Identification and quantification of N-linked oligosaccharides released from glycoproteins: An inter-laboratory study

Smita Thobhani1, Chung-Ting Yuen3, Marc J A Bailey2, and Christopher Jones1,3

2Analytical Science, National Physical Laboratory, Teddington, Middlesex, TW11 0LW, UK; and 3National Institute for Biological Standards and Control, Blanche Lane, South Mimms, Hertfordshire, EN6 3QG, UK

Received on March 7, 2008; revised on September 5, 2008; accepted on September 28, 2008

As characterization of glycosylation is required for the licensing of recombinant glycoprotein therapeutics, technique comparability must be assessed. Eleven UK laboratories (seven industrial, two regulatory or government, two academic) participated in an inter-laboratory study to analyze N-glycans present in four mixtures prepared by PNGase F cleavage of commercial glycoproteins: human α1-acid glycoprotein (Hα1), bovine α1-acid glycoprotein (Bα1), bovine pancreatic ribonuclease B (RNaseB), and human serum immunoglobulin G (hIgG). Participants applied their routine glycan mapping methodology using predominantly chromatography and mass spectrometry to identify and quantify components. Data interpretation focused on the relative amounts of different glycan structures present, the degree of sialylation, antennary and the galactosylation profiles, fucosylation and bisecting GlcNAc content, and the number of glycan components identified. All laboratories found high levels of sialylation for Hα1 and Bα1 (Z-numbers 271 ± 24 and 224 ± 18, respectively), but varying ratios of di-, tri-, and tetra-antennary chains. The Z-score for hIgG glycans had high variability as values obtained from mass spectrometric and chromatographic methods clustered separately. The proportion of the major penta-mannosyl chain from RNaseB was between 29 and 62%. Proportions of fucosylated and bisected GlcNAc chains from hIgG were between 58 and 96% and 9 and 23%, respectively. Mass spectrometric approaches consistently identified more glycan species, especially when both N-glycolylneuraminic acid (Neu5Gc) and N-acetylneuraminic acid (Neu5Ac) were present. These data highlight the need for well-characterized reference standards to support method validation and regulatory guidance on selection of approaches. Pharmacopoeial specifications must acknowledge method variability.

Keywords: glycan analysis/glycoproteins/IgG/N-glycans/sialylation

Introduction

An increasing number of biopharmaceutical products are glycoproteins, including monoclonal antibodies (mAbs), glycoprotein hormones, and proteins from the coagulation cascade. As the variability inherent in protein glycosylation creates a heterogeneous product, regulatory guidelines, such as those from the International Conference on the Harmonization (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use (Anonymous), require characterization of the glycosylation. Consistency batches are used to quantify lot-to-lot variation, and comparability of glycosylation must be demonstrated if manufacturing processes change. If glycosylation is a critical quality parameter, uniformity of appropriate aspects of the glycosylation profile will be required for lot release of individual batches. Key aspects include the degree of sialylation, or fucosylation, and galactosylation (ratio of G0, G1, and G2) of glycans chains in mAbs. The wide range of methodologies and analytical platforms, and the challenge of characterizing heterogeneous mixtures of complex glycans create comparability issues that need to be assessed to support development of appropriate standards for regulatory purposes.

Whilst glycosylation can be analyzed at the level of the intact glycoprotein, or through coupled HPLC-mass spectrometric analysis of peptide/glycopeptides mixtures, the favored approach remains separation, identification, and quantitation of released oligosaccharides. Under the Human Proteome Organization (HUPO) Human Disease Glycomics/Proteome Initiative (HGPI), Wada and colleagues (2007) compared methods for profiling glycans applicable for diagnostic purposes. Twenty laboratories, of which over 80% were academic, analyzed two human glycoproteins: serum immunoglobulin G (IgG) and transferrin obtained from healthy individuals. Matrix-assisted laser desorption-ionization (MALDI) mass spectrometry (MS) of permethylated glycans (six laboratories) was shown to yield consistent quantitation which correlated to that from chromatography of fluorophore-labeled glycans (five laboratories). Although the study was thorough and critical, the choice of unique noncommercial glycoproteins hinders further comparability studies. Data analysis concentrated on quantification of a relatively small number of glycans relevant to diagnosis. The participation of specialist carbohydrate institutes may not reflect the status of glycoprotein analysis in the biopharmaceutical industry. As participating laboratories both released and analyzed N-glycans, the sources of uncertainty cannot be isolated.

The study reported here was designed (i) to isolate variability in analytical approaches by distributing released N-linked glycans, (ii) to focus on key factors affecting clinical efficacy of biopharmaceutical products, and (iii) to compare glycan profiles rather than a subset of components. Participants, drawn principally from the UK biopharmaceutical industry, were asked to
analyze these samples using their regular in-house procedures. Use of abundant commercial glycoproteins allows glycos to be selected on the basis of pharmaceutical relevance and sufficiently well-characterized glycan reference material to be prepared for future work and for parallel analyses to assess the influence of cleavage and glycan isolation approaches on the overall analytical error.

