Introduction

The quality of the product is a most important aspect of manufacturing for products of therapeutic applications, even more so for biologics. The complex structures of those products make the quality determinants harder to quantify. Ensuring the manufacturing process produces these biologics of consistent quality is critical for the manufacturer as well as for the public. Manufacturing processes should be designed to robustly deliver consistent yields and product quality. Increasingly cells and viruses with even higher degrees of complexity than proteins are employed as therapeutic agents. Although gene and cell therapy products are still few, maintaining and monitoring the quality of these products will only become more crucial. The vast knowledge and experience of therapeutic protein manufacturing will help set the path forward for the development of manufacturing processes for these new products.

Panel 4.1. Quality of a Therapeutic Product

- Structural identity
- Functional/biological activities
- Impurities
- Product homogeneity
  - Natural variability
  - Chemical variants

What constitutes product quality? Generally speaking, a product is established by its structural identity (Panel 4.1). Sometimes, structural variants of the product arise in the manufacturing process.[1] These must be controlled in the drug substance so as to restrict their levels to within acceptable limits. Most therapeutic proteins are glycoproteins whose glycans vary somewhat between individual protein molecules, but whose glycan heterogeneity must be within a certain bound. Inevitably, the product will have some contaminants, which must also fall within a

set limit. Finally, the product must be tested and found biologically functional and safe before product release. The general quality requirements and characterization methods are detailed in Table 4.1.

The productivity of cell culture bioprocesses is currently very high. For many proteins, such as IgG, the titer in the reactor has exceeded the highest level in its native environment *in vivo*. At such high concentrations, physical interactions with other components may differ from those in older generation processes or those *in vivo*. These molecular interactions may pose challenges in protein aggregation or host cell protein contamination. The high productivity achieved nowadays is accomplished with a high cell concentration and often high-stress culture conditions such as high sugar content, high salt and waste metabolite concentration, and low temperature in the production bioreactor. Under more extreme culture conditions, the secreted product may be subject to chemical modifications as well as enzymatic modifications caused by enzymes released from dead cells. Some higher-order protein structural changes, such as aggregation or fragmentation, may even take place during the product recovery process. For some products, the productivity of the producing cell line has reached or even exceeded the secretion rate of professional secretors in the human body (e.g., liver cells and antibody- and insulin-secreting cells). It is unclear whether these hyperproducing cells, having been pushed to the limit of some of their functions related to protein synthesis and secretion, have exposed their product to a higher probability of deviation (e.g., a higher degree of heterogeneity in a glycan profile or a higher rate of amino acid misincorporation).

The first task in ensuring consistent product quality in manufacturing is the identification of the properties of a given product that play a key role in its safety and clinical efficacy. This is undertaken while the product is being developed, and the knowledge is applied while developing the process. In the manufacturing process, the product is made in bioreactors, purified through downstream steps, stored as a drug substance, and then converted to a drug product for use in clinical application. The operating conditions in production and downstream recovery that may lead to unsatisfactory product quality should then be identified and the controlled range of critical process parameters defined. After the product is launched, the product quality data is continually collected, compiled, and analyzed in order to better control product quality during the manufacturing process. The process of quality enhancement and process improvement thus continues throughout the product's life cycle.

This chapter will first discuss those chemical and physical properties of protein biologics which are important quality attributes. This will be followed by a discussion on Quality by Design (QbD), an approach to process development whereby the manufacturing process and controls are proactively designed to deliver the intended product. The last section

Table 4.1. Characterization of a Protein Biologic

Class	*Attributes*	*Assays*
Primary sequence and identity	Intact mass, peptide map, N-terminal sequencing, MS-based sequence identification (mutations, misincorporations)	SDS-PAGE and WB, HPLC, mass spectrometry
Higher-order structure	Secondary (α-helix structure), quaternary structure (e.g., disulfide bonds)	Circular dichroism (CD), differential scanning calorimetry (DSC), Fourier transform infrared spectroscopy (FTIR), analytical ultracentrifugation (AUC), disulfide bond analysis, free-thiol analysis
Purity/impurity	Residual host cell proteins (HCPs), residual DNA, other process impurities such as insulin, IGF-1, Protein A, excipients, methotrexate, and beta-glucan	Hydrophobic interaction chromatography (HIC), ELISA, HCP by LC-MS/MS
Structural variants	Aggregates / high and low molecular weight species	Size-exclusion chromatography (SEC), SDS-PAGE, capillary electrophoresis (CE)
	Charge variants; acidic and basic	Cation exchange chromatography (CEX), capillary zone electrophoresis (CZE), HIC, capillary isoelectric focusing (cIEF)
	Degradation and chemical modification, hot spots (deamidation, oxidation, isomerization, glycation, succinic acid modification, leader sequence, etc.)	Peptide map-MS/MS
	Glycoforms (N-linked glycan, sialylated species, high mannose, afucosylation, galactosylation, o-sialylation)	HILIC, CZE, UPLC, HPLC-MS/MS
Functional	Ligand binding kinetics	ELISA, surface plasmon resonance (SPR)
	MoA-based, cell-based function assay	Bioassay
	Fc effector function assay (specific for Fc-containing products) e.g., FcRN binding	Bioassay
	Complement dependent cytotoxicity (CDC), antibody dependent cell-mediated cytotoxicity (ADCC)	Bioassay
Safety	Endotoxin, bioburden, sterility, mycoplasma	
General tests	Affinity (Protein A) HPLC, protein content by UV280, visible and sub-visible particles, excipient, pH, appearance, osmolarity, residual moisture, clarity, color	

emphasizes the importance of maintaining consistent product quality throughout the entire life cycle of the product.

Quality of Protein Product

Identity of Protein Product

In protein production through recombinant cells, the identity of the product gene is dictated by the gene sequence that encodes the product and by the downstream purification process. In the early stages of product development, the chemical structural identity of the product is affirmed at both the gene sequence level and the protein level. During manufacturing, the product is routinely tested to ensure its structural integrity as well as to detect the presence of structural variants.

Molecular and immunochemical identity and quantification

The identity of a protein product is typically established by its molecular weight and chemical structure. The purified product protein is typically run on polyacrylamide gel electrophoresis (PAGE) to assess its molecular weight and purity in terms of contamination by other proteins. The purified protein solution is run on a polyacrylamide gel in an electrical field, which separates individual proteins based on their mobility as determined by their size or molecular weight (Figure 4.1). The post-electrophoresis gel is stained with Coomassie Blue or Silver Stain to visualize the migration of the protein from its initial position. Typically, the protein sample loaded into PAGE is treated with sodium dodecyl sulfate that denatures the protein, dissociates non-chemically cross-linked protein subunits, and extends the protein to better reflect its size. Since larger proteins migrate at a slower rate, the molecular weight of a protein molecule can be determined by calibrating to a mixture of protein markers. PAGE also serves as a tool for checking protein purity. A highly purified product protein should show only the gel band(s) of the product without any contaminating band(s) of impurity. Some proteins consist of subunit proteins that are bound together by disulfide bonds (e.g., IgG). On a native (non-denatured) protein gel and a reduced gel (in which a reducing chemical breaks the disulfide bond), the protein thus displays different bands as the reduced monomers migrate at different rates down the gel. This type of analysis is often employed to detect low molecular weight species or fragments generated by enzymatic degradation.

Western Blot is another method used to identify the product protein. After the purified product protein is run on PAGE, the protein is blotted to a membrane. The membrane is then treated with an antibody against the product protein. The antibody, which had been previously labeled with an enzyme, isotope, or fluorogenic tag, binds to the protein band.

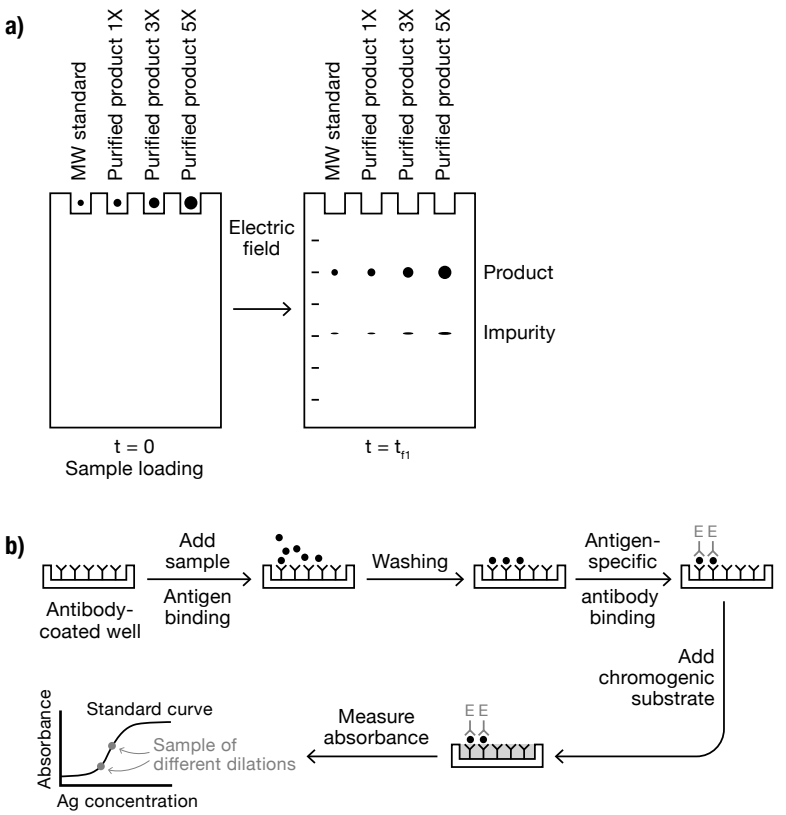

Figure 4.1. (a) Polyacrylamide gel electrophoresis for determining product protein molecular weight and purity. (b) Enzyme-linked immunosorbent assay (ELISA) for determining antigen concentration.

When the unbound antibody in the membrane is washed off and a chromogenic substrate of the antibody-conjugated enzyme is applied, the product protein bands appear in the correct molecular weight location.

Immunological method ELISA, which utilizes antibodies against the product protein, is widely used for measuring protein concentration. There are a number of different ways to implement ELISA. In one scenario, the antibody against the product protein is coated on a 96-well or 384-well assay plate (Figure 4.1b). The sample solution containing the product at different dilutions is then added to allow for binding to the antibody. Then a second antibody that recognizes a different epitope of the product and has been conjugated to an enzyme is added to bind to the product that was captured by the antibody coated on the plate. After a washing step, the amount of the enzyme of the second antibody remaining on the coated surface is quantified by measuring the rate of the enzymatic reaction. The enzymes commonly used typically convert a chromogenic substrate to a product with a visible signal. For example, horseradish peroxidase (HRP) converts a substrate to a colored molecule

that can be detected using a spectrophotometer. A standard curve constructed with a serially diluted, purified antigen is used to determine the concentration of the antigen in the sample. Such "sandwich ELISA" is used for quantification of antibody products and many other proteins. Another method specific to antibody quantification utilizes affinity chromatography. Protein A, which preferentially binds to antibodies, is immobilized on a column support. At neutral pH, the antibody molecule binds to Protein A through the Fc region while host-cell proteins, cell culture media components, and buffer flow through the column. Captured antibodies elute at acidic pH and are detected by UV absorbance at 280 nm. A calibration curve is derived from standards and the corresponding peak areas using linear regression analysis. Sample and control product concentrations are calculated from the calibration curve.

Peptide sequence identity

The most accurate check of protein identity is based on the primary structure, or the protein's amino acid sequence. This is generally done by the classical Edman degradation using an automated protein sequencer or by liquid chromatography followed by mass spectrometry. In Edman degradation, complex proteins with multiple chains are separated into single chain molecules. Large proteins are then enzymatically digested into fragments of less than 50 amino acids long. Next, the target protein or peptide is immobilized onto a solid and subjected to chemical hydrolysis from the protein's N-terminus. The released amino acid is made to fluoresce and identified in HPLC. By cycling through the terminal amino acid degradation process, the entire sequence of the protein can be established.

