Differences in the glycosylation of recombinant proteins expressed in HEK and CHO cells

Amelie Croset, Laurence Delafosse, Jean-Philippe Gaudry, Christian Arod, Loïc Glez, Christophe Losberger, Damien Begue, Ana Krstanovic, Flavie Robert, Francis Vilbois, Laurent Chevalet, Bruno Antonsson

ARTICLE INFO

Article history:
Received 21 February 2012
Received in revised form 7 June 2012
Accepted 11 June 2012
Available online 16 July 2012

Keywords:
Glycobiology
Glycoprotein
Isoform
Posttranslational modification
Sialic acid

ABSTRACT

Glycosylation is one of the most common posttranslational modifications of proteins. It has important roles for protein structure, stability and functions. In vivo the glycostructures influence pharmacokinetics and immunogenicity. It is well known that significant differences in glycosylation and glycostructures exist between recombinant proteins expressed in mammalian, yeast and insect cells. However, differences in protein glycosylation between different mammalian cell lines are much less well known. In order to examine differences in glycosylation in mammalian cells we have expressed 12 proteins in the two commonly used cell lines HEK and CHO. The cells were transiently transfected, and the expressed proteins were purified. To identify differences in glycosylation the proteins were analyzed on SDS-PAGE, isoelectric focusing (IEF), mass spectrometry and released glycans on capillary gel electrophoresis (CGE-LIF). For all proteins significant differences in the glycosylation were detected. The proteins migrated differently on SDS-PAGE, had different isof orm patterns on IEF, showed different mass peak distributions on mass spectrometry and showed differences in the glycostructures detected in CGE. In order to verify that differences detected were attributed to glycosylation the proteins were treated with deglycosylating enzymes. Although, culture conditions induced minor changes in the glycosylation the major differences were between the two cell lines.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

The properties of a protein are largely determined by its amino acid sequence. However, protein characteristics are also modified and regulated by a large number of post-translational modifications, taking place during or after synthesis of the polypeptide chain (Walsh et al., 2005). One of the most common post-translational modifications is glycosylation and approximately half of all human proteins are estimated to be glycoproteins (Wong, 2005). Protein glycosylation is a conserved mechanism that occurs in yeast, plants and animals (Lommel and Strahl, 2009). Simple forms of glycosylation have also been identified in bacteria (Nothhaft and Szymanski, 2010). Glycosylation are enzymatic processes where glycans are added to specific amino acids in the polypeptide chain. Two types of glycosylation, N-linked and O-linked, occur in proteins. N-linked glycosylation starts as a co-translational process in the endoplasmic reticulum (ER), where a branched presynthesized oligomeric glycan structure is attached to the nitrogen of an asparagine in the polypeptide chain (Yan and Lennarz, 2005). The glycan structure is then trimmed before the protein is transferred to the Golgi apparatus (GA) where the glycan structure is further modified. O-linked glycosylation is a post-translational process occurring in the GA (Peter-Katalinic, 2005). Single monosaccharides are attached to the hydroxyl group of serine or threonine residues. The glycan structures are subsequently built up by addition of individual monosaccharides. In addition to the two major glycosylation pathways, glycans can also be attached to arginine, tyrosine, hydroxylysine, hydroxyproline and tryptophan residues (Spiro, 2002). These modifications are less common and restricted to specific proteins. Protein glycosylation can be very heterogeneous, resulting in a large number

Abbreviations: HEK, human embryonic kidney cells; CHO, Chinese hamster ovary cells; TCA, trichloroacetic acid; PEI, polyethylenimine; IEF, isoelectric focusing; PNGase F, peptide N-glycosidase F; APTS, 8-amino-1,3,6-pyrenesulfonic acid.

* Corresponding author. Tel.: +41 22 414 9802.
E-mail address: bruno.antonsson@merckgroup.com (B. Antonsson).
1 These two authors contributed equally to the work.
2 Present address: Institut Curie, Centre Universitaire, Orsay, France.
3 Present address: National Research Council – Biotechnology Research Institute, Montréal, Quebec, Canada.
4 Present address: Novartis Pharma AG, Basel, Switzerland.

0168-1656/– see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jbiotec.2012.06.038
of isoforms of the protein (Hua et al., 2011). Potential glycosylation sites can be either occupied or unmodified, and each site can be occupied by a different glycan structure in different protein molecules.

Although, glycosylation was long considered as an unimportant protein decoration, it is now clear that it has important functions. Glycosylation affects protein folding, stability, solubility, protein–protein interactions and in vivo bioavailability, biodistribution, pharmacokinetics and immunogenicity (Kaushik et al., 2011; Kayser et al., 2011; Li and d'Anjou, 2009; Oberg et al., 2011; Opanasopit et al., 2001). In general the activity of a protein is determined by its primary amino acid sequence however, there are several examples where glycosylation affects and regulates the activity (Rajagopalan et al., 2010; Straumann et al., 2006; Su et al., 2010). Impaired or changed protein glycosylation is associated with a number of human pathologies, including, rheumatoid arthritis, Leroy disease, and leukocyte-adhesion deficiency type II (Gornik and Lauc, 2008; Koscielak, 1995). Alterations in protein glycosylation are frequently associated with different cancers, although the function or implication of those changes mostly remains unclear (Dennis et al., 1999; Orntoft and Vestergaard, 1999).

Most proteins used for in vitro and in vivo studies today are produced as recombinant proteins in various cell culture based expression systems. Widely used expression systems are, in addition to bacteria, where little post-translational modifications occur, yeast, insect cells and different mammalian cell lines. It is well known that the glycosylation structures are different between mammalian, yeast and insect cells (Brooks, 2006). However, it is much less known what differences exist between mammalian cell lines. Two frequently used mammalian cell lines are HEK and CHO. HEK is a cell line originally derived from Human Embryonic Kidney tissue. HEK cells are easy to grow and transfact, and transient transfection frequently gives good expression levels, which has made this cell line widely used in research. CHO is derived from Chinese Hamster Ovary. This cell line is frequently used for expression after stable transfections, it has a good long-term stable gene expression with high expression levels. For transient expression CHO is more difficult to transfact and the expression levels are frequently lower compared to transiently transfected HEK cells.