The study highlighted the commonly used approaches in the biopharmaceutical community and added focus in the development of standards and reference materials in this area, including the elaboration of monographs on oligosaccharide analysis for both the European and United States Pharmacopoeias.

Results

Study design, choice of model glycoproteins, and sample preparation

The study comprised two components: a questionnaire and an experimental aspect. The questionnaire provided information on the glycan analysis methodology employed, whilst the experimental part assessed comparability of the participants’ ability to identify individual components and to quantify sialylation, the antennary structures present, mannosyl content and the different galactosylated, fucosylated, and bisecting N-acetylgalactosamine (GlcNAc) profiles. Laboratories were prepared with four samples of purified N-glycan mixtures (approximately 25 µg of each) with free reducing termini for derivatization and a detract ladder, a commonly used reference material. The identities of the glycoproteins used to prepare the N-glycans were not disclosed to the participants until after the study.

Four N-glycan mixtures were prepared by PNGase F digestion of human α1-acid glycoprotein (Hα1), bovine α1-acid glycoprotein (Bα1), bovine pancreatic ribonuclease B (RNaseB), and human immunoglobulin G (HIgG). Hα1 is heavily glycosylated and contains Neu5Ac terminated complex-type bi-, tri-, and tetra-antennary N-linked glycans. On the other hand, Bα1 contains, principally, biantennary complex chains carrying both Neu5Ac and Neu5Gc, which are not expressed in man and are unwelcome in biopharmaceuticals. The high levels of sialylation and multiantennary chains mimic those from recombinant glycoproteins, such as erythropoietin and glycoprotein hormones, in which the degree of sialylation is a key assay, as sialylation controls biological activity through modulation of the product’s in vivo half-life. RNaseB contains neutral N-glycans of the high-mannose-type, which are typically present on glycoproteins produced from insects, fungal (yeast), plant and mammalian systems. Serum HIgG derived glycans were included as antibodies are the fastest developing class of therapeutic glycoproteins. The conserved N-glycan chains in the Fc domain, which include highly fucosylated, complex bi-antennary glycan chains with varying degrees of galactosylation and some minor non-fucosylated and/or with bisecting GlcNAc residues, have a role in modulating the downstream functions of antibodies (Arnold et al. 2007). For example, correct glycosylation of the Fc region enables effector ligands to bind hIgG and activate antigen clearance mechanisms: a process known as “antibody-dependent cell cytotoxicity” (ADCC) (Jeffers et al. 1998; Jefferis 2005). Some hIgG molecules also possess sialylated glycan chains associated with other antibody domains. Figure 1 illustrates some typical N-glycan structures found on the glycoproteins described above.

Study methodology and data analysis approach

Eleven UK laboratories took part in the study, including seven biopharmaceutical manufacturers and analytical contract laboratories. Data obtained from the participants were anonymized and laboratories designated by a letter in no particular order. A standardized excel spreadsheet with a common glycan nomenclature was provided to aid quantitative reporting of the observed glycans as relative percentages. The spreadsheet was divided into five sections for reporting high-mannose chains, hybrid chains, and complex bi-, tri-, and tetra-antennary chains. Each section listed the major N-glycans that could occur. Values for the different glycan structures were reported as relative percentages to the total amount of glycan components found. Laboratories performed glycan analysis using their routine methods and techniques. Normalization of values received from participants was not performed. Samples refer to the materials analyzed and not the number of laboratories taking part in the study.

Chromatographic and MS analyses were the principal strategies employed by the participants, as shown in Table I. Seven of the eleven laboratories carried out chromatographic analysis, using either 2-aminobenzoic acid (2-AA; 3 labs), 2-amino benzamide (2-AB; 3 labs), or 4-aminobenzoic acid (4-AA; 1 lab) as the fluorophore label to quantify the relative molar amounts of glycans. Samples were also analyzed by MS (6 labs), either as underivatized, permethylated, methylated, fluorophore-labeled material, or as desialylated glycans. Relative abundance of glycans for MS was determined through signal intensity. In general, a combination of techniques was employed in most laboratories to identify glycan structures. One laboratory used capillary electrophoresis (CE) for analysis but these data were not complete and could not be included in our main dataset. High-performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) was not employed by any of the participants in this study.