Amino acid sequence identity can also be established by mass spectrometry. To establish the sequence identity of the protein product, the purified protein is subjected to proteolysis to fragment the protein into smaller size peptides (typically, trypsin is used to cleave the C-side of a peptide bond adjacent to lysine or arginine) (Panel 4.2). This process cleaves the protein to peptides of mostly 15–30 amino acids long (Figure 4.2). The resulting mixture of peptides is then subjected to reverse-phase HPLC or capillary electrophoresis for separation. Next, the peptides are separated into peaks and fed into the mass spectrometer (MS) for detection. Each peptide peak will have a characteristic mass over charge ratio (m/z) of the

Panel 4.2. Protein Identification by Peptide Mapping

- Peptide mapping entails exposure of the protein product to chemical hydrolysis at specific points along the protein backbone
- This generates a series of peptide fragments, which can be separated by various techniques (e.g., RP-HPLC, MS/MS)
- A standardized sample of the product generates a characteristic peptide map
- Peptide map fingerprinting can be used to detect point mutations in a product's gene that could lead to an altered primary structure (i.e., amino acid sequence)

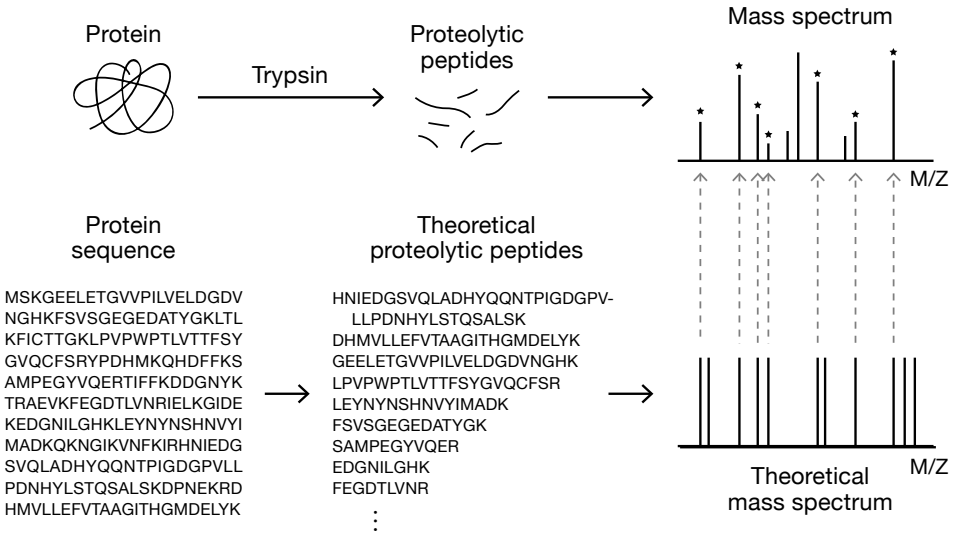

Figure 4.2. Determining the identity of a protein with a known sequence using mass spectrometry after tryptic digestion.

MQAFVGTLTK

B Series	Y Series
MQAFVGTLT	K
MQAFVGTL	TK
MQAFVGT	LTK
MQAFVG	TLTK
MQAFV	GTLTK
MQAF	VGTLTK
MQA	FVGTLTK
MQ	AFVGTLTK
M	QAFVGTLTK

Figure 4.3. Fragmentation of tryptic peptide in MS/MS to generate a contiguous series of b- and y-ions for *de novo* sequence determination.

corresponding tryptic peptide as compared to a reference standard. Further fragmentation using tandem mass spectrometry (MS/MS) further fragments each peptide into a series of peptide fragment ions. Since each fragmentation event breaks a peptide bond and generates two fragments, one on the N-side (y-ion) and one on the C-side (b-ion), these y- and b-ions bear m/z values that are characteristic of the fragmentation pattern of the corresponding peptide. From this pattern, the amino acid sequence can be deduced (Figure 4.3). Mass spectrometry can therefore be used for *de novo* sequencing in which the identity of the protein is unknown.

Mass spectrometry is also used in sequence verification. By searching protein databases to match fragmentation patterns, mass spectrometry-based methods can be readily used to find the identity of a protein or even to identity proteins in a mixture. It is also a sensitive method for protein identity verification. If an amino acid in a peptide is altered or chemically modified in a small fraction of protein molecules, the fragmentation pattern in the mass spectrum will be different from the original. Mass spectrometry can also be used to detect contaminating protein in a purified protein sample. It is thus an important tool in protein identification as well as in chemical structural variants detection.

Higher-order structure identity

The biological function and activity of a protein is dictated by its higher-order (secondary, tertiary, and quaternary) structure. Even if its amino acid sequence is unchanged, a protein molecule may have an altered secondary, tertiary, or quaternary structure. Conversely, the effects of an amino acid change on the functional activity of the protein is dependent on the site of the amino acid change and its effect on the higher order of protein structure. Changes in higher-order structure may result in different biological activity or even immunogenicity (Panel 4.3). In the comparative study of biosimilar protein therapeutics, it is important to demonstrate that the higher-order structure of the biosimilar is the same as the innovator drug. Traditional analytical tools for protein higher-order structure determination include X-ray diffraction and NMR. These methods are of limited use for routine characterization because of the amount of materials needed and the lengthy time required for analysis.

Circular dichroism in the far-UV region is used to reveal changes in protein secondary structure. Due to the chiral nature of proteins, they have different absorption of left-handed and right-handed circularly polarized light. The absorption spectra of a protein in circular dichroism spectroscopy is affected not only by its α-helix and β-sheet content, but also its three-dimensional structure. This type of investigation is thus a powerful tool in revealing secondary structural changes of a protein upon exposure to an agent of change.

Panel 4.3. Higher-Order Structures

- Influence functional activities
- Biosimilars are required to have same higher-order structure as innovator drug
- Typical methods of characterization
  - X-ray diffraction
  - NMR
  - Circular dichroism spectroscopy

Impurities

Therapeutic proteins produced in cell culture undergo extensive purification, virus inactivation, and viral particle clearance before becoming a drug substance for filling. However, impurities inevitably exist in the purified protein (Panel 4.4). The impurities can be from medium components or from molecules released from host cells or lysed host cells. The impurities may be compounds that are co-purified with the product protein or that are not sufficiently discriminated from the product in various product purification operations. They may also arise because of their physical association with the protein. Characterizing and quantifying these impurities is important in assuring the safety of the product. Impurities

Panel 4.4. Impurities

- Host cell protein (HCP)
  - Poses immunogenic concerns
  - May include enzymes that can cause changes in formulated drug products over time (even if the host cells are of human origin)
- Other co-purified media components, e.g., methotrexate (MTX)
- Host cell DNA
- Virus particles
- Endotoxins

that are commonly checked include host cell protein, host cell DNA, and endotoxins.

Host cell protein

The vast majority of therapeutic proteins are produced in non-human mammalian cells. The amount of native proteins secreted by host cells into the culture supernatant is small. However, with the high cell density and long duration of fed-batch cultures, these proteins may reach levels that affect final product purity. As the cell viability decreases in the late stage of culture, cell lysis inevitably leads to the release of host cell proteins. The host cell proteins co-purified with the product protein during recovery pose a risk of eliciting an immunoresponse in patients (Panel 4.4). Additionally, the contaminating host proteins may include some that have enzymatic activities or are prone to spontaneous chemical modification that, when present in the final drug product over a long period of time, may cause a change in product quality. The concern of host cell protein (HCP) contamination is magnified when the dose of the therapeutic protein is high, because the tolerance level of HCP is based on a dose, not the per-unit weight of the product. The detection and quantification of host cell protein contamination is thus required by regulatory agencies. A common analytical method relies on immunoassays, like ELISA, using the antiserum developed against the lysate of the host cell culture. The host cell lysate is used to establish a calibration curve for quantitation.

Structural Homogeneity of the Product

Protein products made in cell culture may not be entirely uniform with respect to their structure (Panel 4.5). Glycans on different molecules of the same protein, or on different glycosylation sites of the same molecule, are somewhat different. In contrast, inhomogeneity may arise from the presence of a small fraction of molecules with variant structures, though the vast majority of the protein molecules have a "normal" structure, including primary, secondary, tertiary, and quaternary structures. For example, even though the fidelity of protein translation is very high, misincorporation of amino acids into the primary sequence occurs at low frequencies.[2] Under some culture conditions, such as shortage of a particular amino acid, the frequency of amino acid misincorporation may increase, leading to detectable quantities of variant proteins.[3] An enzymatic or post-translational modification reaction, such as signal peptide cleavage or C-terminal lysine cleavage, may not be completed, leading to a subpopulation of the product with variant structures.

In addition to the variants caused by the cells' protein processing variability, structural variants may also arise non-enzymatically after proteins

are secreted into the culture medium, or occur in the course of downstream product recovery. For example, disulfide bonds in a protein may be reduced, further oxidized, scrambled, or cross-linked between different molecules.[4]

Many of the protein variants seen in cell culture processes are also seen in the human body. Oftentimes, structural heterogeneity or variability of proteins occur naturally. Nevertheless, the extent of their variability in the final product must be controlled within a specified range.

Panel 4.5. Protein Structural Heterogeneity

- Glycan heterogeneity occurs naturally through biosynthesis
- Protein structural variants arise from:
  - Errors in translation (amino acid misincorporation)
  - Mutations in a copy of the product gene in the host cell
  - Cellular biochemical reactions
  - Chemical or enzymatic modifications after protein secretion in the bioreactor
  - Chemical modifications in downstream processing
  - Structural changes in formulation and storage

Glycosylation Profile

As discussed above, the heterogeneity of glycan structures on proteins bears an important difference with other structural variations of proteins. Intrinsically, all product protein molecules should have the same primary, secondary, tertiary, and quaternary structure. Deviations or variants are anomalies and must be controlled to fall below certain levels. In contrast, heterogeneity in glycan structure is a natural phenomenon that occurs even for glycoproteins synthesized in the human body. When the glycans on a product protein are stripped off the protein and structurally analyzed, different structural classes of glycans are found to have different abundance levels for that protein, giving rise to a glycan distribution profile (Panel 4.6). Such distributions on a protein may vary with respect to the tissue in which it is expressed, the patient's health and age, etc.

Panel 4.6. Glycan Heterogeneity

- Glycan structure on different molecules is heterogeneous
- The glycan profile (i.e., the structural distribution pattern) is affected by:
  - The expression of the enzymes involved in glycosylation
  - The cells' metabolic state
  - Cultural conditions

For glycoproteins produced in cell culture, the glycan profile may vary with the host cell line or even the particular cell clone used, the cultural conditions, and the cells' metabolic state. Glycan structure may also be modified through degradation. For example, the presence of sialidases and other glycosidases, likely released from dead cells, causes the trimming of sialic acid or other sugar moieties from the glycan. The final glycan profile of the product is further affected by the isolation and purification process, as the selectivity of proteins with different types of glycans may differ across separation unit operations. Controlling glycan profiles within a prescribed bound throughout the manufacturing process is important for product release.

Glycan assays

The presence of glycans of different compositions and sizes on protein molecules cause small differences in their motility through gels in electrophoresis (Panel 4.7). Since sialic acid carries a negative charge, glycoproteins with differing sialic acid levels can be separated in isoelectric focusing gel, cation exchange chromatography, or capillary electrophoresis. These methods can be used in characterizing glycans, especially when the glycan structure is not expected to exert a major effect on its clinical efficacy, and the characterization primarily serves the purpose of assuring consistency of product quality.

