

Oligosaccharide sequencing technology

Recent advances allow any well-founded laboratory to carry out rapid and straightforward characterization of subpicomolar levels of *N*-glycans released from one to five microgram quantities of glycoprotein.

Pauline M. Rudd, Geoffrey R. Guile, Bernhard Küster, David J. Harvey, Ghislain Opendakker & Raymond A. Dwek

The glycosylation machinery in a cell is available to all proteins that enter the secretory pathway. Two major classes of oligosaccharides may be attached to proteins that contain the appropriate signals in their peptide chains. *N*-linked sugars may be added to the amide side-chain of some Asn residues, which form part of the triplet AsnXaaSer/Thr, while *O*-linked sugars may be added to the hydroxyl side-chain of some Ser or Thr residues. Glycoproteins generally consist of mixtures of glycosylated variants of a single polypeptide (glycoforms). The composition of the glycoform population is influenced by the 3-dimensional structure of the protein¹.

A full understanding of the implications of glycosylation can only be reached when the glycoprotein is viewed as an entity. While peptide sequencing is routinely available to protein chemists there is also a need for a robust, rapid and automated technology for oligosaccharide sequencing at the subpicomolar level, which is applicable to all glycoproteins.

Earlier developments in sequencing

The first predictive approach to oligosaccharide analysis was based on BioGel P4 gel permeation chromatographic separations of glycan pools and exoglycosidase sequencing of individual sugars² using enzyme arrays³. More recently, the higher resolving power of high-performance liquid chromatography (HPLC), first used for analysing sugars in 1975 (refs 4 and 5), has been exploited⁶⁻⁸. The use of 2- and 3-dimensional glycan analysis in which glycan pools are resolved into individual sugars by successive passages through anion-exchange, hydrophobic and hydrophilic interaction HPLC columns was pioneered mainly by Japanese scientists^{9,10}. The glycan structures are assigned by comparing the 2- or 3-D coordinates of the elution positions of the samples with those of known standards. Other separation techniques include capillary electrophoresis¹¹⁻¹² and polyacrylamide gel electrophoresis¹³. Mass spectrometry is an important method for determining oligosaccharide composition. NMR can determine unambiguously the complete structure of an oligosaccharide

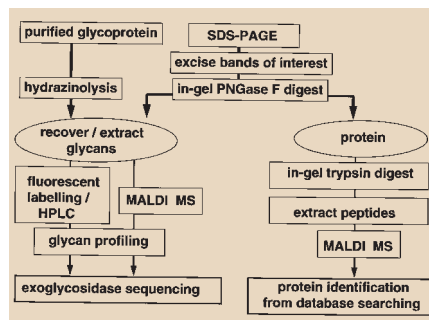


Figure 1 Strategies for glycan analysis and protein identification.

where sufficient material is available. Clearly oligosaccharide analysis is an advancing field to which many groups have contributed²⁻¹⁶.

A strategy for *N*-glycan analysis

The strategy illustrated in Fig. 1 specifically addresses the need for a rapid and robust method of *N*-glycan analysis, which can be applied routinely to microgram levels of glycoproteins with the minimum requirement for specialized equipment or expertise. The steps involved in oligosaccharide analysis include releasing the sugars from the protein, and labelling, separating and analysing the components of the glycan pool.

For samples where the protein is difficult to purify or amounts are limited, the *N*-glycans may be released directly from a band on an SDS-PAGE gel using peptide *N*-glycosidase F (PNGase F)¹⁷.

For detection in HPLC analysis, the non-reducing termini of the glycans are labelled with 2-aminobenzamide (2AB)¹⁸, to allow direct quantitation from the HPLC profiles. In addition, the reaction conditions retain sialic acid residues and the label is compatible with other analytical systems such as matrix-assisted laser desorption ionization mass spectrometry (MALDI MS).

