Accepted Manuscript

Title: On the glycosylation aspects of biosimilarity

Authors: László Hajba, Ákos Szekrényes, Beáta Borza, András Guttman

PII: S1359-6446(17)30511-1
DOI: https://doi.org/10.1016/j.drudis.2018.01.009
Reference: DRUDIS 2142

To appear in:

Please cite this article as: Hajba, László, Szekrényes, Ákos, Borza, Beáta, Guttman, András, On the glycosylation aspects of biosimilarity.Drug Discovery Today https://doi.org/10.1016/j.drudis.2018.01.009

On the glycosylation aspects of biosimilarity

László Hajba[1], Ákos Szekrényes[2], Beáta Borza[2], and András Guttman[2,3]

[1]MTA-PE Translational Glycomics Research Group, Research Institute for Biomolecular and Chemical Engineering, University of Pannonia, Veszprem, Hungary
[2]Horváth Csaba Memorial Institute for Bioanalytical Research, University of Debrecen, Hungary
[3]Sciex Separations, Brea, CA 92821, USA

Corresponding author: Guttman, A. (guttman@mik.uni-pannon.hu)

Teaser: This review summarizes the critical quality attributes of biosimilarity, also introducing the glycosimilarity index as an additional parameter to prove similarity between innovator and biosimilar products.

Highlights

- The critical quality attributes of biosimilarity are reviewed
- Regulatory and statistical considerations are summarized.
- Importance of glycosylation analysis during biosimilar production is detailed.
- Analytical methods to track N-glycosylation patterns are reviewed.
- Glycosimilarity index is introduced.

The recent expiration of several protein therapeutics opened the door for biosimilar development. Biosimilars are biologic medical products that are similar but not identical copies of already-authorized protein therapeutics. Critical quality attributes (CQA), such as post-translational modifications of recombinant biotherapeutics, are important for the clinical efficacy and safety of both the innovative biologics and their biosimilar counterparts. Here, we summarize biosimilarity CQAs, considering the regulatory guidelines and the statistical aspects (e.g., biosimilarity index) and then discuss glycosylation as one of the important attributes of biosimilarity. Finally, we introduced the 'Glycosimilarity Index', which is based on the averaged biosimilarity criterion.

Keywords: average bioequivalence; biosimilarity index; glycosylation pattern; glycosimilarity index.

Introduction: biosimilar market overview

The term 'biologics' covers a range of recombinant protein pharmaceuticals, such as hormones, growth factors, blood products, monoclonal antibody-based therapeutics, fusion proteins, antibody–drug conjugates, and recombinant vaccines. They have become an essential part of the modern pharmaceutical palette and represent a rapidly growing section of the pharmaceutical industry [1]. As of May 2017, 74 antibody-based molecules had been approved by regulatory authorities all over the word. Furthermore, there are 70 antibody-based molecules currently in Phase III and 575 in Phase I/II clinical trials [2]. Spending on biologics development between 2013 and 2014 increased by 32.4%, whereas spending on small-molecule drugs only increased by 6.8% over the same timeframe. Based on this trend, an annual increase of 16% is expected from 2015 to 2018. The Global & USA Biosimilar Market Analysis from 2015 estimated that the USA will be the largest market for biosimilars, contributing 4–10% of the global biologics market by 2020, depending on the number of biosimilars approved by the US Food and Drug Administration (FDA) [3].

Eli Lilly's insulin (developed by Genentech, Inc.) was one of the first recombinant protein drugs on the market, introduced during the 1980s [4], followed by dozens of biotherapeutics for the treatment of different diseases [5,6]. In recent years, the patents and regulatory data protection periods for many of those original recombinant protein-based biotechnological medicines started to expire, opening the door for other companies to develop their own versions, dubbed 'biosimilars' [4]. As the generally accepted nomenclature suggests, biosimilars are similar but not

exactly the same as the original innovator products. Biosimilars, in most instances, are produced using different cell lines and manufacturing and purification processes; therefore, the final product is not identical to the originator [7]. During biosimilar development, the following important issues should be addressed: (i) verification of similarity; (ii) interchangeability of biosimilars and original innovator product; (iii) unique naming of the biosimilar for differentiation; (iv) regulatory framework; (v) regulatory guidelines to assist manufacturers in product development; (vi) intellectual property (IP) rights; and, last but not least, (vii) safety [1].

Here, we summarize the attributes, including statistical aspects (e.g., biosimilarity index) of biosimilarity in light of regulatory guidelines. We then discuss glycosylation, one of the important attributes of biosimilarity, ,focusing particularly on the control of glycosylation processing and the analysis of the resulting glycosylation pattern during all stages of biosimilar production.

Attributes of biosimilarity

Main characteristics of biosimilarity

The exact definition proposed for biosimilars in accordance with European Medicines Agency (EMA), FDA, and WHO guidelines is as follows: 'Biosimilars are copy version of an already authorized biological medicinal product with demonstrated similarity in physicochemical characteristics, efficacy and safety, based on a comprehensive comparability exercise' [8]. Thus, a biosimilar should be highly similar to the reference medicinal product based on its physicochemical and biological properties. Comparable clinical safety and efficacy of the biosimilar and the reference product should also be confirmed [9]. Immunogenicity is one of the most-cited safety concerns of biosimilars; therefore, the corresponding clinical data are crucial before approval [10]. Any observed differences have to be properly justified because they could have potential impact on the clinical performance [9]. Whereas biosimilars are copy versions of the innovator products, biobetters are biologics with structural changes, better functional targeting, or an improved formulation that can result in improved clinical safety and efficacy compared with the innovator product [11].

During biosimilar development, a step-wise approach is recommended, starting with comprehensive physicochemical and biological characterization with proper and preferably orthogonal analytical methods [12]. The clinical data needed to support biosimilarity are influenced by the evidence obtained in the preclinical *in vivo* studies, namely during physicochemical and biological characterization [13]. Differences in molecular structure (e.g., post-translational modifications, higher order structures, etc.) are often assessed through rigorous analytical testing. Molecular structure parameters represent the CQAs relevant to the clinical outcomes between a proposed biosimilar and the reference product, referred to as an 'analytical similarity assessment' [14]. A step-wise approach is adopted to confirm biosimilarity and interchangeability, including: (i) structural and functional characterization; (ii) animal studies, such as toxicity, pharmacokinetics (PK), pharmacodynamics (PD), and immunogenicity; and (iii) clinical studies to assess and validate safety, efficacy, and immunogenicity issues in humans [15].

Regulatory requirements

According to FDA Guidelines, proper analytical methodologies with good sensitivity and specificity should be applied during the structural characterization of therapeutic proteins. The biosimilar product and the reference product should be compared in terms of the following criteria: (i) primary structures, such as amino acid sequence; (ii) secondary, tertiary, and quaternary structures; (iii) post-translational modifications, such as glycosylation, phosphorylation, and so on; (iv) degradation hotspots, such as protein deamidation and oxidation; and (v) chemical modifications, such as PEGylation. Biosimilar developers should conduct comprehensive structural head-to-head characterization on multiple lots of their biosimilar candidates and the reference products to understand the lot-to-lot variability in context with manufacturing process parameters. Lots selected for the analyses should be representative to demonstrate biosimilarity between the candidate and reference products. Selection of the representative lots and the number of lots to be analyzed should also be properly justified [13]. The regulatory agencies require accurate characterization of biosimilars based on the above-listed criteria. Physicochemical characterization of the innovator product and its biosimilars have been reported in several case studies [16–18]. Improvements in product quality are desirable and can be achieved with changes in the manufacturing process; therefore, the postchange product can show improvements in clinical efficacy and/or safety compared with the prechange product [19]. In the International Conference on Harmonisation (ICH) Q5E Guidelines, comparability is defined so that the CQAs of the pre- and postchange products are not identical but highly similar and have no harmful effect upon the safety or efficacy of the biosimilar product [19,20].