Panel 4.7. Glycan Assays

- The sialic acid content of glycans affects the pI of the protein and can be assayed by isoelectric focusing or by anion exchange chromatography
- Binding assays using different lectins generate different binding patterns depending on the glycan structure
- N-glycans can be enzymatically stripped off the protein and analyzed in HPLC after being fluorescently labeled with 2AB

The glycan profile can also be determined by utilizing the varying binding affinities of different glycans to different lectins. The extent of a glycosylated protein binding to a lectin can be detected by biotin or digoxigenin-labeled lectins. Using a combination of lectins, often made into array format, one can gather information about a glycoprotein's carbohydrate profile. However, its resolution power is limited and it does not give complete structural information.

To structurally characterize the glycosylation profile on a protein, glycans are detached from the protein and subjected to chromatographic separation. For N-glycans, the cleavage of the glycan is achieved through a treatment with enzyme peptide-N-glycosidase F (PNGase F). The free glycans are then fluorescently labeled with 2-aminobenzamide (2AB) at the reducing terminal N-acetylglucosamine (GlcNAc). Following a cleanup step to remove protein and excess labeling reagent, the fluorescently labeled glycans are separated by liquid chromatography such as hydrophilic interaction chromatography (HIC). The relative % peak areas of the glycans are calculated and provide relative abundance of the various glycan species. However, there is no equivalent enzymatic treatment that releases O-glycans from serine or threonine. The release of intact glycan remains a major challenge in structural analysis of O-glycans. Chemical treatment can release these glycans, but it also alters the sugar structure, making it difficult to label and increase the sensitivity of detection. As a result, O-glycan analysis has lagged behind that of N-glycans. Structurally, O-glycans on a protein can be much more complex than N-glycans. The degree of complexity of O-glycans is dictated by the glycosylation gene expressed in the cell or tissue. Many of the O-glycosylated proteins are on the cell surface. Some expressed proteins in CHO cells are O-glycosylated, but, despite the large repertoire of enzymes involved in O-glycosylation, the number expressed in CHO cells is relatively small and the corresponding O-glycans are less structurally complex.

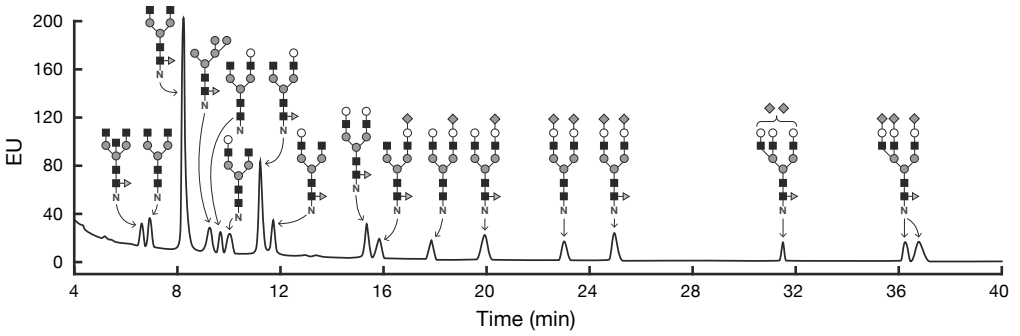

Figure 4.4. A chromatogram of eluted N-glycans.

When marked with a fluorescent label, the relative area/peak height of the glycan gives the relative abundance level of different glycans in the overall protein (Figure 4.4). However, when multiple glycosylation sites are present in a protein, it does not provide information on the distribution among the sites. It is possible to use protease to digest the protein to peptides, isolate and purify the glycosylated peptide, and then use mass spectrometry to determine the *m*/*z* ratio and determine the glycan distribution for each glycosylation site. However, if there are more than two sites the analysis can be rather complex. And even then, such analysis does not provide information on which glycans on one site are co-expressed with other glycans at another site of the same protein molecule. Such linked-glycan distribution on protein molecules will be readily available if a single-molecule-based analytical method is available.

Protein Structural Variants

Some protein variants bear alterations of their amino acid sequence. These alterations in the primary sequence may occur biologically through rare errors in protein or transcript processing or through chemical modification (Panel 4.8).[1] Some events, mostly occurring non-biologically, alter the higher-order protein structure, such as disulfide bond scrambling or the formation of aggregates.

Often, variants of a protein product are grouped based on routinely used protein characterization methods. Size (or aggregate) variants are those which appear as a large molecular weight peak based on size-exclusion chromatography. Charge variants are the molecules which are eluted from ion exchange chromatography and appear as separate and smaller peaks, or which form a separate band

Panel 4.8. Protein Structural Variants

Primary Sequence Variants

- Errors in protein expression
  - Amino acid misincorporation
  - Uncleaved leader sequence
  - C-terminal or N-terminal truncation
  - N-terminal pyroglutamate
- Abnormalities in biological events
  - Uncleaved C-terminal lysine
- Chemical modifications
  - Lysine succination
  - Deamidation of asparagine
  - Lysine glycation
  - Methionine oxidation
  - Aspartate isomerization
- Higher-order structural variants
  - Disulfide bond management
  - Aggregation

in isoelectric focusing.[5] As discussed above, the glycosylation pattern of a glycoprotein is normally heterogeneous but may be variable during the manufacturing process. Some consider glycan profile deviation from the norm or the prescribed range to be a structural variation. On the whole, structural variants are present as a relatively small fraction of the protein products. These variants may arise intracellularly before the protein molecule is secreted, or may be modified chemically or biochemically after secretion. The impact of these variants on the immunogenicity, efficacy, and potency of the protein has been described in literature.

Primary sequence variants

Primary sequence variants of a product protein may arise from a mutation in the DNA sequence which encodes the product gene in the producing cell. The mutation rate in the protein coding sequence during somatic cell replication is rather low. If any mutation is present in the product gene, it is likely to have originated from a single cell and occur in only one of the multiple copies of sequences that encode the product. A test to ensure the absence of such anomalies in the producing cell is typically carried out before the cell bank of a cell line is established.

Mistakes in the primary sequence are unlikely to be caused by transcription errors, as the frequency of such errors is low. Unspliced RNA species are seen in RNAseq at low frequencies, and the probability of a similar occurrence in the product transcript is therefore also very low. Such a transcript, if translated, will most likely generate a misfolded protein and be degraded. A product with a primary sequence variant is more likely to arise from errors in translation. If an amino acid is deficient in the medium, a mismatched amino acid may be used to charge the tRNA, leading the incorrectly charged amino acyl-tRNA to carry the wrong amino acid into the elongating protein molecule.

Other amino acid sequence variants may occur post-translationally. At a low frequency, the leader sequence that directs the nascent protein to translocate into the endoplasmic reticulum may not be cleaved, producing a variant. Many proteins have lysine or arginine in their C-terminus. The basic amino acids in the C-terminus are liable to cleavage by carboxypeptidase. The cleavage of lysine at the C-terminus of the heavy chain of many IgG molecules by carboxypeptidase is often not complete, leaving some molecules with uncleaved C-terminus lysine.

Chemically modified structural variants

Many structural variants arise through chemical modifications of amino acids without the involvement of cellular enzymes. These chemical modifications occur mostly after protein molecules are secreted into the extracellular environment. However, an amino group in lysine or a N-terminus amino acid may react with succinic acid or fumaric acid in

the intracellular environment, although this does not appear to occur at high frequency.

Deamination, or the removal of an amide group, may occur in asparagine to form a succinimide or be further hydrolyzed to aspartate or iso-aspartate (Figure 4.5a). C-terminus glutamine can be cyclized, losing its amide group and forming pyroglutamate (Figure 4.5b). Glycation, the condensation of glucose or other reducing sugars to the ε-amino group of a lysine or the amino group of an N-terminus amino acid, occurs spontaneously in culture fluid (Figure 4.5c). Methionine residue may be oxidized by reacting with reactive oxygen species to form methionine sulfoxide (Figure 4.5d).

Higher-order structural variants

In additional to various chemical modifications that alter the amino acids in the primary structure, the disulfide bonds in a protein may be structurally modified in different ways. Disulfide bonds have a relatively low dissociation energy. Some disulfide bonds are located in the structurally flexible regions of a protein and are susceptible to chemical modifications that may lead to cleavage, trisulfide formation, etc. Disulfide bonds may also scramble their arrangement, often causing major conformational changes in a protein molecule and inducing aggregation. Protein aggregation may also arise from other mechanisms. For example, β-sheet secondary structures in a protein exert intra- or inter-molecular interactions that contribute to the tertiary and quaternary structure. Misalignment of intermolecular interactions may occur in the cell culture environment, especially in the sometimes extreme cultural conditions of late stage fed-batch culture. Some downstream processing conditions, such as low pH, hasten aggregate formation. Chemical modifications of disulfide bonds can be detected by peptide fingerprinting using mass spectrometry. Aggregation or protein size variants are discernible by size exclusion chromatography. This kind of chromatography separates proteins based on differences in their hydrodynamic volumes. Molecules with larger hydrodynamic protein volumes elute earlier than molecules with smaller volumes. The protein sample is loaded onto a size exclusion column, separated isocratically and the eluent is monitored by UV absorbance. Calculating the percentage of each separated component as compared to the total integrated area provides the relative abundance.

Charge Variants

The presence and abundance levels of various structural variants can be characterized by peptide fingerprinting in mass spectrometry or other instrumental analysis. However, for routine assessment, they are often assessed by more readily accessible assay methods, such as size exclusion chromatography and ion exchange chromatography.

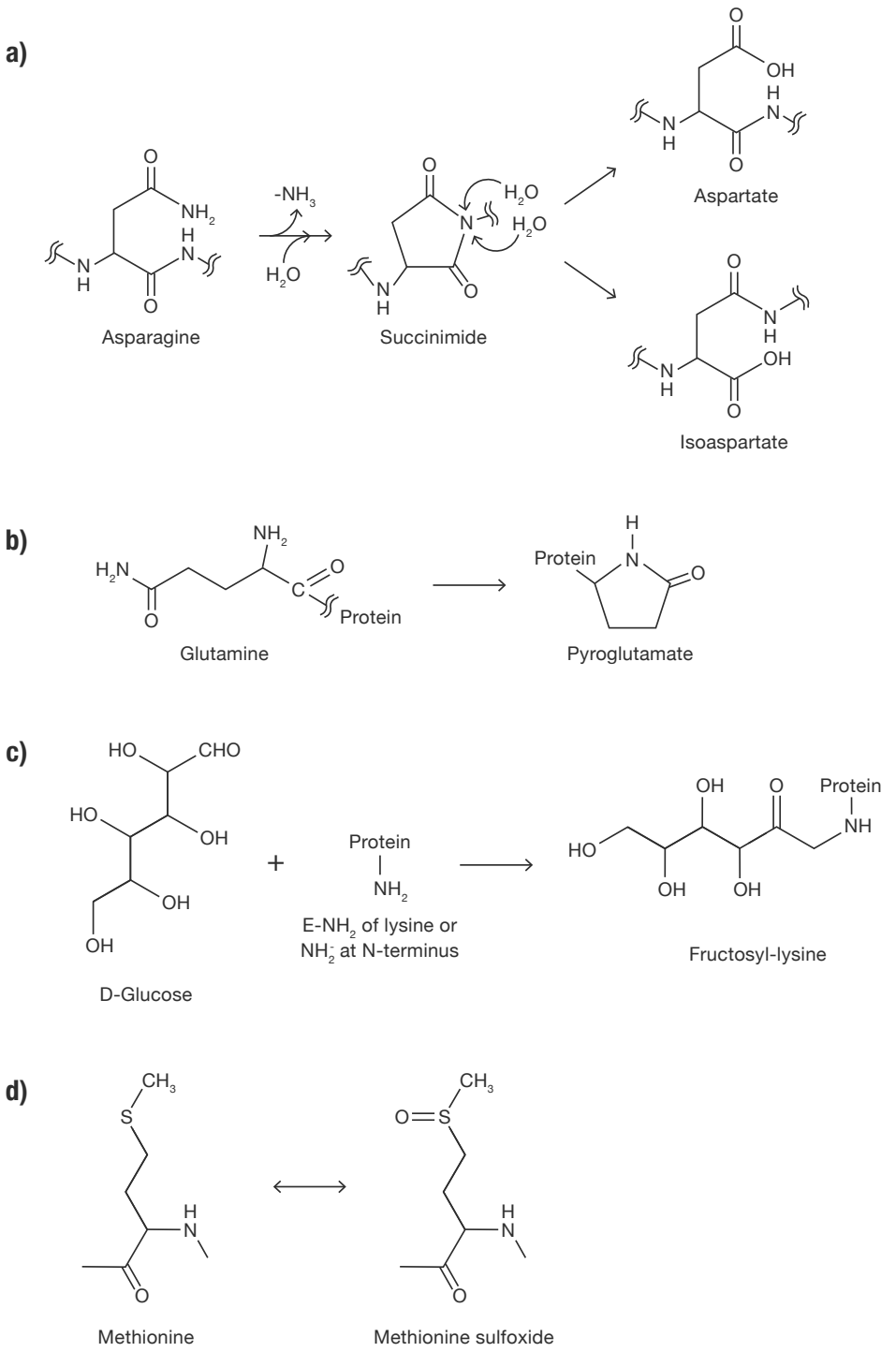

Figure 4.5. Chemical modifications that generate protein structural variants. (a) Deamidation of asparagine. (b) Pyroglutamate formation from glutamine. (c) Glycation of lysine. (d) Methionine oxidation.

