1. Introduction

ICH Q6B is a useful guide for determining appropriate analytical testing of a biotherapeutic. This document describes an approach to streamlining this testing by selecting a short list of the most powerful analytical methods early in development and the use of universal methods. Universal methods are designed to provide useful data for the majority of protein therapeutics. Some of these universal methods include size exclusion chromatography for the determination of aggregation and degradation products, ion exchange HPLC for the determination of charge variants and oligosaccharide profiling for glycosylated therapeutics. This allows for more testing earlier without spending time and money optimizing methods for a particular drug.

The authors discuss how to determine what testing is required in early development, how to know when more testing is required and when more optimization of the universal methods is warranted. Data will be shown to demonstrate how the same assay can be used from initial lot-to-lot comparisons through assay validation and a full characterization of the therapeutic. For example, N-linked oligosaccharide profiling by HPLC is a relatively quick and simple means of evaluating the glycosylation of a biotherapeutic using chromatographic peak areas. Understanding the heterogeneity of the glycosylation early in product development from lot-to-lot comparisons, along with bioassay, in vivo and pre-clinical testing, is an effective way of defining the window for the product’s critical quality attributes. Further along in development, the N-linked profiling assay can be used to quantitate the heterogeneity of the glycosylation and the oligosaccharide peaks can then be collected and identified by mass spectrometry, glycosidase sequencing and other methods.

2. Guiding principles for characterization throughout drug development

The methods will change, assays may be added, and some assays eliminated during product development. However, as a drug moves from Pre-Clinical to Clinical to Marketing Approval there are certain points to keep in mind starting with the earliest lot to ensure a solid characterization package for your drug.

Scientific Literature: For many drugs there is very little information on how the structure of the molecule impacts the function of the molecule. However, when there is information on
the molecule or similar molecules, this information can be used to select the quality attributes of your molecule most likely to affect the efficacy or safety of your drug. These attributes are worth monitoring over the course of development so that they can be used to define the design space for your drug—variation that is acceptable because it has been demonstrated to have no affect on the safety and efficacy of the drug.

**Retains:** Assays evolve and new ones may be developed throughout the characterization process. It is important to be able to retrospectively analyze your drug. This is especially true for any clinical or key pre-clinical lots. The more lots you have to test, the easier it is to confidently set assay specifications.

**Reference Material:** Beyond retains, it is important to establish Reference Materials. These are drug lots that are set aside to demonstrate that future lots are comparable to previous material and that the drug is not changing. These can also ensure that an assay is delivering consistent results over time.

**Stability:** Drug lots that are not comparable to historical lots, whether these changes are coincident with decreased biological activity or not, are extremely valuable for understanding the relationship between activity and drug heterogeneity—what heterogeneity is acceptable and what heterogeneity negatively impacts the safety and/or efficacy of the drug product. It is important to identify assays that detect these critical quality attributes and the early in development. When changes in the product do not occur naturally, it is possible to force these changes using accelerated stability/degradation studies where the drug is stressed and forced to degrade.

**Release Tests/Assay Validation:**
It is always important to keep in mind that some assays will become Release Tests and will need to be validated. It is wise to remember this when choosing and designing assays to avoid trouble in the future when they need to be validated and transferred to a QC laboratory.

### 3. Pre-clinical evaluation of sample heterogeneity

Numerous lots of drug are prepared throughout development of a therapeutic. These are first used in animal studies, followed by preclinical studies and finally clinical studies. Throughout the course of these studies, valuable information regarding the heterogeneity of the drug can be determined early on using some basic analytical assays. These studies can help to define a window of acceptable product variation. “If a consistent pattern of product heterogeneity is demonstrated, an evaluation of the activity, efficacy, and safety (including immunogenicity) of individual forms may not be necessary” (ICHQ6B). For example, oligosaccharide profiling, IEX-HPLC and SEC-HPLC are universal assays usually requiring little methods development and are relatively quick and easy to perform. These assays are also easily transferred to a QC environment, can be validated and used as lot release assays. Performing these assays early in the development process helps to identify the extent of heterogeneity routinely seen in a sample. Performed routinely, these assays can identify stability issues (oxidation/deamidation or aggregation/degradation during storage) or critical process parameters (glycosylation changes with process change). Understanding the product variants at an early stage aids in optimizing the types of assays required for the determination of product identity later on in development. For example, if charge variants are present in a sample, the IEX-HPLC assay may need to be further optimized for resolution of these variants. Further, these variants can be isolated and subjected to peptide
mapping techniques so that the site of oxidation can be determined and the peptide map optimized to resolve the variant peptides. It is often possible to force degradation of the drug using accelerated stability or degradation pathway studies where the drug is stressed by changing parameters such as the pH of the formulation and storage temperature. This helps to identify critical quality attributes of the drug and determine whether the assays are able to detect these critical changes. It also allows for determining which changes in the drug impact the biological activity.