An amide-silica HPLC column was selected because the system could be optimized to minimize interaction of the column with the fluorescent label. High-resolution separations of both charged and neutral *N*- and *O*-glycans were achieved in a single run. The system may be used preparatively, as well as analytically. A volatile buffer system (acetonitrile and ammonium formate) was selected so that individual glycans could be recovered free of buffer salts. A major advantage of this system is that the elution positions of the glycans may be used

to predict structure, as positive incremental values are obtained for the addition of all monosaccharide residues to basic core oligosaccharides^{19,20}.

Assignments may be confirmed by sequential exoglycosidase digestion using highly specific enzymes. The products of such digestions are oligosaccharide fragments from which the enzyme(s) have removed specific terminal residues. As a result of the high resolution of the HPLC system, a rapid sequencing method has now been developed. In this approach, five aliquots of the entire glycan pool are sequenced simultaneously using a standard panel of enzyme arrays, after which the products are analysed using the same HPLC system. Any unusual structures are highlighted since these will not be digested to a common trimannosyl *N*-linked glycan core by any of the arrays. These structures may be isolated from preparative HPLC runs for further analysis.

Predictive HPLC glycan profiling

A novel, sensitive and reproducible HPLC technology has been developed to resolve subpicomole levels of heterogeneous mixtures of sugars fluorescently labelled with 2-aminobenzamide in their correct molar proportions¹⁸.

N- and *O*-linked oligosaccharides released from glycoproteins are resolved on the basis of differences in hydrophilicity, which reflects arm specificity, linkage and monosaccharide composition. In contrast to earlier technology based on gel filtration², structures of both the neutral and charged components of an entire *N*-glycan pool may be predicted from a single HPLC run (Fig. 2a, b).

The structures of the glycans are predicted from their HPLC elution positions determined in glucose units with reference to a dextran ladder. Assignments are based on the elution positions of known glycans from which incremental values for the addition of monosaccharide residues to oligosaccharide cores have been calculated (Fig. 2b). The quality of the resolution and the ability to predict glycan structures allow the interpretation of the full oligosaccharide profiles of glycoproteins, which may contain 30 or more structures at each glycosylation site (Fig. 3).

A major advantage of all HPLC systems is that they can be used preparatively. This means, for example, that the glycans in individual peaks can be collected, the