Proteins carry many charged side chains in their amino acids. At a biological pH, most proteins are weakly charged. Changes in the net charge of a protein molecule may cause changes in its stability and activity (Panel 4.9). A charged variant may have an altered binding affinity to its target receptor or antigen.[6] It is thus important that the basic or acidic charge variants of a product protein be characterized and controlled within a specified range. Charge variants can be divided into acidic and basic charge variants. The former is eluted ahead of the main product peak in cation exchange chromatography (CEX), or afterwards in anion exchange chromatography (AEX), meaning that they have higher levels of a negatively charged acidic group (Figure 4.6). Basic charge variants are those which are eluted after the main product group in CEX or before in AEX.

Operationally, charge variants are defined by their elution profile in AEX or CEX. Typically, at a set pH, charged variants are separated on a cation or anion-exchange column and eluted using a salt gradient. The eluent is monitored by UV absorbance. The charged variant distribution is evaluated by determining the peak area of each variant as a percent of the total peak area. It is important to note that the elution profile of a protein in AEX and CEX is determined by a molecule's binding affinity in chromatography. Binding affinity is affected not only by the charge of a molecule but also by protein conformation and many other factors. This is in contrast to isoelectric focusing (IEF). IEF probes a protein's isoelectric point (pI), or the pH at which the protein has no net charge. A variant with a different isoelectric point forms a distinct band that differs from the standard product protein in the IEF gel. Some changes that do not affect the pI of a protein, such as the oxidation of methionine and the rearrangement of disulfide bonds, may change the binding affinity of the protein and cause it to appear as a separate minor peak in an ion exchange chromatogram. Different structural variants that appear as acidic and basic charge variants are listed in Panel 4.10.

Panel 4.9. Protein Charge Variants

- Proteins consist of many different amino acids comprising weak acidic and basic groups
- Identification and monitoring of charge variants throughout manufacturing is critical for the production of safe and effective drugs
- Charge variants impact the stability and activity of the protein, which may also cause immunogenic responses

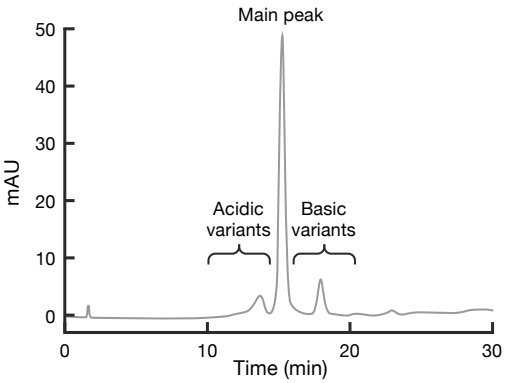

Figure 4.6. Elution of protein charge variants in cation exchange chromatography.

Panel 4.10. Effect of Structural Variation on Charge Variants

Structural Variants Contribute to:

- Increase of acidic species
  - Deamidation of asparagine
  - Leader sequence removal
  - Glycation of amino group in the N-terminus or ε-NH_2 of lysine
  - Glycosylation, sialic acid content
  - Amino acid misincorporation/mutation
- Increase of basic species
  - C-terminal lysine/arginine uncleavage
  - Succinic acid modification of NH_2 group
  - Amino acid misincorporation/mutation
  - In some cases, high molecular weight species
  - Some glycan attributes, such as high mannose
- No effect on charge variants
  - Methionine oxidation
  - N-terminal pyroglutamine
  - Aggregation
  - Disulfide bond rearrangement

Biological Activities

Functional Assay

Structural characterization as discussed above evaluates the quantity of various anomalies in the product. Some of the changes pose safety concerns or may affect the efficacy of the product. For example, disulfide bond scrambling is likely to elicit an immunogenic response. Some other changes, such as glycation and uncleaved C-terminus lysine, are seen in different proteins and at various levels in the blood circulation, and may not pose a major safety concern. In product characterizations, different structural variations are assayed separately and reported as a percentage of the whole product. When different alterations occur concurrently in a molecule, the effect on safety or efficacy may be different from when those changes occur in different molecules. However, the current methods of protein characterization do not reveal the combination of different structural variants and their distribution profile. There are large gaps in completely relating structural changes to functional changes. Hence, full characterization of a protein product must include the measurement of functional activities in a physiologically relevant manner.

A protein product may function through its catalytic activities, such as glucocerebrosidase's main function is to cleave the specific glycosidic bond. A product protein may also function through binding to its ligand in the body fluid, binding to receptors on the cell surface that trigger a downstream signaling response, or by blocking the competitive binding by its biological counterpart. A functional assay may involve quantifying the enzymatic activity of the product, measuring the binding strength of the product to the ligand, or using effector cells for a bioassay. In all cases, the appropriate assay to be used is chosen based on the understanding of the mechanism of action (MOA) of the molecule.

Binding assay

Since mAbs are the largest class of therapeutic proteins, binding assays between an antigen and an antibody are a commonly practiced functional assay. The binding kinetics are measured by surface plasmon resonance (SPR) (Figure 4.7). First, the ligand proteins are immobilized

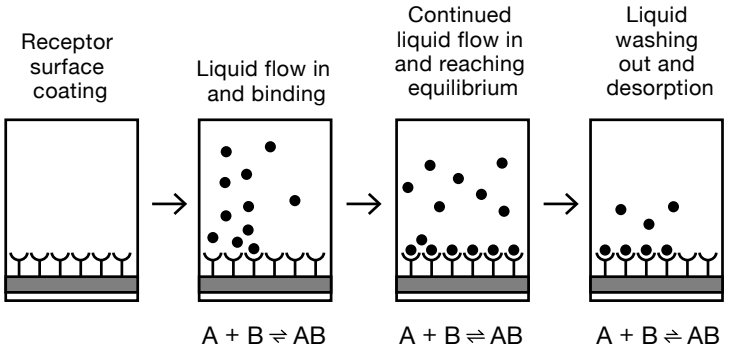

Figure 4.7. Adsorption and desorption kinetics of ligands in surface plasma resonance (SPR).

onto the sensor chip. The antibody protein is then flowed into the ligand-coated binding reaction chamber to allow for the measurement of antibody association to the ligand. This is followed by a period of dissociation in which a solution without the antibody is flowed through to measure the kinetics of dissociation. The process is repeated after regeneration of the surface immobilization of the ligand. The surface resonance increases as antibodies and ligands form complexes in the association phase and decreases in the dissociation phase. Using the binding kinetic equations, the rate constant of binding and dissociation can be determined.

Bioassays

Bioassays that quantify the biological activity of the protein using cultured cells are the most relevant evaluation of the product short of using animals. It enables the effect of changes in the protein on its biological activities to be quantitatively evaluated. But an effective bioassay is also difficult and time-consuming to develop. It relies on a good understanding of the mechanism of actions of the drug. An example is shown in Figure 4.8a and b that depicts the assay of antibody-dependent cellular cytotoxicity (ADCC) and complement-mediated cytotoxicity (CMC). In ADCC, an assay of two types of cells, effector cell and target cell, are used. The target cell expresses the antigen that is recognized by the antibody. The binding of the antibody to the target cell recruits the killer cell, which exerts cytotoxic action on the target cells. In CMC, the antibody binding to the antigen on the surface of the target cell recruits complements. The binding of the complement to the target cell then elicits complement-mediated cell lysis. With an effective bioassay, one can evaluate the activity at different doses of the antibody (Figure 4.8c). Consistent kinetic behavior in a bioassay assures the quality of the product. If a product batch is shown to have an altered behavior as shown in the response curve of a bioassay, further investigation is called for.

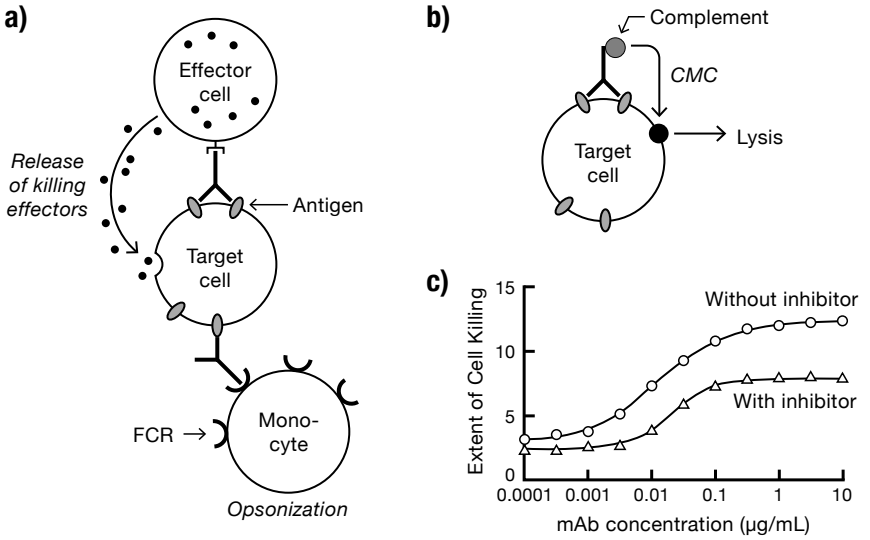

Figure 4.8. Bioassay of antibody-mediated target cell killing. (a) Antibody-dependent cellular cytotoxicity assay. (b) Complement-mediated cytotoxicity assay. (c) Killing results at different doses with and without an inhibitor that binds to target cells.

A common requirement in binding assays and bioassays is the availability of reference products and assay materials. For binding assays, the reference materials are the product and ligand molecules. For bioassays, the assay cells must be standardized, banked, and made readily available. Biological assays are often subject to a higher degree of variability. These additional challenges make bioassays less frequently used than chemical and structural characterizations.

Quality by Design (QbD) of Cell Culture Products

An important objective of process and product development is generating an understanding of process and product quality, and defining a strategy to monitor and control product quality during manufacturing. The variability of protein structures underlines the importance of identifying the properties whose variability affects the product's safety and efficacy, and the need to control those properties within a defined range. In the early 2000s, the FDA published a series of documents on cGMP in the twenty-first century and began a drive to put QbD into drug development and manufacturing (Figure 4.9).[7] Many of these documents, including ICH Q8 (R2) (Pharmaceutical Development), ICH Q9 (Quality Risk Management), ICH Q10 (Pharmaceutical Quality System), and

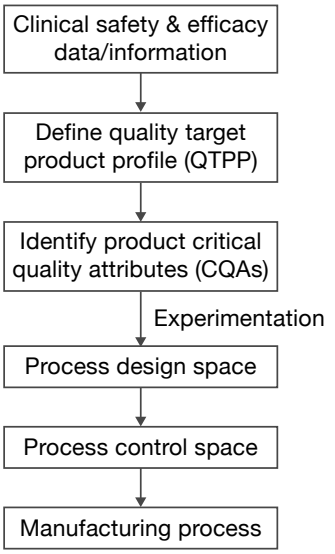

Figure 4.9. Incorporation of QbD in process development and manufacturing.

