Comprehensive Characterization of the N and O-Linked Glycosylation of a Recombinant Human EPO

Matthew A. Lauber, Stephan M. Koza, and Erin E. Chambers
Waters Corporation, Milford, MA, USA

APPLICATION BENEFITS

■ Two facile strategies to elucidate information about both the N and O-linked glycosylation of EPO
■ Unprecedented HILIC separations of high antennarity released N-glycans and intact protein glycoforms
■ MS compatible HILIC to enable detailed investigations of sample constituents
■ ACQUITY UPLC® Glycoprotein BEH Amide Column (300Å, 1.7 µm stationary phase) is QC tested via a glycoprotein separation to ensure consistent batch-to-batch reproducibility

INTRODUCTION

The immunoglobulin G (IgG) modality has paved the way for many efficacious protein-based therapies. At the same time, numerous highly effective patient therapies have also been made possible by the production of recombinant, human hormones and enzymes. For example, erythropoiesis stimulating therapeutics, like epoetin (EPO) alpha, have long been available for the treatment of anemia. Such a therapy for increasing patient red blood cell counts was first made possible by the commercialization of Epogen, which has been available in the US market since its approval by the FDA in 1989. And now, because the landscape of the biopharmaceutical industry continues to evolve and Epoen patents expired in 2013, EPO drug products are targets for being developed into both international and domestic-market biosimilars.

Epoetin alpha has a relatively small primary structure, yet it has 3 sites of N-glycosylation and 1 site of O-glycosylation (Figure 1). Because of its glycosylation, epoetin alpha has a molecular weight between 30 and 40 kDa even though its protein mass amounts to only 18 kDa. Interestingly, the glycosylation of epoetin is very much tied to its potency and serum half life. Two attributes of its glycan profile that are known to show positive correlations with in vivo activity include antennarity and sialylation. As a result, it is critical for the glycosylation of an epoetin therapeutic to be well characterized. In addition, the significance of epoetin glycosylation suggests that detailed glycan profiling would be a path toward establishing a viable epoetin biosimilar.

Figure 1. Sequence and structural information for recombinant, human epoetin alpha (rhEPO).
**EXPERIMENTAL**

**Sample description**

A recombinant, human epoetin alpha expressed from CHO cells (PeproTech, Rocky Hill, NJ) was reconstituted in 50 mM HEPES NaOH pH 7.9 buffer to a concentration of 2 mg/mL.

N-glycans were released from rhEPO and labeled with RapiFluor-MS using a GlycoWorks RapiFluor-MS N-Glycan Kit and the instructions provided in its care and use manual (p/n 715004793). RapiFluor-MS labeled N-glycans were injected as a mixture of 90 µL SPE eluate, 100 µL dimethylformamide, and 210 µL acetonitrile.

To facilitate analysis of O-glycosylation, rhEPO was N-deglycosylated using the rapid deglycosylation technique outlined in the care and use manual of the GlycoWorks RapiFluor-MS N-Glycan Kit (p/n 715004793).

**Method conditions (unless otherwise noted)**

**Column Conditioning**

New (previously unused) ACQUITY UPLC Glycoprotein BEH Amide, 300Å, 1.7 µm Columns should be conditioned via two or more sequential injections and separations until a consistent profile is achieved. The care and use manual of the column can be referred to for more information (p/n 720005408EN).

**LC conditions for RapiFluor-MS Released N-Glycans**

- **LC system:** ACQUITY UPLC H-Class Bio System
- **Sample temp.:** 10 °C
- **Analytical column temp.:** 60 °C
- **Flow rate:** 0.4 mL/min
- **Injection volume:** 10 µL
- **Column:** ACQUITY UPLC Glycoprotein BEH Amide, 300Å, 1.7 µm, 2.1 x 150 mm (p/n 176003702, with Glycoprotein Performance Test Standard)
- **Fluorescence detection:** Ex 265 nm / Em 425 nm, 2 Hz
- **Sample collection/Vials:** Sample Collection Module (p/n 186007988) Polypropylene 12 x 32 mm Screw Neck vial, 300 µL volume (p/n 186002640)