Sample stability and assessment of intra-laboratory data variability

To assess sample stability during distribution and storage, two lots of three of the samples (RNaseB, HIgG, and Hα1) were sent from NIBSC to the USA and back by ordinary air mail post, a journey lasting 10 days. These samples were then fluorophore-labeled and analyzed by normal-phase Amide-80 HPLC against two lots of the same batches of material which had been stored at −20°C. No difference was observed in the Amide-80 HPLC profiles between the stability test and reference samples (Figure 2). The HPLC profiles show consistency in preparing aliquots of the samples and their stability during transfer. The study also shows a high level of consistency in derivatization and glycan profiling when carried out in a single laboratory. Typical CV% for the retention times and peak areas is about 3%. Since the most labile feature of the glycans was considered to be sialylation, the profile was determined using anion exchange chromatography (AEX) and variation in Z-number was determined. Z-number is a parameter characteristic of all charged glycans defined as the sum of the percentage contents of neutral, mono-, di-, tri-, and tetrasialylated glycans multiplied by the number of charges:

\[ Z = (\% \text{ neutral} \times 0) + (\% \text{ mono-} \times 1) + (\% \text{ di-} \times 2) + (\% \text{ tri-} \times 3) + (\% \text{ tetra-} \times 4). \]
Results showed highly consistent Z-numbers for samples with moderate to high contents of acidic N-glycans: Ha1 (Z = 290.5, 2.7% CV) and hlgG (Z = 39.5, 2.0% CV) indicating high stability and consistency of the N-glycans. However, relatively high variation was observed for RNaseB (Z = 3.5, 14.5% CV) and the reason is likely that it contains mainly neutral N-glycans (~97%) and the presence and detection of minor acidic N-glycan components are uncertain.

**Acidic glycan (sialylation) profile**

Participants generally used some form of AEX to determine the proportions of various acidic glycans. Each participant (labs A to J) showed similar trends in differentiating between the different levels of sialylation in the four N-glycan mixtures. The large majority of the glycans from Ha1 and Bα1 were sialylated, with values of 99.6 ± 0.6% and 98.8 ± 3.1% (mean ± standard deviation (SD)), respectively. The glycans from RNaseB contained virtually no acidic species, whilst a significant proportion of those from hlgG were sialylated (22.6 ± 9.8%), which is in accord with the literature (Raju et al. 2000). Data from this sample showed the most variation between the 10 participants returning data.

Where MS was employed, the percentage of sialic acid-containing glycans varied relatively little from the chromatographic data for the highly sialylated materials Ha1 and Bα1 (Table II). On the other hand, the data for the hlgG glycans suggested that the two approaches give different values (mean ± SD): 10.1 ± 3.0% for MS (n = 3) compared to 28.0 ± 5.3% for chromatography (n = 7), where n is the number of laboratories. This variation may reflect analyses...
of native oligosaccharides, where charges cause differential ionization efficiency as well as fragmentation complication. The types of MS (MALDI or ESI) employed could also be another factor. In a recent study, MALDI-TOF MS analysis of permethylated glycans yielded sialylation levels comparable to those obtained by chromatographic methods (Wada et al. 2007).

The samples from Hα1 and Bα1 contain differing amounts of mono-, di-, tri-, and tetrasialylated glycans, and laboratories were asked to establish the distribution of these (Figure 3). The Hα1 sample (chart A) contained glycans decorated with one to four sialic acid residues, with the monosialylated as the minor species. Most participants found trisialylated glycans as the major species (44.0 ± 8.4%): the two exceptions were labs C and I, which reported disialylated species as the major components. Both laboratories analyzed fluorophore-labeled material: one group by MS and the other by HPLC so no conclusion could be drawn. For the Bα1 sample (chart B), all participants observed disialylated structures as the major class (62.1 ± 11.1%). Analyses from lab F showed that the sialylation profile varied between normal-phase HPLC and DEAE approaches, with the normal-phase technique returning lower sialylation values (data not shown).

Five participants (labs A, D, H, G, and I) using mass spectrometry reported that sample Bα1 contained both Neu5Ac and Neu5Gc residues. These participants detected ions 16 Da units greater for Neu5Gc than for Neu5Ac. Although the study protocol did not request sialic acid residues to be released and identified, published approaches using both reversed-phase chromatography and HPAEC exist (Hara et al. 1987). The evidence from this dataset is inconclusive whether labeling of the glycans, with a fluorophore or by methylation, aided the detection of Neu5Gc compared to underivatized glycans. In general, better resolution and signal-to-noise ratio of spectra were generated from derivatized glycans for the profiling of the highly sialylated N-glycans of Hα1 and Bα1.

Monosialylated structures predominated over disialylated ones in the hlgG sample (Figure 4); however, large quantitative variation was observed. Mean ± SD calculated relative values for mono- and disialylated glycans were 15.7 ± 6.6% and 6.9 ± 3.9%, respectively. No pattern was established between chromatographic and mass spectrometric methodologies with respect to the ratio of sialylated structures.