Panel 4.11. Elements of Pharmaceutical Development

1. **Quality Target Product Profile (QTPP)**
   A prospective summary of the ideal quality characteristics of a drug product, taking into account the safety and efficacy of the product
2. **Critical Quality Attribute (CQA)**
   A physical, chemical, biological, or microbiological property or characteristic that should be within an appropriate limit, range, or distribution to ensure the desired product quality (safety, identity, purity, biological activity, and stability). Ideally, this considers the mechanism of action of the protein.
3. **Design Space**
   The multidimensional combination and interaction of input variables and process parameters that have been demonstrated to provide assurance of quality
4. **Control Strategy**
   A planned set of controls, derived from current product and process understandings, that ensures process performance and product quality

ICH Q11 (Development and Manufacture of Drug Substance), provide high-level direction with respect to the scope and definition of QbD as it applies to the biologics industry. These documents also introduced a number of items that are listed in Panel 4.11.

QbD is an approach to process development wherein the manufacturing process and controls are proactively designed to deliver the intended product. QbD elements include the following: a Quality Target Product Profile (QTPP) that identifies the Critical Quality Attributes (CQAs) of the drug product; process design and understanding, including identification of Critical Process Parameters (CPPs), linking CPPs to CQAs; a control strategy that includes specifications for the drug substance and drug product as well as controls for each step of the manufacturing process; and process capability and continual improvement. QbD tools include prior knowledge, risk assessment, mechanistic models, design of experiments (DoE), data analysis, and process analytical technology.

The implementation of a QbD-based pharmaceutical development starts with defining the characteristics of the drug product that are required to achieve safety and clinical efficacy. This set of product characteristics is referred to as the Quality Target Product Profile (QTPP). Next, the physical, chemical, biological, or microbiological features of the drug

product and drug substance that must be controlled within a bound in order to ensure the QTPP is met are identified. These features, called Critical Quality Attributes (CQAs), are variables that can be measured and quantified. How these CQAs are affected by process conditions in different operations of manufacturing is established from knowledge of the process or is determined experimentally. The information is used to define a design space of process variables within which CQAs of the product can be expected to fall into the permissible range. A control space is then established within which the process will be operated, thus ensuring that the CQAs of the resulting product meet the specifications. With a QbD-based approach, risk analysis is employed to assess various process aberrations that can cause CQAs to stray out of bounds, and mitigating measures are taken to reduce the risk. The implementation of QbD starts from the beginning of process development and continues through characterization, establishment, and commercial manufacturing. It is continually practiced throughout the product's life cycle so as to continually improve the process and enhance product quality.

The first task in implementing QbD is the development of knowledge of the relationship between process parameters (e.g., raw materials, cell inoculum), culture conditions, and product quality. Such relationships are inevitably complex, with multiple interactive input variables and process parameters that affect both productivity and product quality. Critical to understanding the relationship between process parameters and product quality are analytical tools which quantify various quality variables. Advances in analytical techniques have enabled sophisticated evaluation of physicochemical properties, biological activity, immunochemical properties, purity, and impurities.

Regulatory documentations (e.g., ICH Q6B) stipulate test procedures and acceptance criteria for biotechnological/biological products. Shown in Table 4.2 are some quality attributes with the category of quality as defined by regulatory agencies, the analytical methods often used to characterize the clinical product, and the material generated during process development studies.

QTPP and CQAs

QbD-based product and process development first considers the desired or target quality characteristics of the final drug product (QTPP) according to how it will be administered to a patient and what properties it must have to be safe and effective. An example of a QTPP for a recombinant antibody drug product is shown Table 4.3. The QTPP will be used to determine the CQAs of the product. For each characteristic, a qualitative target description or a quantitative measure is given. At

Table 4.2. Protein Structural Variants and Categories of Quality Attributes Affected

Product quality attribute	*Method(s) of characterization*	*Attribute category (ICH Q6B)*
Aggregation	SEC-HPLC	Identity, purity, stability
	Gel electrophoresis	Identity, purity, stability
	Analytical ultracentrifugation	Identity, purity
	SEC-HPLC with MALS	Identity, purity
C-terminal lysine	Ion exchange chromatography	Identity, purity, stability
	Isoelectric focusing	Identity, purity, stability
	Peptide mapping with MS	Identity, purity, stability
Deamidated isoforms	Ion exchange chromatography	Identity, purity, stability
	Isoelectric focusing	Identity, purity, stability
	Peptide mapping with MS	Identity, purity, stability
Glycosylation	Monosaccharide composition analysis	Identity
	Oligosaccharide profile	Identity
	Sialic acid content	Identity
	Galactose content	Identity
	Fucose content	Identity
Oxidation	Peptide mapping with MS	Identity, purity, impurity

Table 4.3. QTPP Listed in A-Mab Example from ICH Q9

Product attribute	*Target*
Dosage form	Liquid, single use
Protein content per vial	500 mg
Dose	10 mg/kg
Concentration	25 mg/mL
Mode of administration	IV, diluted with isotonic saline or dextrose
Viscosity	Acceptable for manufacturing, storage and delivery without the use of special devices (for example, less than 10 cP at room temperature)
Container	20R type 1 borosilicate glass vials, fluoro-resin laminated stopper
Shelf life	≥ 2 years at 2–8 °C
Compatibility with manufacturing processes	Minimum 14 days at 25 °C and subsequent 2 years at 2–8 °C, soluble at higher concentrations during UF/DF
Biocompatibility	Acceptable toleration on infusion
Degradants and impurities	Below safety threshold, or qualified
Pharmacopoeial compliance	Meets pharmacopoeial requirements for parenteral dosage forms, colorless to slightly yellow, practically free of visible particles and meets USP criteria for sub-visible particles

first glance, the list does not seem to present much information on the characteristics of the drug substance that are related to safety and efficacy. A closer examination will reveal the quality requirements of the product. The drug product must be stable as a liquid under refrigerated conditions at a concentration of 25 mg/mL for at least two years. The drug substance must also be stable for 14 days at room temperature and at higher concentrations during processing. A long stability is required since the degradants and impurities are below the safety threshold at the end of the shelf life. These product specifications point to the importance of protein stability. Hence the propensity of the drug product to aggregate, and the level of charge variants will likely impact the QTPP. The dose and concentration to be administered intravenously is high. This suggests that the level of host cell proteins and other impurities will likely impact the safety of the product.

Table 4.4 is a partial list of the quality attributes of a recombinant antibody product. Some attributes, such as pH and excipient concentration, are related to the composition and strength of the drug product. A number of attributes related to impurities or structural variations must be controlled at low levels to ensure the degradants and impurities aspects of the QTPP are met.

The list of quality attributes derived from the QTPP can be rather long, as seen in the example shown in Table 4.4. However, not all the attributes identified are critical and must be controlled within a range to minimize the risk of negatively affecting clinical outcomes (Panels 4.12 and 4.13). The task of identifying the CQAs among all the attributes typically takes place during product and process development, and before the manufacturing process is established and process and clinical data is available. The selection relies heavily on the knowledge of past products and processes, as well as the understanding of the mechanism

Table 4.4. Examples of Product Quality Attributes

Product variants	*Process-related impurities*	*Adventitious agents*	*Composition and strength*
• Aggregation • Conformation • C-terminal lysine • Deamidated isoforms • Disulfide bonds • Fragmentation • Glycation • Glycosylation • Oxidation • Thioether link	• Host cell DNA • Host cell proteins • Leached protein A • Selective agent (e.g., MTX) • Cell culture medium components (e.g., insulin) • Purification buffer components • Leachables	• Viral purity • Microbial purity • Endotoxins	• Product concentration • Potency • Osmolality • pH • Particulates • Clarity • Color • Volume • Excipient concentration

Panel 4.12. Identifying CQAs from QTPP-Derived Attributes

- CQAs are attributes that may affect the clinical outcome and must be controlled within bounds
- CQAs are identified before commercial manufacturing starts
- Identification relies on knowledge of past processes and the MOA of the product
  - Some molecular features of the product may be important for a particular MOA
- May use scoring or semi-quantitative mechanisms to rank CQAs

Panel 4.13. Establishing Quantitative Bounds for CQAs

- Deciding the value of CQAs is initially quite difficult, as process and clinical data is often insufficient
- May use functional assays (binding assay, bioassay)
- May use external conditions or purified variants to enrich variants under test conditions
- Employ various lots during testing
- Important to maintain a comprehensive library of samples generated, especially at larger scales as the process evolves
- Significant changes in attribute levels may occur *in vivo*

of action (MOA) of the product. A literature review on the MOA of the new candidate drug is performed and a knowledge database on experience from similar products is surveyed. For example, if the mechanism of a monoclonal antibody product is to block receptor sites, attributes that enhance antibody-dependent cellular cytotoxicity (ADCC) and complement-dependent cytotoxicity (CDC) may not be important for efficacy. As a result, fucosylation at the Fc region may not be considered a CQA. On the other hand, if the MOA is ADCC, the level of fucosylation of the glycan in the Fc region might be selected as a CQA.

There has been an increasing effort to employ a scoring system to select CQAs. Each member of a multifunctional team scores the attributes. Each attribute is given an impact score based on its perceived importance to the fulfillment of the QTPP and an "uncertainty" score that accounts for the extent of the variability of the attribute. The selection of CQAs is then based on the overall score. As the product advances and experiences from clinical and *in vitro* studies accumulate, knowledge of the degradation pathways and structure–activity relationships of the product are used to refine CQAs. A prudent approach is to adopt a risk assessment procedure, as will be discussed later in this chapter. In general the identification of CQAs and definition of their bounds is an iterative process that is refined as the product matures through commercialization life cycle.

CQA range and validation

After the selection of CQAs, the appropriate value or value range must be specified for each CQA. Deciding the value of CQAs is a difficult task, especially in the early stage of product development when sufficient clinical and manufacturing data is unavailable. In the early stages of process development and QbD establishment, data from *in vitro* binding assays and bioassays are used to evaluate the effects of different ranges of

CQAs on the performance of the product. Since the drug substance production process typically produces various variants within a range, in order to test the conditions out of typical range, extreme stress conditions are used to accelerate the generation of some structural variants, such as degradation products and aggregates. Some variants, including charged variants, can be enriched or purified for testing *in vitro* or in small animals. The data from *in vitro* and *in vivo* studies are augmented with historical data from the established drug product in the initial definition of CQA values. When refining CQAs for manufacturing, it is important to incorporate a large number of product lots produced under different settings so as to better identify CQAs and set release specifications. In addition to development and clinical manufacturing lots, lots with "extreme" variant levels generated under extreme conditions may also be included.

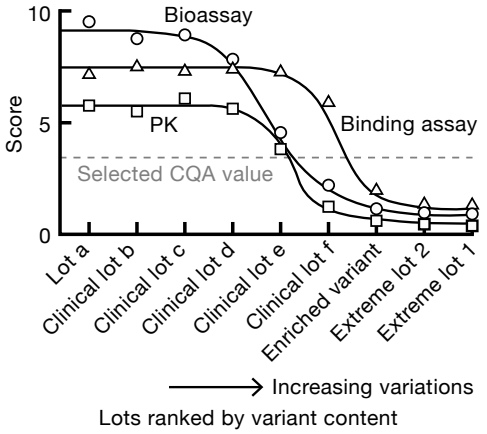

Figure 4.10. A hypothetical instance of using quality scores of different assays to select CQA specifications.

A hypothetical scenario using binding assays, bioassays, and PK to understand the impact of variants on the efficacy of the product is shown in Figure 4.10. In the early stage of product development, various assays can be used to set the ranges of CQAs. As clinical lots for clinical trials become available and clinical studies proceed, the data can be used to support the ranges of CQAs selected for clinical manufacturing lots. It is important to generate a comprehensive library of samples generated, especially at larger scales as the process evolves, so that the final validated bioassay can be used to characterize as many development and clinical lots as possible before final CQAs and specifications are set.

A consideration in evaluating acceptable ranges for product attributes is that significant modification of the product may occur when it is administered to the patient. For example, deamidation of proteins is known to occur in the blood stream. Predicting such attribute changes *in vivo* using a model or by measurement will generate data that supports the selection of a wider range of CQA values.

CQAs in Different Stages of Manufacturing

A manufacturing process of therapeutic proteins can be largely divided into four stages: drug substance manufacturing (consisting of production, downstream recovery, and purification); drug substance handling; drug product manufacturing; and drug product handling (Figure 4.11). Each stage further consists of different unit operations. The bulk of production occurs in the drug substance manufacturing stage. However, it

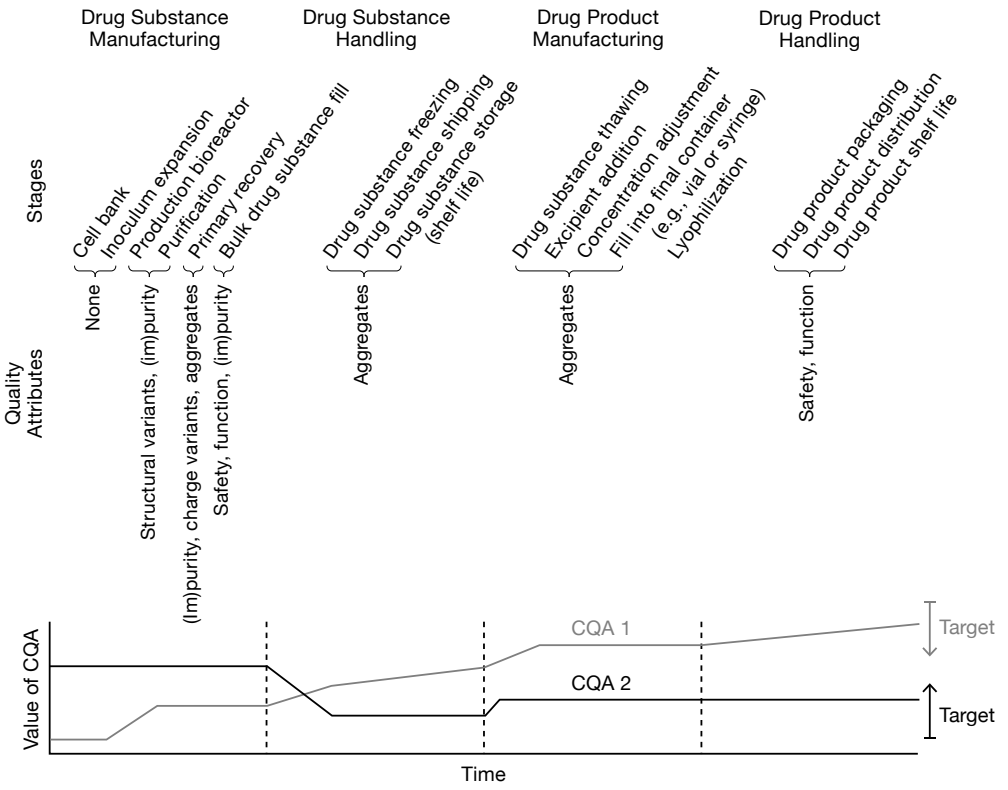

Figure 4.11. Key quality attributes vary across stages of manufacturing.

is the final drug product that directly impacts safety and efficacy in patients, and the final drug product that the QTPP is designed for. The value of the CQAs may not be constant throughout the manufacturing process. As the product stream advances through different stages of manufacturing, the purity and possibly the homogeneity of the product will increase, and the CQA value should approach the final value of the drug product. For example, if protein aggregation is a CQA and is set to below a threshold value, as the product goes through harvest and purification the fraction of aggregates should decrease. A CQA may become significant only after a certain stage of the manufacturing process. For example, a particular type of product variant may arise that is due to exposure to low pH conditions; it thus appears in downstream processing but not in the production bioreactor. A comprehensive understanding of CQAs in different stages of manufacturing facilitates the setting of design targets or specifications around CQAs which the product needs to meet.

Quality Considerations before Process Development

Increasingly, quality assessment is being incorporated into decision-making on various matters related to the production of the drug, sometimes even before process development starts, since some decisions will

affect product quality robustness in future manufacturing (Panel 4.14). For example, sometimes multiple candidate molecules are considered for the same drug, such as in the case of antibodies with somewhat different primary sequences in the complementarity determining region (CDR) that bind to the same antigen. Product quality, stability, and platform fit are considered when designing and selecting lead monoclonal antibodies and other protein therapeutic candidates. For example, the propensity of these molecules to form protein aggregates may differ. Selecting one with a low propensity for aggregation will reduce yield loss in process and possibly impact long-term product stability. Another example is the viscosity of the protein solution, which may impact the delivery for the protein therapeutic.

Panel 4.14. CQAs in Early Development and Manufacturing

- CQAs considered even before process development
  - Selection of candidate molecules (e.g., propensity to aggregate)
  - Selection of cell line (e.g., glycosylation stability)
  - Can also adopt a scoring system in decision-making
- CQA values change in different stages of manufacturing

Producer cell lines used in biomanufacturing have all been clonally selected. Different cell clones are often somewhat variable not only in terms of productivity but also in many properties that affect product quality. Products produced by different clones can have different features that are considered CQAs, such as fucosylation or propensity to aggregation. These quality factors are considered in the selection of clones for manufacturing. Shown in Table 4.5 is an example of a scoring table that assesses factors that affect the stability of the cell line (integration site, copy number, epigenetic stability), productivity (growth rate, titer, specific productivity), and attributes of the product quality (sequence variants, aggregation, structural variants, glycosylation, modifications to the CDR). Different weights are assigned to different characteristics and the overall score is used to select the best clone to ensure "manufacturability" and reduce the risk of CQAs falling out of range.

Quality Risk Management

Quality risk management (QRM) is a systematic process for the assessment, control, communication, and review of risks to the quality of the drug product across the product's life cycle. Although the quantitative risk may be calculated in different ways, the evaluation fundamentally includes an identification and assessment of hazards throughout the totality of product characterization, process development, and clinical experience. To initiate a QRM process when developing a product, a cross-functional team which assembles background information and data on the product and process and identifies available tools and timelines is typically formed (Figure 4.12). Risk assessment is then performed to identify the hazards, or process failure modes, that may impact the

Table 4.5. A Hypothetical Example of Cell Line Screening by Quality Score

Clone	***Stability Attributes***				***Productivity Attributes***					
	Transgene		*Integration site*		*Specific productivity*		*Titer*		*Growth rate*	
	Copy # / integrity	*Score*	*Quality*	*Score*	*pg/cell/ day*	*Score*	*g/L*	*Score*	*h^{-1}*	*Score*
1	1/+	10	+	8	25	7	4.1	6	0.03	8
2	3/++-	7	+--	3	17	5	4.5	7	0.028	8
3	5/+----	5	++---	4	15	4	3.9	6	0.032	8
4	3/+++	8	+--	3	21	6	2.9	5	0.026	5
5	1/-	4	+	2	12	2	2.5	5	0.025	5

Clone (cont.)	***Quality Attributes***						***Total Score***
	Aggregation		*Structural variants*		*Afucosylated glycans*		
		Score		*Score*		*Score*	
1	++	4	++	3	2.1	5	7
2	+	5	-	7	1.5	6	9
3	+++	2	-	7	1.7	6	8
4	-	6	+	4	3.5	2	7
5	-	6	+	4	3.2	2	6

CQAs. The probability that a given failure mode will occur is then evaluated. Finally, risks are evaluated against criteria and categorized, often using a semi-quantitative scoring system or by categorizing them as low, medium, and high risk. Many tools and types of risk assessments can be used in QRM.

Once risks are identified and classified, the team determines if the identified and analyzed risks are acceptable. If not, the likelihood and impact of a hazard are investigated and better understood, and then actions are identified to reduce or eliminate the risk. These actions to reduce the severity and probability of harm often include increasing the detectability of hazards. In some cases, residual low-level risks are deemed acceptable.

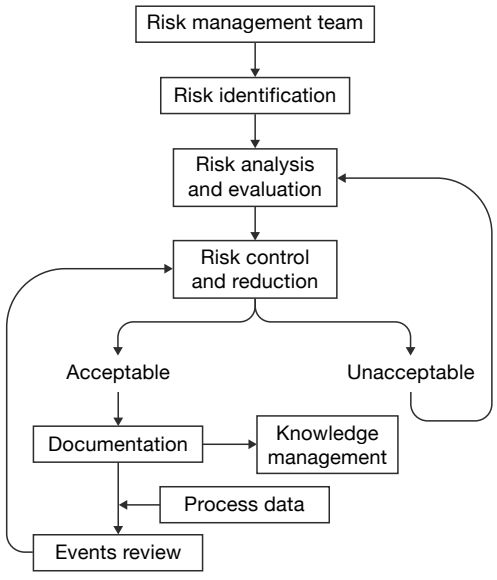

Figure 4.12. A general flow chart of quality risk management.

The output of the QRM process is typically a document that summarizes the risk assessment analysis as well as the risk control actions taken. However, this document is only a snapshot of the knowledge available at the time of its completion. As risks deemed unacceptable are mitigated, another round of risk identification, analysis, and evaluation is completed to ensure evaluation of any new risks introduced or increased significance of existing risks. In addition, risks should be periodically reviewed to ensure new knowledge and experience are considered.

Risk Mitigation

Upon identification of risks, means of mitigation are devised (Table 4.6). One may take a rigorous risk management approach using tools such as Ishikawa (fishbone) diagrams, cause-and-effect analyses, or failure modes and effects analysis (FMEA).

When investigating the relationship between the process variable in question and the outcome of the quality attributes, it is always wise to start with a review of the historical data. For example, the duration of the process may be identified as a factor that can potentially affect the CQA host cell protein content. An investigation of the historical data can be initiated to determine the incidence frequency that the HCP exceeds the set value for the process time periods in question. To facilitate risk and benefit analyses, data on productivity and other process variables are also analyzed. A decision regarding a risk management plan for the identified process variables can then be made. After the investigation, a smaller number of process parameters with a high impact on the CQAs is

Table 4.6. Example of Risk Assessment of Process Parameters

Process Parameter in Production Bioreactor	***Failure Mode***	***Risk Level****							***Risk Mitigation***
		Quality attributes affected				*Product attributes affected*			
		Aggregation	*α-Fucosylation*	*HCP*	*DNA*	*Product yield*	*Viability at harvest*	*Turbidity at harvest*	
Inoculum density	Low density	L	M	M	M	H	M	M	DOE
Inoculum viability	Low viability	M	M	H	H	H	M	M	DOE
Inoculum age	Old inoculum culture	M	M	M	L	H	L	L	Mining knowledge database
Medium lot	Mixed medium lots	L	M	L	L	L	L	L	Not required
⋮	⋮	⋮				⋮			⋮

*H: high, M: moderate, L: low

selected for more detailed experimental characterization, and to generate a model that quantifies their effects, singularly and interactively, on the CQAs. This is performed on the unit operations that contribute to the variability of the CQAs.

Risk assessments are initiated during the product and process development stages, and are used in defining the operating (control) space of the process parameters. The assessment continues and is updated throughout the process characterization and performance verification stages. Even after routine manufacturing, it continues throughout the product life cycle management stages.