The development process of a biosimilar product starts with an extensive literature search to understand the mechanism of action (MoA) of the reference medicinal product for all of the existing indications and to gain knowledge about any reported structure- or function-related adverse events or immunogenicity issues. Using the collected information, CQAs of the drug can be determined using various risk assessment approaches and their criticality can be assessed. Once the CQAs are determined, a comprehensive analytical biosimilarity approach can be designed (Figure 1). In the next step, CQAs such as aggregation level, charge variant profile, post-translational modifications, and biological activity, are analyzed in several lots of the reference product to assess the quality target product profile (QTPP). This step is followed by the selection, design, and development of the expression system (expression construct and host cell line) and other steps of the manufacturing process to provide a firm foundation to generate the biosimilar product. The CQAs are continuously monitored during manufacturing process

development to provide feedback for process parameter changes [5]. Reverse engineering, extensive iterations in process development, and analytical characterization are all included in biosimilar development to produce a product that is highly similar to the reference material. Given the usually limited information on the originator's manufacturing processes, this process can involve the use of a new cell line, newly developed cell culture conditions, purification steps, and, in some instances, a more-contemporary formulation to that used by the originator at the time of market entry [21].

The predefined biosimilarity criteria should address issues such as the degree of similarity and sensitivity to small deviations. Appropriate statistical methods, such as equivalency test, quality range approach, and the introduction of the so-called 'biosimilarity index', as well as clinical study designs, are all required by the FDA to assess biosimilarity with a high level of confidence [15].

Statistical considerations and the biosimilarity index

The FDA suggests using a three-tier approach for analytical similarity assessment. It provides statistical methodology guidance to deliver the totality-of-the-evidence for demonstrating biosimilarity. The identified CQAs with relevant clinical outcomes are classified into the three tiers based on their risk (i.e., high, moderate, and low risk in terms of clinical safety and efficacy). Different statistical approaches for the different tiers are recommended, namely equivalency test for CQAs from Tier 1 (high risk), quality range approach for CQAs from Tier 2 (moderate risk), and descriptive raw data and graphical presentation for CQAs from Tier 3 (low risk). Chow *et al.* thoroughly described the different statistical approaches in relation to the above [22].

The biosimilarity index, or average bioequivalence (ABE) index is a disaggregated, probability-based (PB), scaled, and weighted criterion for the assessment of biosimilarity that can be expressed by Equation 1 [23], where T (test product, biosimilar) and R (reference product, innovative artifact) are the parameters of interest (e.g., PK data). Thus, the biosimilarity index (p_{PB}) between the test and the reference product can be defined as:

$$p_{PB} = \left(1 - \delta < \frac{T}{R} < 1 + \delta\right) \quad [1],$$

where $0 < \delta < 1$ is the biosimilarity limit for the PB method.

To apply the local biosimilarity index (calculated based on a specific product parameter) between the test and the reference product, the following steps have to be followed, as described in [24–26]: (i) assess the average biosimilarity between the test product and the reference product based on a set of biosimilarity criteria. The criteria could be based on mean, ratio, or variability of the test product parameter. The ratio of means of a given study endpoint should fall within the biosimilarity limit of 80–125% for *in vivo* and 90–111% for *in vitro* studies [27,28]; (ii) calculate the local PB biosimilarity index using Equation 1; and (iii) declare the test product highly similar if the 95% confidence lower bound of the biosimilarity index (p_{PB}) is larger than a prespecified number of p_0, where p_0 can be obtained based on an estimate of reproducibility probability for a study comparing a reference product to itself.

The proposed biosimilarity index has the following advantages: (i) it is robust with respect to selected biosimilarity criteria for the assessment of bioequivalence; (ii) the index takes variability into consideration and sensitive to the variation of the test product; and (iii) it allows the assessment of the degree of similarity, and, thus, provides an answer to the question 'how similar is similar' [29,30].

The biosimilarity (or ABE) index can be also described by a moment-based method with the relative distance (rd) based on Equation 2 [31]:

$$rd = \frac{d(T,R)}{d(R,R)} \quad [2],$$

where $d(T,R)$ is the difference between the test (T) and reference (R) products, and $d(R,R)$ is the difference between several batches of the reference product itself. Typically, a one-sided 95% confidence interval is applied for rd-based determination of biosimilarity and/or bioequivalence. If the one-sided 95% upper confidence limit is less than the biosimilarity limit, then the test product is considered to be biologically similar to the reference product. In this case, based on the traditional average bioequivalence criterion, the biosimilarity index can be defined as (Equation 3):

$$rd = M_T/M_R \quad [3],$$

and in the case of scaled average bioequivalence criterion (Equation 4):

$$rd = (M_T - M_R)/\sigma_R \quad [4],$$

where M_T is the mean of the test product parameters (e.g., PK data, PD data, and analytical data, such as integrated peak area), M_R and σ_R are the mean and standard deviation of the reference product parameters, respectively [24].

Glycosylation of biosimilars

Regulation of glycosylation during biosimilar production

Most recombinant protein biologics and their biosimilars (e.g., monoclonal antibodies, fusion proteins, antibody-drug conjugates, etc.) are typically glycosylated. Therefore, glycosylation is a CQA and probably one of the most

challenging aspects to demonstrate biosimilarity for therapeutic glycoproteins, given the noncoded properties of the glycosylation machinery. Glycosylation of biosimilars could have a significant impact on the biological activity, stability against proteolysis, PK profile, serum half-life, effector functions, and immunogenicity of the drug products [21,32]. Glycosylation particularly influences the two major effector functions of therapeutic monoclonal antibodies, namely antibody-dependent cell-mediated cytotoxicity (ADCC) and complement-dependent cytotoxicity (CDC). For example, terminal galactose (Gal), *N*-acetylglucosamine (GlcNAc), and mannose (Man) residues affect CDC activity, while terminal sialic acid, mannose, core fucose (Fuc), and bisecting GlcNAc affect the ADCC function [33–35]. Therapeutic monoclonal antibodies, such as anti-CD20 rituximab (Rituxan®), anti-Her2 trastuzumab (Herceptin®), antitumor necrosis factor-α (anti-TNF-α) infliximab (Remicade®), and anti-RhD (HyperRHO®) are capable of ADCC effector function. The possible antitumor mechanism of rituximab and alemtuzumab (Campath-1H) is CDC [33]. The addition of sialic acid to the terminus of the conserved Fc region glycan structures transfers IgG antibodies into anti-inflammatory mediators, rendering them capable of suppressing autoantibody-driven inflammation [36]. In addition, the relative levels of high-mannose glycans can affect some PK properties of therapeutic antibodies. In particular, therapeutic monoclonal antibodies containing high mannose (HM) glycans at the Fc region are cleared more rapidly in humans than are other glycoforms [37]. Glycan residues, such as α1,3-galactose, β1,2-xylose, α1,3-fucose, and *N*-glycolylneuraminic acid (Neu5Gc), have a highly negative impact on the safety and/or immunogenicity of monoclonal antibodies [38]. Therefore, these potentially immunogenic species should be rigorously monitored [39,40]. In a recent study by Jaquez and coworkers, the correlation between the percentage of afucosylated and HM glycans as well as FcγRIIIa binding activity and ADCC activity were calculated in the case of infliximab and its biosimilars [41].