The overall charge for samples Hα1, Bα1, and hlgG could be evaluated using the Z-number. The Z-values for laboratories A to J, depicted in Figure 5, showed a higher mean Z-number for the Hα1 sample than for the Bα1 sample: 271 ± 24 and 224 ± 18, respectively, with the mean value for hlgG being 29 ± 14. The greatest variation in the Z-number was observed for hlgG, with a CV of 47.3% as compared to 8.9 and 8.1% for the Hα1 and Bα1 samples, respectively. This observation could be accounted for by the presence of a mixture both sialylated and nonsialylated glycans for hlgG, which proved more challenging to quantify by MS alone in this study.

### Table II. Comparison of the percentage of sialylated glycans (mean ± SD) for Ha1, Ba1, and hlgG analyzed by MS and HPLC

<table>
<thead>
<tr>
<th></th>
<th>Hα1 (%)</th>
<th>Bα1 (%)</th>
<th>hlgG (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS</td>
<td>99.6 ± 0.8</td>
<td>96.7 ± 5.8</td>
<td>10.1 ± 3.0</td>
</tr>
<tr>
<td>HPLC</td>
<td>99.7 ± 0.5</td>
<td>99.7 ± 0.4</td>
<td>28.0 ± 5.3</td>
</tr>
</tbody>
</table>

**Bi-, tri-, and tetra-antennary profiles (Ha1 and Ba1)**

The data on the antennary profiles received from participants were limited. Only results where sufficient information was reported are discussed (six laboratories in total). Figure 6 represents the antennary profile for the highly sialylated Hα1 and Bα1 samples. Material was analyzed as underivatized (labs B
Characterization of N-glycans

Fig. 3. Bar chart showing the relative quantities of sialylated structures found in the Hα1 (chart A) and Bα1 (chart B) samples. The symbol * indicates where MS only was used.

Fig. 4. Bar chart showing relative abundance of the sialylated structures found in hIgG. The symbol * indicates where MS only was used.

Fig. 5. Calculated Z-number for the sialylated materials Hα1, Bα1, and hIgG. The symbol * indicates where MS only was used.

and H), methylated (labs A and G), fluorophore-labeled (labs D and I) or desialylated (neutral) (lab B), with HPLC and MALDI-TOF MS, the major techniques employed by participants. Sample Hα1 was found to contain a mixture of bi-, tri-, and tetra-antennary glycans, with the triantennary structure found as the most abundant species by most participants. On the other hand, for Bα1, bi-antennary chains were the predominant class and tetra-antennary structures were not observed. “Other” refers to structures that were either of hybrid type or were not identified. All SDs for the mean percentage of different antennary structures were below ± 10%. The CVs for sample Bα1 showed good agreement within the dataset with a maximum value of 4.2%.
which we consider surprising as this material was predicted the smallest value at 10.5%. In general, the variability is large, the largest due to its relative low abundance, with M6 having the smallest deviation; however, the CV for M4 was largest variation (SD) observed was for M5 (\(29.0 \pm 6.5\)). Laboratories using MS techniques (labs A, D, H, I) observed that the RNaseB sample also contained high-mannose and/or hybrid oligosaccharides. A small percentage of triantennary structures (1.5 ± 0.9%) was also reported by these participants. The proportion of species containing different numbers of terminal galactosyl residues (G0, G1, and G2) is often employed to monitor the glycosylation pattern of recombinant monoclonal antibodies, and also to monitor disease status in rheumatoid arthritis, which is linked to changes in IgG glycosylation (Youings et al. 1996). The bi-antennary structures were evaluated, wherever possible, with regard to their galactosylation profile, and results from nine laboratories are displayed in Figure 8. Where laboratories distinguished between the two monogalactosylated isomers, the values were combined and cited as G1.

Mannosyl profile of the neutral glycans (RNaseB)

RNaseB contains mainly neutral oligosaccharides of the high-mannose type. A breakdown (where possible) of the different mannosylated forms found by participants (labs A–I) is shown in Figure 7. Profiling was performed either through normal-phase HPLC or MALDI-TOF MS. Each of the nine laboratories reported the most abundant glycan components to be those with five mannose units (M5; 39.7 ± 10.3%), and those with six mannose residues (M6) the next largest group. Minor glycan species identified by participants contained four mannose units (M4; 2.1 ± 2.0%). Laboratories using MS techniques (labs A, D, H, I) observed that the RNaseB sample also contained high-mannose N-glycans lacking an N-acetylglucosamine residue at the reducing terminus. The appearance of these species in the glycan mixture could be accounted for by the presence of contaminating endo-glycosidase Endo-H activity in the PNGase F, which had been purified from Flavobacterium meningosepticum. Laboratories B, D, G, H, and I, using MS, also identified the presence of up to 10.5% of complex/hybrid-type “other” N-glycans.

The relative percentages of M4, M5, M6, M7, M8, and M9 components found for RNaseB are summarized in Table III. The largest variation (SD) observed was for M5 (±10.3%), with M4 having the smallest deviation; however, the CV for M4 was the largest due to its relative low abundance, with M6 having the smallest value at 10.5%. In general, the variability is large, which we consider surprising as this material was predicted to be the easiest to identify and has the least dissimilarity as it consists almost entirely of neutral N-glycans. Unfortunately, the source of the observed variation could not be attributed to specific techniques, due to the small dataset.