Knowledge Management

An important component of QRM is the creation of knowledge from data and the subsequent utilization and dissemination of that knowledge (Figure 4.13). Starting from the beginning, prior knowledge of the product and process from public literature, internal institutional documents,

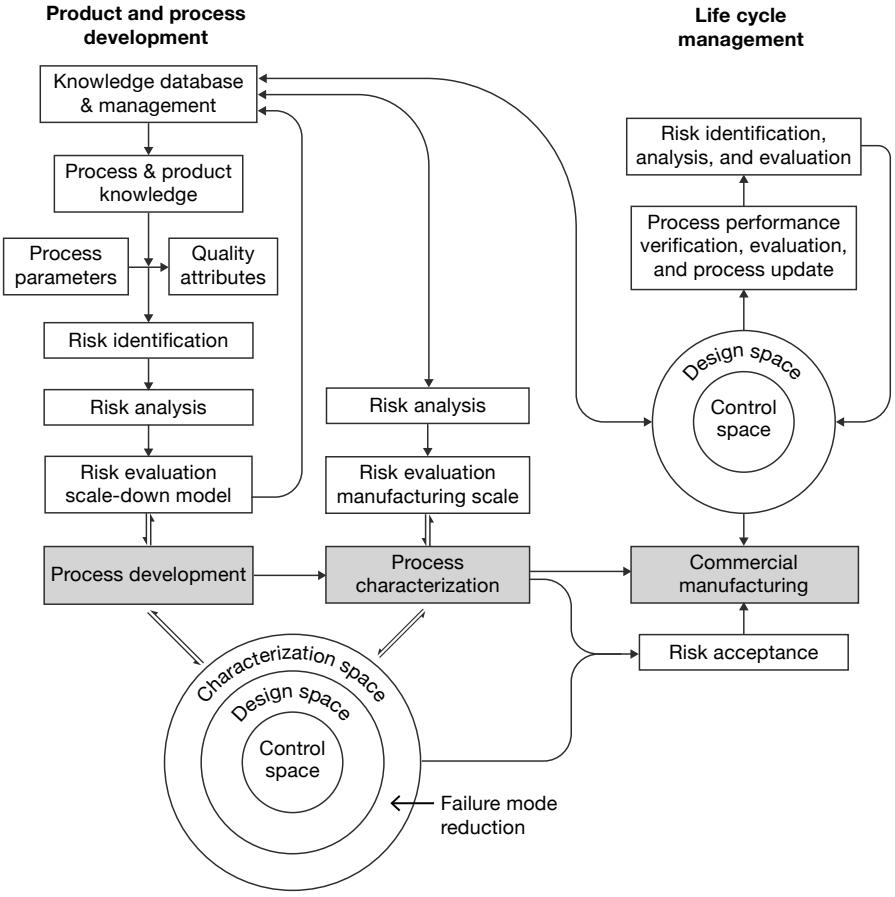

Figure 4.13. Incorporating risk management into process development, manufacturing, and life cycle management.

including prior regulatory submissions, and knowledge databases are captured for QRM. As product development progresses, information on the safety and efficacy of the product is generated from non-clinical and clinical studies, and CQA risk assessments are updated. Process development studies and clinical manufacturing also generate additional knowledge prior to commercialization. Once the product is launched, experience in commercial manufacturing and technology transfers to alternate manufacturing facilities and creates much insight into the relationship between process parameters and CQAs. Compiling, analyzing, storing, and disseminating the knowledge gained from all of the above, a process known as knowledge management, is critical to continued product quality enhancement (Panel 4.15). Many companies utilize spreadsheets, databases, and specialty software to ensure that that knowledge is managed in an efficient way. Some create process libraries that include standardized risk assessments and compilations of data from past process development experiences. Information technology (IT) tools such as Biovia Discoverant® and Tibco Spotfire® can be used as knowledge management tools to aggregate large datasets.

Design Space and Control Strategy

The ultimate objective of QbD is to establish a manufacturing process that generates products which meet the QTPP and CQAs. In the jargon of QbD for pharmaceutical development, this is achieved through defining the process design space. Mapping and defining the design space is an important task in process development (Panel 4.16). However, due to the high attrition rate of protein drug candidates, design space mapping is often not initiated until a candidate has shown promise in clinical studies.

A manufacturing process, and each unit operation in the process, can be described in general terms as having two types of inputs, controlled and uncontrolled, and each of these can have multiple inputs. The inputs include materials inputs such as inoculum and raw materials, as well as process parameter values such as a pH set point. The process then converts the inputs into outputs. Similarly, outputs include materials such as the product, cells, and contaminants, as well as non-material parameters such as product quality attributes (Figure 4.14a). The range of process inputs, including operating parameters, that gives the highest yield and product quality and results in the "optimal" process outcome is called the design space.

When defining a process design space, information from past process development and knowledge of processes can be called upon; the initial operating range of each process parameter obtained from such knowledge is called the knowledge space (Figure 4.14b). The knowledge

Panel 4.15. Knowledge Management in QbD

- Important part of QRM
- Types of knowledge:
  - Literature and internal documents on products and processes
    - Safety, efficacy from clinical/non-clinical studies
  - Manufacturing experiences with CQAs
- Entails data compilation and results analysis

Panel 4.16. Defining the Design and Control Spaces

- The design space is designed for each unit operation
- Defining spaces is time- and effort-intensive, and is performed only on promising products
- Determines the relationship between controllable inputs (material and parametric) and quality output
- Works from past experience to explore process quality
- Augmented by experimentation to refine the process-quality relationship
- Identifies the "optimal range" as the design space
- Spaces intersect among different CQAs
- Define a control within which routine operations will be performed

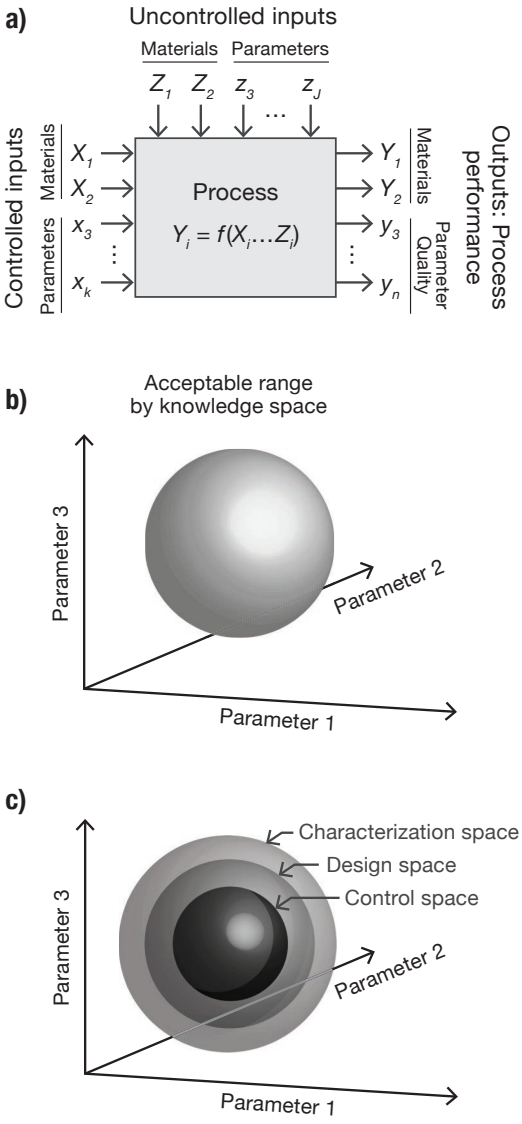

Figure 4.14. (a) Inputs and outputs of a unit operation of process. Some inputs are critical process parameters. (b, c) The knowledge space of critical process parameters facilitates the definition of the design and control spaces. The characterization space is typically larger than the design space.

space is not sufficient to define the design space. In order to map the design space, one has to understand the space beyond its boundary. This space, called the characterization space, spans a wide range of process parameters that reach regions not necessarily optimal for productivity or product quality. Furthermore, the characterization process often includes

a thorough assessment of parameter interactions. Upon completion of characterization, the boundary of the design space is defined such that any combination of inputs within the boundary will produce a product of acceptable quality. The control space, within which manufacturing operations are practiced, is typically within and smaller than the design space, thus assuring optimal process performance. The behavior of CQAs in the design space and control space is carefully studied for individual unit operations of the manufacturing process.

Critical Process Parameters (CPPs)

Design of experiments (DOE) studies are used to identify the criticality of process parameters. A critical process parameter is defined in ICH Q8 as a process parameter whose variability has an impact on a CQA and should therefore be monitored or controlled to ensure that the product has the desired quality (Panel 4.17). A process parameter that does not impact a CQA but may affect process consistency is a non-CPP.

The goal of the evaluation of the impact of process parameters on quality attributes, and the subsequent definition of the design space, is to identify high-risk inputs and include them in the process characterization studies. If necessary, low-resolution screening experiments including several parameters can be first performed to eliminate non-critical inputs from further experimentation.

Panel 4.17. Critical Process Parameters

- CPPs affect CQAs and must be controlled within a range
- The intersection of optimal spaces for different CQAs gives the "optimal" range for each CPP
- The relationship between CPPs and CQAs is used to develop control strategies
- This is unique to each unit operation and the results are linked for overall manufacturing

The design space for a unit operation can be determined by overlaying the permissible regions of each CQA (Figure 4.15). From surface response experiments in DOE studies, the boundary between permissible and impermissible regions can be described by an equation or a model of the process. Next, models for all CQAs are overlaid to identify the region that is permissible to all CQAs. In practice, productivity is also taken into consideration. From such a plot, the normal operating range of the critical process parameters can be determined.[9] Normal operating ranges and acceptable ranges (e.g., action limits) are described in batch records and automation recipes in manufacturing.

Scale-Down Model and Design Space

The experimental characterization of the effect of process parameters on CQAs is carried out in a multiplex manner in a scale-down model (SDM) that is predictive of the manufacturing-scale process (Panel 4.18). It is therefore crucial to understand the capabilities and limitations of the scale-down model for evaluating the process parameters of interest in the manufacturing scale. For the bioreactor step of cell culture process,

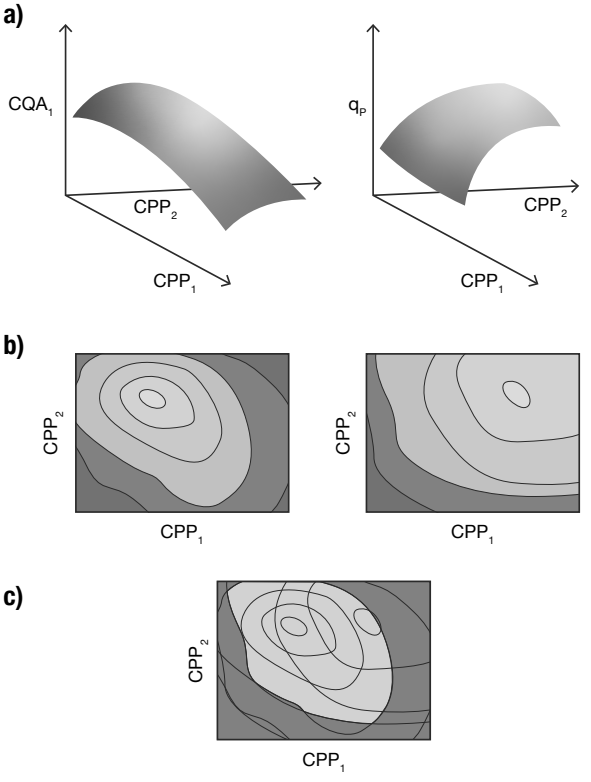

Figure 4.15. (a) Illustration using surface design to superficially evaluate the effect of CPPs on different CQAs and q_p. (b, c) Defining the "optimal" range of CPPs by their intersection.