The proper information about the glycosylation profile of a biosimilar product is well recognized as a CQA that needs to be tightly monitored and controlled [42]. For example, if the MoA is based on ADCC or CDC functions, core fucosylation or galactosylation are chosen, respectively, as CQAs [43]. Heterogeneity of the therapeutic antibody glycoforms increases with the diverse set of host cells used by the industry during the manufacturing process, which can be critical regarding the resulting effector functions of the product. The relationship between the conserved Fc glycosylation and the effector function associated with the clinical activity of therapeutic antibodies makes careful analysis and characterization of glycosylation-related CQAs crucial. The information obtained can help during the manufacturing process to fine tune the glycosylation profile of the therapeutic antibodies in the pipeline [38].

Currently approved therapeutic antibodies are produced by mammalian cell cultures, utilizing Chinese hamster ovary (CHO), human embryonic kidney (HEK), mouse myeloma (NS0), or mouse hybridoma (SP2/0) cell lines [44]. Regula and coworkers performed comprehensive glycosylation analysis of the Human leukocyte receptor IIIa (hFcγRIIIa), derived from HEK and CHO cells [45]. They concluded that the expression cell lines chosen highly affected the glycopatterns of hFcγRIIIa, whereas the protein sequence and glycosylation sites remained unchanged for both cell types. Their results are summarized in Table 1.

Although non-mammalian cell expression systems have several advantages (e.g., higher production yield, simple and inexpensive culturing, etc.) over mammalian cell lines, their utilization in biologics production is hindered because of their limited glycosylation machinery and the potential introduction of immunogenic residues. Glycoengineering is a promising new tool to 'humanize' non-mammalian cell expression systems by methods such as gene knockout, RNA interference, enzyme overexpression [46], and genome editing with RNA-guided CRISPR-Cas9 nuclease systems [47]. As an example, GlycoFi glycoengineering technology is based on a humanized yeast cell line of *Pichia pastoris* that allows the production of human glycoproteins with not only HM, but also complex *N*-glycosylation modifications [48]. Furthermore, the glycosylation pattern of a therapeutic glycoprotein produced by GlycoFi is comparable to mammalian cell expression systems, such as CHO and NS0 cell lines. Thus, the *P. pastoris* expression system represents a promising approach for the production of large-scale batches of therapeutic humanized glycoproteins at lower cost. GlycoDelete is another promising glycoengineering approach, which produces proteins with small, sialylated *N*-glycans and reduces the complexity of the modified mammalian cell expression system compared with native mammalian cell lines [49,50]. Therapeutic proteins expressed in GlycoDelete-modified cells have reduced *N*-glycosylation heterogeneity and clearance compared with those expressed in wild-type cells and, thus, can lead to increased reproducibility in their production. Predictive glycoengineering can also help in the development of biosimilars. For example, the computational method using the Markov chain model of glycosylation can predict the quantitative amounts (e.g., nutrients, inhibitors, and enzymes) by which glycosylation reaction rates must be perturbed to obtain a specific glycosylation profile [51]. Brühlmann *et al.* developed a parallel design-of-experiment (DoE) approach. Using this method, 17 glycosylation modulator media supplements were identified and arranged in five groups based on their effects on glycosylation. After the fed-batch culture in 96-deepwell plates with different medium composition, the best-performing glycosylation modulators were selected using multivariate analysis, principal component analysis, and decision trees [52].

The manufacturing processes of biologics involve many steps, starting with cloning of the desired gene into a complementary DNA vector, continued by transferring it into a suitable host cell (e.g., *Escherichia coli*, yeast, CHO, or NS0), production in a bioreactor, and purification and formulation of the biological drug product [53], (Figure 2). The chosen cell line, cell growth stage, and process parameters, such as nutrient availability, dissolved oxygen, supplements, temperature, pH, operation schedule, and culture modes, can all have significant effects on the glycosylation pattern of the biological drug products [54]. Nagy and coworkers extensively reviewed which process

parameters affect glycosylation of the therapeutic antibody (produced by mammalian cell lines) and to what extent [55]. The presence of sialidases and other glycosidases, released from dead cells, can cause degradation of previously assembled glycosylation structures [56]. Moreover, glycan composition can be selectively modified during the downstream part of the process in the different purification steps. Therefore, lot-to-lot variability in the glycosylation profile is also one of the best indicators of process robustness.

Analytical methods to track N-*glycosylation patterns*

N-glycosylation analysis to characterize therapeutic glycoproteins in the field of process analytical technology is vital [57]. There are several excellent reviews focusing on the glycosylation analysis of therapeutic glycoproteins in detail [58–63]. *N*-glycosylation analysis of therapeutic proteins can be performed at three different levels (intact protein, glycopeptide, and released glycans), as depicted in Figure 3 [64]. Intact glycoproteins can be analyzed by mass spectrometry (MS) via direct infusion or coupling with separation techniques, such as liquid chromatography (LC)-MS or capillary electrophoresis (CE)-MS [65,66]. Such top-level analysis also offers information about glycan pairing on the two heavy chains of an IgG molecule. The middle-up approach (IdeS digestion and reduction of the intact glycoproteins to yield Fc/2, Fd', and LC fragments) allows more-accurate identification of the glycosylation profiles, because the analysis is conducted on subunits of approximately 25-kDa protein fragments. Guillarme and coworkers applied HILIC-MS with a wide-pore stationary phase to qualitatively profile the glycosylation patterns at the protein level and compared the originator monoclonal antibodies to the their biosimilar counterparts [67,68]. In the middle level, after proteolytic digestion of the glycoproteins, LC–MS/MS of CE-MS/MS analysis is usually performed. Middle-level analysis can provide information about glycosylation macro- (site occupancy) and microheterogeneity (variation of glycan structures at a given site). In bottom-level analyses, the glycans are enzymatically or chemically cleaved off from the protein backbone. The released glycans are usually labeled with a fluorescent dye and separated by LC or CE, with the option of MS detection. Bottom-level analysis provides information about the glycosylation profile and, with the help of exoglycosidase-based sequencing, fine details can be obtained about the linkage and positional isomers [69]. Fully automated carbohydrate sequencing by capillary electrophoresis was also recently reported [70].

Tharmalingam *et al.* [71] developed a real-time *N*-glycosylation monitoring (RT-GM) framework to characterize the entire time-course of a fed-batch culture. The sample preparation platform utilized microsequential injection (mSI), which was coupled with an ultraperformance liquid chromatography (UPLC) system based real-time glycan profiling for therapeutic antibody manufacturing. The authors studied manganese (Mn)-induced glycosylation changes in a fed-batch culture with this automated analytical system. Mn-supplemented cultures exhibited higher galactosylation, and consistent fucosylation and mannosylation levels compared with their control cell culture counterparts. The developed real-time *N*-glycosylation monitoring framework can accelerate process development because the information from monitoring the manufacturing steps can be rapidly translated into process control and quick feedback decisions. Although comprehensive knowledge of the glycosylation of the final product is an important CQA, the time-related *N*-glycosylation changes are also important, such as loss of sialic acids during the storage of the therapeutic protein.

Mazzeo and coworkers performed LC-based glycosylation profiling of trastuzumab and its biosimilar counterpart using fluorescent detection. Their results confirmed that the ratio of individual glycans (quantified by integrated peak area) differed between the biosimilar and the reference product [72]. In another case study, Sanchez-De Meloa *et al.* analyzed the glycan profile of trastuzumab and its biosimilar candidates by normal-phase HPLC and MALDI-TOF-MS [73]. They observed that the relative abundance (peak area %) of the individual glycans on the monoclonal antibody candidates were not similar to the reference product. Liu and coworkers used the same normal-phase HPLC and MALDI-TOF-MS approaches for the glycosylation analysis of cetuximab and its biosimilar. The glycan profile of the biosimilar was in good agreement with the innovator product, although previously not-reported aberrant *N*-linked glycans with NeuAcLac motifs were identified in the biosimilar [74]. Montacir *et al.* performed glycosylation analysis of rituximab and its biosimilar at three different levels [protein (intact mass and subunits Fc and HC), glycopeptide, and glycan] with different MS-based analytical methods (e.g., MALDI-TOF-MS). Although the glycosylation sites of both were identical, the carbohydrate profiles of the originator and the biosimilar were qualitatively similar, but quantitatively heterogeneous [75].

