total carbohydrate were derivatized with APTS and one batch was diluted 1000-fold, while the other batch was cleaned up using the spin column approach and also diluted 1000-fold. One microliter of each final preparation was analyzed on the DNA sequencer, and all malt-oligosaccharides with a degree of polymerization ≥ 4 showed the same relative abundance in both electropherograms (not shown). This demonstrates the nonsize-selectivity of the cleanup for
Oligosaccharides consisting of four or more monosaccharide units.

Subsequently, we adapted this cleanup procedure to a 96-well format by filling the wells of a 96-well Durapore-lined Multiscreen plate with the Sephadex G10 resin. The eluate of these plates is collected in another 96-well plate with tapered wells and evaporated to dryness in a vacuum centrifuge equipped for 96-well plates. The labeled glycans are then reconstituted in water.

The 96-well-based cleanup procedure of APTS-derivated glycans described here should also be applicable for capillary electrophoresis of these compounds (Guttman et al., 1996b), thus fully utilizing the potential sensitivity of this methodology.

Internal standardization

To increase the reproducibility and accuracy of the glycan profiling, we added a rhodamine-labeled oligonucleotide mixture to each sample. Two mixtures were used: the commercially available rhodamine-labeled sizing standard Genescan™ 500 and a mixture of rhodamine-labeled 6-, 18-, 30-, and 48-meric oligonucleotides. The former standard was used in the experiment represented in Figure 2, the latter one in the experiment of Figure 3. By reserving one lane of each gel for an APTS-derivatized malto-oligosaccharide ladder, also containing the rhodamine-labeled standard, the electrophoretic mobility of each glycan can be very reproducibly expressed in glucose units.

This kind of internal referencing of carbohydrate analysis profiles has so far only been described in high-performance anion exchange chromatography, where both a pulsed amperometric detector (detection of the malto-oligosaccharides) and a fluorescence detector (detection of labeled analytes) were necessary to obtain this result (Kotani and Takasaki, 1998). Here, the four-color fluorescence detection capabilities of the DNA sequencer obviate this need. This internal standardization principle may also be applicable to the new capillary electrophoresis laser-induced fluorescence detectors that are capable of two-color fluorescence detection.

Optimization of the Applied Biosystems 377A DNA-sequencer system for high-resolution separation of glycans in the size range relevant for N-glycans

The acrylamide percentage of the gel used here (12%) and the other electrophoresis conditions were optimized for maximum resolution of a malto-oligosaccharide reference mixture with degrees of polymerization of 4–20. This is the size range that is most relevant for N-glycan mixtures derived from mammalian and plant tissues. The gel was kept at 23°C with an external cooling bath to minimize the thermal diffusion of the glycans during electrophoresis. This considerably enhanced the resolution of our separations. The standard buffer used for DNA-sequencing gels was used throughout (see Materials and methods). The borate in this buffer forms complexes of different stability with different carbohydrate isomers, thus increasing the chance of electrophoretically resolving these isomers (Le et al., 1997). It should be straightforward to optimize the electrophoresis parameters for other classes of protein-linked glycans, if necessary.

The detector-response curve is linear over more than three orders of magnitude (1 fmol to >1 pmol, $R^2 = 0.9978$) and 1 fmol of labeled chitotetraose (test compound) can be detected with a signal to noise ratio of $> 3$.

Sample cleanup for MALDI-TOF-MS in a 96-well format

MALDI-TOF-MS of underivatized N-linked glycans is a well-established technique (Kuster et al., 1998, 1997; Papac et al., 1998; Colangelo and Orlando, 1999) at about the same level of sensitivity as the DNA sequencer-assisted methodology described here. The two techniques can give complementary information on the analytes. The sequencer technique often resolves the isobaric glycan isomers and provides reliable quantitation of the observed species, whereas MALDI-TOF-MS gives the exact mass of the glycans. Therefore, we adapted the AG-50-WX8 desalting step described by Papac et al. (1998) to a 96-well format (see Materials and methods), completing our high-throughput sample preparation scheme. As we are merely adapting an already-described sample preparation step in a 96-well format, no MALDI-TOF-MS results are shown and we refer to the aforementioned study for experimental details of the MS procedure.