For this study 12 proteins containing different numbers of potential N- and O-linked glycosylation sites were selected and each one was expressed and purified from HEK and CHO cells. The purified proteins were analyzed by SDS-PAGE, gel IEF and/or capillary IEF. To verify the contribution of glycosylation to the differences in protein patterns detected the proteins were treated with deglycosylating enzymes and reanalyzed after the treatment. For the more in depth analyses subsets of the proteins were analyzed by mass spectrometry and released N-linked glycans were analyzed by capillary gel electrophoresis.

2. Materials and methods

2.1. Reagents

The Poros 20 MC resin was from Applied Biosystems. Superdex 75 and Sephadex G-25 resins were from GE Healthcare. The NuPAGE 10% Bis-Tris gels, IEF 3–10 gels, IEF 3–7 gels, MES-SDS running buffer, MOPS-SDS running buffer and BenchMark protein ladder were from Invitrogen. InstantBlue staining solution was purchased from Expeoden. Sialidase, O-glycosidase, β-galactosidase, glucosaminidase, corresponding reaction buffers and denaturation solutions were from QA-Bio. PNGase F was either form QA-Bio or from Roche.

2.2. Prediction of glycosylation sites

N- and O-linked glycosylation sites were predicted using the NetNGlyc and NetOGlyc prediction programs at http://www.cbs.dtu.dk/services/ (Julienius et al., 2005).

2.3. Cell culture conditions

Human Embryonic Kidney (HEK) cells expressing the Epstein–Barr virus Nuclear Antigen (EBNA) were purchased from Invitrogen (Cat. no. R620-07). The HEK293 EBNA cells were adapted to grow in suspension in EX-CELL™ VPRO serum-free medium (SAFC, Cat. no. 14561C) supplemented with 4 mM l-glutamine using the sequential adaptation method (Sinacore et al., 2000). To enhance productivity, a HEK293 EBNA clone over-expressing the EBNA1 (HEK293-6E) established by Dr. Yves Durocher at the National Research Council Canada was used. The cells were maintained in suspension culture in Erlenmeyer flasks in Freestyle™ F17 serum free medium (Invitrogen, Cat. no. 0050092) supplemented with 4 mM l-glutamine.

Chinese Hamster Ovary (CHO) cells were purchased from Invitrogen. The CHO-S cells were already pre-adapted to grow in suspension in ProCHO™ 5 serum free media (Lonza, Cat. no. 12-766Q) supplemented with 4 mM l-glutamine.

Production cultures were carried out in spinners (DasGip) or Erlenmeyer flasks (Corning) depending on the cell line. The cell cultures were maintained at 37 °C in a humidified 5% CO₂ atmosphere. Cell viability was determined by the Trypan blue [0.4% (w/v)] dye exclusion method. Cell density was determined by hemocytometer (Neubauer) counting of viable cells per milliliter of cell suspension. Cells were routinely diluted every 2 or 3 days to between 0.3 × 10⁵/ml and 3 × 10⁶ cells/ml.

2.4. Transfection reagent

A stock solution of JetPEITM (Polyplus Transfection) was stored at −20 °C. A 1 mg/ml stock solution of 25–kDa linear PEI (Polysciences, Warrington, PA) was also prepared in water, acidified with HCl to pH 2 until dissolved, neutralized with NaOH, sterilized by filtration (0.2 μm), aliquoted, and stored at −20 °C.

2.5. Transfection protocols

2.5.1. HEK293 EBNA Freestyle™ 293, PEI mediated transfection protocol

HEK293 EBNA cells were maintained in suspension culture in spinners in EX-CELL™ VPRO serum-free medium supplemented with 4 mM l-glutamine. Three hours before transfection, the cells were recovered by centrifugation and resuspended in fresh DMEM:F12 medium supplemented with 5 g/L d-glucose, 10 mM HEPES, 7.5 mM l-glutamine, 1% fetal bovine serum (FBS) and 4 mL/1 insulin–transferrin–selenium–X supplement (100X) (Invitrogen, Cat. no. 51500-056) at a cell density of 3.0 × 10⁵ cells/ml. Two hundred ml of cell suspension was transferred to a 1L spinner. Seven hundreds and fifty micrograms of DNA was diluted in 50 ml of fresh DMEM:F12 with supplements as above, 15 mg of PEI was added, and the mixture immediately vortexed and incubated for 10 min at room temperature before added to the cells. After 90 min incubation 250 ml of pre-warmed Freestyle™ 293 expression medium (Invitrogen, Cat. no. 12338-018) supplemented with 4 mM l-glutamine and 4 mL/1 ITS–X supplement (100X) was added. The transfected cultures were incubated for 6 days for recombinant protein production, at 37 °C in a 5% CO₂ incubator and a spinner speed of 80 rpm.
2.5.2. HEK293-6E Freestyle™ F17, JetPEI™ mediated transfection protocol

HEK293-6E cells were maintained in suspension culture in Erlenmeyer flasks in Freestyle™ F17 serum free medium (Invitrogen, Cat. no. 0050092) supplemented with 4 mM l-glutamine. Twenty-four hours prior to transfection the cells were diluted to $1 \times 10^5$ cells/ml. The following day, cell density should be around $1.5-2.0 \times 10^6$ cells/ml and viability greater than 97% to ensure optimal transfection. Three hundreds and sixty milliliters of cell suspension was transferred to a 2 L Erlenmeyer flask. Four hundreds micrograms of DNA was diluted in 20 ml fresh Freestyle™ F17 serum free medium supplemented with 4 mM l-glutamine and immediately vortexed. Eight hundreds micrograms of JetPEI™ were diluted in 20 ml of fresh Freestyle™ F17 serum free medium supplemented with 4 mM l-glutamine and immediately vortexed. The two solutions were mixed and immediately vortexed. The mixture was incubated for 15 min at room temperature before added to the cell culture. Twenty-four hours post-transfection, 10 ml of TN1 20% (w/v) stock solution were added to obtain a final concentration of 0.5%. The transfected cultures were incubated for 6 days for recombinant protein production, at 37°C in a 5% CO₂ incubator on a shaker at 80 rpm.