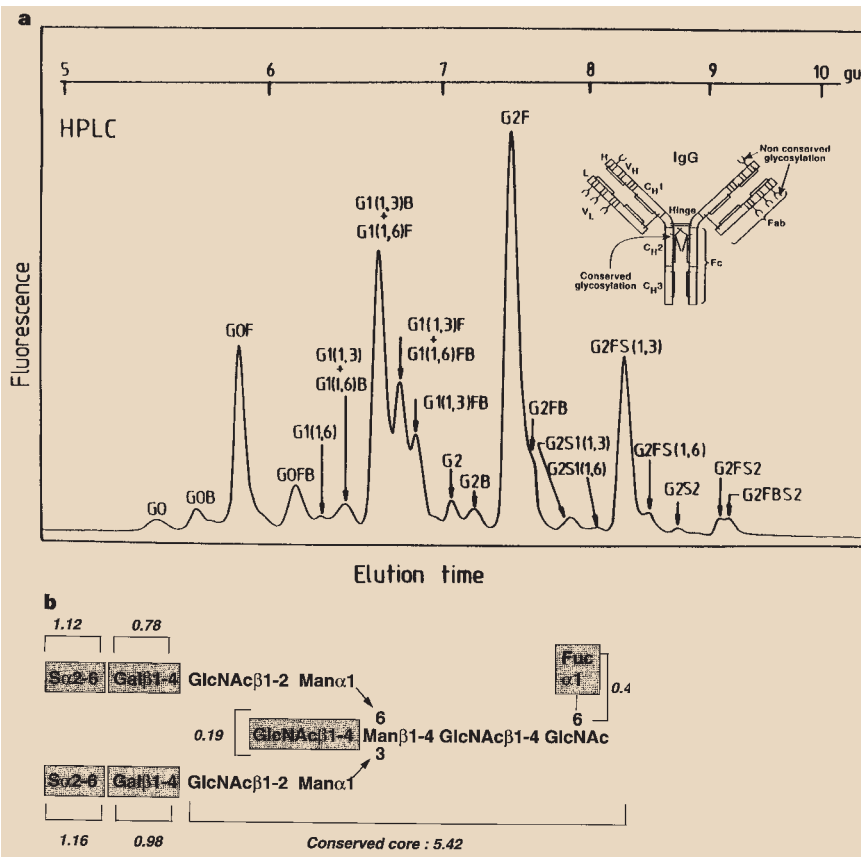


Figure 2a, Normal phase HPLC separation of glycans from normal human serum IgG. Peaks were assigned normal glucose unit values by comparison with a standard dextran ladder, the elution positions of which are shown at the top of the figure. The inserted diagram of IgG indicates the conserved *N*-glycosylation sites in the Fc and shows variable glycosylation sites in the Fab. (Nomenclature for describing oligosaccharide structures: A(1,2) indicates the number of antennae linked to the trimannosyl core; G(0–2) indicates the number of terminal galactose residues in the structure; F: fucose; B: bisecting *N*-acetylglucosamine (GlcNAc); S: sialic acid; Gal: galactose.) b, Complex biantennary sugar from which the set of IgG sugars is derived. The presence or absence of the boxed monosaccharide residues gives rise to the heterogeneity seen in panel a. The elution positions depend on the hydrophilic interactions of each oligosaccharide with the HPLC column and are sensitive to the environment of the monosaccharide residues within the glycan structure. This gives the column its fine specificity. For example, the incremental values for the addition of galactose to the α 1,3 arm differs from the addition of the same residue to the α 1,6 arm. Tables of incremental values are given in refs 19,20. These form the basis of the HPLC predictive technique. To illustrate this, incremental values for the addition of the different monosaccharide residues to the trimannosyl core to obtain the fucosylated, bisected, disialylated sugar are shown in this figure.

volatile buffers removed, and the glycans analysed by mass spectrometry or capillary electrophoresis to validate assignments. In cases where unusual or novel structures are indicated, further information can also be obtained by 2-D chromatography. For example, glycan pools containing sulphated sugars isolated from normal phase separations may be further resolved by weak anion exchange chromatography²¹.

The new technology enables a detailed comparison of glycosylation profiles and has been used, for example, to compare IgG glycans from normal and rheumatoid IgG (ref. 22). Predictive HPLC analysis can be used to gain an overview of the entire glycosylation

of a glycoprotein (Fig. 3). Rapid profiling enables recombinant glycoproteins to be monitored during cell culture. Furthermore, the ability to resolve neutral and acidic glycans simultaneously is particularly useful where acidic sugars are likely to play a functional role as, for example, in the brain²³. Current studies are focused on extending the methodology to *O*-glycans which are resolved by the same system.