During early stage development, there should be an adequate focus placed on determining the appropriate analytical tests for the identity, purity and concentration or activity of the drug.

4. Phase I: Determination of physicochemical properties

Determination of the physicochemical properties of the drug substance is a critical step in the development of suitable identity tests. These properties include the purity, concentration and/or activity, molecular weight and size of the sample as well as charge, hydrophobicity and post-translational modifications (PTMs).

4.1 Purity and molecular weight determination

Purity and molecular weight analyses are most easily evaluated by SDS-PAGE/CE-SDS (native and denaturing) along with Western blot analysis. SDS-PAGE/CE-SDS allows for the visualization of the protein species in a sample and can be quantitative depending on the detection method. Comparison of native versus denatured sample by gel electrophoresis or CE analysis can also confirm the presence of 3-dimensional structure such as disulfide bonding and oligomerization. Further analysis of the gel by Western blotting with various antibodies identifies the presence of drug substance degradation products along with process specific contaminants. If molecular weight variants such as aggregates or degradation products occur in a product, size exclusion chromatography can be a valuable tool for the routine detection and quantitation of these variants. SEC-HPLC is a universal method and can be run with minimal development work. SEC columns can separate over a wide range of molecular weights, and optimization of the column conditions need not take place until later in the product development stage if the method is not separating molecular weight variants identified in the product through orthogonal methods. SEC-HPLC can be preparative and used as an initial step for a two-dimensional analysis of size variants. The variant peaks can be isolated for characterization by MALDI-TOF MS to confirm the size or analyzed by peptide mapping for identity, etc. Once the purity of the sample is ascertained by these methods, the molecular weight of the drug substance should be confirmed by MALDI-TOF analysis. This type of multidimensional analysis is illustrated in work by Kotia et al. (Kotia, 2010). In this work, truncated fragments of a monoclonal antibody heavy and light chain are detected by CE-SDS. These truncated fragments were correlated to heavy and light chain peptide maps and sequence identity was confirmation by N-terminal sequencing. Multidimensional analysis of product degradation utilizing both SEC-HPLC and CEX-HPLC followed peptide mapping with MS-ESI detection is shown in work by Lau et al. and Kim et al. (Lau H., 2010) (Kim, 2010)

4.2 Protein content and extinction coefficient determination

Protein content needs to be determined early in development since many other assays are dependent upon it. It is important to use an assay that is precise, i.e. gives reproducible
results. For instance, the results of potency and bioassays are reported based on protein concentration. Reporting these results based on inconsistent protein determinations yields inconsistent results that may not accurately reflect the activity of the drug. Determination of protein content is not as straightforward as it would seem. Colorimetric assays such as the Bradford assay or bicinchoninic acid assay rely on the use of an external reference standard for protein determination. However, different proteins will have different molar ratios of reactive sites and the same amount of one protein can give a different absorbance than the same amount of the reference standard used. Amino acid analysis can also be used for quantitative analysis of protein content, but this assay also has its difficulties. Different proteins can hydrolyze differently under the same conditions and many amino acids are not stable to acid hydrolysis. Numerous studies have been conducted by the ABRF to determine the precision and accuracy of colorimetric protein assays and quantitative amino acid analysis. These studies report high variability between test sites and variability in the accuracy of measuring different proteins as well. Summarizing the results from the AARG2003 study (Alterman et al.), 28 laboratories were given five protein samples at a concentration of 2.5 mg/mL. Excluding the results from the fetuin sample (average determined concentration = 1.18 mg/mL) the average yield of all four proteins was 1.69 ± 0.24 mg/mL. This indicates a substantial bias (32%) in the determination of protein content, although some of this bias could be due to initial sample handling and salt content. In our hands, we see an assay bias of 13% for our BSA standard (N= 94) with an average yield of 1.75 ± 0.47 mg/mL from a 2.01 mg/mL standard solution.

Theoretical extinction coefficients can be easily determined from the amino acid sequence (e.g. using the ProtParam tool at www.ExPaSy.org). However, the extinction coefficient needs to be confirmed experimentally using a suitably determined protein concentration as protein conformation can have an effect on the measured absorbance.