Introducing the Glycosimilarity Index

It is still challenging to determine and describe the level of similarity between two *N*-glycosylation profiles; however, assessing alikeness at the glycosylation level is one of the most crucial parts of an analytical biosimilarity study when the reference product is glycosylated. Here, we introduce a glycoanalytical profile-based similarity-scoring approach, referred to as 'Glycosimilarity Index', that can be used to calculate the level of similarity between the *N*-glycosylation profiles of any given reference and test items.

To assess the *N*-glycosylation similarity between a biosimilar therapeutic protein and its reference product, first all *N*-glycosylation related attributes (e.g., total fucosylation, galactosylation, sialylation, mannosylation, etc.) and their criticality have to be determined. The FDA suggests a three-tier-based approach to rank the different CQAs is applicable here too. In this case, as a first approximation, we consider tier ranking to represent a criticality risk rating of the *N*-glycosylation quality attributes with regard to their potential impact on activity, PK/PD, safety, and immunogenicity. Glycosylation CQAs most relevant to clinical outcomes will be classified to Tier 1, whereas CQAs that are less (mild-to-moderate) or least relevant to clinical outcomes or immunogenicity will be classified to Tier 2

and Tier 3, respectively, as described above for the overall biosimilarity [22]. To assess glycoanalytical similarity for the respective CQAs between the reference and biosimilar product, the FDA-proposed equivalence tests for CQAs in Tier 1, a quality-range approach for CQAs in Tier 2, and raw data comparison or graphical presentation for CQAs in Tier 3 are also applicable. To determine the criticality and to assess the appropriate tier ranking of the different *N*-glycosylation-related CQAs, the following questions have to be addressed: (i) does the *N*-glycosylation have any direct or indirect effect on the efficacy of the product? If yes how? (e.g., via effector functions, directly/indirectly affecting target binding, influencing protein structure that alters function, etc.); (ii) where are the CQA *N*-glycans located? (Fc, Fab, surface exposed, etc.); (iii) are there any immunogenic sugar residues present? (NGNA, α-Gal, etc.); and (iv) what analytical assay is used for *N*-glycan profiling? (It is important to assess the acceptance range that also depends on the performance of the analytical assay).

First tier ranking along with the assessment criteria is determined for each *N*-glycosylation-related attribute, such as total fucosylation, total mannosylation, total galactosylation, and so on. This allows the calculation of a product-specific glycosimilarity index. First, the reference product(s) is(are) analyzed in triplicates. The profiles are aligned and normalized using the highest *N*-glycan peak before the min, max, mean, and standard deviation are determined for those points in the separation trace (CE or LC). The migration and/or retention time of each point of the profile is converted to glucose units (GU) using a glucose homo-oligomer ladder injected prior and after the samples. A virtual ladder, by using the recently introduced triple internal standard approach, is also appropriate for GU value calculation [76]. The electropherogram and/or chromatogram is than fractionally divided into different GU value sections, based on the major oligosaccharide classes in the profile (neutral, sialylated, HM, and hybrid). Tolerance limits are determined for each GU value interval, based on the tier ranking of the *N*-glycosylation-related attributes in hand. For each attribute, tolerance limits are defined during tier ranking based on the criticality of the attribute, the variation of the reference, and the performance of the glycoanalytical assay. If fucosylation and/or afucosylation is assessed as a Tier 2 attribute with a similarity range of mean ±1.5 ·SD, then each GU value interval that represents fucosylated and/or afucosylated structures will be compared using a mean ±1.5 ·SD tolerance window. Points between the upper and lower limits are considered as 100% similar, regardless of their distances from the mean. The contribution to similarity score of points outside of the limits is calculated using a sum of squared distances to average data [77]. The different attribute similarity scores are calculated using Equation 5:

$$A_i = Y_i + X_i \times \sum_{j=1}^{n} \frac{((x_{uj} - x_{lj}) - \overline{x_j}))^2}{(x_{sj} - \overline{x_j})^2} \qquad [5],$$

where A_i is the % profile similarity of the i^{th} attribute (i=1,...,n), Y_i is the % of points inside the tolerance limits in the i^{th} attribute-related GU section, X_i is the % of points outside of the tolerance limits in the i^{th} attribute-related GU sections, x_{uj} is the upper limit of the j^{th} point, x_{lj} is the lower limit of the j^{th} point, x_j is the mean of j^{th} reference material points and x_{sj} is the j^{th} sample point that falls outside the tolerance limits. The final Glycosimilarity Index is calculated form the combination of the different attribute similarity scores using Equation 6:

$$\mathrm{GI} = \frac{\sum_{i=1}^{n}(A_i + B_i)/2}{n} \qquad [6],$$

where GI is the Glycosimilarity Index, n is the number of attributes and B_i is the % extent similarity of the i^{th} attribute. B_i is considered to be 100% if the summed relative amount of the certain glycoforms related to the i^{th} attribute is within the predefined tolerance limits. If the summed relative amount of these peaks falls outside the range, B_i is calculated using Equation 7:

$$B_i = \frac{((b_{ui} - b_{li}) - \overline{b_i}))^2}{(b_{si} - \overline{b_i})^2} \times 100 \qquad [7].$$

The *N*-glycosylation of the test product is considered to be highly similar to the reference product if the glycosimilarity index falls within the 80–100% range. A graphical representation of the glycosimilarity index is depicted in Figure 4.

Concluding remarks

Recent expirations of the patent protections for many of the original recombinant protein-based biotherapeutics provided the opportunity for large-scale biosimilar development. Regulatory authorities require comprehensive analysis of protein structure, post-translational modifications, and biological activity during the development process of a biosimilar product. These molecular structure parameters represent the CQAs relevant to clinical efficacy and safety. Given that most recombinant protein biologics and, consequently, their biosimilar counterparts are glycosylated, information about their carbohydrate moiety is an important CQA to demonstrate biosimilarity at the glycosylation level (i.e., glycosimilarity). Glycosylation of biosimilars could have a significant impact on the biological activity, effector functions, and immunogenicity of the drug products. The Glycosimilarity Index, established based on the traditional average bioequivalence criterion biosimilarity index, could serve as an additional parameter to support biosimilarity between the innovator and the biosimilar product.

Acknowledgments

The authors gratefully acknowledge the support of the MTA-PE Translational Glycomics Momentum (Lendulet) grant # 97101 of the Hungarian Academy of Sciences and the National Research, Development and Innovation Office (NKFIH) (K 116263) grants of the Hungarian Government. This work was also supported by the BIONANO_GINOP-2.3.2-15-2016-00017 project.