2.5.3. CHO-s ProCHO™ 5, JetPEI™ mediated transfection protocol (standard)

Three hours before transfection, CHO-s cells were centrifuged and resuspended in fresh DMEM:F12 supplemented with 5 g/L d-glucose, 10 mM HEPES, 7.5 mM l-glutamine and 4 mL/L insulin–transferrin–selenium–X supplement (100X) (ITS-X, Invitrogen, Cat. no. 51500-056) at a cell density of 2.5 × 10⁵ cells/ml. Two hundreds milliliters of cell suspension was distributed per 1 L spinner. Two hundreds and fifty micrograms of DNA was diluted in 50 ml fresh DMEM:F12 supplemented with 5 g/L d-glucose, 10 mM HEPES, 7.5 mM l-glutamine and 4 mL/L ITS-X supplement (100X), 750 μg of JetPEI™ was added, and the mixture immediately vortexed. Mixture was incubated for 10 min at room temperature prior to its addition to the cells. Following a 3 h incubation with DNA–PEI complexes, culture medium was collected to 500 ml by the addition of 250 ml of pre-warmed ProCHO™ 5 serum free media (Lonza, Cat. no. 12-766Q) supplemented with 6 mM l-glutamine and 4 mL/L ITS-X-supplement (100X). Transfected cultures were incubated for 6 days for recombinant protein production, at 37°C in a 5% CO₂ incubator and a spinner speed of 80 rpm.

2.5.4. Optimized CHO-s ProCHO™ 5, PEI mediated transfection protocol (optimized)

A modified protocol has been developed from the previous one with an optimized N/P ratio for the formation of polyplexes. CHO-s cells were maintained in suspension culture in spinners in ProCHO™ 5 serum free media (Lonza, Cat. no. 12-766Q) supplemented with 4 mM l-glutamine. Three hours before transfection, CHO-s cells were centrifuged and resuspended in fresh DMEM:F12 supplemented with 5 g/L d-glucose, 10 mM HEPES, 7.5 mM l-glutamine and 4 mL/L insulin–transferrin–selenium–X supplement (100X) (ITS-X, Invitrogen, Cat. no. 51500-056) at a cell density of 2.4 × 10⁶ cells/ml. Two hundreds milliliters of cell suspension was distributed to a 2 L Erlenmeyer flask. Seven hundreds and fifty micrograms of DNA was diluted in 50 ml fresh DMEM:F12 supplemented with 5 g/L d-glucose, 10 mM HEPES, 7.5 mM l-glutamine and 4 mL/L ITS-X supplement (100X), 12 mg of PEI was added, and the mixture immediately vortexed. Mixture was incubated for 10 min at room temperature prior to its addition to the cells. Following a 3 h incubation with DNA–PEI complexes, culture medium was collected to 500 ml by the addition of 250 ml of pre-warmed ProCHO™ 5 serum free media supplemented with 4 mM l-glutamine and 4 mL/L ITS-X supplement (100X). Transfected cultures were incubated for 6 days for recombinant protein production, at 37°C in a 5% CO₂ incubator on a shaker at 80 rpm.

2.6. Protein purification

All the recombinant proteins were expressed as secreted proteins and had a 6-HIS tag at the C-terminus. They were purified by metal affinity chromatography followed by size exclusion chromatography. The culture supernatant was diluted with 1 volume equilibration buffer (50 mM NaH₂PO₄, 600 mM NaCl, 8.7% (v/v) glycerol, pH 7.5) and filtered through a 0.22 μm sterile filter. The sample was loaded at a flow rate of 10 ml/min onto a 4 ml Poros metal affinity column (10 mm × 50 mm) charged with Ni²⁺ ions and equilibrated in equilibration buffer. The column was washed at a flow rate of 20 ml/min with 10 CV equilibration buffer followed by 30 CV equilibration buffer containing 20 mM imidazole. The protein was step eluted with 600 mM imidazole in equilibration buffer at a flow rate of 2 ml/min. Fractions of 2.7 ml were collected and the fractions containing the protein were pooled. After each purification cycle the Poros column was stripped with EDTA, recharged with Ni, washed with elution buffer and re-equilibrated in the equilibration buffer.

The pool from the metal affinity column was either desalted into PBS, pH 7.4 on a Sepharose G-25 column or further purified on a 120 ml Superdex 75 size exclusion column equilibrated in PBS, pH 7.4. Fractions of 5 ml were collected. The fractions containing the protein were pooled and if necessary concentrated on an Ultracel centrifugation concentrator with a MWCO of 3 kDa to have a final concentration between 1.5 and 2 mg/ml. The final sample was filtered on a sterile 0.22 μm filter (Millipore), aliquoted and stored at -80°C. The protein concentration was determined by UV spectroscopy using the calculated extinction coefficient at 280 nm.

The culture medium used for the HEK293-6E cell line contains an unidentified component interfering with purification on metal affinity. Thus, these samples were first captured on an ion exchange resin, Q-Sepharose or SP-Sepharose depending on the pl of the protein, and the proteins were step eluted with the metal affinity equilibration buffer and then purified as described above on metal affinity and SEC.

2.7. SDS-PAGE

The proteins were analyzed on NuPAGE 10% Bis-Tris gels in either a MES SDS running buffer (separation range 2.5–200 kDa) or a MOPS SDS running buffer (separation range 14–200 kDa). The running buffer was selected depending on the molecular weight of the proteins. The proteins were diluted in reducing SDS sample buffer and heated at 95°C for 5 min. Four μg of protein were loaded per well and the electrophoresis was performed at 200 V for 37 min with MES buffer and for 45 min with MOPS buffer. The gels were stained with either InstantBlue or with Simple Blue and distained with distilled water.