Table III. Summary of the high-mannose structures found in RNaseB

<table>
<thead>
<tr>
<th></th>
<th>Min (%)</th>
<th>Max (%)</th>
<th>Mean (%)</th>
<th>Median (%)</th>
<th>SD (%)</th>
<th>CV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M4</td>
<td>0</td>
<td>5.6</td>
<td>2.1</td>
<td>1.9</td>
<td>±2.0</td>
<td>95.6</td>
</tr>
<tr>
<td>M5</td>
<td>29.0</td>
<td>62.5</td>
<td>39.7</td>
<td>38.5</td>
<td>±10.3</td>
<td>25.9</td>
</tr>
<tr>
<td>M6</td>
<td>21.3</td>
<td>28.4</td>
<td>24.7</td>
<td>23.4</td>
<td>±2.6</td>
<td>10.5</td>
</tr>
<tr>
<td>M7</td>
<td>4.4</td>
<td>12.5</td>
<td>8.7</td>
<td>8.6</td>
<td>±2.6</td>
<td>29.7</td>
</tr>
<tr>
<td>M8</td>
<td>6.0</td>
<td>20.4</td>
<td>12.5</td>
<td>13.7</td>
<td>±4.4</td>
<td>35.6</td>
</tr>
<tr>
<td>M9</td>
<td>1.8</td>
<td>10.8</td>
<td>5.7</td>
<td>5.5</td>
<td>±2.6</td>
<td>45.5</td>
</tr>
</tbody>
</table>

\(n = \) number of laboratories.

Galactosylated, fucosylated, and bisecting GlcNAc profiles (hlgG)

Analysis of the hlgG sample showed that approximately 89% of the material consisted of complex bi-antennary structures with varying degrees of galactosylation. Laboratories D (9.0%), H (1.8%), and I (2.1%), which all used MS, also reported the presence of the high-mannose and/or hybrid oligosaccharides. Another source of variation may be due to the determination of the source of the observed variation could not be attributed to specific techniques, due to the small dataset.

Oligosaccharide structures found on hlgG frequently contain additional core-linked fucose and bisecting GlcNAc units. Nonfucosylated therapeutic antibodies are of interest to industry, as this modification enhances antibody-dependent cellular cytotoxic activity (Satoh et al. 2006; Jefferis 2007), whilst the presence of bisecting GlcNAc inhibits the addition of fucose (Ferrara et al. 2006). The prevalence of these modified...
structures was therefore evaluated, where sufficient data were available, and their relative abundance is shown in Figure 8. The occurrence of bisecting GlcNAc was significantly lower than core fucosylation in all cases. Mean percentage ± SD of GlcNAc and fucose found was 16.9 ± 5.8% and 80.9 ± 12.5%, respectively.

The ratios of fucosylated and bisecting GlcNAc species found in the three galactosylated structures are shown in Table V. The proportion of chains which were fucosylated decreased as the level of galactosylation increased, from 0.94 ± 0.04 (G0), 0.90 ± 0.08 (G1) to 0.77 ± 0.18 (G2), but statistical significance was unclear. On the other hand, the observed increase in the mean ratio of species containing bisecting GlcNAc residues to those not containing this modification is less variable. In summary, laboratories showed that hIgG contained highly fucosylated bi-antennary glycan structures bearing low levels of bisecting GlcNAc units.

**Comparison of glycan mapping techniques**

The total number of carbohydrate structures identified by laboratories using different techniques is displayed in Figure 9. Some laboratories employed more than one method, and these data were considered separately and assigned a different

---

**Table IV.** Galactosylated structures in hIgG

<table>
<thead>
<tr>
<th>Structure</th>
<th>Min (%)</th>
<th>Max (%)</th>
<th>Mean (%)</th>
<th>Median (%)</th>
<th>SD (%)</th>
<th>CV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G0</td>
<td>17.8</td>
<td>34.9</td>
<td>25.3</td>
<td>24.2</td>
<td>±6.1</td>
<td>24.2</td>
</tr>
<tr>
<td>G1</td>
<td>26.3</td>
<td>43.3</td>
<td>36.9</td>
<td>38.8</td>
<td>±6.3</td>
<td>17.1</td>
</tr>
<tr>
<td>G2</td>
<td>17.7</td>
<td>38.8</td>
<td>31.7</td>
<td>33.0</td>
<td>±7.8</td>
<td>34.5</td>
</tr>
</tbody>
</table>

*n = number of laboratories.