4.3 Charge and hydrophobicity variants

Charge and hydrophobicity variants are most easily analyzed by ion-exchange and reversed-phase HPLC. These assays are highly sensitive and selective, both being able to detect changes at a single site in the protein sequence. Both assays are an excellent starting point for multi-dimensional analysis as the methods are quantitative and scalable. Variant peaks can be isolated for characterization by mass spectrometry methods or peptide mapping. Hydrophobicity variants can consist of single point amino acid substitutions. Charge variants can include oxidized, deamidated, phosphorylated, and sulfated species as well as variation in the extent of sialylation. More than likely these variants will need to be characterized using in-depth identity tests.

Ion exchange chromatography can easily pick up low levels of charge variants. The assay is very quick and in general, requires only a small amount of purified protein. Running this analysis early on allows for the early identification of charge variants. Early identification of the extent, or percent modification, can be used to set product specifications. These assays can also be used throughout process development to determine which process changes affect product variation. Ion exchange HPLC is an excellent starting point for further investigation into the identification of the variants. Vlasek et al. describe the use of cation exchange chromatography for characterization of a monoclonal antibody in which two lysine variants were separated using CEX-HPLC as well as two acidic variants. (Vlasak, 2009). The initial use of ion-exchange HPLC allowed for isolation and extensive characterization of the variant
peaks, such as preparation of Fab fragments from the variants as well as preparation of heavy and light chains, followed by mass spectrometry analysis, Edman degradation, DSC to evaluate stability and CD spectroscopy for higher order structure analysis.

4.4 Glycosylation
Glycosylation of biotherapeutics is now routinely analyzed early in development and typically drug lots are assayed as they are manufactured for lot-to-lot consistency in their glycosylation. We have already covered this topic in detail in a review article (Higgins, 2010). Although monosaccharide composition analysis was historically used to monitor glycosylation (releasing monosaccharides from oligosaccharides using acid and then quantifying each monosaccharide) this method has now been mostly replaced by oligosaccharide profiling. Oligosaccharide profiling involves releasing intact N- and/or O-linked oligosaccharides from the protein and then analyzing them by HPLC or mass spectrometry. Oligosaccharide profiling can be used to monitor the population of oligosaccharides present on the glycoprotein and most importantly determine whether the heterogeneity is the same from lot-to-lot. Mass spectrometry of the released pool of oligosaccharides can be used to characterize the types of oligosaccharides present (oligomannose, complex, and antennarity). Both HPLC and MS can be used to quantify the relative ratio of different oligosaccharides, however, MS often requires the use of radioactive labels. The ratio is valuable when these methods are used for lot release.

Monosaccharide composition analysis is now mostly utilized in situations where additional testing is warranted: process qualification lots, lots for reference material or comparability testing. It is also used when a particular monosaccharide is critical to a drug (e.g. the amount of fucose) or could impact safety (a monosaccharide not commonly found on human proteins but found in the expression system used to produce the drug). Sialic acid analysis can also be a useful assay since the level of sialylation is often a critical quality attribute as it affects the plasma clearance of many glycoproteins.

5. Phase II/III: Development and characterization of identity tests

5.1 Determination of primary amino acid sequence and peptide mapping
The primary amino acid sequence of a protein is determined using a combination of analytical techniques. Complete sequence information is often difficult to obtain. The N and C terminal amino acids need to be confirmed. N-terminal sequencing can be performed by Edman degradation or MALDI-PSD analysis. MALDI-PSD analysis has the advantage in that it can also be used to identify the C-terminal amino acids in a protein or a peptide. Otherwise, C-terminal sequencing involves the use of a specific C-terminal protease followed by separation and identification of the released amino acids by HPLC. Full sequence determination is generally performed using a combination of peptide mapping following peak identification by Edman degradation or MS/MS.

Development of a good peptide map is critical for the development of appropriate identity tests. Once the purity of the product has been established the identity of the protein should be determined by peptide mapping. Peptide mapping involves cleaving a protein into smaller peptides, generally using enzymes that cleave the protein at specific amino acids to generate peptides that can be predicted from the peptide sequence. These peptides are then separated by reversed-phase HPLC followed by detection using UV and/or MS. At the start, if the protein is known to contain N-linked oligosaccharides or disulfide bonds, comparative
peptide maps of deglycosylated and non-reduced protein should be generated to evaluate the extent of variability in these regions and to begin identifying the disulfide linked peptides and the glycopeptides. A good peptide map will resolve most of the peptides in the mixture to baseline. Well resolved peaks can then be identified by N-terminal sequencing using Edman degradation or LC-MS. Both methods are commonly used.