2.8. Gel IEF

IEF was performed on gels with a pH separation range of either 3–10 or 3–7. Ten μg of protein diluted in one volume of IEF loading buffer were loaded per well and the separation protocol was 1 h at 100 V, 1 h at 200 V and 30 min at 500 V. At the end of the focusing step the gels were incubated for 1 h in 12% TCA, washed three times during 10 min in distilled water, stained with InstantBlue.
2.9. Capillary IEF

The samples containing 250–500 μg of protein were desalted into water on a centrifugation concentrator (Amicon) through addition of 1 ml water followed by centrifugation at 4000 x g for 10 min. The desalting step was repeated 3 times and the samples were finally concentrated to approximately 50 μl. The analysis reagents were added to the desalted sample. For protein 1, 0.35% methylcellulose, 4 M urea, 4% 8–10.5 Pharmalyte, pl markers 8.18/10.45 and for protein 4, 0.35% methylcellulose, 4 M urea, 1% 5–8 Pharmalyte/3% 8–10.5 Pharmalyte, pl markers 5.12/9.77. The samples were analyzed on an iCE280 analyzer system (ProteinSimple) equipped with a 50 mm length, 100 mm ID capillary. The separation was performed using 100 mM NaOH as catholyte and 80 mM H3PO4 as anolyte solution. The separation was performed at 1500 V for 1 min followed by 5 min at 3000 V for protein 1, and 15 min at 3000 V for protein 4. The electropherograms were acquired through measuring the absorbance at 280 nm. The pl of the different isoforms was calculated by the software using the two standards included in each run and the % of each isoform was calculated. The peak areas were determined and the results presented and compared in bar graphs.

2.10. Sialidase treatment

70 μl of protein solution (70 μg) were mixed with 20 μl of reaction buffer and 2 μl of sialidase. The sample was incubated at 37 °C for 4 h. The optimal reaction time under conditions used had been determined previously.

2.11. Deglycosylation

In order to obtain an as complete a removal of glycans as possible the proteins were incubated with a mixture of deglycosylating enzymes. 70 μl of protein solution (70 μg) were mixed with 20 μl of reaction buffer, 2.5 μl denaturation solution and 1 μl of each of PNGaseF, O-glycosidase, sialidase, β-galactosidase, and glucosaminidase. The sample was incubated at 37 °C for 24 h. The optimal reaction time under conditions used had been determined previously.

2.12. Sample preparation for liquid chromatography–mass spectrometric analysis

All proteins were analyzed in their reduced form before and after treatment with PNGase F. For removal of N-linked glycans 10 μg of protein was incubated with 5 U PNGase F in PBS pH 7.5 for 16 h at 37 °C. The protein was reduced by incubation in 4 M guanidine HCl, with10 mM DTT for 1 h at 56 °C. Samples were stored at 4 °C after reduction prior to LC–MS analysis.

2.13. Liquid chromatography–mass spectrometric analysis (LC–MS)

LC–MS analyses were performed by reversed phase chromatography on a capillary HPLC system (Ultimate 3000, Dionex, Sunnyvale, CA, USA) coupled to a Sciex QSTAR Ultima mass spectrometer (Applied Biosystems, Foster City, CA, USA). Five μg of protein were loaded onto a Dionex ProSwift RP-10R monolithic capillary column, 1 mm × 5 cm, operated at a flow rate of 30 μl/min running a 7 min gradient from 30% to 64% acetonitrile in 0.05% formic acid. The mass spectrometer was operated in the positive ion mode using an electrospray ion source. Mass spectra were recorded between m/z 500 and m/z 2000. A solution of horse heart myoglobin was infused for 3 min as an internal mass calibration standard using the instrument’s syringe pump during the sample loading/desalting step. The mass spectrometer was operated using the Analyst software and spectra were deconvoluted using Analyst’s Biotools.

2.13.1. Capillary gel electrophoresis glycoanalysis

The sample preparation and analysis of N-glycans was performed according to the published procedures (Papac et al., 1998; Rapp et al., 2011).

3. Results

The aim of this study was to examine whether the glycosylation pattern of proteins expressed in the HEK and CHO cell lines is significantly different. A set of 12 proteins of different sizes, between 9.5 kDa and 52 kDa, containing various numbers of potential N- and/or O- glycosylation sites were selected and expressed in HEK293 EBNA and CHO-S cells (Table 1). The two cell lines were transiently transfected and grown in standard conditions usually applied for production of recombinant protein as described in Section 2. The proteins were purified and stored in PBS, pH 7.4 at −80 °C until analyzed. The proteins were first analyzed on SDS-PAGE to verify whether differences in the apparent molecular weight, indicative of differences in posttranslational modifications, could be detected (Fig. 1A). For most of the proteins differences were detected between the HEK and the CHO derived protein batches, even on this relatively insensitive analytical method. Either clear differences in the apparent molecular weight (proteins 1, 5, 6, 10, 11) or more or less focused bands (proteins 2, 4, 7, 8, 9, 12) were seen. Protein 3, which was included as a negative control containing no predicted glycosylation sites, showed no difference in migration pattern between the two cell lines. In order to verify that the differences detected were associated with glycosylation the samples were treated with a mixture of five deglycosylating enzymes to remove both N- and O-glycosylation as described in Section 2. When reanalyzed after the deglycosylation treatment several proteins migrated at a significant lower apparent molecular weight (proteins 4, 5, 6, 7, 8, 9, 10, 11, 12). Other proteins did not show large changes in the molecular weight but the bands got more focused and the smears were less pronounced (proteins 1, 2). In general the differences disappeared or at least got less significant, suggesting that the differences in the apparent molecular weights were indeed due to differences in the glycosylation (Fig. 1B). For example protein 5, which has 3 potential N-glycosylation sites, migrated at around 40 kDa and showed a clear difference in the migration pattern. After deglycosylation it migrated close to the theoretical molecular weight of 20 kDa and no significant difference was detected between the protein from the two cell lines. The negative control, protein 3, showed no change after the treatment. The proteins, where significant differences still remained after the treatment, all contained potential O-linked glycosylation, which is less efficiently removed. These results suggest that the differences in glycosylation between the two cell lines are substantial.

