Journal of Chromatography A, 1335 (2014) 81–103

Contents lists available at ScienceDirect

Journal of Chromatography A

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

Review

# Modern chromatographic and mass spectrometric techniques for protein biopharmaceutical characterization

Koen Sandra*, Isabel Vandenheede, Pat Sandra

*Research Institute for Chromatography (RIC), President Kennedypark 26, 8500 Kortrijk, Belgium*

ARTICLE INFO

*Article history:*
Received 11 October 2013
Received in revised form
27 November 2013
Accepted 29 November 2013
Available online 10 December 2013

*Keywords:*
Chromatography
Mass spectrometry
Biopharmaceutical
Monoclonal antibody
Therapeutic protein

ABSTRACT

Protein biopharmaceuticals such as monoclonal antibodies and therapeutic proteins are currently in widespread use for the treatment of various life-threatening diseases including cancer, autoimmune disorders, diabetes and anemia. The complexity of protein therapeutics is far exceeding that of small molecule drugs; hence, unraveling this complexity represents an analytical challenge. The current review provides the reader with state-of-the-art chromatographic and mass spectrometric tools available to dissect primary and higher order structures, post-translational modifications, purity and impurity profiles and pharmacokinetic properties of protein therapeutics.



**Contents**



## 1. The protein biopharmaceutical landscape

Since the commercial introduction in 1982 of recombinant human insulin for the treatment of diabetes, hundreds of protein biopharmaceuticals, classified as either therapeutic proteins or monoclonal antibodies (mAbs), have been approved by the regulatory agencies [1]. Today the global protein therapeutics market is worth over 100 billion dollar, thereby evolving toward a total pharmaceutical market share of 20%. It is expected that, within the current decade, more than 50% of the new drug approvals will be biologics [1–3]. Despite the fact that therapeutic proteins are presently dominating monoclonal antibodies in terms of overall sales, the latter are the fastest growing class of therapeutics [3]. Nowadays, around 30 monoclonal antibodies are marketed, nine displayed blockbuster status in 2010 and five of the ten top-selling biopharmaceuticals in 2009 were monoclonal antibodies

* Corresponding author. Tel.: +32 56 204031.
*E-mail address:* koen.sandra@richrom.com (K. Sandra).

0021-9673/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.chroma.2013.11.057

namely infliximab (Remicade), bevacizumab (Avastin), rituximab (Rituxan/Mabthera), adalimumab (Humira) and trastuzumab (Herceptin) [1,3,4]. The antibody fusion protein etanercept (Enbrel) together with first and second generation erythropoietin's (EPO) (Epogen/Aranesp) and next-generation insulin (Lantus) and granulocyte colony stimulating factor (G-CSF) (pegfilgrastim – Neulasta) completed the top ten in 2009 [1]. While these top selling biopharmaceuticals are successfully being applied in the treatment of diseases with a high incidence such as cancer, autoimmune disorders, diabetes and anemia, a diverse set of biomolecules have also been introduced dictated toward rare genetic diseases. The monoclonal antibody eculizumab (Soliris) introduced in 2007 for managing ultra-rare paroxysmal nocturnal hemoglobinuria (PNH) and the replacement enzymes acid-α-glucosidase (Myozyme and Lumizyme) and idursulfase (Elaprase) for the treatment of, respectively, Pompe disease and Hunter's syndrome are good examples of so-called orphan drugs [1].

Over the years biopharmaceuticals have substantially been engineered to optimize their efficacy and safety profiles [5]. Following its original introduction in 1989, second and third generation variants of EPO appeared on the market in 2001 and 2007, respectively. Compared to the original product, representing the recombinant version of human EPO, serum half-life has substantially been improved by introducing two additional N-glycosylation sites (second generation) and by conjugation to polyethylene glycol (third generation) [6]. Similarly, driven by the need to reduce immunogenicity and increase efficacy, therapeutic antibodies have evolved from purely murine to chimeric, humanized and human sequences [3,7]. The antibody market is further expected to be reshaped by various next-generation formats such as bispecific mAbs, antibody–drug conjugates (ADC), antibody mixtures, antibody fragments (Nanobodies, fragment antigen binding – Fab) and brain penetrant mAbs next to glyco-engineered formats [1,3,7]. Recent years witnessed the introduction of the first bispecific mAb (catumaxomab) and antibody–drug conjugate (brentuximab vedotin) [1,3,4].