Use of the developed procedure for N-glycan profiling and sequencing

Combination of the described approaches with exoglycosidase digestions can give structural information on the glycans under study at the low femtomolar level. In Figure 2, this is exemplified for a mixture of trisialylated triantennary complex- versus N-linked oligosaccharides derived from bovine thyroglobulin. As can be seen in panel 2a in Figure 2, the technology is capable of resolving sialic acid linkage isomers. Panel 3a reveals the resolution of Gal-β-1,4 and β-1,3 linkage isomers. This level of analysis is impossible with mass spectrometric techniques and allows quantitation of the different isomers. In Figure 3, we show the N-glycan profiling of glycans derived from 500 ng human α1-acid glycoprotein. This amount of starting material is sufficient to obtain both the native profile and the results of five exoglycosidase arrays. The resulting profiles after exoglycosidase array digestion are fully consistent with the N-glycan structures reported to be present on human α1-acid glycoprotein (Kuster et al., 1998). These structures are of the bi-, tri-, and tetraantennary type with N-acetyl lactosamine and branch fucosylation modifications. For example, the peak remaining after β-N-acetylhexosaminidase digestion is compatible with a single branch fucosylation of a small percentage of the glycans, because this linkage to fucose precludes the β-N-acetylhexosaminidase activity.

An important asset of the technology developed here is that one can combine and compare the results with those obtained from MALDI-TOF-MS of the same analytes on at least the same level of sensitivity. This allows one to characterize the N-glycans present on picomole amounts of any glycoprotein with an unprecedented level of detail, on both a qualitative and a quantitative (relative quantities of isomers, for example) basis.

The possibility of obtaining structural information on N-glycans derived from picomole amounts of glycoproteins with the DNA-sequencing equipment present in most life science laboratories should help open up the field of modern glycobiology to a broader range of research groups than is the case now.
Fig. 2. Sequencing of 2 pmol of a mixture of trisialylated, triantennary N-glycans. A 2-pmol sample of the Glyko standard oligosaccharide mixture A3, derived from bovine thyroglobuline, was derivatized, cleaned up, and split in five equal samples for exoglycosidase array digestion before loading on a 36-cm 12% polyacrylamide sequencing gel and electrophoresis using the ABI 377 DNA sequencer. Peaks depicted in blue represent APTS-derivatized carbohydrates. Peaks depicted in red (at scan ± 6075 and scan ± 7800) represent components of the rhodamine-derivatized Genescan 500™ internal standard. In Panels a, the electropherograms obtained after digestion with different exoglycosidase arrays are shown. Panels b summarize the corresponding structures, as inferred from the known intact oligosaccharide structure and from the specificity of the exoglycosidases used. Panel 1a: malto-oligosaccharide sizing reference standard. Only the relevant size range of Glc₃ to Glc₁₂ is shown. Panels 2a,b: intact A₃ oligosaccharide mixture. Panels 3a,b: *Arthrobacter ureafaciens* sialidase digest. The origin of the two minor peaks at scan 6750 and 6950 is unclear but probably reflects the presence of an impurity in the original sample. These peaks disappeared completely on β-N-acetylgalactosaminidase digestion (Panel 5a). Panels 4a,b: *Diplococcus pneumoniae* β-1,4-galactosidase digest. Panels 5a,b: sialidase + β-1,4-galactosidase + jack bean β-N-acetylgalactosaminidase digest. Panels 6a,b: sialidase + β-1,4-galactosidase + β-N-acetylgalactosaminidase + jack bean α-mannosidase digest. The minor peaks originating from the Gal β-1,3 containing isomers are indicated with an arrow in the electropherograms.
Materials and methods

N-glycan standard mixture

The A3 N-glycan standard, containing a mixture of trisialylated triantennary complex type N-linked oligosaccharides derived from bovine thyroglobulin was obtained from Glyko, Novato, CA.