5.2 Disulfide and glycopeptide mapping
If a protein contains disulfide bonds, it is necessary to cleave the disulfide bond with a suitable reducing agent such as dithiothreitol followed by blockage of the free cysteine with an agent such as iodoacetamide or iodoacetic acid to prevent the random reformation of the bonds. Reduction of the disulfide bonds is necessary to confirm the amino acid sequence of the peptide, as mixed peptides will not correlate to the known amino acid sequence. Peptides involved in disulfide linkages can be identified by comparison of the reduced and alkylated peptide map with a peptide map generated from the non-reduced protein (see Figure 1). Additional identification of the peptides involved in the disulfide bridging can be confirmed by MALDI-TOF analysis using a reducing matrix. These matrices will break the disulfide bonds on target, generating masses of the disulfide linked peptide along with the masses of each of the released peptides. Examples of the peptide mapping for the determination of disulfide bonding in antibodies can be found in Bloom (1997) and Wypych (2008) (Bloom, 1997).

Fig. 1. Disulfide Peptide Map of Fetuin. The tryptic peptide map is enlarged to show the region of the map containing the most changes. Arrows indicate the generation of new peaks in the peptide map when fetuin is reduced and alkylated prior to trypsin digest. The top (red) chromatogram represents fetuin tryptic peptides generated from non-reduced fetuin. The bottom (black) chromatogram represents fetuin tryptic peptides generated from reduced and alkylated fetuin.

Comparison of peptide maps generated from glycosylated and deglycosylated drug substance can be used to identify glycosylated peptides (see Figure 2). Identification of these peptides can be performed by LC-MS, LC-MALDI or N-terminal sequencing. Further, glycosylated peaks can be isolated and for the purpose of attempting to determine the type of glycosylation on the peptide. Examples of approaches to the identification of glycopeptides by glycopeptide mapping can be found in (Rohrer, 1993) (Ohta, 2002) Ohta (2002) and Rohrer (1993).
5.3 Site-specific glycosylation

Site-specific glycosylation analysis gives detailed data on the relationship between peptide and glycosylation heterogeneity. A review by An et al (2009) describes current methods used. The most challenging part of site-specific analysis is the enrichment of glycopeptides, a necessary step due to the relative low abundance of glycopeptides in comparison to non-glycosylated peptides. Such methods as lectin affinity columns, HILIC (Hydrophilic Interaction Chromatography) or ERLIC (Electrostatic Repulsion-Hydrophilic Interaction Chromatography) columns are employed, to varying degrees of success.

Lectins will bind oligosaccharides and different lectins show a preference for different types of oligosaccharides. Lectin affinity columns can therefore be used to bind glycopeptides and/or separate glycopeptides into groups with differences in their oligosaccharides. HILIC enrichment is based upon the proposed increase in hydrophilicity of glycopeptides versus non-glycosylated peptides due to the polysaccharide moiety. ERLIC operates in a similar fashion as HILIC, with the addition of electrostatic interactions to the binding mechanism. Both HILIC and ERLIC utilize gradients of high organic solvent to low organic solvent, where HILIC is generally performed using a polar chromatographic matrix and ERLIC utilizing an ion exchange matrix. Glycopeptides will bind more strongly than most non-glycosylated peptides, allowing for the removal of non-glycosylated peptides from a mixture. While these methods can enrich the glycopeptide content of a peptide mixture, the separation of the two has not been optimized. Some glycopeptides will be found in the early eluting fractions, and complex mixtures of glycopeptides can be seen to elute over a wide concentration of decreasing organic solvent along with some non-glycosylated peptides.

Due to the heterogeneous nature of glycosylation, one peptide might contain many different glycoforms, leading to challenges in separation and identification.

5.4 Characterization of the oligosaccharide profile

Characterization of the oligosaccharide profile is critical for understanding the heterogeneity of glycosylation. By having a well-characterized oligosaccharide profile, changes in glycosylation are easier to track. A summary of characterization options is shown in Figure 2.
3. Oligosaccharide profiling is commonly done by HPLC coupled with fluorescent detection (after oligosaccharide derivatization with a fluorescent label such as 2-aminobenzoic acid or 2-aminobenzamide), in-line LC-MS (typically using electrospray ionization (ESI) as the MS ion source) or standalone MS (either permethylated, derivatized, or native glycans using either MALDI or ESI).