The glycostructures of proteins frequently contain sialic acid residues, a negatively charged sugar, which results in charged isoforms detectable by isoelectric focusing (IEF). The proteins were analyzed on gel IEF and for several proteins even more striking differences between the two cell lines were revealed (Fig. 2A). Protein 1 has a theoretical pI of 9.7 and was not detected within the separation range of the IEF system used. Proteins 2 and 3 were running at the very top of the separation range and potential differences could not be assessed. For the rest of the proteins differences in the isoform patterns were detected. In general, the proteins expressed in CHO cells had more acidic isoforms, indicating a higher content of sialic acid. For some of the proteins, for example proteins 4, 5, 7, 11, the HEK derived proteins showed more isoforms compared
Table 1

Test proteins. The proteins were expressed in HEK and CHO cells, purified and the purified proteins were analyzed.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Protein name</th>
<th>Mw theoretic</th>
<th>Mw observed</th>
<th>ΔMw</th>
<th>N-linked predicted</th>
<th>O-linked predicted</th>
<th>pI theoretic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MCP1-6HIS</td>
<td>9507</td>
<td>14000</td>
<td>4500</td>
<td>1</td>
<td>0</td>
<td>9.74</td>
</tr>
<tr>
<td>2</td>
<td>SMALL INDUCIBLE CYTOKINE A15 (13T)-6HIS</td>
<td>10,985</td>
<td>12,000</td>
<td>1000</td>
<td>0</td>
<td>1</td>
<td>8.02</td>
</tr>
<tr>
<td>3</td>
<td>PHOSPHOLIPASE A2-6HIS</td>
<td>15,674</td>
<td>20,000</td>
<td>4300</td>
<td>0</td>
<td>0</td>
<td>7.60</td>
</tr>
<tr>
<td>4</td>
<td>CYSTATIN F-D-ATT-6HIS</td>
<td>16,628</td>
<td>27,000</td>
<td>10,000</td>
<td>2</td>
<td>3</td>
<td>8.30</td>
</tr>
<tr>
<td>5</td>
<td>ERYTHROPOIETIN-K116E-ATT-6HIS</td>
<td>20,478</td>
<td>38,000</td>
<td>17,000</td>
<td>3</td>
<td>0</td>
<td>7.88</td>
</tr>
<tr>
<td>6</td>
<td>LOW-AFFINITY NERVE GROWTH FACTOR (1–221)-6HIS</td>
<td>24,269</td>
<td>60,000</td>
<td>35,000</td>
<td>1</td>
<td>16</td>
<td>4.15</td>
</tr>
<tr>
<td>7</td>
<td>PROLACTIN RECEPTOR (1–211)-6HIS</td>
<td>23,340</td>
<td>32,000</td>
<td>6700</td>
<td>2</td>
<td>0</td>
<td>6.30</td>
</tr>
<tr>
<td>8</td>
<td>APOLIPOPROTEIN H (V197L)-6HIS</td>
<td>37,991</td>
<td>52,000</td>
<td>15,000</td>
<td>3–4</td>
<td>2</td>
<td>7.90</td>
</tr>
<tr>
<td>9</td>
<td>FOLLISTATIN-RELATED PROTEIN 1-D-ATT-6HIS</td>
<td>34,763</td>
<td>50,000</td>
<td>15,000</td>
<td>3–4</td>
<td>0</td>
<td>5.30</td>
</tr>
<tr>
<td>10</td>
<td>T-CELL SURFACE GLYCOPROTEIN CD5 (1–355)-6HIS</td>
<td>40,000</td>
<td>50,000</td>
<td>10,000</td>
<td>2</td>
<td>8</td>
<td>7.72</td>
</tr>
<tr>
<td>11</td>
<td>EPHRIN-B1 (1–216)-6HIS</td>
<td>24,400</td>
<td>40,000</td>
<td>16,000</td>
<td>1</td>
<td>6</td>
<td>7.34</td>
</tr>
<tr>
<td>12</td>
<td>PROSAPOSIN-6HIS</td>
<td>52,719</td>
<td>70,000</td>
<td>17,000</td>
<td>5</td>
<td>0</td>
<td>4.86</td>
</tr>
</tbody>
</table>

to the CHO, notably in the more basic region. To verify whether the different isoforms were indeed the result of different amounts of sialic acid in the glycostructures, the proteins were treated with sialidase, an enzyme that selectively cleaves off sialic acid from the glycans of the proteins (Uchida et al., 1979). After treatment, the proteins were reanalyzed on gel IEF (Fig. 2B). The proteins moved into more basic forms and they condensed into fewer isoforms. Furthermore, significant differences between the proteins expressed in HEK and CHO cells were no longer detected. These results confirm that the differences in isoform patterns detected between the proteins expressed in HEK and in CHO are indeed due to different levels of sialic acid residues attached to the glycostructures. Since protein 1 was not detected in the gel IEF separation range the protein was analyzed by capillary IEF where the isoform pattern could be obtained (Fig. 3A). The relative peak areas presented in the bar graph show that the protein derived from two different HEK cell lines, HEK293 EBNA and HEK293 6E, showed almost identical pattern with approximately 50% as the most acidic isoform and only around 20% in the most basic isoform. A clear difference was seen compared to CHO where 50% was present as the most basic isoform and only 15% as the most acidic form. For this protein the CHO derived protein had more of the basic isoforms compared to the HEK derived protein. However, all the isoforms in the CHO derived protein were slightly more acidic compared to the corresponding HEK isoform. After treatment with sialidase the acidic isoforms disappeared, over 80% of the protein condensed into the most basic isoform and significant differences between the HEK and CHO derived protein were no longer detectable (Fig. 3B). Protein 4 was also analyzed by capillary IEF and consistently with the results of gel IEF the protein showed a large number of isoforms (Fig. 3C). The isoform pattern for the protein expressed in HEK293 EBNA and HEK293 6E were close to identical, whereas the CHO derived protein was significantly shifted to the acidic isoforms. As seen in the bar graph the protein derived from the two HEK cell lines had 50% of the protein in acidic and 50% in basic isoforms. Whereas, the protein from CHO cells had 80% in the acidic and only 20% in the basic isoforms. Also, for this protein treatment with sialidase significantly reduced the more acidic isoforms and the entire isoform profile was shifted to the basic region. Both HEK and CHO derived protein condensed into a few isoforms in the basic region (Fig. 3D). However, differences in the isoform distribution still remained after the treatment, with more basic isoforms in the HEK derived protein. This might indicate that the different isoforms in this protein are not only a result of sialic acid modification.