**Table V.** Ratios of fucose and bisecting GlcNAc residues found in hIgG

<table>
<thead>
<tr>
<th>Structure (n = 8)</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>SD</th>
<th>CV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fucose</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G0/G0total</td>
<td>0.88</td>
<td>1.00</td>
<td>0.94</td>
<td>±0.04</td>
<td>4.3</td>
</tr>
<tr>
<td>G1/G1total</td>
<td>0.73</td>
<td>1.00</td>
<td>0.90</td>
<td>±0.08</td>
<td>9.1</td>
</tr>
<tr>
<td>G2/G2total</td>
<td>0.46</td>
<td>0.96</td>
<td>0.77</td>
<td>±0.18</td>
<td>23.1</td>
</tr>
<tr>
<td>GlcNAc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G0/(G0fb+G0b)</td>
<td>3.56</td>
<td>9.69</td>
<td>5.37</td>
<td>±2.13</td>
<td>39.7</td>
</tr>
<tr>
<td>G1/(G1fb+G1b)</td>
<td>4.08</td>
<td>9.09</td>
<td>5.50</td>
<td>±1.81</td>
<td>33.0</td>
</tr>
<tr>
<td>G2/(G2fb+G2b)</td>
<td>3.72</td>
<td>11.06</td>
<td>6.22</td>
<td>±2.46</td>
<td>39.5</td>
</tr>
</tbody>
</table>

*n = number of laboratories.
Comparison of the number of glycans found with MS alone and HPLC

<table>
<thead>
<tr>
<th>Type of Analysis</th>
<th>Hα1</th>
<th>Bα1</th>
<th>RNase B</th>
<th>hIgG</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS</td>
<td>17 ± 6.3</td>
<td>23 ± 20.1</td>
<td>14 ± 5.3</td>
<td>17 ± 10.8</td>
</tr>
<tr>
<td>HPLC</td>
<td>5 ± 2.3</td>
<td>4 ± 0.6</td>
<td>6 ± 1.4</td>
<td>7 ± 1.3</td>
</tr>
</tbody>
</table>

Fig. 9. Total number of glycans identified for Hα1, Bα1, RNase B, and hIgG by MS, HPLC, and CE.

Why measure quantitative glycosylation profiles?

As protein glycosylation is not template-derived, it can vary quantitatively depending on the cell line and growth conditions, and hence, potentially, between product batches or nominally identical products. The variation in glycosylation can influence biological activity and clinical efficacy. Hence, accurate identification of glycan components and their quantification is of high importance to the biopharmaceutical industry. EU guidelines offer simplified procedures for the licensing of “biosimilar” products where manufacturers can prove that their product is...
essentially similar to an already licensed reference product, which implies a requirement that the glycosylation profile is quantitatively similar. In addition, there is growing interest in comparative glycomics and the understanding of how glycosylation changes in response to development, disease, and other factors affecting an organism. This work requires accurate identification of the glycan species present and quantitative data, which are often obtained from different laboratories. Thus, understanding inter-laboratory data comparability is central to judging the validity of the conclusions.

Comparability of methodology and data
Oligosaccharide analysis performed in the biopharmaceutical industry exploits a range of approaches. In general, organizations employed at least two techniques, chromatography and MALDI-TOF MS, with AEX frequently exploited to investigate the sialic acid/charge profile of glycans. In some instances, analysis by chromatography was coupled offline or online to MS. HPAEC is also widely used, although none of the UK laboratories in this study used this approach. These techniques typically quantify different parameters such as fluorescence, an electrochemical redox reaction, or ion current. The response factors may or may not be the same for different molecular species, and for mass spectrometry, factors such as ionization efficiency and ion suppression may be important for quantification. Chromatographic approaches typically have limited resolution, which defines the number of discrete species which can be resolved, but which may separate isomers in a way mass spectrometry cannot. In this study, we sought to understand inter-laboratory variability in the analysis aspect of the work, particularly that which appeared to correlate with methodology of choice. Preliminary experiments had indicated high reproducibility in intra-laboratory analyses. To achieve this, we distributed cleaved glycan mixtures and asked participants to identify and quantify the components. Future studies are planned to assess the effects of sample preparation, glycan cleavage, and isolation on the overall uncertainty of the procedure. The information provided can help to define methods to improve inter-laboratory comparability, through the use of standard glycan mixtures, for example.

More laboratories reported analyses of the RNaseB and hIgG samples than the highly sialylated materials, which may reflect the focus of individual laboratories on, for example, mAb products, or the limitation of the technique. The variability observed between the sialylated, high-mannose, and hlgG structures clearly emphasizes the need to optimize technologies especially when elucidating the profiles of mixtures of neutral and acidic glycans. This was particularly noticeable for the hlgG sample, in which MS detected a lower percentage of sialylated glycans compared to AEX. This is probably due to reduced ionization of charged glycans (due to the charge or suppression effects) resulting in a negative bias compared to neutral oligosaccharides. A range of values was reported for the widely used G0/G1/G2 ratio for hlgG. The variation here was due, in part, to the level of information extracted and the method chosen for identifying the glycans. There appeared to be a correlation, not statistically significant, between G0/G1/G2 and the degree of fucosylation. These observations highlight the need for standardized techniques and the use of appropriate reference materials to enable data comparison amongst specialist and nonspecialist laboratories.