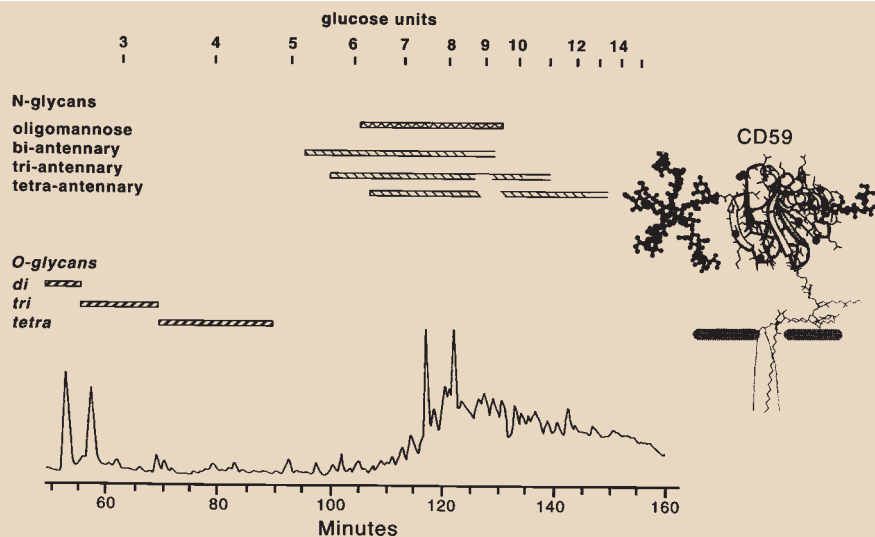
Simultaneous analysis of glycan pools

Further structural information can be obtained rapidly by simultaneously sequencing all of the oligosaccharides in a glycan pool. Enzymatic analysis of oligosaccharides

using highly specific exoglycosidases is a powerful means of determining the sequence and structure of glycan chains. Until recently, this has involved the isolation of single sugars from the glycan pool. It is often impractical and always time consuming to purify each individual sugar to the necessary 80% purity.

The new approach involves the simultaneous digestion of aliquots of a total pool of fluorescently labelled oligosaccharides with a series of multiple enzyme arrays²⁴. The products are analysed by normal phase HPLC or by MALDI MS, and peaks are assigned from a knowledge of the specificity of the enzymes and the incremental values of individual monosaccharide residues¹⁹. The rapid profil-

Figure 3 Normal phase HPLC profile of fluorescently labelled glycans released from human erythrocyte CD59. The elution positions of a ladder of glucose oligomers are shown in the scale above the oligosaccharide profile. This scale is used to assign glucose unit values to the peaks in the chromatogram^{19,20}. The bar chart above the chromatogram has been constructed using the elution positions of standard sugars and can be used to obtain a rapid overview of glycosylation profiles. The hatched bars indicate (i) the elution range for *N*-linked oligomannose sugars (GlcNAc2Man5–9) and (ii) the elution ranges for *N*-linked bi-, tri- and tetra-antennary complex glycans. The open-ended bars indicate that further substitutions (including lactosamine extensions) result in sugars eluting with larger gu values; (iii) the elution ranges for di-, tri- and tetra-saccharides typically associated with *O*-glycans.



ing and simultaneous analysis of the 36 major glycans attached to human erythrocyte CD59 has been carried out in this manner²⁰.

The glycosylation of microgram amounts of a protein can now be determined within 2 to 3 days. For example, the glycans attached to human neutrophil gelatinase B have been simultaneously sequenced using multiple enzyme arrays (Fig. 4a). The purified protein sample may be submitted to hydrazinolysis to release the sugars, or the glycans can be released directly from SDS-PAGE gels, simplifying the purification as illustrated below.

Sugars from protein in SDS gels

A novel 'in-gel release' method has been developed to obtain N-linked oligosaccharides directly from 1-D gels of glycoproteins available only in microgram amounts¹⁷. The pools of charged and neutral sugars can be characterized using the new HPLC technology and neutral and sialylated sugars may also be analysed by MALDI MS. Moreover, the protein remaining in the gel can be used for peptide sequencing following further in-gel digestion with trypsin.

For example, gelatinase B is an inducible metalloproteinase involved in the remodelling of the extracellular matrix. Depending on the producer cell type and the post translational modifications various molecular forms of human gelatinase B have been detected. Purified enzyme preparations of natural gelatinase B are available only in microgram amounts. Purification usually involves substrate affinity chromatography on gelatin-sepharose. However, purified preparations of neutrophil gelatinase B may contain several molecular variants. These cannot be separated by conventional means but can be resolved on SDS gels.

Figure 4b shows a gel of gelatinase B and alongside it the zymogram that indicates which bands contain active enzyme. The 92K band was excised and the glycans released from the gel piece with PNGase F and analysed by MALDI MS and HPLC. The data were consistent with the HPLC analysis of the N-glycans of neutrophil gelatinase B (Fig. 4a).