With the patents of the first generation therapeutic proteins expired, the last decade experienced the approval of the first biosimilar versions [1,2,8]. Furthermore, the knowledge that the top-selling monoclonal antibodies will become open to the market in the coming years has resulted in an explosion of biosimilar versions in development [1,9,10]. The biosimilar market holds great potential but is simultaneously confronted with major hurdles. This stems from the fact that, opposed to generic versions of small molecules, exact copies of recombinant proteins cannot be produced due to differences in the cell clone and manufacturing processes used. Even innovator companies experience lot-to-lot variability [11] and process changes can have drastic effects as experienced by Genzyme in an attempt to upscale the production of acid-α-glucosidase (Myozyme) from a 160 L to a 2000 L fermentor. The glycosylation profile of the newly produced enzyme had changed substantially and authorities considered it as not being similar. The product is now marketed separately as Lumizyme [8]. In Europe, 14 biosimilars have been approved including two recombinant human growth hormone (hGH – somatropins), seven recombinant G-CSF (filgrastims) and five recombinant EPO products [1]. Interesting, glycoprofiles of the latter products appeared to be sufficiently similar to the reference medicines to allow their approval by European regulators. Thus far, no follow-on biologics have appeared on the US market, however, new regulations currently in effect will facilitate the approval over the coming years [1,2].

It is clear that biopharmaceuticals have reshaped the pharmaceutical landscape. It is a highly dynamic and rapidly evolving sector challenging to keep up with. Readers interested in the biopharmaceutical market trends are referred to the excellent yearly and four-yearly overview articles in Nature Biotechnology [1,12–18].

## 2. Characteristics and features of protein biopharmaceuticals

Opposed to small molecule drugs, protein biopharmaceuticals are large, heterogeneous and subject to a variety of enzymatic and chemical modifications during expression, purification and long-term storage. Their complexity can perfectly be illustrated by the humanized monoclonal antibody trastuzumab (trade name Herceptin) on the market since 1998 for the treatment of HER2 positive metastatic breast cancer and recombinantly produced in Chinese Hamster Ovarian (CHO) cells (Fig. 1). This tetrameric antibody is composed of two heavy and two light polypeptide chains connected through four interchain disulfide bridges. Twelve intrachain disulfide bridges, four within each heavy and two within each light chain, furthermore guarantee its structural integrity. The expected formula based on the cDNA sequence used to transfect the host cell is $C_{6460}H_{9972}N_{1724}O_{2014}S_{44}$ (1328 amino acids) corresponding to an average molecular weight of 145,422 Da (calculation based on estimated atomic weights from organic sources [19]). This is roughly 250 times larger than the chemically synthesized worldwide top-selling drug atorvastatin (Lipitor) with molecular formula $C_{33}H_{35}FN_2O_5$. Interestingly, the expected molecular weight of trastuzumab has never been experimentally observed due to co- and post-translational modifications (PTMs) taking place in the cell. During its passage through the endoplasmatic reticulum and Golgi apparatus, complex bi-antennary glycans are sculptured onto a conserved N-glycosylation site (with Asn-Xxx-Ser/Thr consensus sequence) located in the constant region of the heavy chain. A dozen of different sugar species have been measured, of which four are highly dominant (G0, G0F, G1F and G2F) (Fig. 1). Given the combination of two heavy chains and two N-glycosylation sites in a functional molecule, various glycoforms exist at the protein level typically annotated as e.g. G0F/G0F, G0F/G1F, G1F/G1F, etc. [9,20,21]. The non-glycosylated and singly glycosylated variants have as well been observed at low percentages (observations by the authors of this review). During cell culture production, host cell carboxypeptidases act on the antibody resulting in the removal of lysine residues from the C-terminus of the heavy chain [22]. In the case of trastuzumab manufacturing, this action is almost driven to completion with 99% cleavage of the two heavy chain lysines [9,23]. As a result, and depending on the production batch [21], the molecular formulas of the two main species are $C_{6560}H_{10132}N_{1728}O_{2090}S_{44}$ (G0F/G0F) and $C_{6566}H_{10142}N_{1728}O_{2095}S_{44}$ (G0F/G1F) with average molecular weight values of 148,057 and 148,219 Da, respectively. Other modifications reported on the antibody are cyclization of the heavy chain N-terminal glutamic acid with the formation of pyroglutamic acid, oxidation of methionine residues and deamidation/isomerization of asparagine and aspartate residues. Their origins are likely chemically driven during manufacturing and storage and their relative percentages range from 0.5% to 14% [23,24]. Taking all this together, a substantial number of species occur despite the fact that only one protein is actually cloned. Additionally, some minor aggregation (<1%) has been observed upon reconstituting trastuzumab [25]. All these characteristics together with their stabilities have to be revealed during development and subsequently need to be closely monitored in a routine environment following marketing to guarantee quality. In addition, biosimilar development demands for a comparison between originator and follow-on version characteristics. Not all of these modifications reported are considered as critical quality attributes though. Trastuzumabs N-terminal pyroglutamic acid formation and C-terminal lysine truncation are usually not regulatory concerns. Glycosylation, aggregation and aspartate isomerization, on the other hand, are