**Mobile phase A:** 50 mM ammonium formate, pH 4.4 (LC-MS grade; from a 100x concentrate, p/n 186007081)

**Mobile phase B:** ACN (LC-MS grade)

<table>
<thead>
<tr>
<th>Time</th>
<th>Flow Rate (mL/min)</th>
<th>%A</th>
<th>%B</th>
<th>Curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.4</td>
<td>25</td>
<td>75</td>
<td>6</td>
</tr>
<tr>
<td>35.0</td>
<td>0.4</td>
<td>46</td>
<td>54</td>
<td>6</td>
</tr>
<tr>
<td>36.5</td>
<td>0.2</td>
<td>100</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>39.5</td>
<td>0.2</td>
<td>100</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>43.1</td>
<td>0.2</td>
<td>25</td>
<td>75</td>
<td>6</td>
</tr>
<tr>
<td>47.6</td>
<td>0.4</td>
<td>25</td>
<td>75</td>
<td>6</td>
</tr>
<tr>
<td>55.0</td>
<td>0.4</td>
<td>25</td>
<td>75</td>
<td>6</td>
</tr>
</tbody>
</table>

**MS conditions for RapiFluor-MS Released N-Glycans**

- **MS system:** Xevo G2-XS QTof
- **Ionization mode:** ESI+
- **Analyzer mode:** Resolution (~40 K)
- **Capillary voltage:** 2.2 kV
- **Cone voltage:** 75 V
- **Source temp.:** 120 °C
- **Desolvation temp.:** 500 °C
- **Source offset:** 50 V
- **Desolvation gas flow:** 600 L/Hr
- **Calibration:** NaI, 1 µg/µL from 100–2000 m/z
- **Acquisition:** 700–2000 m/z, 0.5 sec scan rate
- **Lockspray:** 300 fmol/µL Human glufibrinopeptide B in 0.1% (v/v) formic acid, 70:30 water/acetonitrile every 90 seconds
- **Data management:** MassLynx® Software v4.1

**LC conditions for Intact Protein HILIC of N-Deglycosylated rhEPO**

- **LC system:** ACQUITY UPLC H-Class Bio System
- **Sample temp.:** 10 °C
- **Analytical column temp.:** 45 °C
- **Flow rate:** 0.2 mL/min
- **Fluorescence detection:** Ex 280 nm/Em 320 nm (Intrinsic fluorescence), 10 Hz
In this application note, we demonstrate the use of two facile strategies that can be used to detail the N and O-linked glycosylation of a recombinant, human epoetin (rhEPO). In this work, rhEPO N-glycans were rapidly released, labeled with GlycoWorks RapiFluor-MS and profiled by hydrophilic interaction chromatography (HILIC) using sensitive fluorescence and mass spectrometric detection. Then, in a second, parallel analysis, N-deglycosylated rhEPO was interrogated by intact protein HILIC to elucidate information on O-glycosylation.

**RESULTS AND DISCUSSION**

**Released N-Glycan analysis of rhEPO using RapiFluor-MS labeling and HILIC profiling**

The glycosylation of recombinant, human epoetin (rhEPO) has been investigated many times before.4-5, 8-13 In large part, these previous studies have required relatively involved techniques. With this work, it was our objective to establish two facile and complementary, LC based approaches for the analysis of EPO, one capable of providing information about N-glycosylation and the other information about O-glycosylation.

A profile of the N-glycans from rhEPO can be readily obtained with a new sample preparation strategy involving the novel glycan labeling reagent, RapiFluor-MS. This sample preparation, based on the GlycoWorks RapiFluor-MS N-Glycan Kit, allows an analyst to rapidly release N-glycans and label them with a tag that provides enhanced sensitivity for fluorescence and electrospray ionization mass spectrometric (ESI-MS) detection.14 In previous applications, RapiFluor-MS has been predominately used in the analysis of different IgG samples.14-16 Nevertheless, using the protocol from the GlycoWorks RapiFluor-MS N-Glycan Kit, an analyst can successfully prepare samples from even heavily glycosylated proteins, such as rhEPO.

RapiFluor-MS labeled N-glycans have proven to be amenable to hydrophilic interaction chromatography (HILIC). Accordingly, HILIC-fluorescence-MS of RapiFluor-MS has emerged as a very powerful tool for detailing the N-glycosylation of proteins.14
To this end, a sample of RapiFluor-MS N-glycans derived from rhEPO was profiled using HILIC. A recently introduced widepore amide column, the ACQUITY UPLC Glycoprotein BEH Amide, 300Å, 1.7µm Column, was selected for this work to obtain high resolution N-glycan separations. This column was purposefully designed to facilitate HILIC separations of large molecules, such as glycopeptide and glycoproteins. However, the widepore particle architecture has also been shown to increase the peak capacity of highly branched, tri- and tetra-antennary N-glycans by 10–20%, making it an ideal choice for the HILIC profiling of EPO N-glycans, which typically exhibit high antennarity. Figure 2A shows the HILIC fluorescence and base peak intensity (BPI) MS chromatograms of the RapiFluor-MS N-glycans resulting from 0.4 µg of rhEPO. Even with this relatively limited amount of sample, high signal-to-noise chromatograms are obtained. The sensitivity of the fluorescence trace allows for accurate, relative quantitation across the profile. The signal-to-noise of the MS chromatogram is also particularly noteworthy, though it should be noted that MS sensitivity decreases as N-glycan structures become larger. Nevertheless, the quality of these particular data is made possible by use of the RapiFluor-MS reagent in combination with the Xevo G2-XS QTof, a new generation MS instrument with improved transmission efficiency and sensitivity. This QTof technology provides unprecedented sensitivity as well as high mass resolution, as can be observed in the collection of mass spectra in Figure 2B that have been used to support the assignment of various N-glycan species.

Figure 2. HILIC profiling of released N-glycans from rhEPO. (A) Fluorescence and (B) base peak intensity (BPI) chromatograms for RapiFluor-MS labeled N-glycans from rhEPO. Chromatograms obtained for glycans from 0.4 µg protein using an ACQUITY UPLC Glycoprotein BEH Amide, 300Å, 1.7 µm, 2.1 x 150 mm Column. (C) MS spectra for four example N-glycan species. N-glycan assignments are listed according to Oxford notation. “+Ac” denotes an acetylation, such as the previously reported O-acetylation of sialic acid residues (Neu5Ac).
The chromatographic and MS-level selectivity afforded by this analysis simplifies making N-glycan assignments such that the species of the rhEPO N-glycan profile were easily mapped (Figure 3).

The rhEPO analyzed in this study exhibits an N-glycan profile comprised primarily of tetra-antennary, tetradsialylated N-glycans (FA4G4S4) with varying N-acetyl lactosamine extensions. However, the profile also shows a highly abundant peak that corresponds to a disialylated, biantennary N-glycan (FA2G2S2). Given that the ratio of tetra-antennary to biantennary N-glycans has a positive correlation with the in vivo activity of an EPO, this analysis has clearly produced valuable information. Other information that can be readily obtained from this N-glycan analysis includes the degree of sialylation and the extent to which structures are modified with lactosylamine extensions. Overall, these results demonstrate that a very-information rich N-glycan profile can indeed be obtained from a comparatively simple RapiFluor-MS N-glycan preparation and a corresponding HILIC-fluorescence-MS analysis.

<table>
<thead>
<tr>
<th>RT (min)</th>
<th>Species</th>
<th>MW&lt;sub&gt;Mono, Theo&lt;/sub&gt; (Da)</th>
<th>Observed m/z</th>
<th>z</th>
<th>MW&lt;sub&gt;Mono, Obs&lt;/sub&gt; (Da)</th>
<th>Mass error (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.21</td>
<td>FA2G2S1</td>
<td>2388.9201</td>
<td>1195.4659</td>
<td>2</td>
<td>2388.9172</td>
<td>1.2</td>
</tr>
<tr>
<td>18.12</td>
<td>FA2G2S2</td>
<td>2680.0155</td>
<td>894.3492</td>
<td>3</td>
<td>2680.0258</td>
<td>-3.8</td>
</tr>
<tr>
<td>22.24</td>
<td>FA3G3S3</td>
<td>3336.2432</td>
<td>1113.0924</td>
<td>3</td>
<td>3336.2554</td>
<td>-3.7</td>
</tr>
<tr>
<td>23.68</td>
<td>FA4G4S4 + Ac</td>
<td>4034.4813</td>
<td>1345.8372</td>
<td>3</td>
<td>4034.4898</td>
<td>-2.1</td>
</tr>
<tr>
<td>24.15/24.60</td>
<td>FA4G4S3</td>
<td>3701.3754</td>
<td>1234.7966</td>
<td>3</td>
<td>3701.368</td>
<td>2.0</td>
</tr>
<tr>
<td>25.52</td>
<td>FA4G4S4</td>
<td>3992.4708</td>
<td>1331.8309</td>
<td>3</td>
<td>3992.4709</td>
<td>0.0</td>
</tr>
<tr>
<td>25.7</td>
<td>FA4G4Lac1S4 + Ac</td>
<td>4399.6135</td>
<td>1467.5425</td>
<td>3</td>
<td>4399.6057</td>
<td>1.8</td>
</tr>
<tr>
<td>26.16/26.66</td>
<td>FA4G4Lac1S3</td>
<td>4066.5076</td>
<td>1356.5104</td>
<td>3</td>
<td>4066.5094</td>
<td>-0.4</td>
</tr>
<tr>
<td>27.34</td>
<td>FA4G4Lac1S4</td>
<td>4357.6030</td>
<td>1090.4097</td>
<td>4</td>
<td>4357.6097</td>
<td>-1.5</td>
</tr>
<tr>
<td>27.95</td>
<td>FA4G4Lac2S3</td>
<td>4431.6397</td>
<td>1108.9143</td>
<td>4</td>
<td>4431.6281</td>
<td>2.6</td>
</tr>
<tr>
<td>28.97</td>
<td>FA4G4Lac2S4</td>
<td>4722.7352</td>
<td>1181.6909</td>
<td>4</td>
<td>4722.7345</td>
<td>0.1</td>
</tr>
<tr>
<td>29.66</td>
<td>FA4G4Lac3S3</td>
<td>4796.7719</td>
<td>1200.2004</td>
<td>4</td>
<td>4796.7725</td>
<td>-0.1</td>
</tr>
<tr>
<td>30.50</td>
<td>FA4G4Lac3S4</td>
<td>5087.8674</td>
<td>1272.976</td>
<td>4</td>
<td>5087.8749</td>
<td>-1.5</td>
</tr>
<tr>
<td>31.77</td>
<td>FA4G4Lac4S4</td>
<td>5452.9996</td>
<td>1364.256</td>
<td>4</td>
<td>5452.9949</td>
<td>0.9</td>
</tr>
</tbody>
</table>

*Figure 3. LC-MS data supporting the identification of various released N-glycan species. “+Ac” denotes an acetylation, such as the previously reported O-acetylation of sialic acid residues (Neu5Ac).*
Profiling the O-Glycosylation of Intact rhEPO using a Widepore Amide HILIC Separation

O-linked glycans can be challenging to characterize due to the paucity of high fidelity mechanisms to release them from their counterpart proteins. Released glycan analysis is an attractive approach for the characterization of N-glycans because of the simplicity and effectiveness of PNGase F deglycosylation. In place of using an analogous, universal glycosidase, analysts have resorted to releasing O-linked glycans by chemical means, such as alkaline beta elimination or hydrazinolysis. These release mechanisms can be challenging to implement and can very often produce artifacts, known as peeling products.

Rather than attempt a released O-glycan analysis of rhEPO, we looked to develop an alternative characterization strategy. A novel workflow was devised that first involved subjecting the rhEPO to rapid deglycosylation using GlycoWorks Rapid PNGase F and 1% RapiGest™ SF surfactant. In a 10-minute preparation, a sample of N-deglycosylated intact rhEPO was obtained that could then be profiled via a HILIC separation with an ACQUITY UPLC Glycoprotein BEH Amide Column. Figure 4 presents the chromatogram obtained in this analysis using intrinsic fluorescence detection and intact protein HILIC techniques that have been described in previous work. The N-deglycosylated rhEPO analyzed in this study resolved into a series of approximately 10 peaks. Online ESI-MS provided highly detailed information, allowing for proteoforms of rhEPO to be assigned to the various chromatographic peaks. The two most abundant LC peaks were found to be represented by deconvoluted masses of 18893.8 and 19185.3 Da, which are consistent with N-deglycosylated rhEPO that has a C-terminal arginine truncation as well as trisaccharide and tetrasaccharide O-linked glycan modifications, respectively. More specifically, the mass shift observed for the lighter species is indicative of a glycan modification comprised of 1 hexose, 1 N-acetylhexosamine, and 1 N-acetylneuraminic acid. Meanwhile, the mass shift observed for the heavier species suggests a glycan modification comprised of the same structure with an additional N-acetyl neuraminic acid.

---

Figure 4. HILIC-fluorescence-MS analysis of N-deglycosylated, intact rhEPO. (A) Fluorescence chromatogram demonstrating O-linked glycan heterogeneity and occupancy. Chromatograms obtained from 0.7 μg protein using a 2.1 x 150 mm ACQUITY UPLC Glycoprotein BEH Amide, 300Å, 1.7 μm, 2.1 x 150 mm Column. (B) Deconvoluted mass spectra corresponding to three of the major rhEPO proteoforms. Peak identifications, in addition to those denoted here, are tabulated in Figure 5.
Further investigation of the LC-MS data also showed that the proteoform of rhEPO that is aglycosylated with respect to the O-linked glycan eluted with a retention time of approximately 8.2 min. Moreover, these LC-MS data indicated there to be at least two additional O-linked glycoforms and even more C-terminal truncation proteoforms (Figure 5). Here, it is seen that this workflow can indeed be used to rapidly profile the O-linked glycosylation of an rhEPO, such that information is gained about both occupancy and heterogeneity.

<table>
<thead>
<tr>
<th>RT (min)</th>
<th>Species</th>
<th>$\text{MW}_{\text{Avg, Theo}}$ (Da)</th>
<th>$\text{MW}_{\text{Avg, Obs}}$ (Da)</th>
<th>Mass Error (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.0</td>
<td>N-deglycosylated, –C-term GDR</td>
<td>18066.5</td>
<td>18065.2</td>
<td>-1.3</td>
</tr>
<tr>
<td></td>
<td>N-deglycosylated, –C-term DR</td>
<td>18123.6</td>
<td>18122.4</td>
<td>-1.2</td>
</tr>
<tr>
<td>8.2</td>
<td>N-deglycosylated, –C-term R</td>
<td>18238.7</td>
<td>18237.4</td>
<td>-1.3</td>
</tr>
<tr>
<td>9.3</td>
<td>N-deglycosylated, –C-term R +Hex1HexNAc1Neu5Ac1+Ac</td>
<td>18937.3</td>
<td>18936.2</td>
<td>-1.1</td>
</tr>
<tr>
<td>9.5</td>
<td>N-deglycosylated, –C-term GDR +Hex1HexNAc1Neu5Ac1</td>
<td>18723.1</td>
<td>18722.3</td>
<td>-0.8</td>
</tr>
<tr>
<td></td>
<td>N-deglycosylated, –C-term DR +Hex1HexNAc1Neu5Ac1</td>
<td>18780.1</td>
<td>18779.1</td>
<td>-1.0</td>
</tr>
<tr>
<td>9.7</td>
<td>N-deglycosylated, –C-term R +Hex1HexNAc1Neu5Ac1</td>
<td>18895.2</td>
<td>18893.8</td>
<td>-1.4</td>
</tr>
<tr>
<td>9.9</td>
<td>N-deglycosylated, –C-term R +Hex1HexNAc1Neu5Ac2+Ac</td>
<td>19228.5</td>
<td>19227.3</td>
<td>-1.2</td>
</tr>
<tr>
<td>10.0</td>
<td>N-deglycosylated, –C-term R +Hex1HexNAc1Neu5Ac1 + O</td>
<td>18911.2</td>
<td>18910.0</td>
<td>-1.2</td>
</tr>
<tr>
<td>10.2</td>
<td>N-deglycosylated, –C-term GDR +Hex1HexNAc1Neu5Ac2</td>
<td>19014.3</td>
<td>19013.7</td>
<td>-0.6</td>
</tr>
<tr>
<td>10.5</td>
<td>N-deglycosylated, –C-term R +Hex1HexNAc1Neu5Ac2</td>
<td>19186.5</td>
<td>19185.3</td>
<td>-1.2</td>
</tr>
<tr>
<td>10.8</td>
<td>N-deglycosylated, –C-term R +Hex1HexNAc1Neu5Ac2 + O</td>
<td>19202.5</td>
<td>19201.2</td>
<td>-1.3</td>
</tr>
</tbody>
</table>