96-well deglycosylation procedure

The protocol was elaborated in detail by Papac et al. (1998). Briefly, the PVDF membrane at the bottom of the wells of a Multiscreen-IP plate (Millipore, Bedford, CA) was wetted with 100 µl methanol and washed three times with 300 µl of water and once with 50 µl of RCM buffer (8 M urea, 360 mM Tris, pH 8.6, 3.2 mM EDTA). The glycoprotein was loaded in the wells, containing 10 µl RCM buffer. Subsequently, additional RCM buffer was added to a minimal volume of 50 µl. The protein was bound to the membrane with a gentle vacuum. This step was followed by two washing steps with 50 µl RCM buffer. The bound protein was then reduced by the addition of 50 µl of 0.1 M dithiothreitol in RCM buffer and incubation at 37°C for 1 h. The reducing solution was removed by vacuum and the wells were washed three times with 300 µl of water. Carboxymethylation was performed by addition of 50 µl of 0.1 M iodoacetic acid in RCM buffer and incubation for 30 min at room temperature in the dark. After removal of this solution, three washes with 300 µl of water followed. The remaining protein binding capacity of the wells was blocked.
by incubation with 100 µl 1% polyvinylpyrrolidone 360 in water at room temperature for 1 h. Again three washing steps were performed as described above, followed by the addition of 1.25 Oxford Glycosystems units of PNGase F (Oxford Glycosystems, Abingdon, UK) in 20 µl of 10 mM Tris-acetate pH 8.3. Digestion is complete after a 3-h incubation at 37°C, after which the solution was transferred to a tapered-well microtiter plate.

96-well AG-50-WX8 cation exchange

If MALDI-TOF-MS of the analytes is required, a fraction of the deglycosylation mixture is treated with 150 mM acetic acid for 3 h at room temperature to ensure complete conversion of glycosylamines to the reducing saccharides. Subsequently, this mixture is applied to the wells of a Multiscreen-Durapore membrane-lined 96-well plate (Millipore), filled with AG-50-WX8 resin in the proton form (Biorad, Hercules, CA, USA). Plates are packed using a 100 µl Multiscreen Column Loader system (Millipore). In two rounds of resin loading, swelling and gentle centrifugation, microcolumns of about 300 µl packed resin are easily and reproducibly obtained. The cation exchange resin removes the protein and salt present in the deglycosylation mixtures with sufficient efficiency to allow direct MALDI-TOF-MS as described elsewhere (Papac et al., 1998).

APTS derivatization reaction

We have found it unnecessary to remove the PNGase F prior to derivatization with APTS, as this practice does not lead to the appearance of contaminant peaks in the size range of 3–25 glucose units. The deglycosylation mixture was evaporated to dryness at the bottom of the tapered well microtiterplate using a Savant vacuum centrifuge equipped for plates. Subsequently, a 1 µl 1:1 mixture of 20 mM APTS (Molecular Probes, Eugene, CA, USA) in 1.2 M citric acid and 1 M NaCNBH3 in DMSO was added to each well. After careful vortexing and short centrifugation of the plate, it was incubated upside down at 37°C overnight, tightly wrapped in parafilm. The following morning, the reaction was quenched by the addition of 10 µl of water. The malto-oligosaccharide size reference ladder was prepared as described (Kobata, 1994) and also labeled with APTS.

96-well Sephadex G10 postderivatization cleanup

The wells of a Multiscreen-Durapore membrane-lined 96-well plate (Millipore), were packed with Sephadex G10 (Pharmacia, Uppsala, Sweden) using the same procedure as described for AG-50-WX8 to reach a column height of 1.2 cm. It is essential that the microcolumns are centrifuged to dryness just prior to sample loading. After loading, the resin beds were eluted 4 times by addition of 10 µl of water and a 10-s centrifugation at 750 × g in a table-top centrifuge equipped for handling 96-well plates (Universal RF-30, Hettich, Tuttingen, Germany). Centrifugation conditions may need some optimization depending on the properties of the centrifuge used. The eluate was collected in another tapered-well microtiterplate and evaporated to dryness. Successful cleanup is hallmarkled by the detection of only faint fluorescence of the eluate on imaging on a standard UV-light box. After evaporation, the derivatized glycans were reconstituted in 5 µl of water.