Fig. 3. A flowchart of possible approaches for characterization of the oligosaccharide profile.

When working from an HPLC fluorescent profiling method, fractions must be collected for further characterization. After fractions are collected, glycoforms can be identified either through glycosidase sequencing, which relies on highly specific enzymes and shifts in retention times to confirm linkage and composition, or through further mass spectrometric analysis. An example of HPLC profile characterization of the N- and O-linked oligosaccharides of fetuin using fraction collection followed by MALDI-TOF-MS identification is shown in Figure 4. Additional mass spectrometric analysis may include post source decay using MALDI-TOF-MS, MSn using such fragmentation techniques as electron transfer dissociation (ETD) or collision induced dissociation (CID), or glycosidase sequencing using shifts in mass to identify losses from the parent mass. A review by Geyer and Geyer (2006) contains further information. Often, multiple approaches are needed to fully characterize an oligosaccharide profile, as seen in Qian et al (2006).

While MS-MS and MSn are powerful tools, analysis of mass spectrometric data can be complicated and time-consuming, requiring database searches and knowledge of glycobiology to successfully narrow down potential structures. While several groups have sought to automate the process, in the authors’ experience the results often contain inaccurate structures, not likely to be found in nature, which makes data analysis more complex.
Fig. 4. N- and O-linked Oligosaccharide Profile of Fetuin. The selected fractions were collected and analyzed by negative ion mode MALDI-TOF-MS to determine peak identity. Linkages were not confirmed; however, the results are strongly correlated to the literature on fetuin. Symbols: ■ N-acetylglucosamine (GlcNAc); □ N-acetylgalactosamine (GalNAc); ● galactose; ○ mannose, △ fucose and ♦ sialic acid.

6. Marketing approval: Lot release assays

Many assays can be used to characterize a drug during development. Some will continue to be used to troubleshoot production problems and demonstrate comparability through process changes and a subset of these assays will become release assays. The list will include release assays common to most, if not all, biotherapeutics such as concentration, purity, peptide mapping and oligosaccharide profiling. Additionally, any assays determined during process development which monitor critical quality attributes specific to the drug will become a release assay. These assays could include quantification of sialic acid, assays used to evaluate the truncation of the amino acid sequence and assays which identify post-translational medications such as phosphorylation or gamma carboxylation.

Unlike assays used for characterization, the release assays will need to be validated and transferred to QC laboratories. Some of these assays such as concentration or activity are so important for moving the product into the clinic that they will need to be validated early and used for lot release on earlier lots. Other assays, like peptide mapping and oligosaccharide profiling might not be validated until later (but before marketing approval). This allows more time for the extensive work often required to identify the peptides/oligosaccharides in the peaks and collect data on any variation seen in these separations from many drug lots before setting specifications.

7. Conclusions

An approach to characterization of a therapeutic has been presented in which analytical testing is simplified by initially focusing on the assays required early in development for all therapeutics (purity, molecular weight, activity and/or concentration). Next, the drug is evaluated using assays that are routinely used for release testing such as peptide mapping and oligosaccharide profiling. Incorporating these assays earlier in development allows for better tracking of lot-to-lot variation in the product and collection of more data before setting assay specifications. Finally, assays will need to be added or existing assays modified to track variation in the drug that is detected during development (changes such as truncation of the protein or post-translational modifications). Tracking any heterogeneity in the drug is most critical if it is likely to affect the safety or efficacy of the drug.
The authors believe that it is important to use universal methods. Universal methods do not require extensive assay development work and are sensitive to certain physicochemical properties of the drug such as molecular weight or charge. It is better to obtain more data earlier in development using these universal methods than to spend time early in drug development optimizing assays specific for your protein before understanding the heterogeneity naturally present in your drug and how it is likely to degrade.

8. References


Drug discovery and development process aims to make available medications that are safe and effective in improving the length and quality of life and relieving pain and suffering. However, the process is very complex, time consuming, resource intensive, requiring multi-disciplinary expertise and innovative approaches. There is a growing urgency to identify and develop more effective, efficient, and expedient ways to bring safe and effective products to the market. The drug discovery and development process relies on the utilization of relevant and robust tools, methods, models, and validated biomarkers that are predictive of clinical effects in terms of diagnosis, prevention, therapy, and prognosis. There is a growing emphasis on translational research, a bidirectional bench to the bedside approach, in an effort to improve the process efficiency and the need for further innovations. The authors in the book discuss the current and evolving state of drug discovery and development.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: