Journal of Chromatography A, 1498 (2017) 128–146

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Orthogonal liquid chromatography–mass spectrometry methods for the comprehensive characterization of therapeutic glycoproteins, from released glycans to intact protein level

Eric Largy[a], Fabrice Cantais[a], Géry Van Vyncht[a], Alain Beck[b], Arnaud Delobel[a,*]

[a] Quality Assistance sa, Technoparc de Thudinie 2, 6536, Donstiennes, Belgium
[b] Centre d'Immunologie Pierre Fabre (CIPF), 5 Av. Napoléon III, BP 60497, 74164, Saint-Julien-en-Genevois, France

ARTICLE INFO

Article history:
Received 24 August 2016
Received in revised form 4 January 2017
Accepted 28 February 2017
Available online 27 March 2017

Keywords:
High-resolution mass spectrometry
Hydrophilic interaction chromatography
Mixed-mode chromatography
Porous graphitic carbon chromatography
Protein glycosylation
Therapeutic proteins

ABSTRACT

Proteins are increasingly used as therapeutics. Their characterization is challenging due to their size and inherent heterogeneity notably caused by post-translational modifications, among which glycosylation is probably the most prominent. The glycosylation profile of therapeutic proteins must therefore be thoroughly analyzed. Here, we illustrate how the use of a combination of various cutting-edge LC or LC/MS(/MS) methods, and operating at different levels of analysis allows the comprehensive characterization of both the *N*- and *O*-glycosylations of therapeutic proteins without the need for other approaches (capillary electrophoresis, MALDI-TOF). This workflow does not call for the use of highly specialized/custom hardware and software nor an extensive knowledge of glycan analysis. Most notably, we present the point of view of a contract research organization, with the constraints associated to the work in a regulated environment (GxP). Two salient points of this work are i) the use of mixed-mode chromatography as a fast and straightforward mean of profiling *N*-glycans sialylation as well as an orthogonal method to separate *N*-glycans co-eluting in the HILIC mode; and ii) the use of widepore HILIC/MS to analyze challenging *N*/*O*-glycosylation profiles at both the peptide and subunit levels. A particular attention was given to the sample preparations in terms of duration, specificity, versatility, and robustness, as well as the ease of data processing.

1. Introduction

Proteins are increasingly used as therapeutics, and form a large subset of the best-selling drugs [1]. These include monoclonal antibodies (mAbs) and proteins produced by recombinant DNA technology (fusion proteins, cytokines and growth factors). Their characterization, from development to post-marketing quality control, is challenging due to their size and inherent heterogeneity notably caused by post-translational modifications (PTMs) [2–4]. Among these, glycosylation is probably the most prominent in terms of occurrence and alteration of protein functions [5–8]. Protein glycans are linked on asparagines in Asn-*X*-Ser or Asn-*X*-Thr motifs ($X \neq$ Pro) for *N*-glycosylation, or on any serines and threonines for *O*-glycosylation.

An adequate glycosylation is critical for therapeutic proteins in terms of structure and binding [9–16], solubility and stability [8,17,18], safety [19], and pharmacokinetics and dynamics [8,20]. Consequently, the glycosylation profile of therapeutic proteins is considered a critical quality attribute, for both innovators and biosimilars, and must be thoroughly analyzed [21–24]. Its early characterization by sensitive assays helps taking informed decisions in latter development stages. For instance, the degree, nature and linkage of sialylation can significantly influence the safety and efficacy of protein-based drugs, notably related to their half-life [25], structural stability [26]s, anti-inflammatory properties [27], and immunogenicity [28–30]. In particular, they participate in the prevention of protease-mediated degradation and the stabilization of the correct conformations of glycoproteins [18], their clearance [25], and the recognition and binding to cells [31]. Finally, protein glycosylation of therapeutic proteins can be engineered to achieve improved pharmacological properties, for instance improved antibody-dependent cell-mediated cytotoxicity (ADCC) activity [9,32,33,16,34].

Therapeutic proteins are typically produced in different expression systems, whose glycosylation machineries function through sequential and competitive steps, hence creating micro- (glycans

* Corresponding author.
E-mail address: arnaud.delobel@quality-assistance.be (A. Delobel).

http://dx.doi.org/10.1016/j.chroma.2017.02.072
0021-9673/

nature for a given site) and macro heterogeneities (number and location of sites) of glycosylation [4,7,35,36]. For instance, production in SP2/0 cells yields a fraction of bisecting glycans, as well as galactose-α-1,3-galactose motifs, whereas CHO cells do not [19]. In addition, the branching nature of glycans creates another level of complexity compared to linear nucleic acids and amino-acids sequences [7]. This results in a challenging analytical puzzle that requires a number of orthogonal analytical techniques at different levels of analysis to be solved in order to comply with regulatory requirements on this increasingly scrutinized PTM [36–40]. Hence, the conference of harmonization (ICH) documents Q5E and Q6B require for glycoproteins the determination of the carbohydrate content (neutral sugars, amino sugars, and sialic acids), and, to the extent possible, the oligosaccharide pattern (antennary profile) and the glycosylation site(s) of the polypeptide chain [38,39]. On the specific matter of mAbs, the EMA states that glycans structure should be characterized, in particular in terms of mannosylation, galactosylation, fucosylation, and silalylation, and that the distribution of main glycan structures (e.g. G0, G1, G2) should be determined [41].

A variety of methods have been developed to tackle these challenges, including – but not limited to – MALDI-TOF, hydrophilic interaction liquid chromatography-ultra high performance liquid chromatography (HILIC-UPLC), reversed-phase high performance liquid chromatography (RPLC), high-performance anion-exchange chromatography (HPAEC) or capillary electrophoresis (CE), often coupled to fluorescence, amperometric, and mass spectrometric detection method (for insightful reviews and comparisons of methods see references [42–46]). A number of separation and detection methods, at all levels of analysis, are included in the pharmacopeia (Eur. Ph. 2.2.59, USP ⟨212⟩, JP 2.64). These techniques require themselves diverse and adapted sample preparation steps, such as enzymatic digestions (subunit or peptide generation; glycans release), as well as chemical treatments (glycans release, glycans labeling), and clean-up. Ruhaak et al. have comprehensively reviewed glycan labeling/derivatization and associated applications [47].

HILIC-UPLC of fluorescently labeled glycans is a common method in glycan analysis, is often used as a reference for alternative methods development [42], and is amenable to automation for high-throughput analysis [48,49]. Notable recent improvements in this area are the use of alternative glycan labels with improved fluorescence and/or MS sensitivity, and the development of stationary phases with wide-pore hybrid silica bonded with amide ligand for efficient separations at the intact and middle-up levels of analysis [50]. It constitutes an orthogonal approach to RP-LC and is particularly of interest in the frame of this work as it differentiates analytes based on their hydrophilic amino acid residues and glycans. Porous graphitic carbon for unlabeled *N*- and *O*-glycans [42,51–58], and reversed-phase for hydrophobically-tagged glycans have also been extensively reported [59].

In the case of mAbs, displaying simple to moderately complex glycosylation profiles, typical mass spectrometry workflows (direct ESI–MS or coupled to LC or CE, and MALDI-TOF) at the intact, middle-down and bottom-up levels can be performed to characterize glycosylation alongside other PTMs as well as molecular weight and amino-acid sequence verification [21,58,60–65]. The use of nanoLC-chip/MS(/MS) has been proposed to detect low abundance glycoforms, both at the intact and peptide levels [57,58,66]. MALDI-TOF allows high-throughput analysis of released glycans, however it often requires possibly tedious derivatization methods, typically permethylation, to stabilize the otherwise labile sialic acid [67,47]. A remarkable improvement in labeling techniques was brought by Wuhrer and colleagues under the form of a dimethylamidation that leads to a discrimination of the functionally different α2,3- and α2,6-sialic acids [68,69]. A stable isotope labeling for QTOF detection combined with a more classical HILIC-UPLC with fluorescence detection of 2-AA labeled *N*-glycans of mAbs was also reported [70].

CE-based techniques have also successfully been employed; glycans are typically labeled with APTS for CE with laser-induced fluorescence (LIF) detection [47,71]. Characterization of antibodies, with an emphasis on their main glycoforms, was performed *via* capillary zone electrophoresis coupled to tandem mass spectrometry by a sheathless interface (CESI–MS/MS) [23,72]. Alternatively, quantitative ESI–MS/MS detection can be supported by glycan tagging [73]. An automated high-throughput method on CE-LIF was developed by Reusch et al. to determine the glycosylation of Fc subunits of IgGs, directly from fermentation broth, using a DNA analyzer [74]. This method operates at the glycan level after protein A purification of the IgG, enzymatic cleavage and APTS derivatization; peak assignment by HILIC-UPLC–MS/MS and fluorescence detection is combined with peak fractionation and subsequent CE-LIF analysis of the MS-characterized fractions. It follows an earlier study at the peptide level involving protein A purification followed by tryptic digestion, purification on hydrophilic interaction SPE, and ESI–MS analysis [75]. The sample preparation may be eased by the use of magnetic beads for glycan capture prior to APTS labeling and CE-LIF analysis [76].

Another approach is the use of high-performance anion-exchange chromatography (HPAEC) coupled to pulsed amperometric detection (PAD), which is a sensitive technique at the glycan level [77,78]. Although it does not require glycan derivatization or clean-up, it suffers from a certain lack of chemical information – similarly to other retention-time based methods (e.g. HILIC-UPLC of 2-AB labeled *N*-glycans). ESI–MS hyphenation is possible to circumvent this issue, but requires a microfluidic ion-suppression unit or prototypal small-bore LC columns to achieve better performances [79,80].

Site-specific glycosylation analysis has been shown to be challenging and lengthy, especially for *O*-glycosylated proteins. It is typically achieved by working at the peptide level, which often involves off-line enrichment and/or fractionation of glycopeptides before LC/MS analysis [81–83]. Amino-acid level site and glycans structure determination can be supported by MS/MS techniques such as collision-induced dissociation (CID), higher C-trap dissociation (HCD), or electron-transfer dissociation (ETD) – the latter being particularly useful for labile glycans (e.g. *O*-glycans, see Section 3.6.2) [83] – or a combination of those [81,84]. On the level of whole *N*-glycoproteome, Park and colleagues have proposed a mapping system, which combines methods for tandem mass spectrometry with a database search and algorithmic suite [85], Zhao et al. have developed an online two-dimensional porous graphitic carbon(PGC)/RPLC platform [54], and enzymatic ^{18}O labeling is routinely used to determine *N*-glycosylation sites, and has also recently been used to label release glycans [86,87]. Microarrays (e.g. lectin arrays) are also a common solution for high-throughput glycosylation profiling for complex mixtures [88,89]; mass spectrometry can be used to interrogate the molecules spotted or captured on the array with [90] or without labeling [91,92]. Glycopeptides quantitation methods to determine the extent of glycosylation [93–95], or the concomitant protein quantitation from plasma and site-specific glycosylation analysis [96,97], have been reported. Similarly to other peptide-based quantitation workflows, multiple reaction monitoring (MRM) on triple quadrupole mass spectrometer constitutes a method of choice [94–97]. MALDI-TOF can be used for the relative quantitation of glycopeptides thanks to stable isotope labeling [98]. Finally, mAb subunit and peptide levels profiling by MS can be assisted by ion mobility [99].

Considering these various and complementary approaches, we wish to illustrate herein how the use of a combination of various cutting-edge yet simple LC or LC/MS(/MS) methods, and operating at different levels of analysis (released glycans, peptides,