References

1 Sekhon, B. and Saluja, V. (2011) Biosimilars: an overview. *Biosimilars* 1, 1–11
2 Strohl, W.R. (2017) Current progress in innovative engineered antibodies. *Protein Cell*. Published online August 18, 2017. http://dx.doi.org/10.1007/s13238-13017-10457–13238
3 Singh, S.C. and Bagnato, K.M. (2015) The economic implications of biosimilars. *Am. J. Manag. Care* 21, S331–S340
4 Kresse, G.B. (2009) Biosimilars - science, status, and strategic perspective. *Eur. J. Pharm. Biopharm.* 72, 479–486
5 Nikolov, N.P. and Shapiro, M.A. (2017) An FDA perspective on the assessment of proposed biosimilar therapeutic proteins in rheumatology. *Nat. Rev. Rheumatol.* 13, 123–128
6 Calo-Fernández, B. and Martínez-Hurtado, J.L. (2012) Biosimilars: company strategies to capture value from the biologics market. *Pharmaceuticals* 5, 1393–1408
7 Nowicki, M. (2007) Basic facts about biosimilars. *Kidney Blood Pressure Res.* 30, 267–272
8 Weise, M. *et al.* (2011) Biosimilars - why terminology matters. *Nat. Biotechnol.* 29, 690–693
9 EMA (2014) *Guideline on Similar Biological Medicinal Products*. European Medicines Agency
10 Stevenson, J.G. (2015) Clinical data and regulatory issues of biosimilar products. *Am. J. Manag. Care* 21, S320–S330
11 Anour, R. (2014) Biosimilars versus 'biobetters' - a regulator's perspective. *Generics Biosimilars Initiative J.* 4, 166–167
12 Berkowitz, S.A. *et al.* (2012) Analytical tools for characterizing biopharmaceuticals and the implications for biosimilars. *Nat. Rev. Drug Discov.* 11, 527–540
13 FDA (2015) *Guidance for Industry: Quality Considerations in Demonstrating Biosimilarity of a Therapeutic Protein Product to a Reference Product*, US FDA
14 Chow, S. (2014) On assessment of analytical similarity in biosimilar studies. *Drug Design.* 3, XXX–YYY
15 Karalis, V.D. (2016) From bioequivalence to biosimilarity: the rise of a novel regulatory framework. *Drug Res.* 66, 1–6
16 Fang, J. *et al.* (2016) Advanced assessment of the physicochemical characteristics of Remicade (R) and Inflectra (R) by sensitive LC/MS techniques. *Mabs* 8, 1021–1034
17 Jung, S.K. *et al.* (2014) Physicochemical characterization of Remsima (R). *Mabs* 6, 1163–1177
18 Visser, J. *et al.* (2013) Physicochemical and functional comparability between the proposed biosimilar rituximab GP2013 and originator rituximab. *Biodrugs* 27, 495–507
19 ICH (2004) *Comparability of Biotechnological/Biological Products Subject to Changes in their Manufacturing Process Q5e*, ICH
20 Schiestl, M. *et al.* (2011) Acceptable changes in quality attributes of glycosylated biopharmaceuticals. *Nat. Biotechnol.* 29, 310–312
21 Azevedo, V. *et al.* (2016) Differentiating biosimilarity and comparability in biotherapeutics. *Clin. Rheumatol.* 35, 2877–2886
22 Chow, S-C. (2015) Challenging issues in assessing analytical similarity in biosimilar studies. *Biosimilars* 5, 33–39
23 Chow, S.C. *et al.* (2010) A comparison of moment-based and probability-based criteria for assessment of follow-on biologics. *J. Biopharm. Stat.* 20, 31–45
24 Lu, Y. *et al.* (2013) Statistical designs for assessing interchangeability of biosimilar products. *Drug Design.* 3, 1000109
25 Chow, S.C. *et al.* (2013) Statistical methods for assessing interchangeability of biosimilars. *Stat. Med.* 32, 442–448
26 Hsieh, T.C. *et al.* (2013) The evaluation of biosimilarity index based on reproducibility probability for assessing follow-on biologics. *Stat. Med.* 32, 406–414
27 FDA (2001) *Guidance for Industry: Statistical Approaches to Establishing Bioequivalence*. US FDA
28 Lu, Y. *et al.* (2014) *In vivo* and *in vitro* bioequivalence testing. *J. Bioequiv. Bioavail.* 6, 067–074
29 Chow, S-C. *et al.* (2011) Scientific factors for assessing biosimilarity and drug interchangeability of follow-on biologics. *Biosimilars* 1, 13–26
30 Zhang, A. *et al.* (2013) Statistical considerations in biosimilar assessment using biosimilarity index. *J. Bioequiv. Bioavail.* 5, 209–214
31 Chow, S.C. *et al.* (2013) Scientific considerations for assessing biosimilar products. *Stat. Med.* 32, 370–381
32 Azadi, P. (2015. 16(7)) *Challenges in Comparability Studies of Glycoprotein Biosimilars*, Pharmaceutical Outsourcing
33 Natsume, A. *et al.* (2008) Improving effector functions of antibodies for cancer treatment: enhancing ADCC and CDC. *Drug Design Dev. Ther.* 3, 7–16
34 Raju, T.S. (2008) Terminal sugars of Fc glycans influence antibody effector functions of IgGs. *Curr. Opin. Immunol.* 20, 471–478
35 Iida, S. *et al.* (2009) Two mechanisms of the enhanced antibody-dependent cellular cytotoxicity (ADCC) efficacy of non-fucosylated therapeutic antibodies in human blood. *BMC Cancer* 9, Artn 58
36 Shade, C.K-T. and Anthony, M.R. (2013) Antibody glycosylation and inflammation. *Antibodies* 2, 392–414
37 Goetze, A.M. *et al.* (2011) High-mannose glycans on the Fc region of therapeutic IgG antibodies increase serum clearance in humans. *Glycobiology* 21, 949–959
38 Reusch, D. and Tejada, M.L. (2015) Fc glycans of therapeutic antibodies as critical quality attributes. *Glycobiology* 25, 1325–1334
39 Szabo, Z. *et al.* (2012) Sialic acid speciation using capillary electrophoresis: optimization of analyte derivatization and separation. *Anal. Chem.* 84, 7638–7642
40 Szabo, Z. *et al.* (2012) Ultrasensitive capillary electrophoretic analysis of potentially immunogenic carbohydrate residues in biologics: galactose-alpha-1,3-galactose containing oligosaccharides. *Mol. Pharm.* 9, 1612–1619
41 Lee, C. *et al.* (2017) Glycosylation profile and biological activity of Remicade (R) compared with Flixabi (R) and Remsima (R). *Mabs* 9, 968–977
42 Schiestl, M. *et al.* (2014) The role of the quality assessment in the determination of overall biosimilarity: a simulated case study exercise. *Biologicals* 42, 128–132
43 Narula, G. *et al.* (2016) Glycosylation in mAb therapeutic products: analytical characterization and impact of process. *LCGC N. Am.* 34, 34–46
44 Jefferis, R. (2009) Glycosylation of antibody therapeutics: optimisation for purpose. *Methods Mol. Biol.* 483, 223–238
45 Zeck, A. *et al.* (2011) Cell type-specific and site directed N-glycosylation pattern of Fc gamma RIIIa. *J. Proteome Res.*10, 3031–3039
46 Costa, A.R. *et al.* (2014) Glycosylation: impact, control and improvement during therapeutic protein production. *Crit. Rev. Biotechnol.* 34, 281–299
47 Steentoft, C. *et al.* (2014) Precision genome editing: a small revolution for glycobiology. *Glycobiology* 24, 663–680
48 Beck, A. *et al.* (2010) GlycoFi's technology to control the glycosylation of recombinant therapeutic proteins. *Expert Opin. Drug Discov.* 5, 95–111
49 Meuris, L. *et al.* (2014) GlycoDelete engineering of mammalian cells simplifies N-glycosylation of recombinant proteins. *Nat. Biotechnol.* 32, 485–U210
50 Piron, R. *et al.* (2015) Using GlycoDelete to produce proteins lacking plant-specific N-glycan modification in seeds. *Nat. Biotechnol.* 33, 1135–1137
51 Spahn, P.N. *et al.* (2017) Predictive glycoengineering of biosimilars using a Markov chain glycosylation model. *Biotechnol. J.* 12, ARTN 1600489
52 Bruhlmann, D. *et al.* (2017) Parallel experimental design and multivariate analysis provides efficient screening of cell culture media supplements to improve biosimilar product quality. *Biotechnol. Bioeng.* 114, 1448–1458