Panel 4.18. Scale-Down Models for Process Characterization

- Usually carried out in a multiplex manner
- In scale translation, all parameters cannot be kept constant
- Requires a strong understanding of the capability and limitations of scale-down models
- The most important physical parameters are selected to be kept constant between the production scale and the scale-down model
- A successful scale-down model should reproduce the time profiles of growth, key nutrients, and metabolites, and have the same product quality as the manufacturing scale under standard culture conditions
- Test parameters can be varied over a range according to DOE protocol

the comparison of the scale-down and the manufacturing-scale models should include cell growth, metabolism, productivity, and product quality. Longitudinal analysis of the time series data of all runs in the scale-down model and all runs in the manufacturing scale should be performed, and statistically acceptable criteria defined.

It is important to keep in mind that in scale translation all scale-sensitive parameters cannot be kept constant. If one chooses a critical parameter to be constant when comparing different scales, some other parameters will change. For example, if the aeration rate per reactor volume is kept constant, then the superficial gas velocity will be higher in the larger scale. Furthermore, the hydrostatic pressure at the large scale will be higher, as well as the oxygen transfer driving force at the bottom of the reactor. As a result, the capacity of CO_2 stripping and oxygen transfer differs between the scale-down model and the manufacturing bioreactor. The scale-down model study should focus on the critical process parameters and take into account other uncontrolled parameters that

are also affected by scale. In the definition of the design space, other relevant datasets (such as pilot-scale runs) should also be considered.

Process characterization in scale-down models typically employ design of experiments (DOE) principles to investigate the impact of multiple variables on cell growth, metabolism, productivity, and, most importantly, product quality. The information learned from the experiments depends on how the experiments are designed. Therefore, careful analysis of preliminary data is the first step in planning such experiments.

Process characterization in scale-down models for design space mapping can be carried out first in lower-resolution screening studies to explore several process parameters and identify the most impactful. The tradeoff of such screening studies is that the interactions of parameters may be confounded, or primary effects may be confounded with interactions. The most impactful variables are then studied again, in a more extensive DOE design, to establish a model of the effects and interactions of these parameters using response surface designs (Figure 4.15a).[9]

The determination of CPPs and the design space is performed for each unit operation as well as for the entire drug substance and drug product manufacturing process. The final steps involve connecting the design spaces of all unit operations via linking studies to ensure that the overall control strategy is effective. A table can be constructed that lists all the critical data of the CQAs and CPPs involved in each unit operation for better dissimilation of the information (Table 4.7). For product variants, such as glycosylation, deamidation, oxidation, glycation, C-terminal lysine, and N-terminal pyroglutamate, the production bioreactor has the dominant effect. The purification process can do little to modulate these attributes. The hold steps can potentially increase the levels of some variants, such as deamidation. Therefore, CPPs of bioreactor operation must be established to control the bioreactor output for these CQA(s). pH, temperature, duration of culture, feed rates, timing and components, dissolved oxygen levels, medium age, cell age can impact these CQAs. Not all of these may simultaneously be CPPs for a given product. Additionally protein recovery processes can also impact certain attributes, especially those relating to aggregation or fragmentation. For aggregates, fragments, host cell proteins, and host cell DNA, the cell culture process is the primary source of quality variability, but purification steps can remove these product- and process-related impurities. Therefore, CPPs in downstream process can be leveraged, and linkage studies are used to identify critical upstream and downstream process parameters.

Different unit operations in a process are thus interactive. The material input into a unit operation is affected by its upstream step and may require adjustment of its CPPs in order to control its CQAs. Linkage studies to examine the combined effects on CQAs are performed to ensure that the combination of worst-case inputs allowed by the design

Table 4.7. Tabulation of the Effects of Process Steps on Quality Attributes

Product Quality Attribute	*CQA*	*Production bioreactor*	*Protein A chromatography*	*Low pH virus inactivation*	*CEX*	*AEX*
Aggregation	Yes	Form	Removal	Risk to form	Removal	Removal
Deamidated isoforms	No	Form				
Oligosaccharide	Yes	Form				
Process-related impurities						
HCP	Yes	Form	Removal	Removal	Removal	Removal
DNA	No	Form				Removal
Protein A	No		Form		Removal	Removal
Viral Safety	Yes			Inactivation		Clearance

Cont.

Product Quality Attribute	*Nanofiltration for viral particle removal*	*UF/DF*	*Compounding*	*Filtration*	*Filling, stopper, cap*
Aggregation			Risk to form		Risk to form
Deamidated isoforms					
Oligosaccharide					
Process-related impurities					
HCP					
DNA					
Protein A					
Viral Safety	Clearance				

space for each CQA would produce a product that meets its specifications. Worst-case conditions are run in bioreactors in the laboratory or at pilot scale and purified at the worst-case boundary of the design space.

QbD in Different Stages of the Product's Life Cycle

In the past decade, QbD has taken root in the manufacturing of biopharmaceuticals. The road map it established has facilitated streamlining product development and compressing the process development timeline. Key to this success was the adoption of a risk analysis-based QbD approach, the integration of knowledge management, and the utilization of platform processes. Nowadays, process characterization at small scales makes up the bulk of process development data. Modern-day,

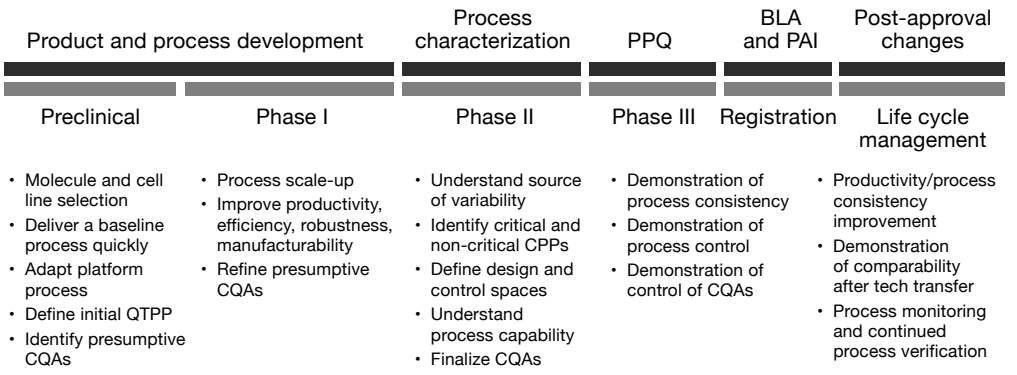

Figure 4.16. QbD-related tasks across the process and product development life cycle.

high-yielding processes result in fewer clinical manufacturing runs, and process performance qualification (PPQ) acceptance criteria are often developed with statistical analyses that include small-scale data.

Figure 4.16 depicts the activities related to QbD during the product's life cycle. In the early stages of product and process development, activities focus on understanding the QTPP of the product and establishing a process to move quickly and safely to clinical trials. As more experience is gained in the process and more knowledge acquired from *in vitro* and *in vivo* studies, understanding of CQAs improves. At the point when, based upon positive clinical trial data, a product is deemed commercialization-ready, the design space mapping studies are completed, critical process parameters are identified, and the control strategy is finalized. The final stage of process development includes Process Performance Qualification (PPQ), which demonstrates that the process is capable of consistently meeting CQA acceptance criteria.

Through the identification of CPPs and their relationship to CQAs, the QbD approach aims to harness the power of Process Analytical Technologies (PAT). The breadth of knowledge generated with multivariate explorations and the development of the design space can be leveraged through use in real-time monitoring and control for enhanced product robustness.

Continued process verification and process changes

During the commercial life cycle of the product, process improvements and technology transfers to new manufacturing facilities may require demonstrations that the product manufactured after such changes is comparable to that produced with the process approved at launch. These types of activities leverage the comprehensive understanding of CQAs and the design space.

The QbD principles used in process and product development are utilized throughout the life cycle of the product. Continued process verification (CPV) is an important continuation of process validation, and generates additional process knowledge beyond the original design space mapping undertaken during process development. CPV can be applied to monitor the robustness of the manufacturing process and allow action to be taken to improve consistency and product quality within the design space.

A CPV may involve identifying trends during process monitoring. Control limits of CQAs are calculated from historical manufacturing data (~30 batches), including the lower control limit (LCL), upper control limit (UCL), and lower and upper specification limits (LSL and USL, respectively) (Figure 4.17). Western Electric rules can be used to determine when the process is trending out of control. For example, when batch data are outside the control limit, six consecutive descending points occur, or nine consecutive points above or below the average occur, action would be taken to move the process back into a state of control. Points that are outside the specification limits are investigated via the Quality System. With continued verification and identification of the cause(s) of incidents which result in points falling outside the specification limit, corrective measures can be taken to narrow the process variability.

The concept of CPV is shown in Figure 4.18. Ideally, batch data is monitored in real time and interfaces with electronic batch records and data aggregation software. Multivariate analysis or data mining of production datasets is then conducted to identify opportunities and propose hypotheses for improvement. Laboratory studies may be conducted to prove these hypotheses via controlled experimentation for possible implementation in manufacturing. Diligence in process monitoring and subsequent action can result in year-over-year improvements to robustness and yield. Importantly, with a QbD-driven process understanding,

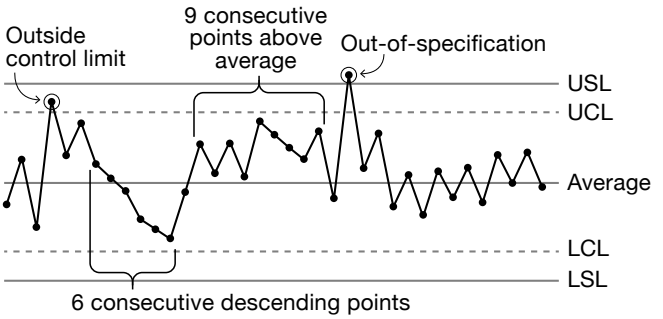

Figure 4.17. Example control chart of CQA records for continual process verification.

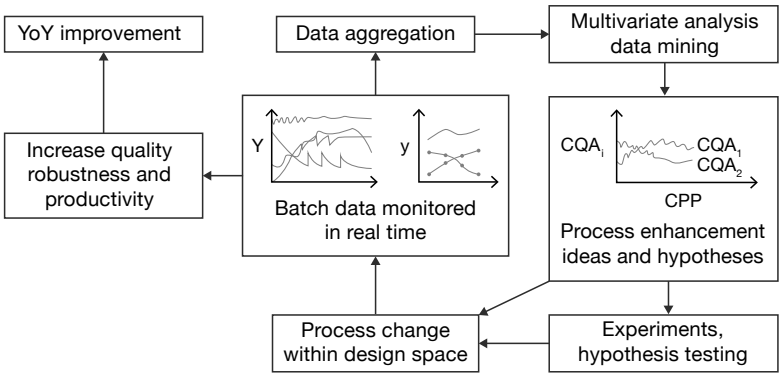

Figure 4.18. Continued process verification provides assurance that the process remains in a state of control.

post-approval process improvements can be performed with parameter values within the design space and potentially without prior health authority approval. This can greatly accelerate the speed at which process robustness adjustments can be implemented in manufacturing. This is fairly straightforward when the proposed changes are within the design space originally defined. However, the industry has been striving to include more complex changes, such as manufacturing site changes, within the scope of a QbD filing. With certain post-approval commitments, this has been possible on occasion, especially when there is sufficient knowledge and experience with legacy products and facilities.

Concluding Remarks

In this chapter we discussed the properties of protein products that contribute to their quality. A sound understanding of these products is important for controlling the product quality during the manufacturing process. For better control of product quality, current product and process development has implemented QbD. The goal of QbD is to reduce product variability, thereby enhancing process development and manufacturing efficiency. This is achieved by designing a robust manufacturing process and establishing clinically relevant specifications. The key elements of QbD include the QTPP, process design and understanding, the control strategy, and continued life cycle improvement.

As the concept of QbD matures, there will be a more consistent approach to QbD and a clearer vision of what its implementation means for all audiences. It is important to understand that QbD is not for reduced regulatory oversight. Rather, it is a broad paradigm of life cycle risk management and process understanding. It provides the opportunity