Glycan components were predominately identified by their molecular weight. Identification was also performed using MS/MS experiments by fragmentation of selected oligosaccharides to provide information on sugar content in terms of hexose, N-acetyllactosamine, and deoxyhexose. In some cases, branching and linkages information could also be deduced when performed in specialist laboratories. Where ambiguity remained, exo-glycosidase digestion experiments were undertaken to confirm the presence, or not, of particular glycans. A comparison of retention times or mass peaks with reference compounds and/or glycan libraries was also exploited to identify glycans. However, in these cases, the greatest variation in glycan identification was observed. The mass spectrometric methods, not unexpectedly, identified larger number of glycan species than chromatographic approaches, particularly when, for example, Neu5Gc was present. The very broad range of approaches available indicates that a larger dataset is required to resolve data differences from various approaches, such as different MS derivatization methods, use of various fluorophore labels, and analysis after desialylation. A larger USP/NIBSC international inter-laboratory study is currently underway, which will provide additional data.

Use of the data from this study
The understanding of inter- and intra-laboratory variability reported in this study provides guidance on the credence which should be given to reported differences in protein glycosylation, whether in academic studies (into the effect of disease state on glycosylation, for example) or in studies supporting biopharmaceutical licensing. This aspect is likely to become increasingly important, as companies seek licensing of glycosylated “biosimilar” therapeutic products (called “follow-on biologicals” in the United States and “subsequent entry biologicals” in Canada). Statistical analysis of trend data from multiple laboratories may highlight significant features which would not be commented on if reported by a single group. Both this study and that of Wada and co-workers (2007) highlight the range of analytical platforms used and demonstrate the need to develop standardized protocols or best-practice guides and platform-independent reference materials that are suitable for the academic, biopharmaceutical, and diagnostic communities. These materials could be used to demonstrate laboratory performance, supporting publication of the work, and/or support identification of glycan components in complex mixtures. Both the European and United States Pharmacopoeia are currently developing monographs on the analysis of protein glycosylation.

Future work
This study focused on analytical aspects, but we recognize that sample preparation and glycan cleavage steps may contribute additional variability to the overall result. Studies to address this issue are planned. Data from the various inter-laboratory studies are being used to support the preparation of pharmacopeial monographs, providing formal guidance on glycan analysis to the biopharmaceutical community, and appropriate procedural reference materials are being developed to aid comparability across techniques and to support method and instrument qualification. The provision of such guidance through well-characterized standardized protocols, both on the choice of technique(s) and on the level of characterization required at
the various stages of glycoprotein drug production will thus be of primary importance for products to gain regulatory approval.

Material and methods

Materials

Glycoproteins human α1-acid glycoprotein (G9885; Hα1), bovine α1-acid glycoprotein (G3643; Bα1), bovine pancreatic ribonuclease B (R7884; RNaseB), and human immunoglobulin G (I4506; IgGh) were all purchased from Sigma-Aldrich (Poole, UK) and used without further purification. PNGase F from *Flavobacterium meningosepticum* (P0704L) was obtained from New England BioLabs (Hitchin, UK) as a kit containing: 10× denaturing solution (5% SDS and 400 mM dithiothreitol), 10× G7 reaction buffer (500 mM Na2HPO4, pH 7.5), 10% detergent solution (NP-40). C18 cartridges (Oasis HLB, 30 mg) were from Waters (Elstree, UK). All other reagents were purchased as reagent grade materials from Sigma-Aldrich and unless otherwise stated were used without further purification. Water was of analytical grade (ELGA, Marlow, UK or equivalent).

Deglycosylation with PNGase F

Cleavage of the N-glycans was carried out using the manufacturer’s (New England BioLabs, NEB) protocol with some modification. Briefly, each glycoprotein (5 mg) was dissolved in water (500 µL) in a 2 mL microcentrifuge tube and denatured with a denaturing solution (55 µL) at 100°C for 10 min. To the cooled solution were added the following: reaction buffer (60 µL), detergent solution (60 µL), and PNGase F enzyme solution (10 µL; 5000 NEB units). The samples were incubated at 37°C for a period of 24 h, and then quenched with 0.5 M acetic acid (300 µL). The released N-glycans were purified as described in the following sections.