53 Dranitsaris, G. *et al.* (2011) Biosimilars of biological drug therapies regulatory, clinical and commercial considerations. *Drugs* 71, 1527–1536

54 Albrecht, S. *et al.* (2014) Therapeutic proteins: facing the challenges of glycobiology. *J. Health Policy Outcomes Res.* 1, 12–17

55 del Val, I.J. *et al.* (2010) Towards the implementation of quality by design to the production of therapeutic monoclonal antibodies with desired glycosylation patterns. *Biotechnol. Prog.* 26, 1505–1527

56 Kuhlmann, M. and Covic, A. (2006) The protein science of biosimilars. *Nephrol. Dialysis Transplant.*21, 4–8

57 Planinc, A. *et al.* (2016) Glycan characterization of biopharmaceuticals: updates and perspectives. *Anal. Chim. Acta* 921, 13–27

58 Suzuki, S. (2013) Recent developments in liquid chromatography and capillary electrophoresis for the analysis of glycoprotein glycans. *Anal. Sci.* 29, 1117–1128

59 Alley, W.R. *et al.* (2013) High-sensitivity analytical approaches for the structural characterization of glycoproteins. *Chem. Rev.*113, 2668–2732

60 Reusch, D. *et al.* (2015) Comparison of methods for the analysis of therapeutic immunoglobulin G Fc-glycosylation profiles-Part 2: mass spectrometric methods. *Mabs* 7, 732–742

61 Reusch, D. *et al.* (2015) Comparison of methods for the analysis of therapeutic immunoglobulin G Fc-glycosylation profiles-Part 1: separation-based methods. *Mabs* 7, 167–179

62 Marino, K. *et al.* (2010) A systematic approach to protein glycosylation analysis: a path through the maze. *Nat. Chem. Biol.* 6, 713–723

63 Hajba, L. *et al.* (2016) Liquid phase separation methods for N-glycosylation analysis of glycoproteins of biomedical and biopharmaceutical interest. A critical review. *Anal. Chim. Acta* 943, 8–16

64 Zhang, P.Q. *et al.* (2016) Challenges of glycosylation analysis and control: an integrated approach to producing optimal and consistent therapeutic drugs. *Drug Discov. Today* 21, 740–765

65 Bush, D.R. *et al.* (2016) High resolution CZE-MS quantitative characterization of intact biopharmaceutical proteins: proteoforms of interferon-beta 1. *Anal. Chem.* 88, 1138–1146

66 Thakur, D. *et al.* (2009) Profiling the glycoforms of the intact alpha subunit of recombinant human chorionic gonadotropin by high-resolution capillary electrophoresis-mass spectrometry. *Anal. Chem.* 81, 8900–8907

67 D'Atri, V. *et al.* (2017) Hydrophilic interaction chromatography hyphenated with mass spectrometry: a powerful analytical tool for the comparison of originator and biosimilar therapeutic monoclonal antibodies at the middle-up level of analysis. *Anal. Chem.*89, 2086–2092

68 Periat, A. *et al.* (2016) Potential of hydrophilic interaction chromatography for the analytical characterization of protein biopharmaceuticals. *J. Chromatogr. A* 1448, 81–92

69 Mittermayr, S. *et al.* (2011) Multiplexed analytical glycomics: rapid and confident IgG N-glycan structural elucidation. *J. Proteome Res.* 10, 3820–3829

70 Szigeti, M. and Guttman, A. (2017) Automated N-glycosylation sequencing of biopharmaceuticals by capillary electrophoresis. *Sci. Rep.* 7, 11663

71 Tharmalingam, T. *et al.* (2015) A framework for real-time glycosylation monitoring (RT-GM) in mammalian cell culture. *Biotechnol. Bioeng.* 112, 1146–1154

72 Xie, H.W. *et al.* (2010) Rapid comparison of a candidate biosimilar to an innovator monoclonal antibody with advanced liquid chromatography and mass spectrometry technologies. *Mabs* 2, 379–394

73 Sanchez-De Melo, I. *et al.* (2015) N-glycosylation profile analysis of trastuzumab biosimilar candidates by normal phase liquid chromatography and MALDI-TOF MS approaches. *J. Proteomics* 127, 225–233

74 Liu, S. *et al.* (2017) Comprehensive N-glycan profiling of cetuximab biosimilar candidate by NP-HPLC and MALDI–MS. *PLoS ONE* 12, e0170013

75 Montacir, O. *et al.* (2017) Comparability study of Rituximab originator and follow-on biopharmaceutical. *J. Pharm. Biomed. Anal.* 140, 239–251

76 Jarvas, G. *et al.* (2016) Triple-internal standard based glycan structural assignment method for capillary electrophoresis analysis of carbohydrates. *Anal. Biochem.* 88, 11364–11367

77 Karlsson, R. *et al.* (2016) Comparison of surface plasmon resonance binding curves for characterization of protein interactions and analysis of screening data. *Anal. Biochem.* 502, 53–63

Author biographies

László Hajba

László Hajba was awarded a PhD in analytical and environmental chemistry from the University of Pannonia, in 2008, after which he carried out research in the field of biospectroscopy and chemometrics. In 2010, he joined the biotechnology research group at the Research Institute of Chemical and Process Engineering, University of Pannonia. In 2013, he started work in the field of microfluidics and bioseparations at the MTA-PE Translational Glycomics Research Group under the supervision of András Guttman.

Ákos Szekrényes

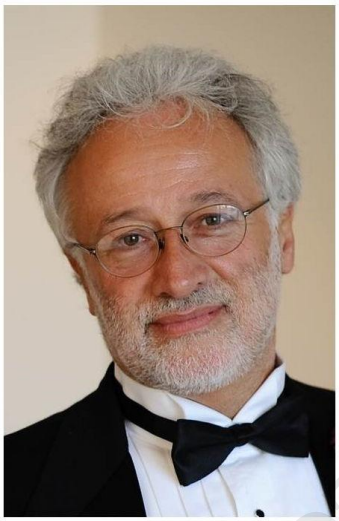
Ákos Szekrényes is currently the head of the analytical development at the biopharmaceutical company mAbxince, where he leads a multidisciplinary analytical team focusing on the development and characterization of biosimilar monoclonal antibodies. He holds a master degree in IT chemistry from the University of Pannonia and a PhD from the Horváth Csaba Laboratory of Bioseparation Sciences at the Medical School of the University of Debrecen, which he carried out under the supervision of András Guttman. His research is focused on novel capillary electrophoresis-based separation techniques and *N*-glycan analysis.

András Guttman

András Guttman is the MTA-PE Lendulet professor of translational glycomics, head of the Horváth Csaba Laboratory of Bioseparation Sciences, and also leads application efforts at Sciex. His work is focused on capillary electrophoresis and CESI-MS based glycomics and glycoproteomics analysis of biomedical and biopharmaceutical interests. Dr Guttman has authored more than 300 scientific publications and 36 book chapters, has edited five textbooks, and holds 23 patents. He is a member of the Hungarian Academy of Sciences, is on the board of several international organizations, serves as editorial board member for various scientific journals, and has been recognized by numerous awards, including the Analytical Chemistry Award of the Hungarian Chemical Society in 2000. He was named as a Fulbright Scholar in 2012, received the CASSS CE Pharm Award in 2013, the Arany Janos medal of the Hungarian Academy of Sciences, the Pro Scientia award of the University of Pannonia, and the Dennis Gabor Award of the Novofer Foundation in 2014. Dr Guttman was also the recipient of the 2017 Dal Nogare Award of the Delaware Valley Chromatography Forum and the Grand Prize of the Swedish Chamber of Commerce.