Fig. 1. Molecular weight analysis by reducing SDS-PAGE. (A) Analysis of the proteins (1–12) expressed in HEK EBNA (H) and CHO (C) cells. (B) Analysis of the proteins after treatment with deglycosylating enzymes to remove N- and O-linked glycosylation. The protein samples were diluted in reducing SDS sample buffer and 4 μg protein was loaded per lane. The proteins were analyzed on 10% gels in MOPS running buffer (proteins 1–4) or in MES running buffer (proteins 5–12). Molecular weight markers (M). After separation the gels were stained with InstantBlue.
To verify that the differences in glycosylation was related to the cell lines and not just interbatch variations three proteins were selected to evaluate the reproducibility of the profiles on SDS-PAGE and IEF. Each protein was expressed 3 times in CHO and 3 times in HEK293 EBNA cells. In addition, the proteins were expressed 2 times in HEK293 6E cells, a variant of the HEK293 EBNA cell line, which normally gives higher protein expression levels. The expressions were performed during three different weeks in order to capture any possible influence of cell passages or variations in manipulations or culture conditions and as can be seen in Table 2 the expression levels showed some variation between the batches. However, the SDS-PAGE pattern was very reproducible between batches as were the differences between CHO and HEK EBNA (Fig. 4A). For two of the proteins (proteins 1 and 9) differences were also detected between the proteins expressed in HEK EBNA and HEK 6E. Protein 1 derived from HEK 6E contained a weak band moving at an apparent lower molecular weight, below the main band in the HEK EBNA samples, at approximately the same position as the major band in the CHO derived samples. For protein 9 a band at a significant lower molecular weight, suggesting less glycosylation, was present in the HEK 6E samples whereas it was absent in HEK EBNA and CHO derived protein. For protein 4 no significant differences between the two HEK cell lines were detected, although a clear difference was seen between HEK and CHO derived protein. On IEF protein 1 was, as seen before (Fig. 2), not detected within the pH analysis range (results not shown). Protein 4 showed the same isofrom pattern as seen in Fig. 2A and the pattern was very reproducible between batches (Fig. 4B). A clear difference was detected between the CHO and the HEK derived protein. No significant difference in the isofrom pattern was seen between HEK EBNA and HEK 6E. Protein 9 showed a slight difference between the CHO and HEK derived protein, with a shift to the more basic isoforms in the HEK derived protein. No significant differences were detected between the HEK EBNA and the HEK 6E. The variation in expression levels between batches did not appear to have any effect on the migration pattern on SDS-PAGE or the isofrom distribution on IEF. Protein 4 expressed in HEK EBNA and HEK 6E showed close to identical patterns, whereas the difference in expression levels was the largest (Table 2). The expression levels in CHO were in between the two HEK cell lines, and a clear difference in the patterns was seen compared to the HEK cell lines. On the other hand protein 9 showed only a small difference in expression levels between the two HEK cell lines, but the largest differences in the SDS-PAGE migration pattern were detected (Table 2, Fig. 4A). Four additional proteins (2, 3, 5 and 11) were also reproduced at larger quantities in CHO and HEK EBNA for further characterization. The SDS-PAGE and IEF profiles were similar to what had been seen earlier with clear differences between the two cell lines, except for protein 3, which has no predicted glycosylation sites (Fig. 4C and D).

To further evaluate the reproducibility and the effects of culture conditions two different transfection protocols were tested for the CHO cells, and the expressed proteins were analyzed and compared. The optimized new transfection protocol generally increased the expression levels, approaching the yields obtained in the HEK cells for some proteins (Table 3). Expression levels could be one factor influencing the glycosylation pattern. However, as shown in Fig. 5, no significant differences in either the SDS-PAGE or the IEF profiles were detected for any of the proteins expressed with the two protocols. Some of the proteins were produced in HEK EBNA at a contract research organization, using the transfection, culture and purification protocols provided by Merck Serono. No significant differences were detected compared to the same proteins produced in HEK EBNA cells at Merck Serono (results not shown).

Table 2

<table>
<thead>
<tr>
<th>Protein</th>
<th>HEK EBNA (mg/l)</th>
<th>HEK 6E (mg/l)</th>
<th>CHO optimized (mg/l)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.1–10.0</td>
<td>22–23</td>
<td>8.2–11.0</td>
</tr>
<tr>
<td>4</td>
<td>1.5–3.5</td>
<td>20–21</td>
<td>5.0–9.0</td>
</tr>
<tr>
<td>9</td>
<td>4.0–5.8</td>
<td>7.1–7.8</td>
<td>1.4–2.8</td>
</tr>
</tbody>
</table>

Fig. 2. Gel IEF analysis. (A) Analysis of the proteins (1–12) expressed in HEK EBNA (H) and CHO (C) cells. (B) Analysis of the proteins after treatment with sialidase to remove sialic acid. The protein samples were diluted in IEF buffer and 10 µg protein were loaded per lane. The proteins were analyzed on pH 3–10 gels as described in Section 2. pl markers (M). After separation the gels were stained with InstantBlue.


341
Combined, these results show that major differences in protein glycosylation exist between CHO and HEK cells, whereas smaller differences were detected between the two HEK cell lines. Minor changes in the transfection and cell culture conditions did not appear to have any significant effect on the glycosylation pattern of the proteins.