Figure 5. LC-MS data supporting the identification of various N-deglycosylated rhEPO proteoforms. “–C-term” denotes the C-terminal truncation of the rhEPO; losses of different residues are noted. Hex, HexNAc, and Neu5Ac stand for hexose, N-acetylhexosamine, and N-acetylneuraminic acid. For example, Hex1HexNAc1Neu5Ac1 corresponds to O-glycosylation involving 1 hexose, 1 N-acetylhexosamine, and 1 N-acetylneuraminic acid. “+O” denotes a mass shift indicative of the addition of an oxygen atom, such as an oxidation or an exchange of Neu5Ac for Neu5Gc. Data supporting identifications of the most abundant rhEPO sequence variant (–C-term R) and its glycoforms are highlighted with bold text. “+Ac” denotes an acetylation, such as the previously reported O-acetylation of sialic acid residues (Neu5Ac).
Several powerful tools have recently emerged for the analysis of glycans that are built upon LC-MS compatible hydrophilic interaction chromatography (HLIC). At the heart of these new glycan analysis workflows is a HLIC column that has been purposefully designed for large molecule separations. With this ACQUITY UPLC Glycoprotein BEH Amide Column, an analyst can achieve higher resolution separations of large, released N-glycans. And when this analysis is paired with RapiFlour-MS labeling, a technique is established that affords not only high resolution but also unprecedented sensitivity. This approach has been successfully applied to obtain highly detailed information about the N-glycosylation of a recombinant, human epoetin alpha (rhEPO). Given that N-glycosylation correlates with the half life and activity of an EPO, such information, with its unparalleled quality, would be invaluable in developing a new EPO therapeutic. EPO is also O-glycosylated; the occupancy and heterogeneity of which could also be critical to demonstrate comparability among different drug substances. Using the ACQUITY UPLC Glycoprotein BEH Amide Column, we have outlined a simple sample preparation and subsequent HLIC separation that is capable of profiling these O-glycan attributes on intact rhEPO. In summary, we have demonstrated the use of two facile strategies that can be used to detail both the N and O-linked glycosylation of recombinant, human epoetin (rhEPO), a molecule which has been perceived to be challenging to characterize due to its relatively complicated glycosylation. Collectively, these tools could be used to accelerate the development of new biosimilars.

References


CONCLUSIONS

Several powerful tools have recently emerged for the analysis of glycans that are built upon LC-MS compatible hydrophilic interaction chromatography (HLIC). At the heart of these new glycan analysis workflows is a HLIC column that has been purposefully designed for large molecule separations. With this ACQUITY UPLC Glycoprotein BEH Amide Column, an analyst can achieve higher resolution separations of large, released N-glycans. And when this analysis is paired with RapiFlour-MS labeling, a technique is established that affords not only high resolution but also unprecedented sensitivity. This approach has been successfully applied to obtain highly detailed information about the N-glycosylation of a recombinant, human epoetin alpha (rhEPO). Given that N-glycosylation correlates with the half life and activity of an EPO, such information, with its unparalleled quality, would be invaluable in developing a new EPO therapeutic. EPO is also O-glycosylated; the occupancy and heterogeneity of which could also be critical to demonstrate comparability among different drug substances. Using the ACQUITY UPLC Glycoprotein BEH Amide Column, we have outlined a simple sample preparation and subsequent HLIC separation that is capable of profiling these O-glycan attributes on intact rhEPO. In summary, we have demonstrated the use of two facile strategies that can be used to detail both the N and O-linked glycosylation of recombinant, human epoetin (rhEPO), a molecule which has been perceived to be challenging to characterize due to its relatively complicated glycosylation. Collectively, these tools could be used to accelerate the development of new biosimilars.