Developing future applications

This method can be used where limited quantities of material and protein purification problems have previously precluded a detailed glycan analysis. Reducing gels allow the straightforward separation of proteins into their component subunits. For example, IgG can be resolved into heavy and light chains¹⁷ and the surface coat proteins of hepatitis B virus or particle can be separated into three major glycoprotein components (S, M and L)²⁵. Also, when gelatinase B from some cell types is purified on gelatin-Sephrose it is contaminated with its natural inhibitor (TIMP-1)²⁶. This enzyme-inhibitor

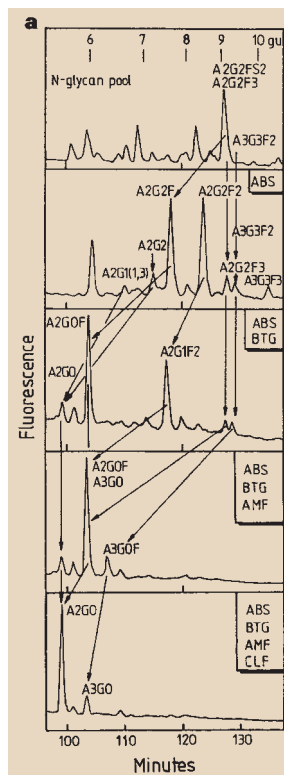
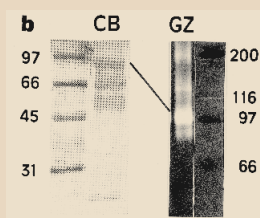


Figure 4a Simultaneous exoglycosidase sequencing of the neutrophil gelatinase B N-glycan pool using the enzyme arrays shown in the panels. Some of the major peaks have been annotated and the arrows indicate the movement of peaks following digestion with the enzyme arrays. The structures of the oligosaccharides are determined from the known incremental values for monosaccharide additions to glycan cores and the specificity of the exoglycosidase enzymes in the arrays. The assignments were consistent with the MALDI MS analysis. (ABS: *Arthrobacter ureafaciens* sialidase, BTG: bovine testis galactosidase, AMF: Almond meal fucosidase, CLF: *Choronia lampas* fucosidase.) b, SDS-PAGE of human neutrophil gelatinase B. Gel electrophoresis was used as the final purification step. Track 1 contains molecular weight markers, track 2 (CB) contains human neutrophil gelatinase B stained with Coomassie blue, track 3 (GZ) shows the result of gelatinase zymography²⁶ used to identify the band containing the full-length active enzyme (92K)²⁷. This band was excised from the gel for glycan profiling using the method described in ref. 17.



complex is stable and SDS-PAGE is used to resolve the two proteins. N-linked oligosaccharides can then be released directly from the gel with PNGase F.

In principle, the method could also be applied to 2-D gels. Glycan analysis will complement protein sequencing in the Proteome Project in which 2-D gels are used for comparing patterns of proteins from cells, tissues or body fluids. Gels of normal and diseased tissue may be compared using computer-controlled imaging techniques, which can identify 'spots' present in one sample but not the other. The possibility of sequencing sugars attached to proteins in such gel 'spots' suggests that this technology may lead to an advance in understanding the role of glycosylation in disease.

State of the art

These developments represent a dramatic advance in the technology available for oligosaccharide analysis. The minimum requirement for a laboratory is for an HPLC system equipped with a normal phase column and a fluorescence detector, a series of glycan standards and a set of endo- and exoglycosidase enzymes. Indeed, many well-founded laboratories may already be in a position to characterize in detail the glycans attached to the glycoproteins they are studying. □

The authors are at the Glycobiology Institute, Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, UK. G.O. is also at Rega Institute for Medical Research, University of Leuven, Minderbroedersstraat 10, B-3000 Leuven, Belgium.