László Hajba

Ákos Szekrényes



András Guttman,

Figure captions

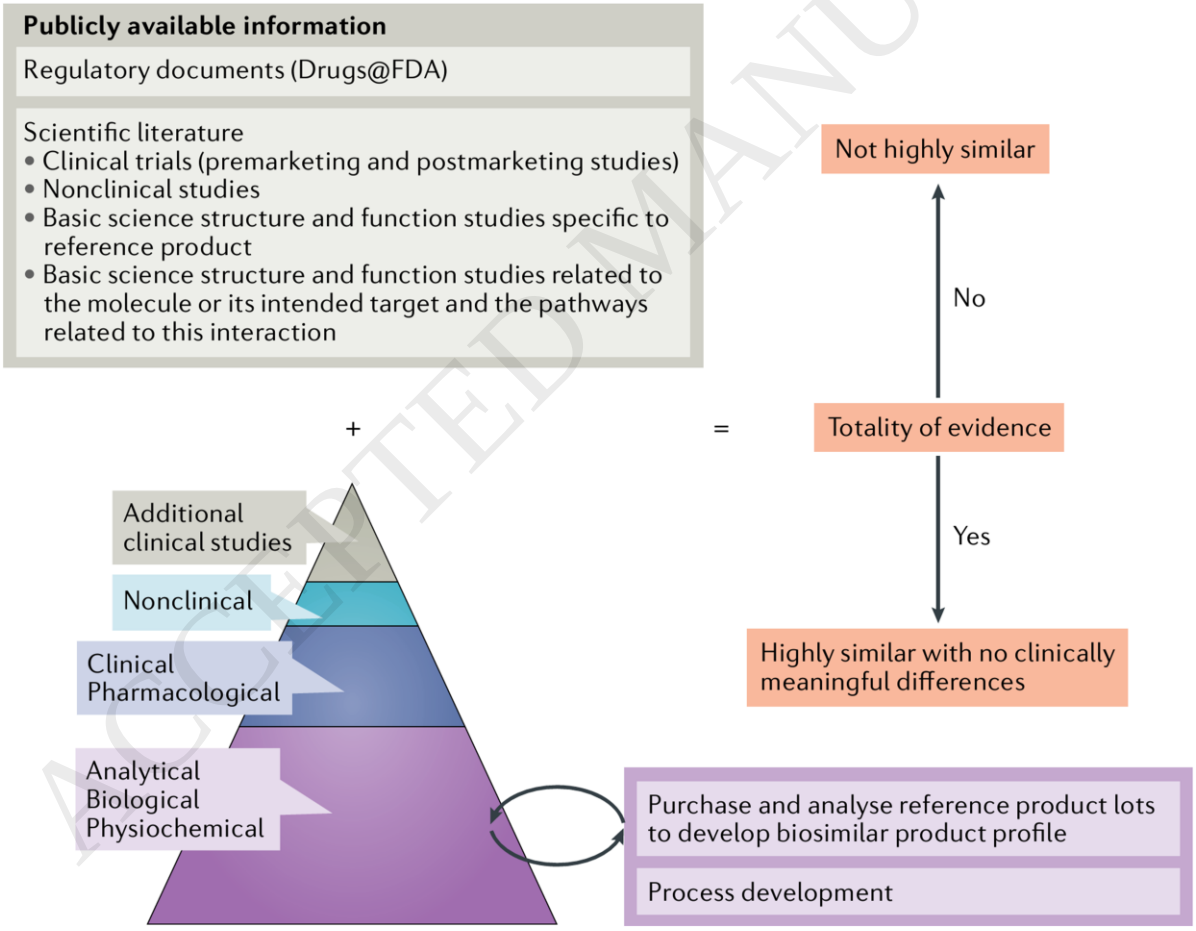
Figure 1. Biosimilar product development scheme. The process

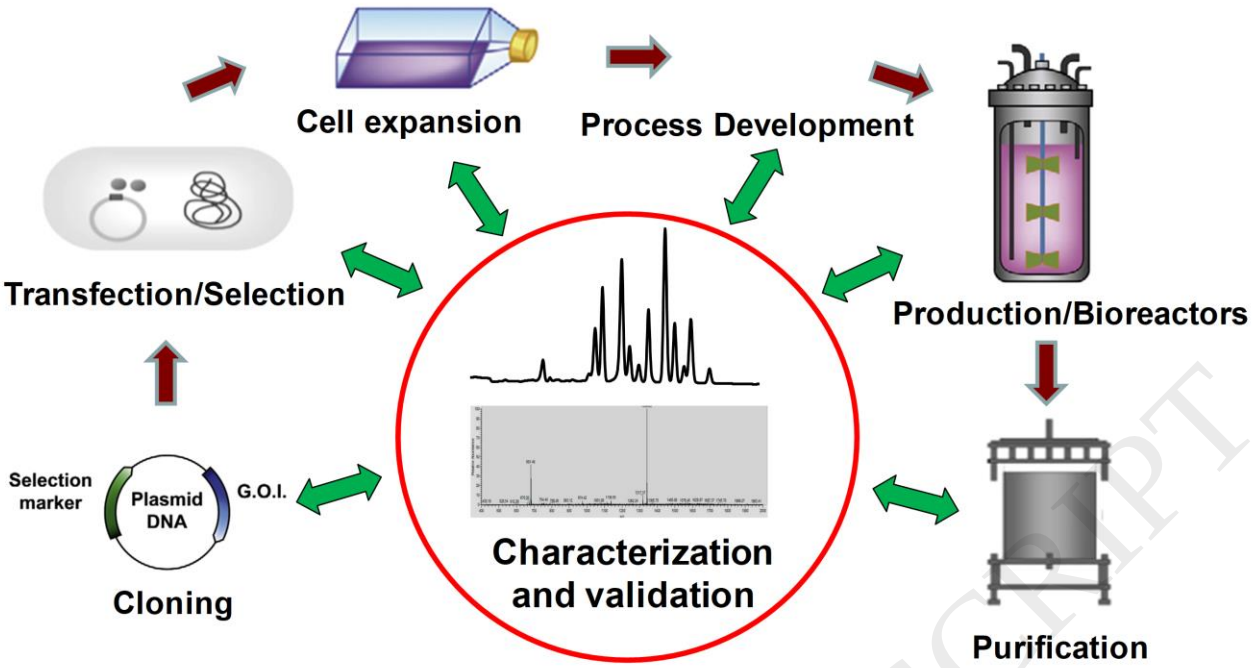
s starts with gathering publicly available information on the reference product (e.g., target and the known biology). The sponsor should analyze the physicochemical properties of the originator product, namely protein structure, post-translational modifications, and biological activity. The next step is the development of an appropriate expression system and manufacturing process for biosimilar production. If regulatory authorities categorize the proposed biosimilar product as 'highly similar', a targeted approach for clinical testing might be possible. Reproduced, with permission, from [5].