Precipitation of protein and purification of released N-glycans

Protein from the reaction with PNGase F was precipitated by the addition of 0.5 M acetic acid (300 µL). After cooling, the solution was vorted and centrifuged (5000 × g, 2 min) and the supernatant transferred onto a C18 cartridge for the removal of the remaining soluble protein/peptides. Prior to use, the C18 cartridge was conditioned as per the manufacturer’s protocol. Supernatant from the PNGase F digest was allowed to pass through at atmospheric pressure into a collection vessel (10 mL disposable polypropylene tube) and the N-glycans eluted with 50 mM acetic acid/5% ethanol solution (0.8 mL). The flow through was lyophilized and stored at −20°C prior to further purification by normal-phase HPLC.

Additional purification of N-glycans by normal-phase HPLC

The N-glycans were purified further by HPLC to remove any remaining soluble peptides and salts using a TSK Amide 80 column (5 µm; 4.6 × 250 mm; Anachem) (Yuen et al. 2002). Hydrolyzed chitin oligomers were used as quality control materials for assessing the column efficiency. Lyophilized N-glycans from the C18 cartridge were taken up in solvent B (100 µL) and injected into the HPLC. HPLC was performed on an inert Gilson binary gradient system (Anachem) fitted with a UV detector. The HPLC mobile phase is aqueous acetonitrile (ACN) containing 5 mM TFA/(10 mM NH4OAc, pH 4.2), final pH ~3 with solvent A containing 80% ACN and solvent B containing 20% ACN. Glycans were eluted with a flow rate of 1 ml/min at room temperature and detected by UV at 228 nm. Elution gradients are 0–5 = 30% B, 7–10 = 100% B, and 12–20 = 30% B. Fractions between 8th and 11th min containing the N-glycans were collected and lyophilized.

Calculating N-glycan content

Since there is great structural heterogeneity in N-glycans, especially in a mixture, accurate account of different structures and composition in each sample are not possible. Therefore, hexose content is used to estimate the total N-glycan in samples so as to simplify the analysis. In addition, only an approximate but consistent amount of N-glycan content was required, since the same batch of sample was distributed to laboratories participating in the study. The concentrated N-glycans were reconstituted into water (500 µL) and the hexose content determined by orcinol staining (Feizi et al. 1994). Briefly, the N-glycans were spotted onto a thin layer chromatography plate using galactose (0.6–0.05 mg/mL) as a calibration curve. The sample spots were sprayed with an orcinol solution (900 mg of orcinol, 25 mL of water, 375 mL of ethanol, 45 mL of 18 M H2SO4) and the plate charred at 110°C until the spots turned violet (~1 min). Hexose content was determined by comparing to the galactose standards. Since the orcinol test does not stain sialic acid and GlcNAc residues, in order to estimate approximately the total N-glycan weight the hexose content was multiplied with the following factors: × 3 for Hα1 and Bα1; × 1.5 for RNaseB; and × 2 for IgGh. Samples were divided into 25 µg glycan aliquots in a microcentrifuge tube, dried down by GyroVap (VA Howe & Co Ltd, UK), and stored at −20°C.

Variability and stability studies

The variability of quality and stability of the cleaved N-glycans were performed by shipping two lots each (25 µg oligosaccharide) of the dried neutral (RNaseB) and moderate to highly acidic (IgGh and Hα1) N-glycans to the USA at room temperature using ordinary airmail post. The materials were returned after 10 days in the post and reductively labeled with a fluorophore. The fluorophore labeled N-glycans were analyzed by normal-phase amide-80 HPLC against two lots of the same batch of material, which had been stored at −20°C. No difference was observed in HPLC profiles between the posted batches and those stored at −20°C (Figure 2).

Funding

Funded under the UK Government’s Department of Trade and Industry (now the Department of Innovation, Universities, and Skills) National Measurement System program, which enabled this project be carried out as part of the Measurements for Biotechnology Programme (2004–2007).

Acknowledgements

NPL and NIBSC would like to thank the following organizations (listed alphabetically) for their participation and time in the inter-laboratory study: Dublin-Oxford Glycobiology Lab, GSK, Intertek, LGC, Lonza Biologies Plc, Ludger Ltd, M-Scan Ltd,
UCB-Celltech and the University of Birmingham. We would also like to thank Dr. Anna Hills for valuable discussions.

Conflict of interest statement
None declared.

Abbreviations
AEX, anion exchange chromatography; Bα1, bovine α1 acid glycoprotein; CE, capillary electrophoresis; CV, coefficient of variation; DHB, 2,5-dihydroxybenzoic acid; GlcNAc, N-acetylglucosamine, Hα1, human α1 acid glycoprotein; HPAEC-PAD, high-performance anion exchange chromatography with pulsed amperometric detection; HPLC, high-performance liquid chromatography; IgG, immunoglobulin G; MALDI, matrix-assisted laser desorption-ionization; MS, mass spectrometry; Neu5Ac, N-acetyleneuraminic acid; Neu5Gc, N-glycolyneuraminic acid; PNGase F, peptide N-glycosidase F; RNaseB, bovine pancreatic ribonuclease B; SD, standard deviation; TOF, time-of-flight.

References