To gain more detailed information about the differences in the glycosylation pattern, three proteins were analyzed on mass spectrometry. The control protein (protein 3) without predicted glycosylation sites showed only one mass peak at the expected mass of 15,674 Da (Fig. 6A). No difference was detected between CHO and HEK derived protein, neither did treatment with PNGaseF change the mass spectra. PNGaseF selectively removes N-linked glycans from the protein. The results for protein 2, which has one predicted O-glycosylation site, is shown in Fig. 6B. The protein from both cell lines showed only a minor mass peak at the
Protein 4 has a potentially more complex glycosylation pattern, with two predicted N-linked and 3 O-linked glycosylation sites. The CHO derived protein showed over 10 mass peaks between 18,690 and 22,429 Da (Fig. 6C). The HEK derived protein showed an even more complex mass peak distribution with a forest of mass peaks between 18,918 Da and 23,667 Da, and the resolution limit was reached. However, protein derived from HEK EBNA and HEK6E showed similar patterns, clearly different from the profile for the protein from CHO cells. After treatment with PNGaseF the mass profile changed and all the samples showed a mass peak at 16,630 Da corresponding to the unmodified protein. In addition, for the CHO derived protein a cluster of 8 peaks, between 17,846 and 19,024 Da, presumably corresponding to the protein with O-linked glycosylation was detected. For the HEK derived protein clear mass peaks appeared, the main peak was at 16,630 Da, corresponding to the unmodified protein, and a cluster of peaks between 18,440 and 18,982 Da, resembling the CHO samples was seen. In addition, low intensity peaks scattered at higher masses were also detectable. The protein derived from HEK EBNA and HEK6E showed no significant differences.

To compare the N-linked glycostructures on the proteins from the two cell lines five proteins were treated with PNGaseF and the released N-glycans were analyzed on capillary gel electrophoresis. For the proteins 5, 11, and 12 one batch produced in CHO and one produced in HEK EBNA were analyzed. As seen in Fig. 7A clear differences between the two cell lines were detected for all three proteins. Protein 4 was produced in CHO, HEK EBNA and HEK6E.

### Table 3

Examples of expression levels (ng/l) in HEK EBNA cells and in CHO cells with the standard and the optimized transfection protocols.

<table>
<thead>
<tr>
<th>Protein</th>
<th>HEK EBNA</th>
<th>CHO standard</th>
<th>CHO optimized</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>8.46</td>
<td>2.07</td>
<td>5.12</td>
</tr>
<tr>
<td>3</td>
<td>10.40</td>
<td>1.69</td>
<td>2.62</td>
</tr>
<tr>
<td>5</td>
<td>4.50</td>
<td>0.93</td>
<td>1.87</td>
</tr>
<tr>
<td>6</td>
<td>1.66</td>
<td>1.26</td>
<td>1.35</td>
</tr>
<tr>
<td>10</td>
<td>3.32</td>
<td>1.02</td>
<td>1.63</td>
</tr>
<tr>
<td>11</td>
<td>4.52</td>
<td>1.25</td>
<td>1.58</td>
</tr>
<tr>
<td>12</td>
<td>4.90</td>
<td>2.07</td>
<td>4.88</td>
</tr>
</tbody>
</table>

Theoretical molecular weight of 10,985 Da. The major mass peak was found at 10,969 Da, which corresponds to the protein with an N-terminal pyro-glutamate. The protein from CHO cells showed in addition to the mass peaks corresponding to the unglycosylated protein a minor mass peak at 11,917 Da, indicating a low degree of glycosylation. The protein from HEK cells showed a more complex glycosylation pattern, with 7 mass peaks, between 11,626 and 13,595 Da. As expected treatment with PNGaseF did not have any effect on the mass peak pattern. Two batches from each cell line were analyzed and the reproducibility was very good between the batches (results not shown). These results are consistent with the pattern on SDS-PAGE, where the HEK sample shows an extensive smear over the main band, which is virtually absent in the CHO sample (Fig. 1A).
Two batches from each cell line were analyzed. The CHO derived material showed a clear different pattern of glycans compared to the HEK derived protein (Fig. 7B). The difference between the two HEK cell lines was minor and presumably within the interbatch variation. The analysis of protein 9 again showed clear differences between the HEK and CHO derived protein (Fig. 7B). However, in addition the glycan pattern between the two HEK cell lines was significantly different. The HEK6E cell line showed a much simpler glycan pattern compared to HEK EBNA. These results are consistent with the SDS-PAGE and IEF results. On SDS-PAGE the HEK6E derived protein is running at an apparent lower molecular weight compared to the HEK EBNA derived protein, suggesting less glycosylation (Fig. 4A).

4. Discussion

Virtually all proteins used in biological and biomedical research today are recombinant proteins expressed and purified from different expression systems. In addition, the number of proteins approved for therapeutic use is increasing rapidly (Durocher and Butler, 2009). All new therapeutic proteins are produced through recombinant procedures, mostly in mammalian expression systems in order to maintain mammalian posttranslational modifications. Recombinant proteins produced in mammalian expression systems are frequently regarded as identical, or at least very similar, to the native protein. Although, differences in glycan structures for specific proteins expressed in various mammalian cell lines have been reported (Van den Nieuwenhof et al., 2000). In an earlier study of the adhesion protein GlialCAM we noticed significant differences in the glycosylation when this protein was produced in HEK293 and CHO-S cells (Gaudry et al., 2008). However, it remained unclear whether this is a general phenomena or specific for this particular protein. There are reports that culture conditions can affect the glycosylation pattern, although the direct causes for the differences mostly remain unclear (Pacis et al., 2011). Changes in culture or production conditions predominantly appear to result in a redistributions of the various isoforms, but less of the appearance of new isoforms (Schiestl et al., 2011).