The analyses referred to in this article were carried out in the Glycobiology Institute using an HPLC system consisting of a Douglas

Scientific DG604 degasser, two Waters 510 pumps, a Waters 717 autoinjector, a Waters temperature control module and a Jasco FP-920 fluorescence detector. The system was controlled via a Waters LAC/E box using Waters Expertease 3.1 software running on a DEC VAX 4000-200 computer. The normal phase, and weak anion exchange HPLC columns (GlycoSep N and GlycoSep C, respectively) were obtained from Oxford GlycoSciences (OGS) UK. Fluorescent labelling with 2-aminobenzamide was carried out using a Signal labelling kit, glycans were released by hydrazine using the GlycoPrep 1000 (OGS) or by PNGase F. Mass spectrometry was carried out using a PerSeptive Biosystems Voyager Elite MALDI TOF mass spectrometer.

- Rudd, P.M. & Dwek, R.A. *Crit. Rev. Biochem. Mol. Biol.* 32, 1–100 (1997).
- Kobata, A. *Anal. Biochem.* 100, 1–14 (1979).
- Edge, C.J. *et al. Proc. Natl. Acad. Sci. USA* 89, 6388–6342 (1992).
- Linden, C. & Lawhead, C.L. *J. Chromatogr.* 105, 125 (1975).
- Palmer, J.K. *Anal. Lett.* 8, 215 (1975).
- Alpert, A.J. *et al. Chromatogr. A.* 676, 191–203 (1994).
- Hase, S. *Methods Enzym.* 230, 225–237 (1994).
- Mellis, S.J. & Baenziger, J.U. *Anal. Biochem.* 134, 442–449 (1983).
- Tomiyama, N. *et al. Anal. Biochem.* 193, 90–100 (1991).
- Takahashi, N., Nakagawa, H., Fujikawa, K., Kawamura, Y. & Tomiyama, N. *Anal. Biochem.* 226, 139–146 (1995).
- Kakehi, K. & Honda, S.J. *J. Chromatogr. A.* 720, 377–393 (1996).
- Hermentin, P. *et al. Anal. Biochem.* 221, 29–41 (1994).
- Jackson, P. *Anal. Biochem.* 196, 238–244 (1991).
- Lennarz, W.J. & Hart, G.W. eds., *Meth. Enzym., Guide to techniques in glycobiology* 230, 1–520 (Academic, San Diego, California, 1994).
- Dwek, R.A., Edge, C.J., Harvey, D.J., Wormald, M.W. & Parekh, R.B. *Ann. Rev. Biochem.* 62, 65–100 (1993).
- Rudd, P.M. & Dwek, R.A. *Cur. Opin. Biotechnol.* 8, ##–## (1997).
- Küster, B., Wheeler, S.F., Hunter, A.P., Dwek, R.A. & Harvey, D.J. *Anal. Biochem.* 250, 82–101 (1997).
- Bigge, J.C., Patel, T.P., Bruce, J.A., Goulding, P.N., Charles, S.M. & Parekh, R.B. *Anal. Biochem.* 230, 229–238 (1995).
- Guile, G.R., Rudd, P.M., Wing, D.R., Prime, S.B. & Dwek, R.A. *Anal. Biochem.* 240, 210–226 (1996).
- Rudd, P.M. *et al. J. Biol. Chem.* 272, 7229–7224 (1997).
- Guile, G.R., Wong, S.Y.C. & Dwek, R.A. *Anal. Biochem.* 222, 231–235 (1994).
- Wormald, M.W. *et al. Biochemistry* 36, 1370–1380 (1997).
- Wing, D.R. *et al. Glycoconjugate J.* 9, 293–301 (1992).
- Iourin, O. *et al. Glycoconjugate J.* 13, 1031–1042 (1996).
- Mehta, A., Lu, X., Block, T., Blumberg, B. & Dwek, R. *Proc. Natl. Acad. Sci. USA* 94, 1822–1827 (1997).
- Opendakker, G., Masure, S., Proost, P., Billiau, A. & Van Damme, J. *FEBS Letts.* 284, 73–78 (1991).
- Masure, S., Proost P., Van Damme, J. & Opendakker, G. *Eur. J. Biochem.* 198, 391–398 (1991).

Acknowledgements: We thank Professor E. M. Southern and Dr Taj Mattu.