Exoglycosidase digestions

Batches of 0.8 µl of the cleaned-up derivatized N-glycans were transferred to 250 µl PCR tubes or tapered-well microtiter plates for treatment with exoglycosidase arrays. In this study, all digestions were done by overnight incubation at 37°C in 10 µl 20 mM sodium acetate pH 5.5 containing the following enzyme mixtures: (1) *Arthrobacter ureafaciens* sialidase (2 µU/ml, Boehringer Mannheim, Germany); (2) sialidase and *Diplococcus pneumoniae* β-1,4-galactosidase (1 U/ml, Boehringer Mannheim); (3) sialidase, galactosidase, and jack bean β-N-acetylhexosaminidase (30 U/ml, Oxford Glycosystems); and (4) sialidase, galactosidase, N-acetylhexosaminidase, and jack bean α-mannosidase (100 U/ml, Sigma Biochemicals, Bornem, Belgium).

Preparation of the samples for gel loading

To each sample, 0.5 µl of the ROX-labeled Genescan™ 500 standard mixture (Perkin Elmer, Foster City, CA, USA) was added. Alternatively, we used a mixture containing 250 fmol each of a rhodamine-labeled 6-, 18-, 30-, and 42-meric oligonucleotide (consisting of repeats of the basic sequence 5′-TAC-3′, synthesized and PAGE-purified by Life Technologies, Merelbeke, Belgium). After addition of the internal standard, 1 µl of deionized formamide was added to facilitate sample loading.

Gel electrophoresis and data analysis

All experiments were performed on an Applied Biosystems 377A DNA sequencer (Perkin Elmer), equipped with an external cooling bath (model RTE 111, NESLAB, Porthsmouth, NH) kept at 23°C (easily connectable to the sequencer according to the ABI PRISM 377 DNA sequencer user bulletin “Modifications for subambient temperature operations”). Due to the high diffusional mobility of carbohydrates, we skipped one gel lane between each two samples, to avoid cross-contamination and to ease the lane tracking process. In the 36-well sequencing format, this allows analysis of 18 samples per run; in the 64-well format, 32 samples can be analyzed in parallel. The gel contained 12% of a 19:1 mixture of acrylamide:biacrylamide (Biorad) and was made up in the standard DNA-sequencing buffer (89 mM Tris, 89 mM borate, 2.2 mM EDTA). Polymerization was catalyzed by the addition of 200 µl of a 10% ammoniumpersulfate solution in water and 20 µl of N,N,N′,N′-tetramethylethylediamine. The gels were of the standard 36 cm well-to-read length throughout the study. Prerunning was done at 3000 V for 1 h. After prerunning the gel, the wells were thoroughly rinsed with the sequencing buffer and 1.8 µl of the samples was loaded. The electrophoresis voltage during separation was 4000 V and data were collected for 5 h (separation of glycans up to 25 glucose units in size). Data analysis was performed using the Genescan 3.1 software (Applied Biosystems) and the window for the lane tracker was set to 7, so that the fluorescence of the whole width of the bands was integrated. We used the same fluorescence-overlap correction matrix as for DNA sequencing using BigDye dye terminators on our machine. The fluorescence of APTS-derivatized carbohydrates and rhodamine-labeled oligonucleotides was readily resolved.
Acknowledgments

Nico Callewaert is a research assistant of the Fund for Scientific Research Flanders (FWO). S. Geysens holds a fellowship from the Institute for the Advancement of Scientific and Technological research in Industry (IWT).

Abbreviations

APTS, 8-amino-1,3,6-pyrenetrisulfonic acid; MALDI-TOF-MS, matrix-assisted laser desorption and ionization time-of-flight mass spectrometry; PAGE, polyacrylamide gel electrophoresis; PNGase F, peptide-N-glycosidase F; PVDF, polyvinylidene difluoride; RFU, relative fluorescence units.

References