Information on glycosylation pattern of proteins expressed in different mammalian cell lines remains very restricted. For production of recombinant proteins a large number of cell lines are used, originating from different tissues and different species. Thus, to further investigate protein glycosylation we produced and analyzed a set of recombinant proteins in HEK and CHO cells, two frequently used cell lines. These two cell lines represent two different species, human and hamster, as well as two tissues, kidney and ovary cells, respectively. Here we detected very significant differences of glycosylation between the proteins expressed in HEK and CHO cells. The differences in the amount of glycans attached to the proteins were so significant that it could be detected on the relatively insensitive SDS-PAGE analysis. Either the proteins migrated at different apparent molecular weights or the bands were more or less diffuse, indication of a more heterogeneous mass distribution. Since glycosylation has been shown to influence several protein functions, including protein–protein interactions and activity, the cell line used for the production of recombinant proteins could have a significant effect on their function in in vitro assays and in vivo animal models. In vivo it has been shown that the same protein expressed in different tissues can have different glycosylation profiles, however these differences have not been linked to any particular changes in the protein activity (West et al., 2010). Furthermore, it is well known that posttranslational modifications, including glycosylation, frequently are changed in tumor cells compared to the originating tissue (Meany and Chan, 2011; Saeland et al., 2011).

Another striking difference between the two cell lines was the isofrom pattern detected on IEF, which showed to be due largely to differences in the sialic acid content. The content of sialic acid has been shown to influence protein activity in vitro and it is known to influence the biological half-life of proteins in vivo (Lee et al., 2002; Otto et al., 2004; Wide et al., 2010). Proteins lacking terminal sialic acids are binding to receptor in the liver and are eliminated, reducing their half-life (Ashwell and Harford, 1982). In general, the proteins expressed in CHO cells had more acidic isoforms compared to the HEK derived protein, which could favor the in vitro half-life of proteins produced in CHO compared to HEK cells. However this is
Fig. 6. Mass spectrometry analysis. Proteins expressed in CHO or HEK cells were analyzed before and after treatment with PNGaseF. Spectra were recorded between m/z 500 and 2000. (A) Protein 3, the negative control with no predicted glycosylation sites, expressed in CHO and HEK EBNA cells. (B) Protein 2, containing one predicted O-linked glycosylation site, expressed in CHO and HEK EBNA cells. (C) Protein 4, containing 2 predicted N-linked and 3 O-linked glycosylation site, expressed in CHO, HEK EBNA and HEK 6E cells. Two batches from each cell line were analyzed.

not an absolute rule, for example protein 1 had more acidic isoforms when expressed in HEK compared to CHO cells.

An additional implication of the large difference in isoform distribution could be for purification of the recombinant proteins, when ionic exchange chromatography, which is based on the charge of the protein, is used in the purification protocol.

The two cell lines are derived from two different species and from two different cell types, thus significant differences in the glycosylation profiles might be expected. Several reasons for

Fig. 7. Capillary gel electrophoresis analysis of N-linked glycans released from the proteins with PNGaseF treatment. After APS labeling the glycans were separated by capillary electrophoresis. (A) The glycan profiles for proteins 5, 11 and 12 from CHO and HEK EBNA derived proteins. (B) Glycan profiles of proteins 4 and 9 from CHO, HEK EBNA and HEK 6E derived proteins. Two batches from each cell line were analyzed.

The differences could be imagined. The panels of natural proteins produced in ovary and kidney cells certainly are different, and the requirement for glycosylation might not be the same. Over 250 different enzymes have been identified to be involved in protein glycosylation (van den Eijnden and Joziasse, 1993). Whether the expression of enzymes participating in protein glycosylation is considerably different in the two cell lines remains unknown, but could be a potential explanation for the differences detected. Most cells presumably have a low expression level of the enzymes involved in protein glycosylation but increased expression levels of specific enzymes have been shown in specific cell types or phases of cell development (Gillespie et al., 1993; Harduin-Lepers et al., 1993).

At the early stage in research, when only small quantities of the proteins are needed, transient expression in HEK cells is frequently used, whereas at later stages, when larger quantities are needed, the expression system is changed to stable expression in CHO cells. Frequently, the recombinant proteins are considered identical and results gathered in in vitro and in vivo assays and models with protein produced in HEK cells are not necessary verified with the CHO
derived protein. In light of the significant differences detected in the glycosylation of all proteins analyzed, different functional results might be obtained depending on the production cell line used. Conversely, if differences in activity is not detected this could indicate that changes in the glycosylation pattern is less important and thus slight differences due to different production conditions might not have a adverse effect on the protein quality. It remains unclear whether glycosylation is the main protein backbone modification impacted by the cell type or if other posttranslational modifications are also significantly different.

5. Conclusion

Although some differences in glycosylation of HEK derived proteins were detected when the proteins were expressed in HEK ERNA cells or in a variant HEK 6E, and when the culture media and conditions were changed, the major differences in the glycosylation was detected between the CHO and HEK cell lines. The results clearly show that proteins expressed in two frequently used mammalian cell lines have very significant differences in their glycosylation pattern. The size and number of the glycostructures varies, as well as the amount of sialic acid, resulting in different IEF isofrom patterns, making the protein population more or less acidic. What effects these differences have on activity, protein interactions, stability, biodistribution and in vivo half-life remains to be determined. Nevertheless, in comparing data generated with mammalian produced recombinant proteins the cell type or cell line used for the production needs to be considered.

Contributions

Amlie Croset and Laurence Delafosse performed most of the experiments and participated in the planning of the experiments.

Ana Krstanovic and Flavie Robert planned and performed the capillary gel electrophoresis glycoanalysis.

Francis Vilbois and Damien Begue performed the LC–MS analysis.

Christophe Losberger performed the capillary IEF.

Loic Gleiz supervised the bioprocessing work.

Jean-Philippe Gaudry and Christian Arold supervised the protein purification.

Laurent Chevalet and Bruno Antonsson planned the experiments and Bruno Antonsson wrote the manuscript.

Acknowledgements

A. Krstanovic and F. Robert thank Erdmann Rapp and René Henning from Glyxera®, and Francesca Cuttillo and Horst Bierau from Merck Serono SpA Italy.

References

glycodelin-A can be produced in human kidney 293 cells but not in Chinese hamster ovary cells. European Journal of Biochemistry 267, 4753–4762.