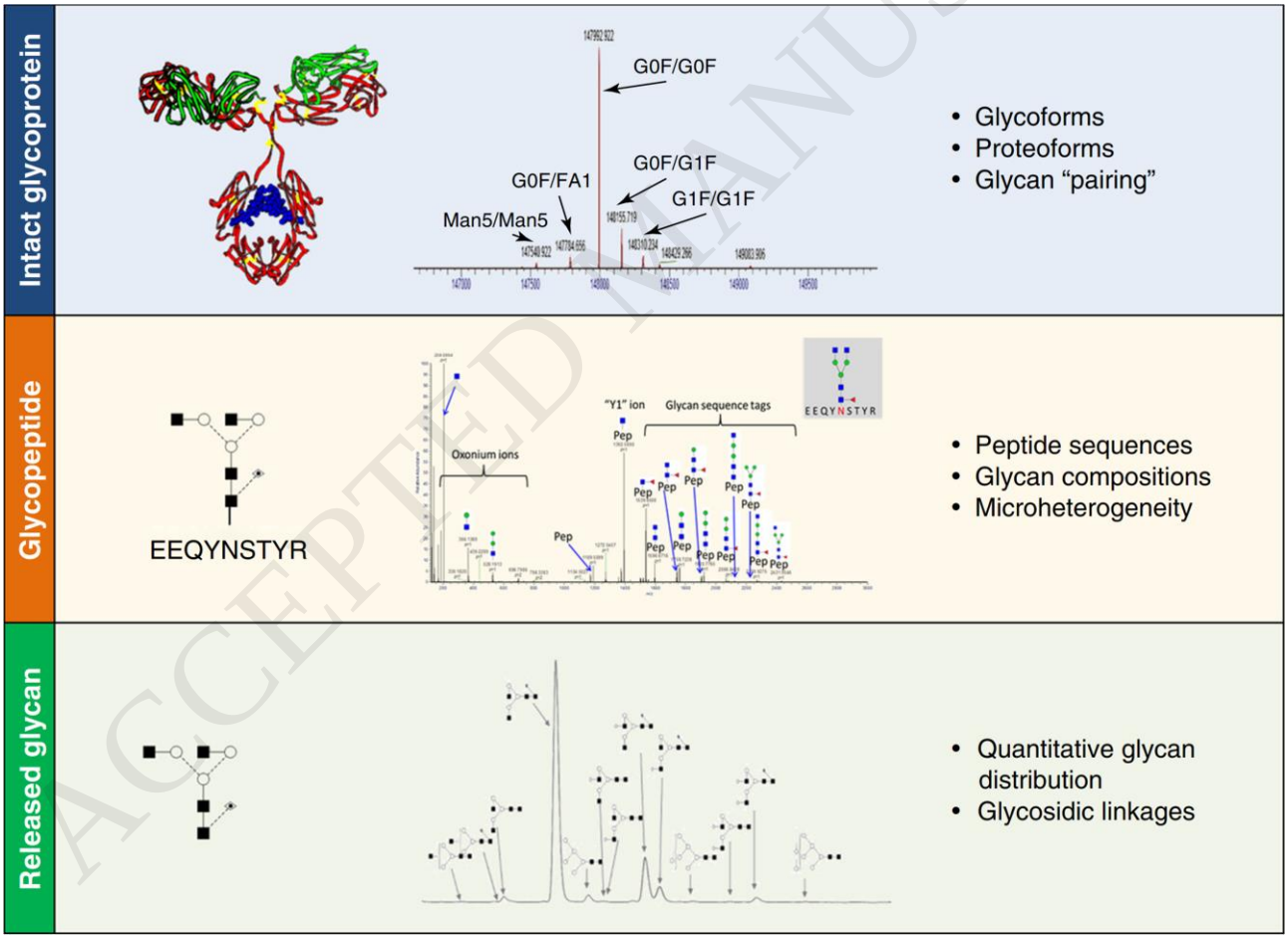
Figure 2. Typical manufacturing process of biologics with the analytical characterization requirement of the steps. The manufacturing process of biologics starts with cloning the relevant gene into a complementary DNA vector and transferring it into a host cell. The appropriate cell line is then selected and expanded in a fermentation medium. In the bioreactor, the selected cell line produces the protein defined by the DNA vector. The complex process ends with the purification of the produced protein therapeutics.

Figure 3. Top-, middle-, and bottom-level *N*-glycosylation analysis. Glycosylation analysis can be performed at these three different levels. In top-level analysis, the intact glycoprotein is analyzed by mass spectrometry (MS) or by liquid chromatography coupled to mass spectrometry (LC-MS). By examining the masses of the different forms of the intact IgG, and comparing with the mass(es) of the deglycosylated IgG, the different glycoforms can be identified. In the middle level, the intact protein is digested to larger glycopeptide fragments and analysed by LC-MS or capillary electrophoresis (CE)-MS. From the middle-level analysis, one can gain information about glycosylation macroheterogeneity (also referred to as 'site specificity'). In the bottom-level analysis of glycosylation, the glycans are either chemically or enzymatically cleaved off the glycoproteins. The released glycans are labeled with a fluorescent dye and analyzed by LC or CE. The bottom-level analysis supports the in-depth microheterogeneity characterization of glycans. Reproduced, with permission, from [64].

Figure 4. Title. (a) An example for the tolerance limits (red-dotted lines) for the Man5 peak in the profile. Tolerance limits are calculated from the Reference Medicinal Product (RMP) inconsistency screening of the different manufacturing lots. For each point of the GU range, the RMP mean and its standard deviation (SD) is calculated to determine the lower and upper tolerance limits. Peak area is considered for the total mannosylation calculation. The similarity score is calculated comparing the test items total mannosylation to the RMP mean ± 1.5 SD range and calculated the same way as for profile similarity as explained in (b). The compositional similarity score considers the similarity scores of each major *N*-glycosylation attribute, such as afucosylation, mannosylation, sialylation, and terminal galactosylation. Similarity scoring system adapted from [77].







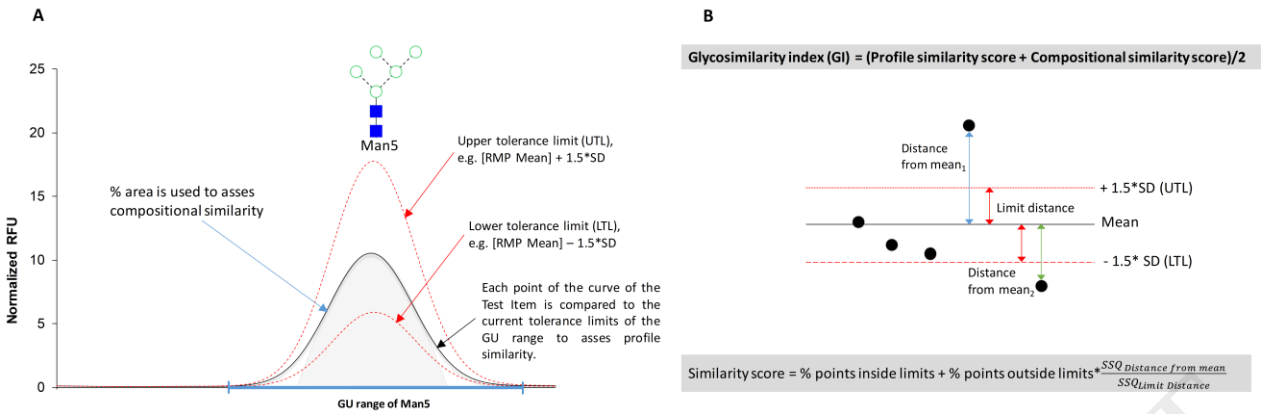


Table 1. Cell type-specific *N*-glycosylation patterns of shFcγRIIIa[a,b]

***N*-glycosylation site**	**Asn-45**		**Asn-74**		**Asn-162**	
Expression host cell	HEK	CHO	HEK	CHO	HEK	CHO
Site occupancy [rel. %]	>95	>95	>95	>95	>95	>95
Complex biantennary and hybrid glycans [rel. %]	>99	~90	n.d.	~5	>95	>95
Complex triantennary glycans [rel. %]	n.d.	~10	n.d.	~5	n.d.	~5
Complex tetraanteannary glycans [rel. %]	n.d.	n.d.	>99	~90	n.d.	n.d.
Core fucosylated glycans [rel. %]	~80	~50	>99	>99	>99	>99
Antennary fucosylated glycans [rel. %]	~15	n.d.	n.d.	n.d.	~70	
Sialylated glycans [rel. %]	~30	~25	~90	>99	~5	~80

[a]After [45].

[b]n.d., not detected.